NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Click to go to the link. ALT + LEFT ARROW to return

Purple text represents the responses from measure developers.

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 0166

Corresponding Measures:

De.2. Measure Title: HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems) Survey

Co.1.1. Measure Steward: Centers for Medicare & Medicaid Services

De.3. Brief Description of Measure: HCAHPS (NQF #0166) is a 29-item survey instrument that produces 10 publicly reported measures:

6 multi-item measures (communication with doctors, communication with nurses, responsiveness of hospital staff, communication about medicines, discharge information and care transition); and

4 single-item measures (cleanliness of the hospital environment, quietness of the hospital environment, overall rating of the hospital, and recommendation of hospital).

Note: The HCAHPS Survey originally included three items about pain which formed a composite measure, Pain Management. CMS discontinued publicly reporting this measure in July 2018. In January 2018, CMS replaced the original HCAHPS pain items with three items that asked about communication about pain. In compliance with the Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act (Pub. L. 115-271) of 2018 (Section 6104), CMS will remove the new communication about pain items from the HCAHPS Survey beginning with October 2019 discharges.

1b.1. Developer Rationale: The HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems) Survey is the first national, standardized, publicly reported survey of patients´ perspectives of hospital care. HCAHPS (pronounced "H-caps"), also known as the CAHPS® Hospital Survey*, is a 29-item survey instrument and data collection methodology for measuring patients' perceptions of their hospital experience. While many hospitals have collected information on patient satisfaction for their own internal use, until HCAHPS there were no common metrics and no national standards for collecting and publicly reporting information about patient experience of care. Since 2008, HCAHPS has allowed valid comparisons to be made across hospitals locally, regionally and nationally.

Three broad goals have shaped HCAHPS. First, the standardized survey and implementation protocol produce data that allow objective and meaningful comparisons of hospitals on topics that are important to consumers. Second, public reporting of HCAHPS results creates new incentives for hospitals to improve quality of care. Third, public reporting enhances accountability in health care by increasing transparency of the quality of hospital care provided in return for the public investment. With these goals in mind, the Centers for Medicare

& Medicaid Services (CMS) and the HCAHPS Project Team have taken substantial steps to assure that the survey is credible, practical and actionable.

S.4. Numerator Statement: The HCAHPS Survey asks recently discharged patients about aspects of their hospital experience that they are uniquely suited to address. The core of the survey contains 19 items that ask "how often" or whether patients experienced a critical aspect of hospital care, rather than whether they were "satisfied" with their care. Also included in the survey are three screener items that direct patients to relevant questions, five items to adjust for the mix of patients across hospitals, and two items (race and ethnicity) that support Congressionally-mandated reports. Hospitals may include additional questions after the core HCAHPS items.

For full details, see the current HCAHPS Quality Assurance Guidelines, V.14.0, pp. 57-65, under the "Quality Assurance" button on the official HCAHPS On-Line Web site at:

https://www.hcahpsonline.org/globalassets/hcahps/quality-assurance/2019_qag_v14.0.pdf

S.6. Denominator Statement: The target population for HCAHPS measures include eligible adult inpatients of all payer types who completed a survey. HCAHPS patient eligibility and exclusions are defined in detail in the sections that follow. A survey is defined as completed if the patient responded to at least 50% of questions applicable to all patients.

S.8. Denominator Exclusions: There are a few categories of otherwise eligible patients who are excluded from the HCAHPS sample frame. As detailed below in sec S.9, these exclusions include patients excluded due to state regulations, no-publicity patients, and specific groups of patients with an admission source or discharge status that results in difficulty collecting patient experience data through a survey instrument.

De.1. Measure Type: Outcome

S.17. Data Source: Instrument-Based Data

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Aug 3, 2005 **Most Recent Endorsement Date:** Jan 07, 2015

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a health outcome measure include providing empirical data that demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service; if these data not available, data demonstrating wide variation in performance, assuming the data are from a robust number of providers and results are not subject to systematic bias. For measures derived

from patient report, evidence also should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

Evidence Summary

- Brief background: This is a PRO-PM of patient experience of care evaluating time spent in a hospital setting
- Logic model of measure maps HCAHPS measure domains to structures, processes and outcomes of hospital care based on patient perception and experience
- Evidence suggesting patient value and meaningfulness include:
    - Solicitation of patient feedback in the development of the instrument
    - Focus group testing of inpatient hospital participants, who indicated that they would consider changing hospitals in response to comparisons of HCAHPS scores
    - Independent patient expressions of values, preferences and needs for inpatient care aligning with survey domains
    - Multiple patient focus group confirmations also cited
    - Patients relying on HCAHPS scores over word of mouth reports
- Evidence demonstrating relationship between outcome and healthcare structure, process, intervention or service include:
    - HCAHPS improvement year over year, especially amongst initially low performing hospitals
    - Cultural competency improvement efforts leading to HCAHPS score improvement
    - Developer cites four studies where hospital managers share best practices to improve HCAHPS scores
    - Developer cites AHRQ guides in improvement of patient experience of care

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- Developer included a new logic model outlining explanation of structures, processes and outcomes of hospital care, with a new diagram
- Developer updated "Value and Meaningfulness" section with more complete description of patient input through literature review, one-on-one meetings and focus groups
- Developer provided updated examples of empirical data demonstrating the relationship between services and interventions that improve HCAHPS performance

Questions for the Committee:

- Is there at least one thing that the provider can do to achieve a change in the measure results?
- If derived from patient report, does the target population value the measured outcome and finds it meaningful?
- The evidence provided by the developer is updated, directionally the same, and stronger compared to that for the previous NQF review. Does the Committee agree there is no need for repeat discussion and vote on Evidence?

Guidance from the Evidence Algorithm

Measure assesses performance on a patient-reported outcome (Box 1) → Empirical data suggest a structure, process, intervention or service may improve measure performance (Box 2) → PASS

Preliminary rating for evidence: ☒ Pass ☐ No Pass

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- Developer provides data gap analysis of 4,300 hospitals by measure domain, reporting means between 52.36 – 82.05, and standard deviations between 4.74 – 10.72

Disparities

- Developer provides analysis of performance variation based on race and on the 7 different language offerings of HCAHPS to provide insights into disparities. Generally, non-English-preferring Black, Hispanic, API and AI/AN patients reported worse experiences than their English-preferring counterparts, except for Russian-preferring White patients.

Question for the Committee:

- Does the Committee agree that there is an ample performance gap and disparities data presented?

Preliminary rating for opportunity for improvement: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence: *For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure.*

- Developer included a new logic model outlining explanation of structures, processes and outcomes of hospital care, with a new diagram
- Developer updated "Value and Meaningfulness" section with more complete description of patient input through literature review, one-on-one meetings and focus groups
- Developer provided updated examples of empirical data demonstrating the relationship between services and interventions that improve HCAHPS performance
- The developer provided updated evidence for the measure and it does appear that the target population finds the measure to be meaningful.

1b. Performance Gap: *Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?*

- Developer provides data gap analysis of 4,300 hospitals by measure domain, reporting means between 52.36 – 82.05, and standard deviations between 4.74 – 10.72 Disparities • Developer provides analysis of performance variation based on race and on the 7 different language offerings of HCAHPS to provide insights into disparities. Generally, non-English-preferring Black, Hispanic, API and AI/AN patients reported worse experiences than their English-preferring counterparts, except for Russian-preferring White patients.
- There is a gap analysis provided by the developer and I do think that clinicians and facilities can use the data to improve their processes.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications **and** Testing

2b. Validity: Testing**;** Exclusions**;** Risk-Adjustment**;** Meaningful Differences**;** Comparability**;** Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Complex measure evaluated by Scientific Methods Panel? ☒ **Yes** ☐ **No**

Evaluators: NQF Scientific Methods Panel

Methods Panel Review (Combined)

Evaluation of Reliability and Validity (and composite construction, if applicable):

- This measure was reviewed by the Scientific Method's Panel and passed with high reliability and validity ratings.
- **Reliability**
  - Testing included score-level and data element testing
  - SMP described the methods used as follows: "Cronbach alpha was used to evaluate the reliability of the composite measures. An ICC and signal-to-noise ratio was used to estimate hospital-level reliability, these are acceptable for evaluating reliability (precision) of hospital scores using the top box approach to scoring." Further, "The results of measure score reliability testing were in general good with 300 surveys per hospital. Hospital level item specific reliabilities were also very good in both top-box score and linear mean score forms."
  - Developer reported top-box scores by domain. Hospital-level reliabilities of 10 HCAHPS measure mean scores ranged from 0.83 to 0.93. All 10 exceeded the threshold of 0.80 and 9 out of 10 (all but Discharge Information) exceeded the very good/0.85 standard.
  - Internal consistency reliability coefficients (Cronbach's alpha) were presented for each of the six multi-item measures. Three of six multi-item measures had internal consistency reliability estimates of 0.80 or higher (Communication with Nurses, Communication with doctors) and three had estimates of 0.68-0.69 (Responsiveness of Hospital Staff, Communication about Medicines, Discharge Information).
  - SMP concluded that the measure has a "uniformly high hospital-level reliability estimates for each measure, and this set of measures should be considered reliable (precise)."
  - Note: Measure developer presented results of Cronbach's alpha testing for multiitem domains of the survey but did not perform any data element level reliability testing for the four single-item domains.

- **Validity**
    - Testing included score-level and data element testing
    - SMP members described the approach as follows: "Both item-level top-box scores and composite scores were correlated with the global rating of provider at patient and hospital level. Hospital-level factor analysis was conducted to identify underlying factors. The developer also compared possible hospital-level composite item groupings to the composites found in the individual-level factor analysis. These analyses were done appropriately and thoughtfully."
    - Generally, the SMP agrees with submitter's conclusion that "the pilot study analyses provide support for the construct validity of the HCAHPS, confirming the intended factor structure. A 2016 discriminant validity analysis confirmed the factor structure. Both the 2003 pilot study and the 2016 analyses found strong correlation of each of the multi-item measures and single-item measures with the two overall measures."
    - One SMP member noted that: "I didn't understand why hospice, nursing home, etc., patients are excluded for evaluation of hospital stay. It seems this would introduce bias into the denominator that would then not be detected in performance measure computation. What is a given hospital disproportionately discharges to nursing home, hospice, etc.?"
    - For risk adjustment: "The approach is generally acceptable, and the developers should be given credit for a careful analysis of the potential effects of race/ethnicity on measure scores. Education and preferred language are included in the adjustment model. Race/ethnicity and SES are not, but the rationale is acceptable."

Questions for the Committee regarding reliability:

- The Scientific Methods Panel is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The Scientific Methods Panel is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

Preliminary rating for reliability:	☒ **High**	☐ **Moderate**	☐ **Low**	☐ **Insufficient**
Preliminary rating for validity:	☒ **High**	☐ **Moderate**	☐ **Low**	☐ **Insufficient**

Measure Number: 0166

Measure Title: HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems) Survey

Type of measure:

☐☐ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐☐ **Outcome** ☒☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**

☐ **Assessment Data** ☐ **Paper Medical Records** ☒☐ **Instrument-Based Data** ☐ **Registry Data**

☐ **Enrollment Data** ☐☒ **Other: Patient Survey**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**

☐ **Population: Community, County or City** ☐ **Population: Regional and State**

☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**

   Reviewer A: No concerns.

   Reviewer B: None

   Reviewer C: None

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☒ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical <u>VALIDITY</u> testing** of <u>patient-level data</u> conducted?

   ☐ **Yes** ☐ **No** NA

6. **Assess the method(s) used for reliability testing**

Reviewer B: High reliability using Spearman-Brown reliability method for 300 completed surveys demonstrated very good to excellent reliability. Internal Consistency Reliability of Multi-item measures using Cronbach's alpha showed sufficient reliability

Reviewer F:

a. Methods were appropriate (Cronbach's alpha, inter-abstractor reliability)

Submission document: Testing attachment, section 2a2.2

Reviewer A: Testing methods are appropriate

REVIEWER D: The method used for assessing measure score reliability was appropriate, reporting both ICC and hospital-level reliability.

In addition, hospital-level reliabilities of the HCAHPS survey items that are combined into composite measures were also appropriately evaluated in two forms: top-box scores and linear mean scores.

Reviewer C: Cronbach alpha was used to evaluate the reliability of the composite measures. An ICC and STN was used to estimate hospital-level reliability, these are acceptable for evaluating reliability (precision) of hospital scores using the top box approach to scoring.

Reviewer E: The methods used for reliability testing were generally acceptable, using standard and well-accepted methods, at both data element and measure score levels.

7. **Assess the results of reliability testing**

Reviewer B: All analysis demonstrated reliability

Reviewer F:

a. Hospital-level inter-abstractor reliabilities of the 10 HCAHPS measure top-box scores ranged from 0.83 (Communication about Medicines and Discharge Information) to 0.93 (Quietness and Recommend Hospital), exceeding the threshold for good reliability. Hospital level (Spearman-Brown) reliabilities ranged from 0.74-0.91.

b. Internal consistency of the various scales (some of which were 2-items) never fell below 0.65, so acceptable to good at the data element level.

Submission document: Testing attachment, section 2a2.3

Reviewer A: Test sample was adequate to generalize for widespread implementation. High confidence that measure results are reliable with 300 completed surveys using top-box and linear-mean scoring.

REVIEWER D: The results of measure score reliability testing were in general good with 300 surveys per hospital.

Hospital level item specific reliabilities were also very good in both top-box score and linear mean score forms.

Reviewer C: The hospital level reliability estimates were acceptable for each measure.

Reviewer E: The results of reliability testing were acceptable, although the results presented for measure score reliability conveyed a bit of a mixed message. One test, referred to as ICC, presented results on a scale that is not the usual one for ICC. Results of this method of testing generally fell below a declared threshold of .05 needed to identify meaningful differences among practices. On the basis of that one test, the results at the measure score level would seem to be unreliable. But, another metric of "reliability" was presented generated by the Spearman-Brown formula applied to the ICC values, indicating that results at the hospital level were sufficiently reliable. The first test apparently indicates the percent of overall variance accounted for by true differences in performance rather than measurement error, with findings for most measures in the set falling below a 5% threshold. If this interpretation is correct, then less than 5% of the variance in observed scores is accounted for by some meaningful underlying performance dimension. With sufficient sample sizes, though, even this very weak "signal" can be reliable.

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   Submission document: Testing attachment, section 2a2.2

   ☐☒ **Yes**

   ☒☐ **No**

   ☐ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

    ☐☒ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒☐ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    Reviewer A: Well documented testing process and test results. Survey is sufficiently reliable at the recommended 300 completed surveys.

    Reviewer B: No concerns, as noted all scores demonstrated high reliability

    REVIEWER D: Testing results of both data elements and measure scores reliability were very good.

    Reviewer C: Due to the uniformly high hospital-level reliability estimates for each measure, this set of measures should be considered reliable (precise).

    Reviewer E: See response to item 7 above.

    Reviewer F:

    a. **All measures met or exceeded minimum reliability standards.**

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Reviewer F:

    there are a few categories of otherwise eligible patients who are excluded from the sample frame:

    "No-Publicity" patients who request that they not be contacted; Court/Law enforcement patients; those with a foreign home address; those discharged to hospice care (home or facility); those excluded because of state regulations; those discharged to nursing homes and skilled nursing facilities

    I didn't understand why hospice, nursing home, etc., patients are excluded for evaluation of hospital stay. It seems this would introduce bias into the denominator that would then not be detected in performance measure computation. What is a given hospital disproportionately discharges to nursing home, hospice, etc.?

Submission document: Testing attachment, section 2b2.

Reviewer A: None.

No concern

Reviewer E: None.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

Reviewer B: None, all analysis using prior survey results demonstrated differences in hospital scores

Submission document: Testing attachment, section 2b4.

Reviewer A: None.

No concern.

Reviewer C: None, this was appropriately evaluated.

Reviewer E: The developers find that a substantial number of practices are either significantly above or below the national mean in scores on essentially all the measures derived from this survey. They have not been able, or not attempted to, show that observed differences are meaningful.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**

Reviewer B: None
Submission document: Testing attachment, section 2b5.
Reviewer A: No concerns.
REVIEWER D: No concern.
Reviewer C: None.
Reviewer E: N/A

15. **Please describe any concerns you have regarding missing data.**

Reviewer B: Agree with submitters that the patient-mix adjustment model accounted for any bias in missing survey responses

Reviewer F:

a. Same as above with regard to exclusions

Submission document: Testing attachment, section 2b6.

Reviewer A: No concerns.

REVIEWER D: No concern.

Reviewer E: None.

16. **Risk Adjustment**

16a. **Risk-adjustment method** ☐ **None** ☒ **Statistical model** *Patient-mix adjustment via linear regression and survey mode adjustment based on randomized mode experiments.* ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒☐ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☒ Yes ☐ No ☐ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☒ Yes ☐ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☒ Yes ☐ No

16d. **Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☒ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion?
☐☒ Yes ☐ No NA

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☒ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration)
☒ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☒☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

Reviewer A: Appropriate.

Reviewer B: Robust patient mix adjustment model addresses critical aspects of risk adjustment for the population

Reviewer C: The risk adjustment approach is adequate**.**

REVIEWER D: Risk adjustment approach was appropriate.

Reviewer E: The approach is generally acceptable, and the developers should be given credit for a careful analysis of the potential effects of race/ethnicity on measure scores**.** Education and preferred language are included in the adjustment model. Race/ethnicity and SES are not, but the rationale is acceptable.

Reviewer F:

Patient-mix adjustment via linear regression and survey mode adjustment based on randomized mode experiments. My understanding is that social and other risk adjustments are not required but voluntary at the provider/site level and are applied on the back-end by CMS. The clinical factors in the HCAHPS patient-mix adjustment model include patient's age, gender by service line, and self-reported overall health. The social risk factors in the HCAHPS patient-mix adjustment model include patient's self-reported education and primary language spoken at home. The submitter did not apply any ordering when including risk factors in the patient-mix adjustment model; all patient-mix adjustment factors were entered simultaneously.

VALIDITY: TESTING

17. **Validity testing level: ☒☐ Measure score ☒☐ Data element ☐☒ Both**
18. **Method of establishing validity of the measure score:**
    - ☐ **Face validity**
    - ☒ **Empirical validity testing of the measure score**
    - ☐ **N/A (score-level testing not conducted)**
19. **Assess the method(s) for establishing validity**
20. **Reviewer A**: Appropriate.

    Reviewer B: Discriminant Validity Analysis: Patient-Level Average Inter-Item Correlations among HCAHPS Multi-item Measures, July 2016 – June 2017 discharges and construct validity Using the 2003 Three-State Pilot Study data

    Submission document: Testing attachment, section 2b2.2

    REVIEWER D: Both item-level top-box scores and composite scores were correlated with the global rating of provider at patient and hospital level.

    Hospital-level factor analysis was conducted to identify underlying factors. The developer also compared possible hospital-level composite item groupings to the composites found in the individual-level factor analysis. These analyses were done appropriately and thoughtfully.

Reviewer C: The authors focused on construct validity using confirmatory factor analysis and correlational analyses. While this approach is unorthodox for NQF validity testing, it is reasonable to perform this approach for patient reported measures of hospital experience.

Reviewer E: The developers rely on correlations among measures in the survey to establish measure score-level validity – a modest level of correlation (neither too high nor too low) is viewed as acceptable evidence of validity. There is no evidence presented linking measure scores to any independent measure of quality of care at the clinic level.

21. **Assess the results(s) for establishing validity**

Reviewer B: Both construct and discriminant analysis supported validity

Submission document: Testing attachment, section 2b2.3

Reviewer A: Test sample adequate and results demonstrate sufficient validity.

Reviewer C: The results supported the construct validity for the measures.

REVIEWER D: Validity test results at both item and score level were mostly very positive.

Reviewer E: Results are generally acceptable, showing moderate correlations among scores and between specific domain scores and overall ratings of care.

Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?** *NOTE that data element validation from the literature is acceptable.*

Reviewer F: Yes, described in great detail with supplemental text and reference citations.

Domains of care (note pain management is being removed from measure)

Communication with doctors (Q5, Q6, & Q7)

Communication with nurses (Q1, Q2, & Q3)

Responsiveness of the hospital staff (Q4, Q10, & Q11)

Communication about medicines (Q12, Q13, & Q14)

Cleanliness and quiet of physical environment (Q8 & Q9)

Discharge information (Q15, Q16, & Q17)

Submission document*: Testing attachment, section 2b1.*

☐☒ **Yes**

☒☐ **No**

☐ **Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

☐☒ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒☐ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

Reviewer B: Agree with submitters conclusion that "the pilot study analyses provide support for the construct validity of the HCAHPS, confirming the intended factor structure. A 2016 discriminant validity analysis confirmed the factor structure. Both the 2003 pilot study and the 2016 analyses found strong correlation of each of the multi-item measures and single-item measures with the two overall measures."

REVIEWER D: Extensive tests were completed, and most results were very positive.

Reviewer C: Construct validity is weaker than other forms of validity, but the findings presented support continued use of these measures.

Reviewer E: The validity of this version of CAHPS and other versions rests largely on assessments of face validity. There is no information presented linking the CAHPS scores to any separate, independent measure of quality of care at the hospital level. The patterns of correlations do demonstrate adequate validity of the measure at the individual patient or data element level and do provide weak evidence for validity at the measure score level.

Reviewer F: Extensive data supporting performance of the measures in expected directions, as predicted.

ADDITIONAL RECOMMENDATIONS

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**

Reviewer E: This set of measures, like all the other CAHPS measure sets, claims to be a set of outcome measures. These are not outcome measures. They do not reflect the state of a patient after treatment; they use the patient report to provide data on care processes. These are process measures, not outcome measures, even though the data come from patient surveys. A satisfaction survey would be an outcome measure, but these are "experience of care" surveys using the patient as a data source about care processes. Since users like CMS make distinctions in their P4P programs between process and outcome measures, often assigning greater weight to outcome measures, this is a very important distinction and the NQF endorsement process should make clear that these are not outcome measures.

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Specifications: *Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?*

- No concerns are noted.
- No concerns about reliability specifications.

2a2. Reliability testing: *Do you have any concerns about the reliability of the measure?*

- No concerns are noted.
- No concerns about reliability testing at this point.

2b2. Validity testing: *Do you have any concerns with the testing results?*

- No concerns are noted.
- No concerns.

Validity- Threats to Validity: *Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data). 2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?*

- The developers find that a substantial number of practices are either significantly above or below the national mean in scores on essentially all the measures derived from this survey. They have not been able, or not attempted to, show that observed differences are meaningful.
- No additional comments

Other Threats to Validity: *Other Threats to Validity (Exclusions, Risk Adjustment). 2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure? 2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?*

- Yes, acceptable results are indicated and a risk adjustment strategy is included in the measure. It is noteworthy that missing data limits the ability to characterize the respondents relative to education and self-rated health when assessing their health care experience. The developers state, "while there was evidence of differential nonresponse overall, and evidence that those with lower response propensity had less positive evaluations of care, there was no evidence that nonresponse weighting based on available data improved the accuracy of hospital scores beyond what could be achieved with PMA (patient mix adjustment)."
- No concerns.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- The HCAHPS Survey is administered in four modes: mail, telephone, mixed (mail with telephone follow-up), or Interactive Voice Response
- Measure developer describes only that the HCAHPS Survey implementation, data submission and oversight "is improved on an annual basis"
- Developer did not address the time and costs of survey administration, which is performed by CMS-approved vendors at the expense of the measured organization

Questions for the Committee:

- Does the Committee feel that the measure developer has provided sufficient information to determine how feasible the survey is to administer?

Preliminary rating for feasibility: ☐ **High** ☐ **Moderate** ☒ **Low** ☐ **Insufficient**

RATIONALE:

- A low rating is assigned because measure developer has not evaluated the burden on plans associated with measure implementation in the form of fees from retention of an approved CAHPS vendor to administer the surveys.

- Based on the information submitted there is low confidence or certainty that the criterion is met.
- **Note:** this is not a must pass criteria per NQF's current rules.

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility: *Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?*

- The HCAHPS Survey is administered in four modes: mail, telephone, mixed (mail with telephone follow-up), or Interactive Voice Response. Developer did not respond to time and cost burdens for administering the survey.
- This is where my biggest concerns lie - this measure is simply too burdensome for patients, facilities, and the clinicians/staff to administer, interpret, and then act upon. I agree with the other reviewer that it is not actually an outcome measure, but rather speaks to processes within a facility.

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No** ☐ **UNCLEAR**

OR

Planned use in an accountability program? ☐ **Yes** ☐ **No**

Accountability program details

- Public reporting of HCAHPS in Hospital Compare, Hospital Inpatient Quality Reporting Program
- HCAHPS used for payment in Hospital Value-Based Purchasing

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured, and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- All hospitals that participate in public reporting on Hospital Compare (~4,500) receive a Preview Report prior to each quarterly public reporting that contains all of their HCAHPS scores and HCAHPS information for them to view prior to the data being publicly reported.

- All IPPS hospitals that participate in Hospital Value-Based Purchasing (~3,000) receive additional reports on an annual basis from CMS that contain their scores on the HCAHPS domain used in the Hospital VBP for pay-for-performance program.
- Reports include information on how to interpret HCAHPS scores.
- Feedback on HCAHPS is regularly collect from survey vendors and self-administering hospitals via conference calls and site visits. CMS meets with provider groups, hospital associations and patient advocacy groups to hear and address HCAHPS concerns. Comments are also accepted through the federal rulemaking process. Developer cites other feedback mechanisms as well.

Additional Feedback:

***Question for the Committee**:*

- Is there anything that the Committee wishes to discuss related to the current use of the measure?

Preliminary rating for Use: ☒ Pass ☐ No Pass

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- Developer cites a study showing a 2.8% improvement in national HCAHPS scores between 2007 and 2011.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

Potential harms

- Allegation that HCAHPS pain management questions created pressure for overprescribing opioids
- While not able to empirically establish this as the case, CMS took the precaution of removing the pain management domain beginning in October of 2019 in response to a mandate from the SUPPORT Act.

Additional Feedback:

N/A

***Question for the Committee**:*

- Do the benefits of the measure outweigh any potential unintended consequences?

Preliminary rating for Usability and use: ☐ High ☒ Moderate ☐ Low ☐ Insufficient

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a. Use: *4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided? 4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users*

been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?

- Public reporting of HCAHPS in Hospital Compare, Hospital Inpatient Quality Reporting Program
- HCAHPS used for payment in Hospital Value-Based Purchasing
- All hospitals that participate in public reporting on Hospital Compare (~4,500) receive a Preview Report prior to each quarterly public reporting that contains all of their HCAHPS scores and HCAHPS information for them to view prior to the data being publicly reported.
- All IPPS hospitals that participate in Hospital Value-Based Purchasing (~3,000) receive additional reports on an annual basis from CMS that contain their scores on the HCAHPS domain used in the Hospital VBP for pay-for-performance program.
- Reports include information on how to interpret HCAHPS scores.
- Feedback on HCAHPS is regularly collect from survey vendors and self-administering hospitals via conference calls and site visits. CMS meets with provider groups, hospital associations and patient advocacy groups to hear and address HCAHPS concerns. Comments are also accepted through the federal rulemaking process. Developer cites other feedback mechanisms as well.
- This measure is in use. However, the ongoing feedback from users about the burden that this measure entails has not been adequately addressed by the developer.

4b. Usability: *4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations? 4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.*

- Developer cites a study showing a 2.8% improvement in national HCAHPS scores between 2007 and 2011. Unexpected findings (positive or negative) during implementation Potential harms
- Allegation that HCAHPS pain management questions created pressure for overprescribing opioids
- While not able to empirically establish this as the case, CMS took the precaution of removing the pain management domain beginning in October of 2019 in response to a mandate from the SUPPORT Act.
- The measure does seem to offer the opportunity for the user to improve care, but the information can be difficult and burdensome to collect and act upon by the facility.

Criterion 5: Related and Competing Measures

Related or competing measures

- The following measures are all related, though not necessarily competing:
    - NQF 0005 CAHPS Clinician and Group Surveys V3.0
    - NQF 0006 CAHPS Health Plan Survey V5.0
    - NQF 0166 Hospital CAHPS Survey
    - NQF 0258 CAHPS In-Center Hemodialysis Survey
    - NQF 0517 CAHPS Home Health Care Survey
    - NQF 1741 CAHPS Surgical Care Survey
    - NQF 2548 Child Hospital CAHPS Survey
    - NQF 2967 CAHPS Home- and Community-Based Services Survey

Harmonization

N/A

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

Related and Competing: *Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?*

- Related or competing measures The following measures are all related, though not necessarily competing:

NQF 0005 CAHPS Clinician and Group Surveys V3.0
NQF 0006 CAHPS Health Plan Survey V5.0 o NQF 0166 Hospital CAHPS Survey
NQF 0258 CAHPS In-Center Hemodialysis Survey
NQF 0517 CAHPS Home Health Care Survey
NQF 1741 CAHPS Surgical Care Survey
NQF 2548 Child Hospital CAHPS Survey
NQF 2967 CAHPS Home- and Community-Based Services Survey Harmonization N/A

- There are a number of related CAHPS measures for other settings.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: June/13/2019

- **No NQF members have submitted support/non-support choices as of this date**

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

HCAHPS_NQF_-0166-_4-24-19_--_NQF_evidence_attachment_Sep2017_-4-.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

Yes

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:**

Measure Title:

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here:

Date of Submission:

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☒ Outcome:

PRO

☒ Patient-reported outcome (PRO):

Experience with care

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

NARRATIVE EXPLANATION of HCAHPS Logic:

The boxes represent structures, processes and outcomes of hospital care. Ovals represent HCAHPS measure domains; please see below.

On the left side of the model are three boxes that represent major structural determinants:

Adequacy of Supply and Skills of Staff;

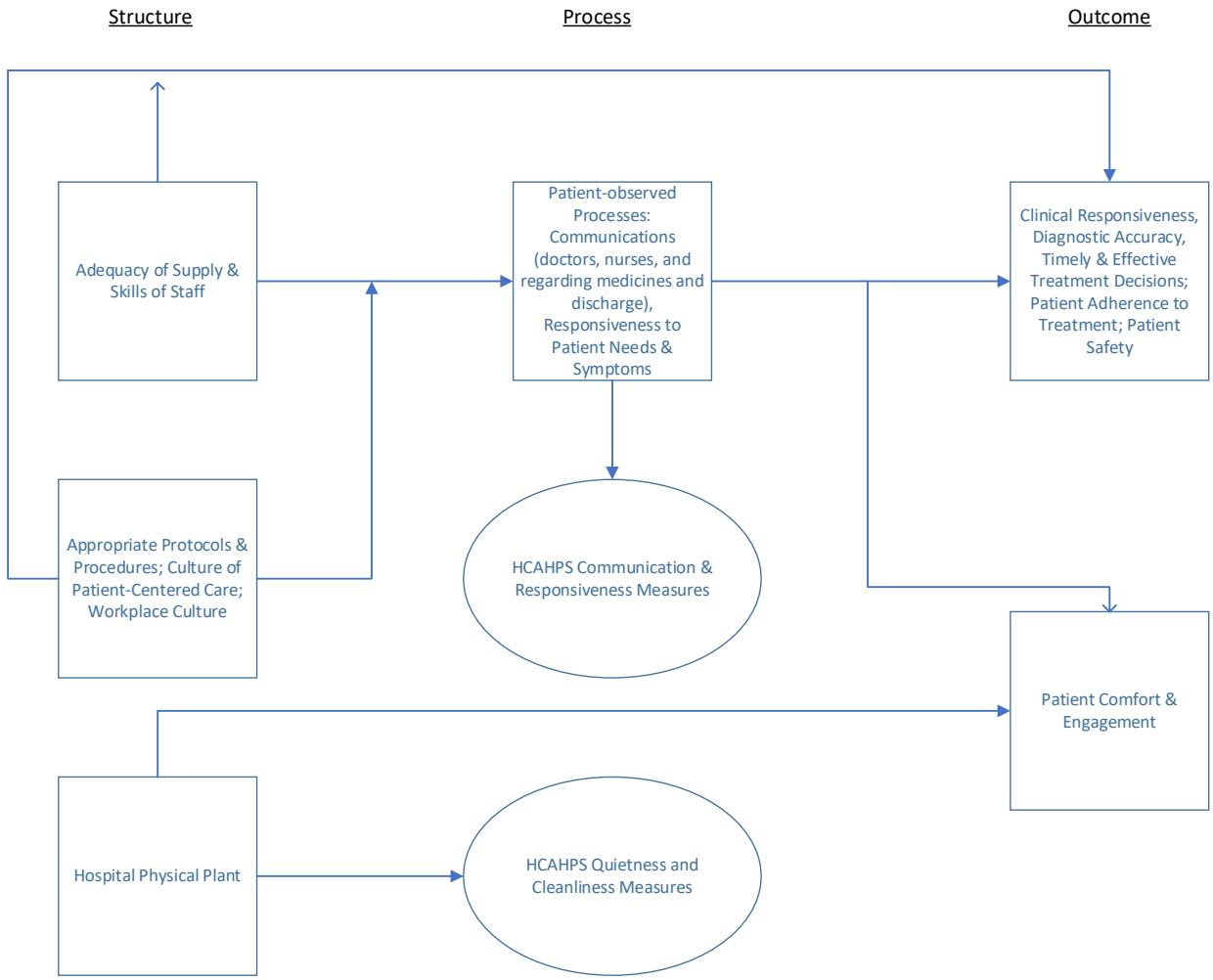
Appropriate Protocols & Procedures; Culture of Patient-Centered Care; Workplace Culture; and

Hospital Physical Plant.

The central box characterizes communications and responsiveness of hospital staff as the generalized processes that are most visible to patients; hence these processes are linked directly to the corresponding HCAHPS measures: HCAHPS Communication & Responsiveness Measures, and HCAHPS Quietness and Cleanliness Measures. (While the match of processes to measures is evident, for legibility we do not break them out into separate boxes and arrows).

On the right side of the graphic we see the outcomes that the structures and processes promote, both clinical outcomes (Clinical Responsiveness, Diagnostic Accuracy, Timely & Effective Treatment Decisions; Patient Adherence to Treatment; Patient Safety) and patient experience as an outcome of value in itself (Patient Comfort & Engagement).

Note that the arrows passing through central box (Communication, Responsiveness etc.) indicate that the effect of the structure on the left on the outcome to the right is partially mediated through the processes in the center. The structure may also have an additional effect on outcomes passing through processes not visible to the patient, such as execution of procedures with technical skill, as indicated by the arrows from structure (Hospital Physical Plant) to outcome (Patient Comfort & Engagement) not passing through the central box. Nonetheless, if some of the structural characteristics (e.g. adequate staffing) measured indirectly through HCAHPS process measures also have effects through unmeasured process pathways, the measures gain broader significance.



1a.3 Value and Meaningfulness: **IF** this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

Each stage of HCAHPS development process has considered the value of the HCAHPS items to patients. The development of the questionnaire included a process to solicit input on the content of the questionnaire and the methods for sampling, data collection, and analysis. Input from stakeholders was garnered through a variety of venues, including a literature review, one-on-one meetings, and focus groups with target consumers (Darby et al. 2005). In focus groups of recent hospital inpatients, participants reported that a very high proportion of the items being considered for the CAHPS Hospital Survey to be so important they would consider changing hospitals in response to information about them (Sofaer et al. 2005). Edgman-Levitan and Cleary (1996) asked patients what was important to them and what affects them. Key dimensions patients identified were respect for patients' values, preferences, and expressed needs; coordination of care; information, communication, and education; physical comfort and pain management; emotional support and alleviation of fear and anxiety; involvement of family and friends; and transition and continuity to the home or community. Multiple focus groups conducted with Department of Veterans Affairs (VA) patients from different geographic regions confirmed these dimensions of quality as being most important to patients (Edgman-Levitan and Cleary 1996). During the course of its development, CMS provided three opportunities for the public to comment on HCAHPS, resulting in well over one thousand comments (CMS.gov). This process resulted in an initial HCAHPS survey.

The second phase involved streamlining the survey. To determine which items to keep in the shortened CAHPS hospital questionnaire, we examined their importance from three perspectives: the degree to which they were

indicators of the composite; their relationships to patients' overall evaluation of the hospital; and their relative ranking, according to what patients and their loved ones told us in focus groups (Keller et al. 2005).

The results of focus groups are valuable because certain themes emerge consistently. But they are limited in part because they cannot provide information about how generalizable the results are. Results from survey data can address this limitation. CAHPS measures, developed to complement more technical quality measures, are measures for which the patients are the best or only source of information and/or perspective, such as the degree to which patients felt that their care was patient-centered (Anhang Price et al. 2014). Several studies provide evidence that patients value the CAHPS measures and find them meaningful. For example, Safran et al. (2001) found that patients who reported the poorest-quality relationships with their physicians were three times more likely to voluntarily leave the physicians' practice than patients with the highest-quality relationships. Patients also use information from patient experience measures to make decisions about their healthcare providers and plans. One study found that seeing publicly reported quality information was a determinant of choosing higher quality-rated health plans, although the weight given to quality information also depended on other features, such as cost and provider choice (Faber et al., 2009).

Consumers may weigh HCAHPS scores more in their hospital choice decision than other sources of quality information, demonstrating the value of this information to consumers. An experimental study comparing the effects of HCAHPS scores and word-of-mouth narratives on consumer's hospital choice (Huppertz and Carlson, 2010) showed that when HCAHPS scores and word-of-mouth narratives reinforced one another, they had strong effects on consumers hospital choices; however, when HCAHPS scores and word-of-mouth narratives were inconsistent, patients more heavily weighted information from HCAHPS scores in their hospital choice decision.

References

Anhang Price R, Elliott MN, Zaslavsky AM, Hays RD, Lehrman WG, Rybowski L, Edgman-Levitan S, Cleary PD. (2014) Examining the role of patient experience surveys in measuring health care quality. Med Care Res Rev 71 (5):522-54.

CMS (Centers for Medicare & Medicaid Services). https://www.cms.gov/Medicare/Quality-Initiatives-patient-assessment-instruments/hospitalqualityinits/hospitalhcahps.html

Darby C, Hays R, Kletke, P. Development and Evaluation of the CAHPS Hospital Survey. Health Services Research, 40 (6 Pt. 2), 1973-1976.

Edgman-Levitan, S. & Cleary, P. D. (1996). What Information Do Consumers Want and Need? A synthesis of research to date, plus interviews with health plan managers and consumer advocates. Health Affairs, 15(4), 42-56.

Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: How do consumers use quality-of-care information? A systematic review. Med Care. 2009; 47 (1): 1–8.

Huppertz, J. W., & Carlson, J. P. (2010). Consumers' Use of HCAHPS Ratings and Word-of-Mouth in Hospital Choice. Health Services Research, 45 (6 Pt. 1), 1602-1613.

Keller, S., O'Malley, A. J., Hays, R.D., Matthew, R.A., Zaslavsky, A.M., Hepner, K.A., & Cleary, P.D. (2005). Methods used to streamline the CAHPS Hospital Survey. Health Services Research, 40 (6 Pt. 2), 2057-2077. PMCID: PMC1361248

Safran DG, Montgomery JE, Chang H, Murphy J, Rogers WH. (2001). Switching doctors: Predictors of voluntary disenrollment from a primary physician's practice. J Fam Practice. 50 (2):130–6.

Sofaer, S., Crofton, C., Goldstein, E., Hoy, E., & Crabb, J. (2005). What do consumers want to know about the quality of care in hospitals? Health Services Research, 40 (6 Pt 2), 2018–2036. doi:10.1111/j.1475-6773.2005.00473.x

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

Public reporting of HCAHPS scores have resulted in quality improvement initiatives aimed at patient experiences with care (Faber et al. 2009), -and there is evidence of improvement in HCAHPS scores (Elliott et al. 2010). HCAHPS data from the second and third HCAHPS implementation years (2008-2009), when hospitals participated on a voluntary basis, showed that all but one HCAHPS measure improved significantly nationally (Elliott et al. 2010). Five years later, meaningful and statistically significant improvement continued for all hospitals, especially among initially low-performing hospitals, reducing some gaps in hospital performance (Elliott et al. 2015). Focusing on 7-year (2007 to 2013) change in doctor communication, Mann et al. (2016) also report evidence of improvement and narrowing of the gap between the lowest- and highest-quartile hospitals. They recommend the further narrowing of the gap may be improved through sharing of best practices.

HCAHPS public reporting has been used to promote specific quality improvement activities tied to cultural competency efforts. Cultural competence is defined by the National Quality Forum (2008) as the ongoing capacity of healthcare organizations and professionals to provide high-quality, safe, patient and family centered, evidence-based, and equitable care. Cultural competency quality improvement activities show notable promise for improving all HCAHPS scores, but particular promise for hospitals with significant racial/ethnic/language minority patient populations, which may reduce disparities in patient experiences (Weech-Maldonado et al. 2012).

In addition, there are several published guides that instruct hospitals on how to improve patient experience of care as measured by the HCAHPS Survey. These resources include information for hospital managers on how to improve HCAHPS scores (Brady 2009), a study of HCAHPS best practices in critical access hospitals, compiled in part from patient focus groups (StratisHealth 2017), tactics for hospitals seeking to improve their HCAHPS scores (Studer et al. 2010; Cook et al. 2014), and a growing series of guides and podcasts created and maintained by the federal Agency for Healthcare Research and Quality (AHRQ) to assist healthcare providers, including hospitals, in the improvement of patient experience of care (AHRQ, 2019).

References

AHRQ (Agency for Healthcare Research and Quality). Creative Strategies To Improve Patient Care Experience. April 2019. https://www.ahrq.gov/cahps/news-and-events/podcasts/index.html

Brady, C. HCAHPS Basics: A resource guide for healthcare managers. HCPro. 2009.

Cook K, Ketelsen K & Kennedy, B. The HCAHPS Handbook 2nd Edition: Tactics to Improve Quality and the Patient Experience. Fire Starter Publishing. 2014.

Elliott MN, Lehrman WG, Goldstein EH, et al. Hospital survey shows improvements in patient experience. Health Affair. 2010. 29 (11):2061–2067.

Elliott, M.N., Cohea, C.W., Lehrman, W.G., Goldstein, E.H., Cleary, P.D., Giordano, L.A., & Zaslavsky, A.M. (2015). Accelerating improvement and narrowing gaps: Trends in patients' experiences with hospital care reflected in HCAHPS public reporting. Health Services Research, 50(6), 1850-1867.

Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: How do consumers use quality-of-care information? A systematic review. Med Care. 2009; 47 (1): 1-8.

Mann, R.K., Siddiqui, Z., Kurbanova, N. & Qayyum, R. (2016). Effect of HCAHPS reporting on patient satisfaction with physician communication. Journal of Hospital Medicine, 11 (2), 105-110.

National Quality Forum. Endorsing a Framework and Preferred Practices for Measuring and Reporting Culturally Competent Care Quality. Washington DC: National Quality Forum; 2008.

StratisHealth (2017). A study of HCAHPS best practices in high performing Critical Access Hospitals. National Rural Health Resource Center. https://www.ruralcenter.org/resource-library/study-of-hcahps-best-practices-in-high-performing-cahs

Studer Q., Robinson B.C. & Cook K. The HCAHPS Handbook: Hardwire Your Hospital for Pay-For-Performance Success. Studer Group. 2010.

Weech-Maldonado, R., Elliott, M.N., Pradhan, R., Schiller, C., Hall, A. & Hays, R.D. (2012). Can hospital cultural competency reduce disparities in patient experiences with care? Medical Care, 50, S48.

1a.3. SYSTEMATIC REVIEW(S) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

1a.4.2 What process was used to identify the evidence?

1a.4.3. Provide the citation(s) for the evidence.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

The HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems) Survey is the first national, standardized, publicly reported survey of patients´ perspectives of hospital care. HCAHPS (pronounced "H-caps"), also known as the CAHPS® Hospital Survey*, is a 29-item survey instrument and data collection methodology for measuring patients' perceptions of their hospital experience. While many hospitals have collected information on patient satisfaction for their own internal use, until HCAHPS there were no common metrics and no national standards for collecting and publicly reporting information about patient experience of care. Since 2008, HCAHPS has allowed valid comparisons to be made across hospitals locally, regionally and nationally.

Three broad goals have shaped HCAHPS. First, the standardized survey and implementation protocol produce data that allow objective and meaningful comparisons of hospitals on topics that are important to consumers. Second, public reporting of HCAHPS results creates new incentives for hospitals to improve quality of care. Third, public reporting enhances accountability in health care by increasing transparency of the quality of hospital care provided in return for the public investment. With these goals in mind, the Centers for Medicare & Medicaid Services (CMS) and the HCAHPS Project Team have taken substantial steps to assure that the survey is credible, practical and actionable.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

NOTE: For the complete response, please see, "ADDITIONAL, A.1" (HCAHPS Survey, NQF 0166 Appendix A.1: Supplemental Materials.)

Distribution of HCAHPS Top-Box Scores

(Data: 1Q16-4Q16 discharges, ~4,300 hospitals, ~3.1 million completed surveys)

	Mean	Std Dev	Percentile											Inter-Q Range
			100%	99%	95%	90%	75%	50%	25%	10%	5%	1%	0%	
Communication with Nurses	80.34	5.90	100	96	90	88	84	80	77	74	71	65	16	7
Communication with Doctors	82.05	5.80	100	98	92	90	85	82	78	76	74	69	29	7
Responsiveness of Hospital Staff	68.82	9.71	100	95	87	82	75	68	62	58	55	49	20	13
Communication about Medicines	65.32	7.80	100	90	79	75	69	64	61	58	55	48	1	8
Cleanliness	74.50	8.46	100	96	89	85	80	74	69	65	62	56	7	11
Quietness	62.79	10.72	100	90	82	77	69	62	56	50	46	41	3	13
Discharge Information	87.25	4.74	100	97	93	92	90	88	85	82	79	72	6	5
Overall Rating	72.77	9.01	100	93	87	84	78	73	67	62	58	49	24	11
Recommend Hospital	72.14	9.95	100	94	88	84	79	73	66	60	56	48	9	13
Care Transition	52.36	7.79	100	74	65	62	56	52	48	44	40	34	0	8

For historical HCAHPS scores, pleases see: http://www.hcahpsonline.org/SummaryAnalyses.aspx

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

N/A

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

NOTE: For the complete response, please see, "ADDITIONAL, A.1" (HCAHPS Survey, NQF 0166 Appendix A.1: Supplemental Materials.)

CMS purposely calculates disparities information at the HCAHPS Survey item level, which role up into the HCAHPS measures. Disparities are calculated on a regular basis to populate the congressionally-mandated disparities reports produced by the Agency for Healthcare Research and Quality (AHRQ). The disparities results can be found as an attachment in the Evidence section, under Evidence, 1.A.

We used 2014-2015 HCAHPS data to investigate differences in inpatient experiences by preferred language within racial/ethnic groups. HCAHPS is a survey of recently-discharged patients' experiences of hospital care in the United States and includes information on self-reported language preference and race/ethnicity that permits this analysis.32 Specifically, the HCAHPS survey asks which of seven languages the patient primarily speaks at home. Sample sizes used were sufficient to examine all preferred-language groups for which HCAHPS provides translations. As such, this dataset allows one to examine preferred language within racial/ethnic groups among even smaller groups, such as Portuguese-preferring Hispanics and Vietnamese-preferring Asian/Pacific Islanders (API).

HCAHPS measures experiences of inpatients of all payer types (Medicaid, Medicare, and all others) who are 18 years or older at admission, stay overnight in the hospital with a principal diagnosis for medical, surgical or maternity care, and are discharged alive.34 Our analysis included the 5,480,308 completed surveys from all 4,517 hospitals in the 50 states and DC that submitted HCAHPS data to the Centers for Medicare & Medicaid Services (CMS) during the eight quarters of calendar years 2014-2015.

We examined six HCAHPS composite measures: Communication with Doctors, Communication with Nurses, Responsiveness of Hospital Staff, Communication about Medication, Discharge Information, and Care Coordination. Three measures were excluded because they do not rely on conversing in a shared language (Quietness, Cleanliness) or are no longer used for incentive payments (Pain Management). Two global measures (Ratings of Hospital and Recommendation of Hospital) were excluded because prior research suggests that such items may elicit different evaluations of the same care from different racial/ethnic and language groups. The survey items comprising five of the six retained composite measures (all but Discharge Information items, which employ yes/no responses) use a standard set of response options: never, sometimes, e, and always. A description of the composite measures is included in the Appendix in Table A.1.

HCAHPS respondents are asked to self-report whether they are of Spanish, Hispanic or Latino origin or descent. They are then asked to select at least one race, with response options of White, Black or African American, American Indian or Alaska Native (AI/AN), Asian, Native Hawaiian or other Pacific Islander, and some other race. Six mutually-exclusive racial/ethnic categories were created using these two items: (1) Hispanic; and non-Hispanic (2) White, (3) Black, (4) API, (5) AI/AN, and (6) multiracial. Following the Office of Management and Budget approach, we classified any patient as Hispanic who endorsed Hispanic ethnicity. Non-Hispanic patients who endorsed exactly one race were classified as that race; those who endorsed Asian plus Native Hawaiian or other Pacific Islander were classified as API; the remaining non-Hispanic patients who endorsed two or more races were classified as multiracial. Our analysis excluded data from multiracial patients (3%), a heterogeneous and difficult-to-interpret group, and patients who did not answer the race item (7%).

Because several languages measured by the survey are associated almost exclusively with a single racial/ethnic group, language was considered within racial/ethnic categories. The HCAHPS survey asks, “What language do you mainly speak at home?” with response options of English, Spanish, Chinese, Russian, Vietnamese, Portuguese, and “some other language.” We included all combinations of preferred language (that is, language spoken at home) and race/ethnicity among our seven languages and five racial/ethnic groups for which at least 400 completed surveys were available nationally: Hispanics (languages included Spanish, English, Portuguese, Other), API (English, Chinese, Vietnamese, Other), Blacks (English, Spanish, Other), AI/AN (English, Other), and Whites (English, Russian, Spanish, Portuguese, Other).

To analyze the types of hospitals utilized by language-within-race/ethnicity groups, we examined key hospital characteristics of bed size (200 or more beds), rural location, profit status (for profit, not-for profit, governmental status), and service line composition (percent medical, surgical, maternity). We also calculated by preferred language within racial/ethnic groups: the average hospital-level proportion of non-English language-preferring patients, the average hospital-level proportion of the matching racial/ethnic group, and the average hospital-level proportion of their same racial/ethnic and language group. Linear regression compared overall, within-hospital and between-hospital patient experiences by preferred language within racial/ethnic groups using standard patient-mix adjustors. Following the CMS approach, we used patient-mix adjusted top-box-scored measures for all composite measures, scoring the most positive response option as 100 and all other responses as 0 prior to averaging non-missing items to create composite scores. The top-box response is “always” for four HCAHPS composites (Communication with Nurses, Communication with Doctors, Responsiveness of Hospital Staff, and Communication about Medication), “yes” for the Discharge Information composite, and “strongly agree” for the Care Coordination composite. To illustrate, the score for a respondent who answered “always,” “always,” “never,” and “sometimes” to four items within a composite would be (100+100+0+0=200)/4 = 50. “Patient mix” refers to patient characteristics not under the control of the hospital that may affect scores of patient experience measures. Patient-mix adjustment accounts for between-hospital differences in the patient population to estimate the scores each hospital would have received if all had

treated the same patients. Standard HCAHPS patient-mix adjustors are patient age; service line (maternity, surgical, and medical [reference category]); self-reported education; self-reported overall patient health; response percentile (a rank-based measure of the latency between discharge date and survey completion that addresses the tendency of later responders to indicate worse care experiences); interactions of maternity and surgical service line with linearly-scored patient age; and preferred language spoken at home. Here, preferred language spoken at home was treated as the primary independent variable, rather than as a patient-mix adjustor.

In Table 3, differences in patient experiences for six HCAHPS measures are shown by language within racial/ethnic group. Generally, non-English-preferring Black, Hispanic, API and AI/AN patients reported worse experiences than their English-preferring counterparts, except for Russian-preferring White patients. Differences between English and non-English preferring patients within the same racial/ethnic group were largest and most consistent (i.e., the findings were both statistically significant and had the same sign) for Care Coordination and smallest and least consistent for Discharge Information and Communication about Medication.

The experiences of White patients were not consistent across measures and language preference. Spanish-preferring and other-language-preferring Whites reported less positive experiences than English-preferring Whites, except for Doctor Communication where experiences were similar for English-preferring and other other-language-preferring Whites. Russian-preferring Whites reported the best experiences among Whites, except for Care Coordination.

Non-English-preferring Black patients reported consistently worse experiences than their English-preferring counterparts, with all differences at least moderate in magnitude (3+ points).

Among Hispanics, Spanish-preferring and other-language-preferring patients reported worse experiences than English-preferring Hispanics except for similar experiences for Communication about Medication for English-preferring and other-language-preferring patients. Portuguese-preferring Hispanics reported worse experiences for only Doctor Communication and Care Coordination. Generally, differences between English-preferring and non-English preferring Hispanics were small (<3 points), except for Care Coordination.

Among API patients, each non-English-preferring group (Chinese-, Vietnamese-, and other-language-preferring) reported worse experiences than English-preferring API, except for Discharge Information. Within API, negative differences compared to English-preferring API tended to be largest for Chinese-preferring API.

Within AI/AN, other-language-preferring AI/AN reported worse care experiences than English-preferring AI/AN; differences were moderate or larger, except for Communication About Medicines.

-- NOTE: The above referenced table could not be copied into section 1b.4. However, this table is included in the Measure Testing Form, 2b3.4a, labelled "Table 3".

From: "Inpatient Care Experiences Differ by Preferred Language within Racial/Ethnic Groups." D.D. Quigley, M.N. Elliott, K. Hambarsoomian, S.M. Wilson-Frederick, W.G. Lehrman, D. Agniel, J.H. Ng, E.H. Goldstein, L.A. Giordano and S.C. Martino. Health Services Research, 1-12. 2019. Published online, 1-6-19: https://doi.org/10.1111/1475-6773.13105

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

N/A

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

Behavioral Health

De.6. Non-Condition Specific*(check all the areas that apply):*

Health and Functional Status : Change, Person-and Family-Centered Care

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

http://www.hcahpsonline.org/

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

Attachment **Attachment:** HCAHPS-_NQF_0166-_Data_Dictionary-_4-9-19.docx

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Attachment **Attachment:** HCAHPS_29-item_11-21-18_--_Mail_Survey_Materials_-English-.pdf

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Patient

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

Yes

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

The HCAHPS Survey originally included three items about pain (a screener and two substantive questions), which formed a composite measure, Pain Management. CMS discontinued publicly reporting this measure in July 2018. In January 2018, CMS replaced the original HCAHPS pain items with three items that asked about communication about pain. In compliance with the Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act (Pub. L. 115-271) of 2018 (Section 6104), CMS will remove the new communication about pain items from the HCAHPS Survey beginning with October 2019 discharges, which will result in a 29-item survey.

The HCAHPS Patient-Mix Adjustment (PMA) model has been updated to incorporate more detailed information about patients' Service Line and Gender. Prior to Quarter 1 2017, the patient-mix adjustment for service line distinguished among the three service line categories: Medical, Surgical, and Maternity. Beginning with Quarter 1 2017 discharges, the patient-mix adjustment will cross patient gender with service line to distinguish among 5 categories: Female Medical, Male Medical, Female Surgical, Male Surgical, and Maternity, which is only female. Female Medical will serve as the reference category for this adjustment. HCAHPS survey results will be adjusted using the new PMA model beginning with January 1, 2017 discharges.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The HCAHPS Survey asks recently discharged patients about aspects of their hospital experience that they are uniquely suited to address. The core of the survey contains 19 items that ask "how often" or whether patients experienced a critical aspect of hospital care, rather than whether they were "satisfied" with their care. Also included in the survey are three screener items that direct patients to relevant questions, five items to adjust for the mix of patients across hospitals, and two items (race and ethnicity) that support Congressionally-mandated reports. Hospitals may include additional questions after the core HCAHPS items.

For full details, see the current HCAHPS Quality Assurance Guidelines, V.14.0, pp. 57-65, under the "Quality Assurance" button on the official HCAHPS On-Line Web site at:

https://www.hcahpsonline.org/globalassets/hcahps/quality-assurance/2019_qag_v14.0.pdf

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

For each question in a multi-item measure, the proportion of responses in the "top" (most positive response) and "bottom" (least positive response) boxes are calculated for a given hospital (completed surveys only). For clarification on which answer values go in each box for each measure go to www.hospitalcompare.hhs.gov. To obtain a hospital's raw score for the top or bottom box category, the mean proportion for all the questions in a given measure is calculated. Note that the middle box is the proportion remaining after the top and bottom boxes have been calculated; see below for details.

The following raw score calculations are performed for each eligible hospital and within each quarter.

- Multi-item Measure Calculation – Communication with Nurses (3 questions):

  Pi1 = Proportion of (item) respondents who said "Never" to question i

  Pi2 = Proportion of respondents who said "Sometimes" to question i

  Pi3 = Proportion of respondents who said "Usually" to question i

  Pi4 = Proportion of respondents who said "Always" to question i

  The index i represents the number of questions in the multi-item measure, here i = 1, 2, 3.

  The bottom box consists of the answer value categories of "Never" and "Sometimes". Bottom Box multi-item measure Score = (P11+P12+P21+P22+P31+P32)/3

  The top box consists only of the answer category "Always".

  Top Box multi-item measure Score = (P14+P24+P34)/3

- Individual Item Example – Cleanliness of Hospital Environment (1 question):

P1 = Proportion of respondents who said “Never” to the question

P2 = Proportion of respondents who said “Sometimes” to the question

P3 = Proportion of respondents who said “Usually” to the question

P4 = Proportion of respondents who said “Always” to the question

The bottom box consists of the answer value categories of “Never” and “Sometimes”.

Bottom Box Individual Item Score = P1 + P2

The top box consists only of the answer category “Always”.

Top Box Individual Item Score = P4

- Global Item Example – Overall Hospital Rating (1 question):

P0 = Proportion of respondents who rated the hospital as 0 (worst hospital possible)

P1 = Proportion of respondents who rated the hospital as 1

P2 = Proportion of respondents who rated the hospital as 2

P3 = Proportion of respondents who rated the hospital as 3

P4 = Proportion of respondents who rated the hospital as 4

P5 = Proportion of respondents who rated the hospital as 5

P6 = Proportion of respondents who rated the hospital as 6

P7 = Proportion of respondents who rated the hospital as 7

P8 = Proportion of respondents who rated the hospital as 8

P9 = Proportion of respondents who rated the hospital as 9

P10 = Proportion of respondents who rated the hospital as 10 (best hospital possible)

The bottom box consists of hospital rating response values from 0 to 6.

Bottom Box Global Item Score = P0 + P1 + P2 + P3 + P4 + P5 + P6

The top box consists of hospital rating response values of 9 and 10.

Top Box Global Item Score = P9 + P10

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

The target population for HCAHPS measures include eligible adult inpatients of all payer types who completed a survey. HCAHPS patient eligibility and exclusions are defined in detail in the sections that follow. A survey is defined as completed if the patient responded to at least 50% of questions applicable to all patients.

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Eligibility for the HCAHPS Survey.

The HCAHPS Survey is broadly intended for patients of all payer types who meet the following criteria:

Eighteen (18) years or older at the time of admission

Admission includes at least one overnight stay in the hospital

An overnight stay is defined as an inpatient admission in which the patient´s admission date is different from the patient´s discharge date. The admission need not be 24 hours in length. For example, a patient had an overnight stay if he or she was admitted at 11:00 PM on Day 1, and discharged at 10:00 AM on Day 2. Patients

who did not have an overnight stay should not be included in the sample frame (e.g., patients who were admitted for a short period of time solely for observation; patients admitted for same day diagnostic tests as part of outpatient care).

Non-psychiatric MS-DRG/principal diagnosis at discharge

Note: Patients whose principal diagnosis falls within the Maternity Care, Medical, or Surgical service lines and who also have a secondary psychiatric diagnosis are still eligible for the survey.

Alive at the time of discharge

Note: Pediatric patients (under 18 years old at admission) and patients with a primary psychiatric diagnosis are ineligible because the current HCAHPS instrument is not designed to address the unique situation of pediatric patients and their families, or the behavioral health issues pertinent to psychiatric patients.

A completed HCAHPS survey is one with responses for at least 50% of the questions that are applicable to all patients (questions 1-10, 12, 15, and 18-22).

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

There are a few categories of otherwise eligible patients who are excluded from the HCAHPS sample frame. As detailed below in sec S.9, these exclusions include patients excluded due to state regulations, no-publicity patients, and specific groups of patients with an admission source or discharge status that results in difficulty collecting patient experience data through a survey instrument.

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

There is a two-stage process for determining whether a discharged patient can be included in the HCAHPS Sample Frame. The first stage is to determine whether the discharged patient meets the HCAHPS eligibility criteria, listed above. If the patient meets the eligibility criteria, then a second set of criteria is applied: Exclusions from the HCAHPS Survey.

Patients who meet the eligible population criteria previously outlined are to be included in the HCAHPS Sample Frame. However, there are a few categories of otherwise eligible patients who are excluded from the sample frame. These are:

"No-Publicity" patients – Patients who request that they not be contacted (see below)

Court/Law enforcement patients (i.e., prisoners); this does not include patients residing in halfway houses

Patients with a foreign home address (the U.S. territories – Virgin Islands, Puerto Rico, Guam, American Samoa, and Northern Mariana Islands are not considered foreign addresses and therefore, are not excluded)

Patients discharged to hospice care (Hospice-home or Hospice-medical facility)

Patients who are excluded because of state regulations

Patients discharged to nursing homes and skilled nursing facilities

"No-Publicity" patients are defined as those who voluntarily sign a "no-publicity" request while hospitalized or who directly request a survey vendor or hospital not to contact them ("Do Not Call List"). These patients should be excluded from the HCAHPS Survey. However, documentation of patients' "no-publicity" status must be retained for a minimum of three years.

Court/Law enforcement patients (i.e., prisoners) are excluded from HCAHPS because of both the logistical difficulties in administering the survey to them in a timely manner, and regulations governing surveys of this population. These individuals can be identified by the admission source (UB-04 field location 15) "8 – Court/Law enforcement," patient discharge status code (UB-04 field location 17) "21 – Discharged/transferred to court/law enforcement," or patient discharge status code "87 – Discharged/transferred to court/law

enforcement with a planned acute care hospital inpatient readmission." This does not include patients residing in halfway houses.

Patients with a foreign home address are excluded from HCAHPS because of the logistical difficulty and added expense of calling or mailing outside of the United States (the U.S. territories - Virgin Islands, Puerto Rico, Guam, American Samoa, and Northern Mariana Islands are not considered foreign addresses and therefore, are not excluded).

Patients discharged to hospice care are excluded from HCAHPS because of the heightened likelihood that they will expire before the survey process can be completed. Patients with a "Discharge Status" of "50 – Hospice – home" or "51 – Hospice – medical facility" would not be included in the sample frame. "Discharge Status" is the same as the UB-04 field location 17.

Some state regulations place further restrictions on patients who may be contacted after discharge. It is the responsibility of the hospital/survey vendor to identify any applicable regulations and to exclude those patients as required by law or regulation in the state in which the hospital operates.

Patients discharged to nursing homes and skilled nursing facilities are excluded from HCAHPS. This applies to patients with a "Discharge Status" (UB-04 field location 17) of:

"03 – Skilled nursing facility"

"61 – SNF Swing bed within hospital"

"64 – Certified Medicaid nursing facility"

"83 – Skilled nursing facility with a planned acute care hospital inpatient readmission"

"92 – Certified Medicaid nursing facility with a planned acute care hospital inpatient readmission"

Hospitals/Survey vendors must retain documentation that verifies all exclusions and ineligible patients. This documentation is subject to review.

Note: Patients must be included in the HCAHPS Survey sample frame unless the hospital/ survey vendor has positive evidence that a patient is ineligible or fits within an excluded category. If information is missing on any variable that affects survey eligibility when the sample frame is constructed, the patient must be included in the sample frame.

For more details, please see the HCAHPS Quality Assurance Guidelines, V14.0, pp. 57-80, located at the "Quality Assurance" button on the official HCAHPS On-Line Web site, at

https://www.hcahpsonline.org/globalassets/hcahps/quality-assurance/2019_qag_v14.0.pdf

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

NOTE: For the complete response, please see, "ADDITIONAL, A.1" (HCAHPS Survey, NQF 0166 Appendix A.1: Supplemental Materials.)

HCAHPS utilizes risk adjustment, not stratification, in reporting outcomes.

Please see below for details regarding S.11.

The information below is taken from a document on our public Web site, HCAHPS On-Line Web site. For more details, and appendices, about the statistical risk model and variables, including the tables that are referenced in the material below, please see the "Mode & Patient-Mix Adjustment Abstract (revised 5/2/08)" paper located via the "Mode and Patient-Mix Adj" button on the official HCAHPS On-Line Web site, at

http://www.hcahpsonline.org/en/mode--patient-mix-adj/

A document containing the patient-mix adjustment coefficients for the April 2018 public reporting of HCAHPS results, based on discharges from July 1, 2016 to June 30, 2017, is located via the "Mode and Patient-Mix Adj" button on the official HCAHPS On-Line Web site, at

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/2018-4_mode_patient_mix-adj.pdf

on the official HCAHPS On-Line Web site, www.HCAHPSonline.org

(Please note: in the document "Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS) of April 30, 2008,"we refer to multi-item scores as "composites," but these are in fact "multi-item measures").

A randomized Mode Experiment of 27,229 discharges from 45 hospitals was used to develop adjustments for the effects of survey mode (Mail Only, Telephone Only, Mixed, or Active Interactive Voice Response) on responses to the CAHPS® Hospital Survey (also known as Hospital CAHPS or HCAHPS). In general, patients randomized to the Telephone Only and Active Interactive Voice Response modes provided more positive evaluations than patients randomized to Mail Only and Mixed (Mail with Telephone follow-up) modes. These mode effects varied little by hospital and were strongest for the Responsiveness, Pain Management, and Discharge Information multi-item measures, the Cleanliness and Quiet items, and the global Rating and Recommendation. The Mode Experiment was also used to develop a model for patient-mix adjustment in order to account for the effect on HCAHPS responses of patient characteristics not under the control of hospitals. Adjustments for the effects of survey mode and patient-mix are necessary for valid comparison of scores across hospitals. After making these adjustments, no adjustments for nonresponse are necessary.

Introduction

The intent of the CAHPS®1 Hospital Survey, also known as Hospital CAHPS or HCAHPS, is to provide a standardized survey instrument and data collection methodology for measuring patients' perspectives of hospital care. In order to achieve the goal of fair comparisons across all hospitals that participate in HCAHPS, it is necessary to adjust for factors that are not directly related to hospital performance but do affect how patients answer HCAHPS survey items. These factors include the mode of survey administration, the characteristics of patients in participating hospitals, and differences between participating and non-participating patients. Collectively, we propose adjustments that are intended to eliminate any advantage or disadvantage in scores that might result from the mode of survey administration or patient characteristics beyond a hospital's control.

In order to ensure that publicly reported HCAHPS scores allow fair and accurate comparisons of hospitals, in 2006 the Centers for Medicare & Medicaid Services (CMS) undertook a Mode Experiment to examine whether mode of survey administration, the mix of patients in participating hospitals, or survey non-response systematically affect HCAHPS survey results and then developed necessary statistical adjustments. This paper summarizes the derivation of these adjustments from that large-scale, randomized mode experiment.

Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS)

The Mode Experiment addressed three important sources of potential bias in hospital-level HCAHPS results. First, hospitals participating in the HCAHPS survey have the option of choosing among four different modes of data collection: Mail, Telephone, Mail combined with Telephone follow-up (also known as Mixed mode), and Active Interactive Voice Response (IVR). If patient responses differ systematically by mode of survey administration, it is necessary to adjust for survey mode.

Second, certain patient characteristics that are not under the control of the hospital, such as age and education, may be related to the patient´s survey responses. For example, several studies have found that younger and more educated patients provide less positive evaluations of healthcare. If such differences occur in HCAHPS data, it is necessary to adjust for such respondent characteristics before comparing hospitals´ HCAHPS results.

Third, if the patients who respond to the HCAHPS survey differ from those who are sampled but do not complete the survey, there is a possibility that patterns of nonresponse may create a bias in reported scores.

Nonresponse bias is a concern if three conditions hold: (1) nonrespondents differ from respondents, (2) nonrespondents and respondents differ in ways that are related to how patients evaluate hospitals using HCAHPS, and (3) these differences persist even after adjusting for survey mode and patient-mix. Only if all three of these conditions hold is it necessary to adjust for survey nonresponse.

The HCAHPS Mode Experiment

To assess the effect of mode of data collection, CMS conducted a large-scale experiment to compare the four allowed modes of HCAHPS data collection: Mail questionnaire only; Telephone interview only; Mixed mode (Mail questionnaire with Telephone follow up if needed); and Active IVR. In the Active IVR mode, live telephone interviewers contact the patients and invite them to participate in an automated IVR interview using their telephone keypads.

A random sample of 45 hospitals from across the United States participated in the HCAHPS Mode Experiment in early 2006. Each hospital provided a sample of discharged patients who met HCAHPS eligibility criteria.2 These samples were randomly allocated to each of the four modes in equal numbers within each hospital and patients were then surveyed accordingly. To assure uniformity in administration, sample selection and surveying for the Mode Experiment were conducted by a single agent, the National Opinion Research Center (NORC) of the University of Chicago. Analysis of Mode Experiment data and construction of the adjustment algorithms were performed by the RAND Corporation for CMS.

Table 1 (below) displays response rates from the HCAHPS Mode Experiment. As can be seen, the response rate was highest for Mixed mode (41.2%) and lowest for IVR (20.7%). Although there was some variation in response rate by hospital (the hospital-level

standard deviation in response rates was 5.6%), the response rate patterns by mode were consistent across hospitals.

(For information about eligibility, please see the HCAHPS Quality Assurance Guidelines, at www.hcahpsonline.org.

Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS))

Table 1: Comparison of Patient Response Rates by Survey Mode in the HCAHPS Mode

Experiment

	MAIL ONLY	TELEPHONE ONLY	MIXED	ACTIVE IVR	OVERALL
Discharges Randomized to Mode	6806	6808	6808	6807	27,229
Cases Determined to be Ineligible in the Field	23				
(0.3%) 928					
(13.6%) 761					
(11.2%) 900					
(13.2%) 2612					
(9.6%)					
Completed Surveys	2239	1607	2489	1220	7555
Response Rate of Eligible Patients (Completes/Eligible1)					
33.0%					
27.3%					
41.2%					
20.7%					
30.7%					
Yield (Completes/ Randomized)	32.9 %	23.6 %	36.6 %	17.9 %	27.7 %

1 “Eligible” is defined as randomized cases minus those determined to be ineligible in the field.

Analysis of the HCAHPS Mode Experiment

CMS estimated mode effects in linear models that include both hospital fixed effects and patient-mix adjustment (PMA)3 for demographic and other patient factors associated with response tendency. For each HCAHPS rating or report item, a linear regression model consisting of mode fixed effects, hospital fixed effects, and patient-mix adjusters was estimated. These linear models generate adjustments for both mode and patient-mix. Because patient-mix adjustment will be employed, we calculate mode adjustments that correspond to the mode effects that remain after patient-mix adjustments.4

Developing the Patient-Mix Adjustment (PMA) Model

Patient-mix refers to patient characteristics that are not under the control of the hospital that may affect patient reports of hospital experiences. The goal of adjusting for patient-mix is to estimate how different hospitals would be rated if they all provided care to comparable groups of patients. In developing the HCAHPS patient-mix adjustment (PMA) model, we sought important and statistically significant predictors of patients’ HCAHPS ratings that also vary meaningfully across hospitals. Adjustors with both of these characteristics will substantially adjust hospital-level scores.

We considered eight candidate PMA variables: service line (medical, surgical, or maternity care), age, education, self-reported health status, language other than English spoken at home, age by service line interactions, emergency room (ER) admission, and percentile response order, also known as “relative lag time,” which is based on the time between discharge and survey completion.5

For the ordinal candidates (age, education, and self-rated health status), we tested whether treating the PMA variable categorically as a series of dummy variables was more predictive of HCAHPS outcomes than a linear form; we used the categorical form only when there was evidence of it being more predictive. We tested the statistical significance of candidate PMA variables in multivariate linear regressions, one for each outcome, using patient-mix adjustors, mode dummies, and hospital dummies as predictors. We calculated the explanatory power of each candidate patient-mix adjustor for hospital-level adjustments (O’Malley et al., 2005).

3 Also known as case-mix adjustment (CMA) in other parts of the CAHPS literature. CMS uses the term patient-mix adjustment here to distinguish this adjustment from severity adjustments for clinical outcomes or payment.

4 These mode adjustments are very similar to the mode adjustments that would be employed in the absence of patient-mix adjustment.

5 Computed as a percentile of all fielded cases within a given hospital and mode, so that the 10th response of 100 fielded cases for the Mail Only mode of Hospital A would be 0.10 and the 40th and last response from that same hospital in that same mode, assuming a 40% response rate, would be 0.40.

Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS) 3

Developing the Mode Adjustments

In making mode adjustments, it is necessary to choose one mode as a reference point. One can then interpret all adjusted data from all modes as if they had been surveyed in the reference mode. Because it is the most commonly used mode in patient surveys, CMS selected the Mail Only mode as the reference mode of survey administration. The choice of mail mode as the reference mode does not indicate that mail mode is preferable to other approved modes in any way.

Surveys conducted in the Mail Only mode are not adjusted further for mode after PMA. Surveys conducted in the other three modes (Telephone Only, Mixed, Active IVR) are adjusted according to the difference in mode effects between that mode and the Mail Only mode, as estimated through linear regression in the HCAHPS Mode Experiment. In particular, the mode effects for each outcome are the coefficients for the mode dummy variables in regression models with three mode dummies, hospital dummies, and the final patient-mix

adjustors. These coefficients estimate the remaining difference between Mail Only mode and each of the other modes after patient-mix adjustment.

Nonresponse Analysis

Logistic regression was used to model response propensity among eligible discharges from hospital indicators, survey mode, and available individual-level administrative variables: age, gender, service line, emergency room admission, and discharge status (sick, left against medical advice, or standard). Nonresponse weights were derived from these models and tested with respect to the extent to which they were associated with patient-mix adjusted scores.

HCAHPS Multi-item measure Scoring

Each of the six HCAHPS composites (Communication with Nurses, Communication with Doctors, Responsiveness of Hospital Staff, Pain Management, Communication about Medicines, and Discharge Information) is calculated as the average of its two or three constituent items. In following previous CAHPS practice, items within a multi-item measure are first individually patient-mix adjusted and then are weighted so as to give each item equal influence within the multi-item measure. Mode adjustments for multi-item measure scores are derived as the unweighted averages of mode adjustments for individual constituent items, so that each item has equal influence on the multi-item measure adjustment.

Mode Adjustment Results

Patients generally provided more best category ("top-box") responses in the Telephone Only and Active IVR modes than in the Mail Only and Mixed modes. Differences between Telephone Only and Active IVR responses were generally small, and only two items differed between Mail Only and Mixed Mode. In particular, Telephone Only responses were more positive than Mail Only for the Communication with Nurses multi-item measure, the Pain Management multi-item measure, the Communication about Medicine

multi-item measure, the Staff Responsiveness multi-item measure, the Cleanliness item, the Quiet item, and the Recommendation item. Active IVR was more positive than Mail Only for the Communication with Nurses multi-item measure, the Discharge Information multi-item measure, and the Quiet item. Mixed Mode was significantly more positive than Mail Only for the Cleanliness item and the Quiet item.

Table 2 (below) presents mode adjustments derived from the HCAHPS Mode Experiment for the best category ("top-box") proportion in models that include patient-mix adjustment. As an example, a patient-mix adjusted score of 84.2% "always" for the Communication with Nurses multi-item measure for a survey conducted by Telephone Only mode would be further adjusted to (84.2% - 4.0% =) 80.2% in order to account for the fact that 80.2% is the corresponding expected score for that multi-item measure had the survey been conducted in Mail Only mode. Here, 4.0% represents the increase in the proportion of patients responding "always" that would be expected from the same patients had they been surveyed by Telephone Only mode (when compared to the reference mode of Mail Only). Similarly, Table 3 (below)presents mode adjustments for the lowest category ("bottom-box") proportions. As an example, a patient-mix adjusted score of 7.2% "never" or "sometimes" for the Communication with Nurses multi-item measure for a survey conducted by Telephone Only mode would be further adjusted to (7.2% - 0.8% =) 6.4% in order to account for the fact that 6.4% is the corresponding expected score for that multi-item measure had the survey been conducted in Mail Only mode. Here, 0.8% represents the increase in the proportion of patients responding "never" or "sometimes" that would be expected from the same patients had they been surveyed by Telephone Only mode (when compared to the reference mode of Mail Only). In this same example, 100.0%- 80.2% (adjusted top-box)-6.4% (adjusted bottom-box)=13.4% would be the fully adjusted score for the "middle-box" category, here corresponding to "usually" for Communication with Nurses.

Table 2: Mode Adjustments of Top Category ("Top-Box") Percentages (after PMA) to Adjust Other Modes to a Reference of Mail

PHONE ONLY MIXED ACTIVE

IVR

Composites

Communication with Nurses

(Always)

-4.0%

-0.3%

-1.8%

Communication with Doctors

(Always)

-1.3%

1.0%

-0.3%

Responsiveness of Hospital Staff (Always)

-4.7%

0.1%

-1.9%

Pain Management (Always) -4.7% -2.3% -3.4%

Communication about Medicines

(Always)

-3.9%

-0.9%

-1.6%

Discharge information (Yes) -1.3% 0.2% -3.2%

Individual Report Items

CLEANLINESS

(Always)

-5.5%

-2.1%

-1.9%

QUIET

(Always) -6.3% -3.1% -10.2%

Global Items

RECOMMEND HOSPITAL

(Definitely Yes)

-4.4%

-1.4%

-2.2%

HOSPITAL RATING (9 or 10) -2.8% -1.8% -1.6%

Table 3: Mode Adjustments of Bottom Category (“Bottom-Box”) Percentages (after

PMA) to Adjust Other Modes to a Reference of Mail

PHONE ONLY MIXED ACTIVE

IVR

Composites

Communication with Nurses

(Always)

-0.8%

-0.5%

-0.6%

Communication with Doctors

(Always)

-2.2%

-1.4%

-1.2%

Responsiveness of Hospital Staff (Always)

-0.2%

-1.9%

-1.4%

Pain Management (Always) -0.6% -0.9% -1.3%

Communication about Medicines

(Always)

0.5%

-1.4%

-1.5%

Discharge information (Yes) 1.3% -0.2% 3.2%

Individual Report Items

CLEANLINESS

(Always)

1.0%

0.4%

0.6%

QUIET

(Always) -1.4% 0.9% 1.4%

Global Items

RECOMMEND HOSPITAL

(Definitely Yes)

0.4%

-0.4%

0.1%

HOSPITAL RATING (9 or 10) 0.9% -1.1% 0.8%

Patient-mix Adjustment Results and Model

All candidate patient-mix adjustors were statistically significant predictors of at least one reported HCAHPS outcome and each had at least as much average explanatory power as PMA variables that have been previously recommended for use in HCAHPS PMA (O'Malley et al., 2005). Age had a significantly nonlinear relationship with 8 of 10 reported outcomes, but education and self-rated health status were well characterized by linear scoring of the ordinal categories. Evaluations of care increased with self-rated health and age (at least through age 74), and decreased with educational attainment. Maternity service had generally more positive evaluations than medical and surgical services. Evaluations were generally lower for those admitted through the ER. Percentile response order (relative lag time) findings showed that late responders tended to provide less positive evaluations than earlier responders.

The final PMA model includes all eight candidate PMA variables as follows: linear self-reported health status, linear education, service line, categorical age, ER admission source, response percentile, service by linear age interactions, and primary language other than English.

Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS) 5

Nonresponse Findings

Although there was evidence of selective nonresponse, the PMA model employed was found to effectively account for any nonresponse bias that could have been addressed through nonresponse weighting. Therefore, no further weighting or adjustment for nonresponse is needed.

Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS) 6

Integrated Patient-mix and Mode Adjustment

Patient-mix and survey mode adjustments are applied sequentially to the raw HCAHPS scores. Survey responses first undergo patient-mix adjustment using the model specified above, adjusting to the unweighted mean of all responding patients in the given public reporting period, which is typically four calendar quarters. It bears mentioning that the exact values of PMA coefficients used for adjustment are not based on the values observed in the HCAHPS Mode Experiment but are re-estimated each reporting period based on the empirical relationship observed between PMA variables and HCAHPS outcomes in that period. Also, please note that although

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

Statistical risk model

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

NOTE: For the complete response, please see, "ADDITIONAL, A.1" (HCAHPS Survey, NQF 0166 Appendix A.1: Supplemental Materials.)

SCORING AND PATIENT-MIX ADJUSTMENTS

Data timeframe

- 12 months of data on a “rolling” basis

Sampling rates

- Monthly samples must be weighted to control for varying sampling rates throughout the year in order to make the combined monthly samples representative of the full population of discharges

Global rating

- Measured by the overall rating of the hospital and the extent to which patients are willing to recommend the hospital (Q18 & Q19)

Domains of care

- Communication with doctors (Q5, Q6, & Q7)
- Communication with nurses (Q1, Q2, & Q3)
- Responsiveness of the hospital staff (Q4, Q10, & Q11)
- Communication about medicines (Q12, Q13, & Q14)
- Cleanliness and quiet of physical environment (Q8 & Q9)
- Discharge information (Q15, Q16, & Q17)

Production of scores—Global ratings

- Overall rating of the hospital

For this item, respondents are asked, “Using any number from 0 to 10, where 0 is the worst hospital possible and 10 is the best hospital possible, what number would you use to rate this hospital?” The scoring on this item will represent the proportion of respondents who gave a rating of 0-7, 8-9, or 10 to the hospital.

The steps to calculate a hospital’s score for “overall rating” follow:

Step 1 – Assign appropriate sampling weight to each case

CMS expects that most hospitals will sample a fixed number of discharges each month to reach the target of 300 completes annually. However, the monthly population of discharges from which these fixed-sized samples are drawn will vary throughout the year. There are more total discharges in some months than others in most hospitals. Thus sampling rates will vary from month to month. To make the combined monthly samples representative of the full population of discharges for the year, it is necessary to adjust for the different monthly sampling rates. Appropriate sampling weights can be assigned to each case to make the combined monthly samples representative of the total population of annual discharges. This will be done as follows:

Calculate the expansion weight for each month (Em).

Em = (Population size for the month) / (Sample size for the month)

Calculate the mean expansion weight for the number of months covered in the score (e.g., 12 months).

E = (Sum of Em) / (number of months)

Calculate the relative weight for each month as the expansion weight for the month divided by the mean expansion weight.

Wm = Em / E

Assign a sampling weight to each case (Wi) based on the month in which the person was discharged and corresponding value of Wm.

Step 2 – Identify relevant cases

Include only cases where survey status is a completed survey.

Include only cases with non-missing values on the overall rating question.

Step 3 – Calculate the proportion of cases in each response category

Proportion of respondents who gave the hospital an overall rating of 0-7:

The numerator is the number of respondents for whom the overall rating (Xi) is 0-7. Each case is weighted by the appropriate sampling weight for the month the person was discharged.

The denominator is the total number of respondents (each weighted by the appropriate sampling weight for the month the person was discharged).

The proportion can be defined as follows:

Let X1i = 1 when Xi is 0-7

= 0 otherwise

P1 = (Sum of WiX1i) / sum of Wi

Proportion of respondents who gave the hospital an overall rating of 8 or 9:

The numerator is the number of respondents for whom the overall rating (Xi) is 8 or 9. Each case is weighted by the appropriate sampling weight for the month the person was discharged.

The denominator is the total number of respondents (each weighted by the appropriate sampling weight for the month the Person was discharged).

The proportion can be defined as follows:

Let X2i = 1 when Xi is 8 or 9

= 0 otherwise

P2 = (Sum of WiX2i) / Sum of Wi

Proportion of respondents who gave the hospital an overall rating of 10:

The numerator is the number of respondents for whom the overall rating (Xi) is 10. Each case is weighted by the appropriate sampling weight for the month the person was discharged.

The denominator is the total number of respondents (each weighted by the appropriate sampling weight for the month the person was discharged).

The proportion can be defined as follows:

Let X3i = 1 when Xi is 10

= 0 otherwise

P3 = (Sum of WiX3i) / Sum of Wi

- Willingness to recommend the hospital

For this item, respondents are asked, “Would you recommend this hospital to your friends and family?” to which they can respond “definitely no,” “probably no,” “probably yes,” or “definitely yes.” A hospital’s score is the proportion of cases in each response category. The approach to the production of a hospital’s score on this item follows the same steps noted for “overall rating of the hospital.”

Production of scores—Domain ratings

There are six domain-level multi-item measures included in the HCAHPS measure: communication with doctors, communication with nurses, responsiveness of hospital staff, communication about medicines, cleanliness and quiet of the hospital environment, and discharge information. The steps to calculate multi-item measure scores follow:

- Communication with doctors

This multi-item measure is produced by combining responses to three questions that ask:

  - “During this hospital stay, how often did doctors listen carefully to you?”
  - “During this hospital stay, how often did doctors explain things in a way you could understand?”
  - “During this hospital stay, how often did doctors treat you with courtesy and respect?”

Respondents can answer "never," "sometimes," "usually," or "always" to each. A hospital's score on the "doctor communication" multi-item measure is the proportion of cases in each response category.

The steps to calculate a hospital's multi-item measure score follow:

Step 1 – Calculate the proportion of cases in each response category for each question

Follow the same steps for calculating the proportion of cases in a response category discussed above for "overall rating of the hospital" to obtain proportions for the first question:

P11 = Proportion of respondents who said "never" to the first question

P12 = Proportion of respondents who said "sometimes" to the first question

P13 = Proportion of respondents who said "usually" to the first question

P14 = Proportion of respondents who said "always" to the first question

Follow the same steps for calculating the proportion of cases in a response category discussed above for "overall rating of the hospital" to obtain proportions for the second question:

P21 = Proportion of respondents who said "never" to the second question

P22 = Proportion of respondents who said "sometimes" to the second question

P23 = Proportion of respondents who said "usually" to the second question

P24 = Proportion of respondents who said "always" to the second question

Follow the same steps for calculating the proportion of cases in a response category discussed above for "overall rating of the hospital" to obtain proportions for the third question:

P31 = Proportion of respondents who said "never" to the third question

P32 = Proportion of respondents who said "sometimes" to the third question

P33 = Proportion of respondents who said "usually" to the third question

P34 = Proportion of respondents who said "always" to the third question

Step 2 – Combine responses from the questions to form the multi-item measure.

Calculate the average proportion responding to each category across the three questions in the multi-item measure:

PC1 = Multi-item measure proportion who said "never" = (P11 + P21 + P31) / 3

PC2 = Multi-item measure proportion who said "sometimes" = (P12 + P22 + P32) / 3

PC3 = Multi-item measure proportion who said "usually" = (P13 + P23 + P33) / 3

PC4 = Multi-item measure proportion who said "always" = (P14 + P24 + P34) / 3

- Communication with nurses

This multi-item measure is produced by combining responses to three questions that ask:

  - "During this hospital stay, how often did nurses listen carefully to you?"
  - "During this hospital stay, how often did nurses explain things in a way you could understand?"
  - "During this hospital stay, how often did nurses treat you with courtesy and respect?"

Respondents can answer "never," "sometimes," "usually," or "always" to each. The steps to calculate a hospital's multi-item measure score for this domain are the same as for "doctor communication."

- Responsiveness of hospital staff

This multi-item measure is produced by combining responses to two questions that ask:

[A screener question identifies patients who needed help getting to the bathroom or using a bedpan]

  - "During this hospital stay, after you pressed the call button, how often did you get help as soon as you wanted?"

- "How often did you get help in getting to the bathroom or in using a bedpan as soon as you wanted?"

Respondents can answer "never," "sometimes," "usually," or "always" to each of the two non-screener questions. The steps to calculate a hospital's multi-item measure score are the same as for "doctor communication," except that only respondents who answered "yes" to the screener question (i.e., they needed help getting to the bathroom or using a bedpan) are included in calculating the proportions for the second question. [The two questions are equally weighted in calculating the multi-item measure, because CMS views them as equally important, even though there will be fewer respondents to the second question.]

- Communication about medicines

This multi-item measure is produced by combining responses to two questions that ask:

[A screener question identifies patients who were given medicine they had not taken before during their hospital stay]

- "Before giving you any new medicine, how often did hospital staff tell you what the medicine was for?"
- "Before giving you any new medicine, how often did hospital staff describe possible side effects in a way you could understand?"

Respondents can answer "never," "sometimes," "usually," or "always" to each of the two (non-screener) questions. The steps to calculate a hospital's multi-item measure score are the same as for "doctor communication," except that only respondents who answered "yes" to the screener question (i.e., they were given medicine they had not taken before) are included in calculating the proportions.

- Cleanliness and quiet of the hospital environment

This multi-item measure is produced by combining responses to two questions that ask:

- "During this hospital stay, how often were your room and bathroom kept clean?" (note addition of quote)
- "During this hospital stay, how often was the area around your room quiet at night?"

Respondents can answer "never," "sometimes," "usually," or "always" to each. The steps to calculate a hospital's multi-item measure score are the same as for "doctor communication."

- Discharge information

This multi-item measure is produced by combining responses to two questions that ask:

[A screener question identifies patients discharged to home]

- "During your hospital stay, did hospital staff talk with you about whether you would have the help you needed when you left the hospital?"
- "During your hospital stay, did you get information in writing about what symptoms or health problems to look out for after you left the hospital?"

Respondents can answer "yes" or "no" to each. The steps to calculate a hospital's multi-item measure score are the same as for "doctor communication," except that only respondents who answered "yes" to the screener question (i.e., they were discharged to home) are included in calculating the proportions.

Patient-Mix Adjustment

Specifications 4.5 and 4.6 provide for the steps to producing raw hospital scores. Final scores shall include a patient-mix adjustment and adjustment for mode effects to better ensure the comparability of scores across hospitals—that is, the purpose of adjusting for patient mix is to estimate how different hospitals would be rated if they all provided care to comparable groups of patients.

- The following variables shall be used in the patient-mix adjustment model for HCAHPS:
  - Service Line and Gender (Female Medical, Male Medical, Female Surgical, Male Surgical, and Maternity)

- Age (specified as a categorical variable)
- Education (specified as a linear variable)
- Self-reported general health status (specified as a linear variable)
- Language other than English spoken at home
- Interaction of age by service

The patient-mix adjustment shall be a regression methodology also referred to as covariance adjustment. As an example:

Let represent the response to item i of respondent j from hospital p (after recoding, if any, has been performed). The model for adjustment of a single item i is of the form:

where is a regression coefficient vector, is a covariate vector consisting of six or more adjuster covariates (as described above), is an intercept parameter for hospital p, and is the error term. The estimates are given by the following equation:

where is the vector of intercepts, is the vector of responses, and the covariate matrix is:

where the columns of are the vectors of values of each of the adjuster covariates, and is a vector of indicators for being discharged from hospital p, p = 1, 2,…P, with entries equal to 1 for respondents in hospital p and 0 for others.

The estimated intercepts are shifted by a constant amount to force their mean to equal the mean of the unadjusted hospital means (to make it easier to compare adjusted and unadjusted means), giving adjusted hospital means:

For single-item responses, these adjusted means are reported. For composites, the several adjusted hospital means are combined using the weighted mean:

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

Sampling Protocol (from HCAHPS Quality Assurance Guidelines, V14.0)

For more details, flowchart, etc. please see the HCAHPS Quality Assurance Guidelines, V14.0, pp. 57-80, located via the

"Quality Assurance" button on the official HCAHPS On-Line Web site, at

https://www.hcahpsonline.org/globalassets/hcahps/quality-assurance/2019_qag_v14.0.pdf

Overview

We describe the process and requirements for selecting a random sample of patients to respond to the CAHPS Hospital Survey (HCAHPS). The HCAHPS sampling protocol is designed to ensure that the patients who participate in the survey are representative of all of the eligible patients who received care within general acute care hospitals. Several HCAHPS sampling protocol illustrations have been included in this chapter.

Note: The HCAHPS Survey is intended to reflect the care received by patients of all payer types, not just Medicare. Therefore, patients of all payer types are eligible for sampling.

The HCAHPS Survey sampling protocol promotes the following:

Standardized administration of the HCAHPS Survey by hospitals/survey vendors

Comparability of resulting data across all participating hospitals

The basic sampling procedure for HCAHPS requires the drawing of a random sample of eligible monthly discharges. Data will be collected from patients in each monthly sample over the 12-month reporting period, and will be aggregated on a quarterly basis to create a rolling 4-quarter data file for each hospital. The most

current four quarters of data are used for public reporting. Hospitals may not switch the type of sampling, mode of survey administration, or survey vendor used, within a calendar quarter. These types of changes can only be made at the beginning of a calendar quarter.

The HCAHPS sampling protocol employs the patient's principal diagnosis at discharge to determine whether he or she falls into one of the three service line categories eligible for HCAHPS: Maternity Care, Medical or Surgical. While V.31 Medicare Severity Diagnosis Related Group (MS-DRG) codes are the preferred method for determining the patient's service line, CMS also allows the following methodologies to be used: V.30 MS-DRG codes; V.29 MS-DRG codes; V.28 MS-DRG codes; V.27 MS-DRG codes; V.26 MS-DRG codes; V.25 MS-DRG codes; V.24 CMS-DRG codes; a mix of V.31, V.30, V.29, V.28, V.27, V.26, V.25, V.24 codes based on payer source; ICD-9 codes (ICD-10 codes anticipated to be implemented October 1, 2014); hospital unit; and New York State DRGs. The method for determining service line must be identified in the XML file, or the HCAHPS Online Data Entry Tool. (For more information see the Data Specifications and Coding chapter.)

In order to use a service line methodology other than those identified above, a hospital/survey vendor must first submit an Exceptions Request Form for approval. (For more information, see the Exceptions Request/Discrepancy Report Processes chapter.)

Proxy responses are not allowed.

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

NOTE: For the complete response, please see, "ADDITIONAL, A.1" (HCAHPS Survey, NQF 0166 Appendix A.1: Supplemental Materials.)

A minimum response rate is not imposed for HCAHPS scores. Annual training and information is provided to data collectors on how to improve response rates. Recently, a podcast was produced and posted on the official HCAHPS On-Line Web site that describes methods known to improve HCAHPS response rates: "Improving Response Rates of HCAHPS Hospital." See: https://www.hcahpsonline.org/en/podcasts/

Specify calculation of response rates to be reported with performance measure results.

HCAHPS can be implemented in four survey modes: mail, telephone, mail with telephone follow-up, or active interactive voice recognition (IVR), each of which requires multiple attempts to contact patients. Hospitals must survey patients throughout each month of the year. IPPS hospitals must achieve at least 300 completed surveys over four calendar quarters. HCAHPS is available in official English, Spanish, Chinese, Russian, Portuguese and Vietnamese versions.

The HCAHPS Survey and its protocols for sampling, data collection, coding and submission can be found in the HCAHPS Quality Assurance Guidelines, V14.0, located at the

"Quality Assurance" button on the official HCAHPS On-Line Web site, at

https://www.hcahpsonline.org/globalassets/hcahps/quality-assurance/2019_qag_v14.0.pdf

Mail Only Survey Administration

Overview

This chapter describes guidelines for the Mail Only mode of the CAHPS Hospital Survey (HCAHPS) administration.

Data collection for sampled discharged patients must be initiated between 48 hours and six weeks (42 calendar days) after discharge. Hospitals/Survey vendors must wait 48 hours to make the first attempt to contact discharged patients. This will allow enough time to pass for the patient to return home and feel settled after his or her hospital stay. Patients must not be given the survey while they are still in the hospital.

Hospitals/Survey vendors will send sampled patients a first questionnaire with a cover letter. A second questionnaire with a follow-up cover letter must be sent to all sampled patients who did not respond to the first questionnaire, approximately 21 calendar days after the first questionnaire mailing.

Note: If after the first mailing the hospital/survey vendor learns that a sampled patient is ineligible for HCAHPS, the hospital/survey vendor must not send the patient the second questionnaire. After the sample has been drawn, any patients who are found to be ineligible must not be removed or replaced in the sample. Instead, these patients are assigned a "Final Survey Status" code of ineligible (2, 3, 4, or 5; as applicable). An Administrative Data Record must be submitted for these patients.

Data collection must be closed out for a sampled patient by six weeks (42 calendar days) following the mailing of the first questionnaire. Patients who receive the HCAHPS Survey must not be offered incentives of any kind. Patients who do not respond to the survey are assigned a "Final Survey Status" code of non-response.

Hospitals/Survey vendors must record and submit lag time for all HCAHPS "Final Survey Status" codes. Additionally, hospitals/survey vendors must include the "Number Survey Attempts – Mail" field in the Patient Administrative Data Record. This field is required when "Survey Mode" in the Header Record is "1 – Mail Only." This field captures the mail wave attempt in which the final disposition of the survey is determined. More information regarding the calculation of lag time and coding of the survey attempts field is presented in the Data Specifications and Coding chapter.

Hospitals/Survey vendors must make every reasonable effort to achieve optimal survey response rates and to pursue contacts with potential respondents until the data collection protocol is completed.

No proxy respondents are permitted in the administration of the HCAHPS Survey, not even for patients who are critically ill, elderly, physically, or mentally impaired. As stated above, a proxy respondent must not answer the survey questions for the patient; however, an individual may assist the patient with reading the survey, writing responses or translation of the survey, but only the patient may provide answers to the survey.

The basic tasks and timing for conducting the HCAHPS Survey using the Mail Only mode of survey administration are summarized below.

Mail Only Survey Administration

Send first questionnaire with initial cover letter to sampled patient(s) between 48 hours and six weeks (42 calendar days) after discharge.

Send second questionnaire with follow-up cover letter to non-respondent(s) approximately 21 calendar days after the first questionnaire mailing.

Complete data collection within six weeks (42 calendar days) of the first questionnaire mailing.

To reiterate, the initial mail-out of the survey must occur between 48 hours and six weeks (42 calendar days) after discharge. Data collection must then be completed no later than six weeks (42 calendar days) after the initial mail-out. To illustrate the timing of survey mail-out, three examples are provided of patients who were discharged from a hospital on July 1.

Example Patient 1:

The first survey is mailed out on July 4 (three days after discharge)

If the patient has not returned the survey by July 25 (21 days after the initial mailing on July 4), a second survey is mailed out

Data collection must be closed out on August 15 for this patient, which is six weeks (42 days) from the July 4 initial mail-out date:

- If the survey is returned on August 15, which is the last day of the survey administration time period for this patient, then the survey is included in the final survey data file and assigned a "Final Survey Status" code of either "1 – Completed survey" or "6 – Non-response: Break off" based on the calculation of percent complete as described in the Data Specifications and Coding chapter

- Lag Time (See the Data Specifications and Coding chapter) for this patient is calculated as 45 days

- If the survey is returned after August 15 (August 16, for example), which is beyond the six weeks (42 days) survey administration time period for this patient, then the survey data are not included in the final survey data file (however, an administrative data record is submitted for this patient) and a "Final Survey Status" code of "8 ? Non-response: Non-response after maximum attempts" is assigned
  - Lag Time for this patient is calculated and entered as the number of days between the patient's discharge from the hospital and the date that data collection activities ended for this patient. Lag time for this patient is calculated as 46 days.

Example Patient 2:

The first survey is mailed out on August 12 (42 days after discharge)

If the patient has not returned the survey by September 2 (21 days after the initial mailing on August 12), a second survey is mailed out

Data collection must be closed out on September 23 for this patient, which is six weeks (42 days) from the August 12 initial mail-out date:

- If the survey is received on September 23, which is the last day of the survey administration time period for this patient, then the survey data are included in the final survey data file and assigned a "Final Survey Status" code of either "1 – Completed survey" or "6 – Non-response: Break off" based on the calculation of percent complete as described in the Data Specifications and Coding chapter
  - Lag Time for this patient is calculated as 84 days
- If the survey is received after September 23, (September 24, for example) which is beyond the six week (42 days) survey administration time period for this patient, then the survey data are not included in the final survey data file (however, an administrative data record is submitted for this patient) and a "Final Survey Status" code of "8 ? Non-response: Non-response after maximum attempts" is assigned
  - Lag Time for this patient is calculated and entered as the number of days between the patient's discharge from the hospital and the date that data collection activities ended for this patient. Lag time for this patient is calculated as 85 days.

Example Patient 3:

The first survey is mailed out on August 12 (42 days after discharge)

If the patient has not returned the survey by September 2 (21 days after the initial mailing on August 12), a second survey is mailed out

If the patient has not returned a survey by September 23, then data collection must be closed out on September 23 for this patient, which is six weeks (42 days) from the August 12 initial mail-out date:

- If the survey is received on September 23, which is the last day of the survey administration time period for this patient, and there is evidence received on September 23 that the patient is deceased (e.g., the words "deceased" written on the survey, etc.) then the survey data are not included in the final survey data file (however, an administrative data record is submitted for this patient) and the "Final Survey Status" code of "2 – Ineligible: Deceased" is assigned
  - Lag Time for this patient is calculated and entered as 84 days

Note: The timing of the survey administration protocol begins with the first mailing and does not restart if another "first mailing" is sent to the patient due to an address correction/update. Therefore, data collection must still be closed out by six weeks (42 calendar days) following the original first mailing.

Production of Questionnaire and Related Materials

The Mail Only mode of survey administration may be conducted in English, Spanish, Chinese, Russian, Vietnamese, or Portuguese. Hospitals/Survey vendors are provided with the HCAHPS questionnaires in English, Spanish, Chinese, Russian, Vietnamese, and Portuguese (HCAHPS Quality Assurance Guidelines, V14.0,

Appendices A through F), and sample initial and follow-up cover letters in English, Spanish, Chinese, Russian, Vietnamese, and Portuguese (Appendices A through F). Hospitals/Survey vendors are not permitted to make or use any other translations of the HCAHPS cover letters or questionnaires. We strongly encourage hospitals with a significant patient population that speaks Spanish, Chinese, Russian, Vietnamese, and Portuguese to offer the HCAHPS Survey in these languages. We encourage hospitals that serve patient populations that speak languages other than those noted to request CMS to create an official translation of the HCAHPS Survey in those languages.

For HCAHPS Survey administration, the OMB Paperwork Reduction Act language must appear in the mailing, either on the cover letter or on the front or back of the questionnaire. (See Appendices A through F for the exact language in English, Spanish, Chinese, Russian, Vietnamese, and Portuguese.) In addition, the OMB control number (OMB #0938-0981) must appear on the front page of the questionnaire.

To reinforce the requirement that no one other than the sampled patient completes the survey, language must be included in the questionnaire, and optionally in the cover letter(s), clearly stating that only the sampled patient may fill out the survey.

Each hospital/survey vendor will submit a sample of their HCAHPS mailing materials (questionnaires, cover letters and outgoing envelopes) with all applicable HCAHPS Quality Assurance Guidelines V13.0 updates for review by the HCAHPS Project Team.

Required for the Mail Questionnaire

The Core HCAHPS questions must be placed at the beginning of the survey. The “About You” HCAHPS questions and any hospital-specific supplemental questions must follow the Core HCAHPS questions (Questions 1-25). The order of the “About You” questions must not be altered and all the “About You” questions must remain together, even if they are placed before or after any hospital-specific supplemental questions. The “About You” questions cannot be eliminated from the questionnaire.

Hospitals/Survey vendors must adhere to the following specifications for questionnaire formatting and the production of mail materials:

Questions and Answer Categories

Question and answer category wording must not be changed

No changes are permitted in the order of the Core HCAHPS questions

No changes are permitted in the order of the “About You” HCAHPS questions, even if they are placed before or after any supplemental questions

No changes are permitted in the order of the response categories for either the Core or “About You” HCAHPS questions

The Core HCAHPS questions must remain together

The “About You” HCAHPS questions must remain together

Question and answer categories must remain together in the same column and on the same page

Response choices must be listed individually for each question, not presented in a matrix format. For example, when a series of questions is asked that have the same answer categories (Never, Sometimes, Usually, or Always), the answer categories must be repeated with every question. A matrix format which simply lists the answer categories across the top of the page and the questions down the side of the page is not allowed, because it has been shown that this format tends to produce inaccurate and incomplete responses.

Response options must be listed vertically (see examples in Appendix A). Response options that are lsted horizontally or in a combined vertical and horizontal format are not allowed.

Formatting

Wording that is underlined in the questionnaire provided in the HCAHPS Quality Assurance Guidelines must be emphasized in the same manner in the hospital’s/survey vendor’s questionnaire

Arrow (i.e., ?) placement in the questionnaire instructions and answer categories that specifies skip patterns must not be changed

Section headings (e.g., YOUR CARE FROM NURSES, etc.) must be included on the questionnaire and must be capitalized

Survey materials must be in a readable font (i.e., Arial or Times New Roman) with a font size of 10-point at a minimum

Other Requirements

All survey instructions written at the top of the questionnaire must be printed verbatim

The text indicating the purpose of the unique identifier ("You may notice a number on the survey. This number is used to let us know if you returned your survey so we do not have to send you reminders.") must be printed either immediately after the survey instructions on the questionnaire or on the cover letter, and may appear on both

Randomly generated, unique identifiers must be placed on the first or last page of the questionnaire, at a minimum. Hospitals/Survey vendors may add other identifiers on the questionnaire for tracking purposes (e.g., unit identifiers, etc.). The patient's name must not be printed on the questionnaire.

The OMB control number (OMB #0938-0981) must appear on the front page of the questionnaire

The OMB language must appear on either the front or back page of the questionnaire or on the cover letter, and may appear on both, in a readable font size at a minimum of 10-point (See Appendices A through F for the exact text in English, Spanish, Chinese, Russian, Vietnamese, and Portuguese); however, the OMB language cannot be printed on a separate piece of paper

The hospital's/survey vendor's return address must be printed on the questionnaire to make sure that the questionnaire is returned to the correct address in the event that the enclosed return envelope is misplaced by the patient

If the hospital's/survey vendor's name is included in the return address, then the hospital's/survey vendor's business name must be used, not an alias or tag line

Note: Hospitals/Survey vendors must include the following copyright statement, preferably on the last page of the survey. The text "the About You questions" may be substituted for "23-29":

"Questions 1-19 and 23-29 are part of the HCAHPS Survey and are works of the U.S. Government. These HCAHPS questions are in the public domain and therefore are NOT subject to U.S. copyright laws. The three Care Transitions Measure® questions (Questions 20-22) are copyright of Eric A. Coleman, MD, MPH, all rights reserved."

Optional for the Mail Questionnaire

Hospitals/Survey vendors have some flexibility in formatting the HCAHPS questionnaire by following the guidelines described below.

Small coding numbers, preferably in superscript, may be included next to the response choices on the questionnaire

It is acceptable to have a place on the survey for patients to voluntarily fill in their name/ telephone number as long as the name/telephone number items are placed after the Core HCAHPS questions. A transition statement must be placed before this item.

Hospital logos may be included on the questionnaire; however, other images and tag lines are not permitted

It is optional to place the title "HCAHPS Survey" on the questionnaire

The phrase "Use only blue or black ink" may be printed on the questionnaire

The name of the hospital may be printed on the questionnaire before Question 1 and in the introduction to Question 21

"Please answer the questions in this survey about your stay at [HOSPITAL NAME]. Do not include any other hospital stays in your answers."

Page numbers may be included on the questionnaire

This is encouraged as a guide to assist patients in responding to all pages of the questionnaire

Color may be incorporated in the questionnaire

The phrase "There are only a few remaining items left" before the "About You" questions may be eliminated

Language such as one of the following may be added in the footer of the survey:

- Continue on next page
- Continue on reverse side
- Turn over to continue
- to continue
- Continue on back
- Turn over

Hospitals/Survey vendors should consider incorporating the following recommendations in formatting the HCAHPS questionnaire to increase the likelihood of receiving a returned survey:

Two-column format that is used in Appendices A through F

Wide margins (at least 3/4 inch) so that the survey has sufficient white space to enhance its readability

Hospitals that choose to use their existing hospital survey in addition to the HCAHPS Survey have two options for mailing: 1) add the hospital's existing survey to the end of the HCAHPS Survey; or 2) send two separate mailings, one containing the HCAHPS Survey and another containing the hospital-specific survey.

Use of Supplemental Questions

Hospitals/Survey vendors may add a reasonable number of hospital-specific supplemental questions to the HCAHPS Survey, following the guidelines described below:

- Hospital-specific supplemental questions may be added to the HCAHPS Survey but only after all of the HCAHPS Survey questions (Questions 1-29). This approach ensures that the survey is conducted consistently across participating hospitals.

Note: Hospital-specific supplemental questions must follow the HCAHPS "About You" questions.

- Supplemental questions must be integrated into the HCAHPS Survey and not be a separate insert
- If a hospital adds supplemental questions to the HCAHPS Survey, the following statement must be placed in the survey immediately before the supplemental questions to indicate a transition from the HCAHPS questions to the hospital-specific supplemental question or questions:
- "[This next question is]/[These next questions are] from [NAME OF HOSPITAL] and [is/are] not part of the official survey."
- Hospitals may include additional transition statements following the required transition statement. Examples of allowable additional transition statements are as follows:
- "Now [NAME OF HOSPITAL] would like to gather some additional detail on topics previously examined. These items use a somewhat different way of asking for your response since they are getting at a slightly different way of thinking about the topics."
- "The following questions focus on additional care you may have received from [NAME OF HOSPITAL]."
- "This next set of questions is to provide [NAME OF HOSPITAL] additional .....

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Instrument-Based Data

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

HCAHPS is available in official English, Spanish, Chinese, Russian, Vietnamese and Portuguese versions. (German will become available in Mail Only mode in Oct. 2019). The HCAHPS Survey and its official translations can be found in the HCAHPS Quality Assurance Guidelines, V14.0, located at the

"Quality Assurance" button on the official HCAHPS On-Line Web site, at

https://www.hcahpsonline.org/en/quality-assurance/

See Appendices A - N.

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

Available at measure-specific web page URL identified in S.1

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Facility

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Inpatient/Hospital

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

N/A.

2. Validity – See attached Measure Testing Submission Form

HCAHPS_NQF_0166-_1-28-19-__Measure_Testing_form.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

Yes - Updated information is included

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 0166
Measure Title: HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems) Survey
Date of Submission: 1/7/2019; 1/24/2019; 1/28/2019

Type of Measure:

☒ Outcome (*including PRO-PM*)	☐ Composite – *STOP – use composite testing form*
☐ Intermediate Clinical Outcome	☐ Cost/resource
☐ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐ claims	☐ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☒ other: Patient survey	☒ other: Patient survey

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

The data used for all measures are from all patients who completed the HCAHPS Survey in 2016, except where noted. The data set has no specific name but consists of all HCAHPS Surveys from calendar year 2016 (21,839,718 eligible patients; 12,839,718 patients sampled; 3,109,558 completed surveys, from 4,345 hospitals across the USA).

1.3. What are the dates of the data used in testing?

The data used for all measures are from 1/1/2016 to 12/31/2016, except where noted.

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	Measure Tested at Level of:
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☐ other:	☐ other:

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

For all measures: more than 4,200 hospitals from across the USA, except where noted.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

Number and Percent of Completed HCAHPS Surveys by Patients' Age, Race/Ethnicity, Education, Language Spoken at Home and Service Line, 2016 (~4,300 Hospitals)

Population Group		TOTAL	
		N	%
Age	18-44	623,421	20.00%
	45-64	892,322	28.63%
	65 and over	1,599,342	51.31%
	Missing	1,778	0.06%
Race/Ethnicity	Non-Hispanic White	2,161,529	69.35%
	Hispanic	291,421	9.35%
	Black	244,519	7.85%
	Asian	73,808	2.37%
	AI/AN	21,625	0.69%
	NHOPI	7,540	0.24%
	Multiracial	95,593	3.07%
	Missing	220,828	7.08%
Education	Less than high school	370,359	11.88%
	High school graduate	894,349	28.69%
	At least some college	1,647,872	52.87%
	Missing	204,283	6.55%
Language at home	English	2,676,451	85.87%
	Spanish	140,753	4.52%
	Chinese	10,091	0.32%
	Russian	4,898	0.16%
	Vietnamese	4,163	0.13%
	Other	40,823	1.31%
	Missing	239,684	7.69%

Population Group		TOTAL	
		N	%
Service Line	Maternity	390,231	12.52%
	Medical	1,480,822	47.51%
	Surgical	1,121,136	35.97%
	Missing	124,674	4.00%
Total		3,116,863	100.00%

Key: AI/AN: American Indian or Alaska Native; Black: Black or African American; Multiracial: Multiple races; NHOPI: Native Hawaiian or other Pacific Islander

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

The data used for most of the submission are from CY 2016. We also cite information from the original testing of HCAHPS in the Three-State Pilot Study of 2003 in section 2b1.3 and analyses related to preferred language and racial/ethnic groups using data from 2014-2015.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

For the HCAHPS Survey, we collect patient-reported data on level of education (8th grade or less; Some high school, but did not graduate; High school graduate or GED; Some college or 2-year degree; 4-year college graduate; More than 4-year college degree) and language spoken at home (English; Spanish; Chinese; Russian; Vietnamese; Portuguese). In the HCAHPS patient-mix adjustment model, Education is specified by the six levels given above, while Language Spoken at home is specified as follows: English; Spanish; Chinese; and, Russian/Vietnamese/Portuguese/Other.

Additionally, we have analyzed differences in patient experiences of hospital care by preferred language within racial/ethnic groups using HCAHPS surveys from 2014-2015. However, race/ethnicity is not used for official patient-mix adjustment of reported outcomes.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☒ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

Outcomes were top-box (most positive response, except 9 or 10 for 0-10 responses) or full-scale patient-level scores for evaluative items or multiple-item measures, as applicable.

ICCs are the ratio of between-hospital to total patient-level variance. The Spearman-Brown Formula was used to convert ICCs into estimated reliabilities at the specified sample size. Variance component models were used

to estimate hospital-level reliability of top-box scores for 10 HCAHPS measures at N=300 completed surveys (as HCAHPS recommends). Mean scores based on the full response scales were also used to test hospital-level reliability, where noted.

Cronbach's alpha (internal consistency) estimates of reliability were calculated for the 6 multi-item measures.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

Hospital-level reliabilities of the 10 HCAHPS measure top-box scores ranged from 0.83 (Communication about Medicines and Discharge Information) to 0.93 (Quietness and Recommend Hospital). That is, all 10 exceeded the threshold of 0.80 (good reliability, with ~0.70 "adequate"), and 8 of 10 (all but Communication about Medicines and Discharge Information) exceeded the very good/0.85 standard.

Hospital-level reliabilities of 10 HCAHPS measure mean scores ranged from 0.83 to 0.93. All 10 exceeded the threshold of 0.80 and 9 out of 10 (all but Discharge Information) exceeded the very good/0.85 standard. Please see below.

Hospital-Level Reliability

Hospital-Level Reliability: The relative amount of true variation in performance, compared to all variation, including measurement error.

- *Unit or hospital-level reliability (Spearman-Brown reliability) is the proportion of variance in hospital scores that reflect true variation among hospitals, compared to total variation*

Hospital-Level Reliabilities of HCAHPS Measure Top-Box Scores at 300 Completed Surveys, 3.1 million completed surveys from 2016 discharges

Communication with Nurses	0.87
Communication with Doctors	0.86
Responsiveness of Hospital Staff	0.92
Communication about Medicines	0.83
Cleanliness	0.87
Quietness	0.93
Discharge Information	0.83
Overall Rating	0.91
Recommend Hospital	0.93
Care Transition	0.86

- *Standards:*
  - *0.7: Adequate (all exceed this)*
  - *0.8: Good (all exceed this)*
  - *0.85 Very good (8 of 10 reach this)*
  - *0.9: Excellent (4 of 10 reach this)*

Hospital-Level Reliabilities of HCAHPS Measure Linear Mean Scores at 300 Completed Surveys, 3.1 million completed surveys from 2016 discharges

Communication with Nurses	0.87
Communication with Doctors	0.86
Responsiveness of Hospital Staff	0.93
Communication about Medicines	0.85
Cleanliness	0.88
Quietness	0.93
Discharge Information	0.83
Overall Rating	0.92
Recommend Hospital	0.93
Care Transition	0.87

Many hospitals have more than 300 completes over the measurement period. HCAHPS reliability is even higher for those hospitals with greater than 300 completes than the values shown for 300 completes.

Below are the Hospital-Level Reliabilities for both Top-Box and Mean Scores of the individual items that make up each of the multi-item measures. Theses reliabilities are for review purposes only, since the individual-level items are not publicly reported.

Hospital-Level Reliabilities of Top-Box Scores of the HCAHPS Survey Items that Are Combined to Form HCAHPS Measures, estimated for 300 Completed Surveys, based on informaiont from 3.1 million completed surveys of patients discharged in 2016

Nurse Courtesy Respect (Q1)	0.81
Nurse Listen (Q2)	0.83
Nurse Explain (Q3)	0.78
Dr Courtesy Respect (Q5)	0.78
Dr Listen (Q6)	0.82
Dr Explain (Q7)	0.81
Call Button (Q4)	0.91
Bathroom Help (Q11)	0.89
Meds For (Q16)	0.74
Side Effects (Q17)	0.80
Help After Discharge (Q19)	0.81
Symptoms (Q20)	0.75
Care Transition Preferences (Q23)	0.83
Care Transition Understanding (Q24)	0.81
Care Transition Med Purpose (Q25)	0.80

Hospital-Level Reliabilities of Linear Mean Scores of the HCAHPS Survey Items that Are Combined to Form HCAHPS Measures, estimted for 300 Completed Surveys, based on 3.1 million completed surveys of patients discharged in 2016

Nurse Courtesy Respect (Q1)	0.83
Nurse Listen (Q2)	0.84
Nurse Explain (Q3)	0.82
Dr Courtesy Respect (Q5)	0.78
Dr Listen (Q6)	0.83
Dr Explain (Q7)	0.83
Call Button (Q4)	0.93
Bathroom Help (Q11)	0.90
Meds For (Q16)	0.77
Side Effects (Q17)	0.84
Help After Discharge (Q19)	0.81
Symptoms (Q20)	0.75
Care Transition Preferences (Q23)	0.85
Care Transition Understanding (Q24)	0.82
Care Transition Med Purpose (Q25)	0.76

Please Note: *CMS does not currently publicly report the individual items that comprise the multi-item measures. This item-level data is not a measure but only a portion of a multi-item measure.*

Internal Consistency Reliability of Multi-item measures

Internal consistency reliability coefficients (Cronbach's alpha) are displayed for each of the six multi-item measures in the table below.

Three of six multi-item measures had internal consistency reliability estimates of 0.80 or higher (Communication with Nurses, Communication with doctors) and three had estimates of 0.68-0.69 (Responsiveness of Hospital Staff, Communication about Medicines, Discharge Information).

Cronbach's Alpha for HCAHPS multi-item measures, 2016 discharges.

	Alpha & Item-Total Correlation
Communication with Nurses	**α = 0.81**
Nurse Courtesy Respect (Q1)	0.74
Nurse Listen (Q2)	0.70
Nurse Explain (Q3)	0.79
Communication with Doctors	**α = 0.85**
Doctor Courtesy Respect (Q5)	0.81
Doctor Listen (Q6)	0.75
Doctor Explain (Q7)	0.82
Responsiveness of Hospital Staff	**α = 0.69**
Call Button (Q4)	0.52
Bathroom Help (Q11)	0.52
Communication about Medicine	**α = 0.68**
Medicine Explain (Q16)	0.52
Side Effects (Q17)	0.52
Care Transition	**α = 0.80**
Preferences (Q23)	0.73
Understanding (Q24)	0.78
Medicine Purpose (Q25)	0.64
Discharge Information	**α = 0.69**
Help After Discharge (Q19)	0.53
Symptoms (Q20)	0.53

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

The results demonstrate that all 10 HCAHPS measures (6 multi-item measures and 4 single-items measures) reliably distinguish hospital performance and are sufficiently reliable for hospital-level public reporting and performance incentives at the recommended 300 completed surveys using top-box and linear-mean scoring.

The 15 individual items ("data elements") that are combined to form the 6 HCAHPS multi-item measures are also sufficiently reliable.

Finally, 3 of 6 HCAHPS multi-item measures demonstrate good (0.80-0.85) internal consistency reliability and 3 demonstrate adequate (0.68-0.69) internal consistency reliability.

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☒ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☒ **Empirical validity testing**

☐ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

We conducted analyses to assess construct validity, testing hypotheses that specific multi-item measures and individual items within the HCAHPS Survey would have moderate, statistically significant positive association with the global measures of patient experience of care: the global hospital rating and the recommendation of hospital to friends and family items. In both the 2003 Three-State Pilot Study and also using completed surveys from all discharges from July 2016 through June 2017, we performed linear regression analyses that predicted the global rating and recommend items from individual multi-item measures and individual item measures using patient level data. We also examined correlations at the hospital level using the 2016-17 data. We found moderate correlations with rating and recommend at the patient level (generally in the 0.3 to 0.4 range) and stronger correlations at the hospital level (0.6 to 0.8); all were statistically significant.

Please see below.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Using the 2003 Three-State Pilot Study data, we conducted hospital level factor analysis to look for distinct constructs that corresponded to the proposed multi-item measures. We found evidence of multiple factors and factor loadings that corresponded to the groupings used for the multi-item measures.

We also conducted patient level discriminant analysis using all 2016-17 discharges, which first involved the patient-level inter-item correlation matrix. Then for every pair of multi-item measures, we tested the hypothesis that the average inter-item correlation was higher within than between multi-item measures. This was confirmed at 0.1 – 0.3 units.

HCAHPS Survey Construct Validity.

Note: From the HCAHPS Three-State Pilot Study, 2003; pp. 4-5 to 4-9. Referenced tables and appendices can be found in the appendices at http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalHCAHPS.html

The HCAHPS Three-State Pilot Study was conducted in 2003. 109 hospitals participated, and 16,691 patients completed the HCAHPS Survey. (Please note: in the Three-State Pilot Study, including Tables 5-13 below, we refer to multi-item scores as "composites," but these are in fact "multi-item measures").

The last four columns of **Table 5** display the relationships of items and composites to both the global rating of the hospital and whether the patient would recommend the hospital to family and friends. Composite scores were calculated as the average of the item scores for that composite. The adjusted R-square (aR2; columns 5 and 7) was the proportion of variance in hospital ratings or recommendations accounted for by the corresponding composite items, taking into account the number of items. The rankings of composites in terms of their relationships to both the global hospital rating (Q52) and the tendency to recommend the hospital (Q53) were (from most to least related): "communication with nurses," "physical comfort," "communication with doctors, "communication about medication," and "discharge information." Multivariate analyses were conducted to determine the unique relationship of each composite, net of the other five, and the unique relationship of each item, net of the other 32, to both criteria. T-values for the regression parameters are displayed in columns 6 and 8 of **Table 5** (Note: tables from the Three-State Pilot Study are found below). We present the t values, noting p-values that are greater than 0.01.

The relationship of items to each criterion was consistent. Within "communication with nurses," the most highly related items had to do with nurses listening to patients and treating them with courtesy and respect. Within "physical comfort," the most highly related items had to do with whether the hospital room was kept clean and whether patients' visitors were helped.

Table 6 displays the correlations of composites with the three 2003 global ratings and patients' reported likelihood of recommending the hospital. While global ratings of doctors and nurses were not retained in HCAHPS, these analyses provide additional evidence of the validity of HCAHPS composites. These correlations

are ordered from highest to lowest. Not surprisingly, the "communications with nurses" composite had the strongest relationship to the global nurse's rating, while the "communications with doctors" composite had the strongest relationship to the global doctor's rating. The two composites most highly related to the global ratings of the hospital and nursing care as well as the probability of recommending the hospital were the "communications with nurses" and the "physical comfort" composites. The two composites with the weakest relationships to the global ratings and the hospital recommendation were "communication about medication" and "discharge information."

The results summarized above focus on internal consistency reliability and correlations with global ratings and willingness to recommend to family and friends. These analyses were conducted to evaluate the items and composites in case composite score algorithms are created at the individual patient level. The main purpose of these analyses was to aid the HCAHPS Analysis Team in their charge to identify items that could be deleted from the pilot study questionnaire for the purpose of shortening the HCAHPS survey before it was implemented nationally. Below are results of a hospital-level factor analysis to identify those composites indicated by hospital-level data. Data accumulated across patients within a hospital provides information about the reliability of the items and composites with regard to measuring care at the hospital level (how well items and composites differentiate between hospitals) and the extent to which items vary by service type.

2016 Updates to Correlational Analyses

As an update to the Three-State Pilot Study, we present the most recent inter-correlations of HCAHPS measures (July 2016 to June 2017 discharges), as well as the correlation of the measures with two overall measures of hospital experience: the global hospital rating and the recommendation of hospital to friends and family items. We also, on pp. 17-18, present updated evidence of the factor structure of HCAHPS.

In the table below, the correlations of the HCAHPS measures with the overall measures range from 0.28 to 0.64. Communication with Nurses measure has the strongest correlation with the overall measures, while Discharge Information has the weakest.

Note: Correlations among the HCAHPS measures are reported on an annual basis. Please see the HCAHPS On-Line Web site for current and historical patient level inter-item correlations. The most recent correlations can be found at: https://www.hcahpsonline.org/globalassets/hcahps/summary-analyses/correlations/report_april_2018_corrs_pain_removed.pdf.

Patient-Level Correlations between HCAHPS Measures, July 2016 – June 2017 discharges

	Communication with Nurses	Communication with Doctors	Responsiveness of Hosp. Staff	Comm. About Medicines	Cleanliness of Hospital Env.	Quietness of Hospital Env.	Discharge Information	Care Transition	Hospital Rating	Recommend the Hospital
Communication with Nurses	1	0.53	0.56	0.50	0.39	0.32	0.27	0.43	**0.64**	**0.58**
Communication with Doctors		1	0.38	0.44	0.27	0.26	0.28	0.40	**0.52**	**0.47**
Responsiveness of Hosp. Staff			1	0.41	0.35	0.32	0.20	0.35	**0.51**	**0.45**
Comm. About Medicines				1	0.33	0.29	0.35	0.45	**0.48**	**0.43**
Cleanliness of Hospital Env.					1	0.28	0.18	0.27	**0.41**	**0.36**
Quietness of Hospital Env.						1	0.13	0.25	**0.35**	**0.29**
Discharge Information							1	0.30	**0.30**	**0.28**
Care Transition								1	**0.47**	**0.45**
Hospital Rating									1	0.76
Recommend the Hospital										1

Next, in the table below, we examine the correlations of the individual survey questions (items) that comprise the multi-item measures with the two overall measures of hospital experience: the global hospital rating and the recommendation of hospital to friends and family items. At the patient level, these correlations range from 0.17 to 0.44. The "Nurse Listen" and "Nurse Courtesy Respect" items have the strongest correlation with the global measures, while the item that asks whether the patient received information about symptoms to look out for after discharge ("Symptoms") has the weakest correlation. At the hospital level, these correlations range from 0.46 to 0.75. The items that ask whether hospital staff took the patient's preferences into account ("Care Transition Preferences") and whether the patient had a good understanding of managing health care needs ("Care Transition Understanding") had the strongest correlation with the two overall measures, while the item that asked whether staff described side effects to new medications to the patient had the weakest correlation.

Correlations between HCAHPS Items within Multi-item Measures and Global Measures at the Patient Level and Hospital Levels (300 or more completed surveys), July 2016 – June 2017 discharges

	Patient-Level		Hospital-Level	
	Rating	Recommend	Rating	Recommend
Nurse Courtesy Respect (Q1)	0.42	0.40	0.75	0.64
Nurse Listen (Q2)	0.44	0.40	0.74	0.61
Nurse Explain (Q3)	0.38	0.36	0.74	0.64
Dr Courtesy Respect (Q5)	0.33	0.32	0.74	0.67
Dr Listen (Q6)	0.36	0.34	0.71	0.64
Dr Explain (Q7)	0.33	0.32	0.69	0.63
Call Button (Q4)	0.35	0.32	0.66	0.51
Bathroom Help (Q11)	0.34	0.32	0.66	0.56
Meds For (Q16)	0.32	0.31	0.67	0.60
Side Effects (Q17)	0.32	0.30	0.56	0.46
Help After Discharge (Q19)	0.20	0.20	0.61	0.56
Symptoms (Q20)	0.17	0.18	0.67	0.65
Care Transition Preferences (Q23)	0.35	0.36	0.84	0.82
Care Transition Understanding (Q24)	0.35	0.37	0.81	0.82
Care Transition Med Purpose (Q25)	0.32	0.33	0.75	0.76

Hospital-Level Factor Analysis (2003)

Much of the variation among respondents' HCAHPS scores is due to individual variation, reflecting characteristics and particular experiences of individual respondents, rather than systematic differences among hospitals. The way these individual characteristics and experiences are related across items is not necessarily the same as the way that different aspects of hospital quality are related. Because only a fraction of the variation in an individual's responses is attributable to the hospital, removing individual-level variability can yield a clearer picture of the relationship between different aspects of quality.

However, to model hospital-level correlations among items it is necessary to account for variation (of individual items) and covariation (of pairs of items) in the data due to sampling. This is done by fitting models relating the sample mean to the sample variance or covariance for each pair of items. Independent models were fitted for each service, yielding service-specific hospital-level covariance matrices. In addition, we pooled the data across services and then fit the same series of models to obtain a consensus hospital-level covariance matrix. Details of both steps of this modeling process are described in Appendix A.

We investigated the structure underlying the between-hospital covariance matrices and estimated the number and structure of possible factors. The principle factor method with squared multiple correlations as initial estimates of communalities and Promax rotation with Kaiser normalization were used to examine the structure of each covariance matrix.

It appeared that four factors were appropriate for each service. The pattern coefficients and patterns of the highest coefficients for each factor for each service are displayed in **Table 7**. The factors are not necessarily ordered by the variance explained by each (magnitude of the corresponding eigenvalues) because they were re-arranged to allow for easy comparison across service types.

The hospital-level factor structures for the "surgery" and "childbirth" services are very similar with the highlighted composite items containing almost the same groups of items. However, the structure for "medical

service" is clearly very different; there only appears to be one dimension of quality (Factor 2) that is consistent with those for "surgery" and "childbirth." Therefore, one might consider grouping the "surgery" and "childbirth" services together for multivariate analysis, and treating "medical service" separately. This would make sense as "surgery" and "childbirth" both involve treatments with a discrete objective, typically leading to a positive outcome; patient experiences on those services might be very different from those for patients on a medical service—that is, to be related to a chronic or emergent condition.

We compared possible hospital-level composite item groupings to the composites found in the individual-level factor analysis. The names of individual-level composites are enclosed in quotes in the following. Factor 2, common to the three services, is the individual-level composite "communication about medication" factor; Factor 3 of "surgery" and "childbirth" is the "communication with doctors" factor; Factor 4 of "surgery" and "childbirth" is close to the "discharge information" factor; and Factor 3 of "other medical service" partly resembles the "pain control" factor. The "physical comfort" factor is subsumed in Factor 1 of "surgery" and "childbirth," and split over Factors 3 and 4 of "other medical service."

We estimated three solutions with five, six, and seven factors, respectively, as the best candidates for the structure of the hospital-level covariance matrix. These five-, six-, and seven-factor solutions are displayed in **Tables 8-10**. Additional statistics for the seven-factor solution are presented in **Tables 11-13**.

The factor analysis with five factors has very high eigenvalues for each of the five factors (after rotation), while the magnitude of the eigenvalues for the 6th and 7th factors in the six and seven factor solutions are much lower (indicating that these additional factors explain less variance). The six factor solution extracts Q17 (room clean) and Q16 (temperature) items from Factor 1 of the five-factor solution to form, together with Q18 (room quiet), a "physical environment" factor. The seven-factor solution additionally extracts Q32 (pain controlled) and Q33 (MD pain help) to form a shortened version of the "pain control" factor. The hospital-level factor analyses separate Q20 (how often bathing) and Q22 (how often bathroom) from the "physical comfort" factor and combine these with Q9 (call for help), to form a "nursing services" factor. The "housekeeping" and "nursing services" factors are both contained in the individual level "physical comfort" factor.

Referring exclusively to the seven-factor solution, Factor 1 might be thought of as a "concern for patient/communication with nurses" factor, Factor 2 as "communication with doctor," Factor 3 as "communication about medication," Factor 4 as "nursing services," Factor 5 as "discharge information," Factor 6 as "pain control" and Factor 7 as "physical environment."

Note: The HCAHPS Survey originally included three items about pain (a screener and two substantive questions), which formed a multi-item measure, Pain Management. CMS discontinued publicly reporting this measure in July 2018. In January 2018, CMS replaced the original HCAHPS pain items with three items that asked about communication about pain. In compliance with the **Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act** (Pub. L. 115-271) of 2018 (Section 6104), CMS will remove the new communication about pain items from the HCAHPS Survey beginning with October 2019 discharges.

Tables 5 to 13 below are part of the HCAHPS Three-State Pilot Study, 2003

(Please note: in the Three-State Pilot Study, including Tables 5-1, we refer to multi-item scores as "composites," but these are in fact "multi-item measures").

Table 5: Correlations of Items with Six Composites, Global Rating of Hospital, and Likelihood of Recommending the Hospital to Family and Friends

Item	Question Label	Integrity of Composites		Relationship of Item and Composite-Level Scores to Hospital Rating and Recommendation			
		Substantial Corr w 2nd Scale	Alpha & Item-Total Correlation	Hosp Rating		Recommended Hosp	
(1) Physical Comfort			α =.81	aR^2=0.44*	t-value=54.18	aR^2=0.33	t-value=37.09
Q20	How often bathing		0.64		6.44		4.56
Q22	How often bathroom	4(.62)	.64		3.75		3.44
Q17	Room clean		.54		31.18		21.43
Q18	Room quiet		.46		11.78		3.34
Q16	Temperature		.43		14.17		6.94
Q24	Privacy		.55		5.28		6.04
Q27	Visitors help	4(.57)	.56		26.29		21.05
(2) Communication with Doctors			α =.88	aR^2=0.31	t-value=31.14	aR^2=0.25	t-value=28.72
Q12	MD listen		.80		10.80		6.88
Q11	MD respect		.73		16.60		15.69
Q13	MD explain things		.76		-4.19		-2.79
Q14	MD enough time		.73		-0.33 #		-0.48 #
Q25	Involve in Tx decisions	1,4 (.53)	.55		7.33		9.85
(3) Communication about Medication			α =.85	aR^2=0.21	t-value=7.82	aR^2=0.16	t-value=4.93
Q39	Taking other Rx		.67		2.29 #		1.31 #
Q38	Purpose of Rx		.70		-2.67		-2.88
Q37	Name of Rx		.67		-1.37#		-0.63#
Q40	Allergic to Rx		.63		5.68		4.13
Q41	Side-effects of Rx		.64		4.43		3.28
(4) Communication with Nurses			α =.88	aR^2=0.50	t-value=93.42	aR^2=0.38	t-value=71.23
Q5	RN listen		.79		30.10		21.66
Q4	RN respect		.73		39.90		33.02
Q6	RN explain things		.71		12.59		8.66
Q7	RN enough time		.72		17.19		12.15
Q9	Help when call button		.66		15.08		11.31
(5) Pain Control			α =.80	aR^2=0.33	t-value=23.66	aR^2=0.26	t-value=22.77
Q32	Pain controlled		.70		5.13		3.37
Q33	Pain help all can		.74		12.43		11.63
Q31	Pain respond quick		.70		4.26		1.70#
Q35	Tests w/o pain	1 (.35) 4 (.34)	.36		0.55 #		6.69
(6) Discharge Information			α =.68	aR^2=0.10	t-value=17.08	aR^2=0.08	t-value=15.22
Q49	Symptoms may have		.55		6.49		8.34
Q47	Activities can't do		.53		1.88#		0.33#
Q48	Help 4u @ home?		.41		9.69		7.16
Q51	Meds – how take?		.38		5.73		2.56#

Item	Question Label	Integrity of Composites		Relationship of Item and Composite-Level Scores to Hospital Rating and Recommendation			
		Substantial Corr w 2nd Scale	Alpha & Item-Total Correlation	Hosp Rating		Recommended Hosp	
Other							
Q28	Introduce self		N/A		11.99		7.07
Q43	Delays in admission		N/A		21.96		20.30
Q44	Living will		N/A		0.50#		5.31

p > 0.01

Table 6: Rank Order (Descending) of Correlations of Six Composites with Global Ratings

Hospital Rating		Nurses Rating		Doctors Rating		Recommend Hospital	
Nurses	.70	Nurses	.81	Doctor	.80	Nurses	.63
Physical Comfort	.65	Physical Comfort	.65	Nurses	.49	Physical Comfort	.63
Doctor	.60	Pain Control	.46	Physical Comfort	.48	Pain Control	.56
Pain Control	.56	Doctor	.48	Pain Control	.46	Doctor	.48
Medication	.47	Medication	.46	Medication	.40	Medication	.45
Discharge Info	.30	Discharge Info	.28	Discharge Info	.28	Discharge Info	.30

Note: Nurses = Communications with nurses; Doctor = Communications with doctors; Medication = Communication about medication.

Table 7: Hospital Level Factor Analysis Stratified by Service

(Note: Rows were ordered first by Surgery, then by Childbirth)

Item	Surgery				Child birth				Other			
	Factor1	Factor2	Factor3	Factor4	Factor1	Factor2	Factor3	Factor4	Factor1	Factor2	Factor3	Factor4
Call Help Frequently (Q9)	0.86	-0.03	-0.08	0.03	0.62	0.10	0.03	0.15	0.41	0.03	0.13	0.34
Spend Enough Time (Q7)	0.84	-0.06	0.01	0.04	0.72	-0.10	0.15	0.09	0.39	-0.05	0.17	0.47
Respect (Q4)	0.83	0.02	-0.07	-0.02	0.80	-0.05	0.05	0.19	0.30	-0.02	0.43	0.30
Listen (Q5)	0.81	-0.04	0.02	0.08	0.88	0.00	0.04	-0.06	0.38	0.00	0.28	0.36
Visitors Help (Q27)	0.70	0.00	0.07	-0.05	0.78	-0.03	0.10	0.01	0.34	0.06	0.32	0.27
MD Pain Help (Q33)	0.63	0.13	0.14	-0.01	0.66	-0.06	0.27	0.12	0.32	0.04	0.20	0.40
How Often Bathroom (Q22)	0.62	-0.01	-0.03	0.14	0.57	0.21	-0.09	-0.20	0.15	0.02	0.13	0.35
Explain Things (Q6)	0.57	0.10	0.14	0.10	0.78	-0.08	0.20	-0.06	0.47	0.16	0.15	0.19
How Often Bathing (Q20)	0.53	0.16	-0.11	0.09	0.56	0.18	-0.05	-0.14	0.08	-0.03	0.12	0.45
Pain Controlled (Q32)	0.51	0.21	0.20	-0.18	0.55	0.04	0.23	0.10	0.26	-0.09	0.14	0.47
Room Quiet (Q18)	0.48	0.03	0.00	-0.06	0.42	0.05	-0.19	0.26	0.13	-0.08	0.61	-0.08
MD Respond Pain (Q31)	0.48	0.21	-0.03	0.07	0.65	0.15	0.08	-0.01	0.20	0.15	0.35	0.14
Room Clean (Q17)	0.44	0.27	0.01	-0.19	0.64	-0.03	-0.29	0.26	-0.10	0.12	0.48	0.18
Introduce (Q28)	0.40	0.26	0.11	-0.11	0.55	-0.11	0.04	0.30	0.18	0.24	0.37	-0.01
Purpose of TX (Q38)	-0.06	0.85	0.00	-0.01	0.41	0.63	-0.04	-0.18	0.09	0.77	0.10	-0.02
Name of TX (Q37)	-0.03	0.78	0.00	0.05	0.43	0.40	0.11	0.09	0.09	0.80	0.06	-0.02
Allergic to RX (Q40)	0.06	0.75	-0.09	0.08	-0.02	0.83	0.02	0.00	-0.01	0.71	0.15	-0.02
Taking other RX (Q39)	0.02	0.73	0.00	0.12	-0.15	0.90	0.10	0.07	-0.03	0.78	-0.02	0.07
RX Side Effects (Q41)	0.03	0.68	0.05	0.08	0.15	0.65	-0.09	0.26	0.09	0.60	-0.14	0.27
Temperature (Q16)	0.18	0.39	0.04	-0.17	0.22	0.08	0.01	0.39	-0.11	-0.03	0.77	0.02
MD Listen (Q5)	-0.04	0.00	0.91	0.03	-0.11	0.04	0.98	0.01	0.95	0.05	-0.14	0.03
MD Respect (Q11)	0.05	-0.08	0.82	-0.13	0.08	0.04	0.79	0.10	0.75	0.07	0.27	-0.31
MD Explain (Q13)	-0.03	0.09	0.77	0.06	-0.03	-0.08	0.80	0.16	0.93	0.07	-0.14	-0.02
MD Enough Time (Q14)	0.06	0.00	0.70	0.06	0.15	0.03	0.70	0.00	0.98	-0.06	-0.08	-0.04
TX Decisions (Q25)	0.32	0.05	0.44	0.13	0.53	-0.06	0.46	-0.08	0.46	0.20	-0.19	0.39
Writing Activities (Q47)	0.01	-0.16	0.05	0.69	-0.08	0.02	0.01	0.83	-0.11	0.29	-0.09	0.45
Help After Discharge (Q48)	0.07	0.08	-0.02	0.64	0.14	0.10	0.05	0.33	-0.12	0.09	-0.04	0.63
Writing Symptoms (Q49)	-0.06	0.06	0.08	0.55	-0.11	-0.01	0.05	0.76	-0.09	0.35	-0.15	0.54
Living Will (Q44)	-0.06	0.03	-0.07	0.47	0.21	-0.18	-0.02	0.23	-0.13	0.04	0.55	-0.09
Tests Without Pain (Q35)	0.25	0.09	0.22	0.07	0.12	0.06	0.53	-0.27	0.32	-0.03	-0.04	0.40
Writing How to take Rx (Q51)	-0.02	0.15	0.02	0.34	-0.09	0.13	0.06	0.59	0.01	0.27	0.00	0.30
Delays in Admission (Q43)	0.29	-0.16	0.07	-0.01	0.11	-0.02	-0.07	0.31	0.33	-0.05	-0.09	0.27
Privacy (Q24)	0.25	0.00	-0.01	-0.03	0.28	-0.01	0.07	-0.13	0.25	-0.03	0.01	0.13

Table 8: Factor Analysis with Five Factors.

Item	Factor1	Factor2	Factor3	Factor4	Factor5
Respect (Q4)	0.83	-0.03	-0.03	0.10	-0.04
Listen (Q5)	0.78	0.03	0.02	0.14	-0.08
Explain Things (Q6)	0.63	0.19	0.07	0.07	-0.04
MD Pain Help (Q33)	0.63	0.15	-0.02	0.11	0.07
Introduce (Q28)	0.62	0.05	0.08	-0.11	0.06
Spend Enough Time (Q7)	0.59	0.07	-0.06	0.32	0.00
Pain Controlled (Q32)	0.58	0.08	0.02	0.12	0.00
Visitors Help (Q27)	0.56	0.12	-0.01	0.24	-0.02
Room Clean (Q17)	0.52	-0.14	0.08	0.09	0.05
Living Will (Q44)	0.52	-0.01	-0.08	-0.38	0.09
Temperature (Q16)	0.49	0.03	0.03	-0.10	0.15
MD Listen (Q5)	0.02	0.90	0.02	-0.02	0.02
MD Explain (Q13)	-0.02	0.82	0.03	0.02	0.06
MD Enough Time (Q14)	0.03	0.80	-0.02	0.09	-0.03
MD Respect (Q11)	0.25	0.76	-0.02	-0.19	-0.02
TX Decisions (Q25)	0.13	0.37	0.05	0.29	0.13
Taking Other RX (Q39)	0.02	-0.01	0.87	-0.10	-0.03
Allergic to RX (Q40)	0.13	-0.05	0.84	-0.14	-0.05
Purpose of TX (Q38)	-0.10	0.08	0.72	0.22	-0.01
Name of TX (Q37)	0.04	0.09	0.63	0.14	0.02
RX Side Effects (Q41)	-0.06	-0.04	0.63	0.19	0.15
How Often Bathroom (Q22)	0.09	-0.05	0.05	0.62	-0.04
How Often Bathing (Q20)	0.04	-0.06	0.05	0.62	0.03
Call Help FreQuently (Q9)	0.32	0.02	0.03	0.48	0.03
Writing Activities (Q47)	0.03	0.02	-0.02	0.00	0.70
Writing Symptoms (Q49)	-0.04	0.02	0.03	0.07	0.69
Help After Discharge (Q48)	0.08	-0.03	-0.04	0.18	0.49
Writing How to take Rx (Q51)	0.26	0.00	0.14	-0.22	0.41
Room Quiet (Q18)	0.32	-0.06	0.02	0.14	0.10
Privacy (Q24)	0.21	0.08	0.02	0.06	-0.07
MD Respond Pain (Q31)	0.33	0.03	0.11	0.33	0.01
Tests Without Pain (Q35)	0.09	0.26	0.06	0.25	-0.07
Delays In Admission (Q43)	-0.13	0.20	-0.14	0.33	0.21

Table 9: Factor Analysis with Six Factors

Item	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
Respect (q4)	0.88	-0.03	-0.06	-0.03	0.03	0.07
Listen (q5)	0.85	0.01	-0.01	-0.07	0.06	0.05
Explain Things (q6)	0.76	0.07	0.12	-0.01	-0.02	-0.06
MD Pain Help (q33)	0.73	-0.02	0.09	0.10	0.03	-0.03
Spend Enough Time (q7)	0.69	-0.06	0.03	0.02	0.24	0.01
Pain Controlled (q32)	0.64	0.01	0.05	0.01	0.05	0.02
Visitors Help (q27)	0.63	-0.01	0.08	-0.01	0.17	0.04
Introduce (q28)	0.59	0.08	0.04	0.06	-0.14	0.10
Living Will (q44)	0.48	-0.08	-0.02	0.11	-0.32	0.05
MD Respond Pain (q31)	0.37	0.11	0.02	0.00	0.28	0.07
Taking Other RX (q39)	-0.03	0.88	0.00	-0.04	-0.09	0.06
Allergic to RX (q40)	0.04	0.84	-0.03	-0.06	-0.12	0.12
Purpose of TX (q38)	-0.02	0.72	0.05	0.01	0.18	-0.07
RX Side Effects (q41)	-0.14	0.63	0.00	0.13	0.20	0.13
Name of TX (q37)	0.21	0.63	0.01	0.07	0.07	-0.20
MD Listen (q5)	0.06	0.02	0.87	0.01	-0.01	-0.01
MD Explain (q13)	0.02	0.04	0.80	0.05	0.03	0.00
MD Enough Time (q14)	0.07	-0.02	0.78	-0.05	0.09	0.03
MD Respect (q11)	0.20	-0.02	0.77	-0.05	-0.16	0.11
Writing Symptoms (Q49)	-0.06	0.03	0.00	0.73	0.06	-0.02
Writing Activities (Q47)	-0.02	-0.02	0.01	0.72	-0.01	0.02
Help After Discharge (Q48)	0.07	-0.04	-0.03	0.50	0.16	0.03
Writing How to take Rx (Q51)	0.21	0.14	-0.02	0.44	-0.23	0.01
How Often Bathing (q20)	0.06	0.05	-0.03	0.00	0.59	0.12
How Often Bathroom (q22)	0.17	0.05	-0.06	-0.05	0.56	0.03
Call Help Frequently (q9)	0.34	0.03	0.03	0.01	0.44	0.13
Room Clean (q17)	0.26	0.08	-0.03	-0.02	0.13	0.44
Room Quiet (q18)	0.05	0.02	0.06	0.01	0.20	0.43
Temperature (q16)	0.24	0.03	0.12	0.09	-0.05	0.38
Privacy (q24)	0.22	0.02	0.08	-0.08	0.05	0.04
TX Decisions (q25)	0.23	0.05	0.33	0.14	0.24	-0.05
Tests Without Pain (q35)	0.30	0.06	0.17	-0.03	0.17	-0.20
Delays In Admission (q43)	-0.13	-0.14	0.21	0.19	0.33	0.05

Table 10: Factor Analysis with Seven Factors

Item	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
Listen (Q5)	0.77	0.02	0.01	0.19	-0.05	-0.02	0.04
Respect (Q4)	0.76	-0.05	-0.04	0.14	-0.02	0.07	0.06
Explain Things (Q6)	0.76	0.15	0.06	0.10	0.02	-0.12	-0.07
Introduce (Q28)	0.57	0.05	0.08	-0.08	0.07	0.00	0.10
Spend Enough Time (Q7)	0.57	0.05	-0.06	0.35	0.03	0.02	0.01
Living Will (Q44)	0.47	-0.02	-0.08	-0.30	0.10	0.07	0.05
Visitors Help (Q27)	0.43	0.09	-0.01	0.25	-0.02	0.17	0.04
MD Listen (Q5)	0.02	0.89	0.02	-0.02	0.01	0.02	-0.01
MD Explain (Q13)	0.03	0.82	0.04	0.03	0.06	-0.07	0.00
MD Enough Time (Q14)	0.01	0.80	-0.02	0.10	-0.04	0.01	0.03
MD Respect (Q11)	0.17	0.79	-0.01	-0.17	-0.05	0.04	0.11
Taking Other RX (Q39)	0.00	0.00	0.88	-0.10	-0.04	-0.02	0.06
Allergic to RX (Q40)	0.06	-0.03	0.85	-0.12	-0.07	0.00	0.12
Purpose of TX (Q38)	-0.05	0.06	0.72	0.21	0.01	-0.03	-007
RX Side Effects (Q41)	-0.17	-0.01	0.63	0.20	0.13	0.01	0.13
Name of TX (Q37)	0.10	0.01	0.63	0.10	0.07	0.10	-0.19
How Often Bathing (Q20)	-0.08	-0.03	0.04	0.66	0.00	0.01	0.12
How Often Bathroom (Q22)	0.06	-0.05	0.04	0.66	-0.04	-0.03	0.02
Call Help FreQuently (Q9)	0.20	0.04	0.03	0.53	0.02	0.04	0.12
Writing Symptoms (NQ49)	0.03	0.01	0.03	0.06	0.73	-0.11	-0.02
Writing Activities (NQ47)	0.06	0.01	-0.02	-0.02	0.72	-0.06	0.01
Help After Discharge (NQ48)	-0.03	-0.04	-0.04	0.16	0.49	0.13	0.04
Writing How To Take RX (NQ51)	0.18	-0.02	0.14	-0.25	0.42	0.13	0.01
Pain Controlled (Q32)	0.14	0.02	0.01	0.05	-0.04	0.69	0.05
MD Pain Help (Q33)	0.30	0.08	-0.02	0.05	0.06	0.59	-0.01
Room Clean (Q17)	0.23	-0.03	0.08	0.18	-0.02	-0.01	0.44
Room Quiet (Q18)	0.07	0.06	0.02	0.24	0.02	-0.08	0.43
Temperature (Q16)	0.13	0.11	0.03	-0.05	0.07	0.17	0.38
TX Decisions (Q25)	0.08	0.34	0.05	0.27	0.15	0.11	-0.05
MD Respond Pain (Q31)	0.06	0.01	0.11	0.32	-0.02	0.34	0.08
Privacy (Q24)	0.13	0.08	0.02	0.07	-0.08	0.09	0.05
Tests Without Pain (Q35)	0.05	0.17	0.06	0.20	-0.05	0.27	-0.19
Delays In Admission (Q43)	-0.23	0.21	-0.14	0.34	0.18	0.07	0.06

Table 11: Eigenvalues of Factors in Seven Factor Solution

Factor	Eigenvalue	Difference	Proportion	Cumulative
1	11.97	10.04	0.68	0.68
2	1.93	0.52	0.11	0.79
3	1.41	0.47	0.08	0.87
4	0.94	0.16	0.05	0.92
5	0.78	0.23	0.04	0.96
6	0.55	0.02	0.03	0.99
7	0.53	0.13	0.03	1.02
8	0.40	0.06	0.02	1.05
9	0.34	0.10	0.02	1.07
10	0.24	0.09	0.01	1.08

Table 12: Correlations between Rotated Factors (inter factor correlations)

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
Factor1	1	0.6041	0.47024	0.61236	0.33151	0.59323	0.27978
Factor2	0.6041	1	0.38062	0.50066	0.4562	0.52278	0.18714
Factor3	0.47024	0.38062	1	0.50721	0.38409	0.40276	0.18003
Factor4	0.61236	0.50066	0.50721	1	0.3566	0.50249	0.15283
Factor5	0.33151	0.4562	0.38409	0.3566	1	0.3432	0.1815
Factor6	0.59323	0.52278	0.40276	0.50249	0.3432	1	0.27008
Factor7	0.27978	0.18714	0.18003	0.15283	0.1815	0.27008	1

Table 13: Variance Explained by Factors

Treatment of Other Factors	Variance Explained						
	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
Eliminate other factors	1.418187	1.595023	1.921565	1.1544753	1.161545	0.683486	0.644435
Ignore other factors	9.279909	7.985192	6.560286	7.755295	4.483049	6.6153	1.997522

*Note: The **Care Transition Measure** (CTM) was originally endorsed by the NQF as NQF #0288. In addition, we tested the validity of the CTM included in the HCAHPS Survey, as follows:*

- *20,670 inpatients discharged January 1, 2012 through March 31, 2012*
- *Random assignment to each of 4 HCAHPS survey modes*
  - *Mail only*
  - *Mixed mode (Mail with Telephone)*
  - *Telephone only (CATI)*
  - *Touch-tone IVR only (TT-IVR)*

Validity of Care Transition Measure

- *Pearson correlations (linear scoring) with 9 existing HCAHPS measures*
  - *Range 0.30 to 0.51*
  - *Highest: Recommend Hospital (0.51), Nurse Communication (0.50), Communication About Medicines (0.49), Overall Hospital Rating (0.48)*

- *PMA and Mode Effect patterns similar to other HCAHPS measures*
- *Like all HCAHPS measures CTM has good psychometric properties*
- *No evidence of a ceiling effect; substantial room for quality improvement*
- *Strong links with Nurse Communication and Communication About Medicines*
- *Lack of redundancy with HCAHPS measures*
- *Transitions in care is a critical aspect of hospital care*
- *Data indicates there is room for quality improvement in this area*
- *Collection and public reporting of CTM will aid quality improvement efforts*

To confirm discriminant validity, in the table below we computed the average patient-level correlations among individual HCAHPS items with the other items within the same multi-item measure, and with items in other multi-item measures. In all instances, the average correlations of items within a multi-item measure (0.27-0.57) were considerably higher than the correlations with items in different multi-item measures (averaging 0.19 to 0.32), providing evidence of discriminant validity.

Discriminant Validity Analysis: Patient-Level Average Inter-Item Correlations among HCAHPS Multi-item Measures, July 2016 – June 2017 discharges

	Communication with Nurses	Communication with Doctors	Responsiveness of Hosp. Staff	Comm. About Medicines	Discharge Information	Care Transition
Communication with Nurses	**0.50**					
Communication with Doctors	0.36	**0.56**				
Responsiveness of Hosp. Staff	0.36	0.28	**0.48**			
Comm. About Medicines	0.33	0.31	0.30	**0.46**		
Discharge Information	0.16	0.17	0.13	0.20	**0.27**	
Care Transition	0.28	0.27	0.26	0.32	0.19	**0.57**
Average Across	***0.30***	***0.31***	***0.26***	***0.32***	***0.19***	***0.26***
Average Within	***0.50***	***0.56***	***0.48***	***0.46***	***0.27***	***0.57***

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

The pilot study analyses provide support for the construct validity of the HCAHPS, confirming the intended factor structure. A 2016 discriminant validity analysis confirmed the factor structure. Both the 2003 pilot study and the 2016 analyses found strong correlation of each of the multi-item measures and single-item measures with the two overall measures.

The 2012 HCAHPS mode experiment also provides similar support for the Care Transition Measure, which was added to the HCAHPS Survey in 2013.

_____

2b2. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section* 2b3

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

The HCAHPS Survey is broadly intended for patients of all payer types who meet the following criteria:

- Eighteen (18) years or older at the time of admission
  - To ensure HCAHPS results are for adults only; there is a distinct and separate survey for pediatric patients.
- Admission includes at least one overnight stay in the hospital
  - A fundamental goal of HCAHPS is to survey adults about their hospital experience as an inpatient (that is, having had at least one overnight stay) in an acute care hospital). Note that other CAHPS surveys are appropriate for out-patient experience of care.
- Non-psychiatric MS-DRG/principal diagnosis at discharge
  - To address concerns about the appropriateness of patient's ability to respond to surveys.

  Note: Patients whose principal diagnosis falls within the Maternity Care, Medical or Surgical service lines and who also have a secondary psychiatric diagnosis are still eligible for the survey.

  Note: MS-DRG codes in the ineligible category include patients with MS-DRG codes for newborn, psychiatric, substance abuse, rehabilitation, or deceased, and MS-DRG codes with no assigned type.
- Alive at the time of discharge
  - A prerequisite for reaching patients. Note that only the patient, not a proxy, is permitted to answer the HCAHPS Survey.

Note: Pediatric patients (under 18 years old at admission) are ineligible because the current HCAHPS instrument is not designed to address the unique situation of pediatric patients and their families.

Note: Patients with a primary psychiatric or substance abuse diagnosis are ineligible because the current HCAHPS instrument is not designed to address the behavioral health issues pertinent to psychiatric patients.

Note: Patients identified with discharge status code (UB-04 field location 17) "30 – Still a Patient or Expected to Return for Outpatient Services" are **not** eligible for the HCAHPS Survey.

Exclusions from the HCAHPS Survey

There is a two-stage process for determining whether a discharged patient can be included in the HCAHPS Sample Frame. The first stage is to determine whether the discharged patient meets the HCAHPS eligibility criteria, listed above. If the patient meets the eligibility criteria, then a second set of criteria is applied: Exclusions from the HCAHPS Survey.

Patients who meet the eligible population criteria outlined above are to be included in the HCAHPS Sample Frame. However, there are a few categories of otherwise eligible patients who are excluded from the sample frame. These are:

- "No-Publicity" patients – Patients who request that they not be contacted (see below)
- Court/Law enforcement patients (i.e., prisoners); this does not include patients residing in halfway houses
- Patients with a foreign home address (the U.S. territories – Virgin Islands, Puerto Rico, Guam, American Samoa, and Northern Mariana Islands are not considered foreign addresses and therefore, are not excluded)
- Patients discharged to hospice care (hospice-home or hospice-medical facility)
- Patients who are excluded because of state regulations

- Patients discharged to nursing homes and skilled nursing facilities

The HCAHPS Survey was designed for the eligible patients, as described above. Most exclusions are due to inability to contact certain categories of patients and allow them to complete the survey within the specified time frame, or in deference to legal restrictions

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

HCAHPS exclusions and eligibility criteria are primarily based on a priori policy goals; they are not based on empirical testing and comparison of categories of patients.

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b4**.**

2b3.1. What method of controlling for differences in case mix is used?

☐ **No risk adjustment or stratification**

☐ **Statistical risk model with _risk factors**

☐ **Stratification by _risk categories**

☒ **Other,** Patient-mix adjustment via linear regression and survey mode adjustment based on randomized mode experiments.

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

As noted in the HCAHPS Quality Assurance Guidelines, V12.0, prior to public reporting, hospitals' HCAHPS results are adjusted for the effects of both mode of survey administration and patient-mix. Generally speaking, HCAHPS adjustments for survey mode are larger than adjustments for patient-mix. The survey mode adjustments that are used in publicly reported HCAHPS results are reported in the paper entitled, "Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS) April 30, 2008," which can be found on the official HCAHPS On-Line Web site at http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/final-draft-description-of-hcahps-mode-and-pma-with-bottom-box-mododoc-april-30-2008.pdf .

In order to derive the mode adjustment coefficients, it was necessary to conduct a randomized mode experiment. The resulting mode adjustment coefficients will not change as a function of the data used in public reporting. For more information on how the HCAHPS mode experiment was conducted and the survey mode and patient-mix adjustments were derived, please see, "Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS Hospital Survey Scores." Elliott, M.N., A.M. Zaslavsky, E. Goldstein, W. Lehrman, K. Hambarsoomian, M.K. Beckett, and L. Giordano. Health Services Research. 2009. 44: 501-518.

The mode experiment data were also used to develop and validate the HCAHPS patient-mix model (which is referred to as "case-mix" elsewhere in the CAHPS literature), as described in the document referenced above. However, in the case of patient-mix adjustment, a randomized experiment is not necessary to accurately estimate the coefficients of the model.

In order to estimate the exact patient-mix coefficients as accurately as possible, we employ the large sample size of each quarterly national publicly reported data set. This approach allows us to detect changes in the association of patient-mix adjustors and HCAHPS measures over time and then adjust accordingly. This approach is consistent with recommended CAHPS practice for case-mix adjustment http://www.ahrq.gov/cahps/.

Note: The patient-mix adjustments applicable to CY 2016 discharges can be found at:

https://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/2017-10_mode_patient-mix-adj_pma.pdf .

Note: The HCAHPS Patient-Mix Adjustment (PMA) model has been updated to incorporate more detailed information about patients' Service Line and Gender. Prior to Quarter 1 2017, the patient-mix adjustment for service line distinguished among the three service line categories: Medical, Surgical, and Maternity. Beginning with Quarter 1 2017 discharges, the patient-mix adjustment included patient gender with service line to distinguish among 5 categories: Female Medical, Male Medical, Female Surgical, Male Surgical, and Maternity, which is only female. Female Medical will serve as the reference category for this adjustment. HCAHPS survey results were adjusted using the new PMA model beginning with January 1, 2017 discharges. Updated survey mode adjustments from a 2016 randomized mode experiment were also incorporated beginning with April 2017 discharges.

Because the Maternity service line is entirely female, CMS has to a degree always employed gender in its HCAHPS service line adjustment. CMS' recent research shows that when the new Service Line and Gender adjustment was implemented, hospitals whose patients are predominantly male experienced a small decrease in their Cleanliness measure score, while those predominantly female experienced a small increase; changes to other HCAHPS measures were smaller. However, by employing the new Service Line and Gender adjustment, scores are more accurate than before. Adjustments contained in Tables 1a, 2a, and 3a are needed to estimate PMA for discharges prior to January 1, 2017. To estimate PMA for discharges on or after January 1, 2017 please see Tables 1b, 2b, found on the HCAHPS On-Line Web site at http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/2017-12_mode-patient-adj_pma.pdf .

Patient-mix adjustment is performed within each quarter of data after data cleaning and before mode adjustment. Coefficients obtained in linear regression models (not reported) estimate the tendency of patients to respond more positively or negatively. The adjustments needed to counter that tendency are obtained by multiplying the patient-mix coefficients by (-1.0). Tables 1a-b and 2a-b below report patient-mix adjustments for the "top-box" (most positive response) and "bottom-box" (least positive response) respectively of the ten publicly reported HCAHPS measures (six multi-item measures, two single-item measures, and two global items), averaged across the four reported quarters.

As an example, patients aged 55-64 were 4.89% (Q2-2016 though Q4-2016) and 4.93% (Q1-2017) or 4.90% for the quarter ((4.89%/4)*3)+((4.93/4)*1) more likely to provide the most positive response ("Always") for items in the Communication with Nurses multi-item measure when compared to the reference group of patients 85 and older. Thus, the corresponding adjustment for patients aged 55-64 relative to patients 85 and older for that multi-item measure is a subtraction of 4.90%, reflected in the "-4.89%" and "-4.93%" entries in Tables 1a and 1b, respectively. Similarly, for each level of decreasing self-rated health status (where 5=poor, 4=fair, 3=good, 2=very good, and 1=excellent), the percentage of patients providing an "Always" response for Communication with Doctors decreased by 4.77% (Q2-2016 though Q4-2016) and 4.79% (Q1-2017) or 4.78% for the quarter ((4.77%/4)*3)+((4.79/4)*1). Thus, a patient in fair health (4) would have a (4-1)*4.78=14.34% lower chance of an "Always" response than a patient in excellent health (1), and the corresponding adjustment for a patient in poor health relative to a patient in excellent health would be +19.12%.

Publicly-reported HCAHPS scores are adjusted to the overall national mean of patient-mix variables across all hospitals reporting in a given quarter (as reported in Tables 3a and 3b). Thus, whether the scores of a given hospital are adjusted upward or downward for a given measure depends not only on these patient-mix

adjustments, but also on the patient-mix of that hospital relative to the national average of these patient-mix characteristics. Specifically, the total patient mix-adjustment for a given hospital is the sum of a series of products, where each product multiplies the adjustment in Tables 1a and 1b (top-box) or Tables 2a and 2b (bottom-box) by the deviation of the hospital's mean on the corresponding patient-mix variable from the national mean on that patient-mix variable (from Tables 3a and 3b).

Four sets of numbers are needed to calculate final patient-mix adjusted scores for a given hospital: (1) Means of HCAHPS outcomes (top-box proportions or bottom-box proportions) for the hospital in question that have been adjusted for survey mode; (2) individual-level patient-mix adjustments from Tables 1 and 2 of this document; (3) that hospital's means on patient-mix variables; and (4) national means on patient-mix variables from Tables 3a and 3b of this document.

Below we provide additional detail regarding the calculation of the response percentile and service line by age interaction variables. A hospital's patient-mix adjustment variable response percentile is calculated as follows: For a given hospital and a given month, all completed surveys are ranked based on their respective "lag times." Lag time is the number of days between a patient's discharge from the hospital and the return of the mail survey, or the final disposition of the telephone or IVR survey. Ranks are averaged in the case of ties. Response percentile is calculated by dividing lag time rank by monthly sample size.

The service line by age interaction variables used in patient-mix adjustment can be calculated by following the steps below for all completed surveys:

1) Create an age variable that can take values from 1 through 8, depending on the age range of the patient. Denote this variable as AGE.
2) Create an indicator variable for whether a survey was from the surgical service line. Let this variable equal 1 if surgical and equal to 0 if not surgical. Denote this variable as SURG.
3) Create an indicator variable for whether a survey was from the maternity service line. Let this variable equal 1 if maternity and equal to 0 if not maternity. Denote this variable as MAT.
4) At this point, every completed survey should have a value from 1 to 8 for AGE, a value of 0 or 1 for SURG, and a value of 0 or 1 for MAT.

The surgical by age interaction variable (Surgical*Age) is equal to the product of SURG and AGE. Similarly, the maternity by age interaction variable (Maternity*Age) is calculated as the product of MAT and AGE. To obtain hospital-level values for these two interaction variables, simply average all the survey-level values just calculated for Surgical*Age and Maternity*Age.

The formula for applying patient mix adjustment is as follows:

If y is the mode-adjusted hospital mean of an HCAHPS outcome (top-box or bottom-box)

a1-a19 are the individual-level adjustments from Tables 1a and 1b or Tables 2a and 2b for the 19 rows other than reference categories (in proportion rather than percentage form)

m1-m19 are the national means for the PMA variables in the same rows in Table 3a and 3b

h1-h19 are the PMA means for the hospital in question in the same form as in Tables 3 a and 3b,

then y'=y+a1(h1-m1)+a2(h2-m2)+…+a19(h19-m19) is the patient-mix and mode-adjusted hospital score for that outcome.

HCAHPS publicly-reported four-quarter hospital averages are weighted proportionately to the number of eligible patients seen by the hospital in each of the quarters. Specifically, each quarter's score has a quarterly weight equal to the quarter's eligible discharge size divided by the total eligible discharge size for the four quarters that make up the reporting period. Quarterly weights are applied after patient-mix adjustment and survey mode adjustment.

For public reporting purposes, HCAHPS scores are rounded to integer percentages. Rounding occurs within top, middle, and bottom-box scores only after patient-mix and mode adjustments have been applied. If the sum of the three scores is not 100%, a further adjustment is made to the middle-box score.

Please note: The information presented here will permit a hospital to closely approximate the effect of patient-mix adjustment on its HCAHPS results. However, exact replication of published HCAHPS results may not be possible because of (1) the effects of data cleaning and (2) small differences between the effects of quarterly patient-mix adjustments and the four-quarter averages presented here. For each future public reporting period, Tables 1a-b, 2a-b and 3a-b will be updated and will be posted on the official HCAHPS On-Line Web site at http://www.hcahpsonline.org/en/mode--patient-mix-adj/.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not Applicable.

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

The risk adjustment method for HCAHPS has been described above in section 2b3.1.1.

The clinical factors in the HCAHPS patient-mix adjustment model include patient's age, gender by service line, and self-reported overall health.

The social risk factors in the HCAHPS patient-mix adjustment model include patient's self-reported education and primary language spoken at home.

We do not apply any ordering when including risk factors in the patient-mix adjustment model; all patient-mix adjustment factors are entered simultaneously.

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☒ **Published literature**

☒ **Internal data analysis**

☐ **Other (please describe)**

2b3.4a. What were the statistical results of the analyses used to select risk factors?

The development of social risk factors was first informed by a literature search on social risk factors related to patient experience and survey response. This process included identification of social risk candidate variables. Next, potential social risk factors were analyzed for exclusion due being potentially endogenous to the patient care received. Finally, examination of patient-mix adjusted outcome differences associated with these risk factors.

Analyses conducted from a random sample of 19,720 patients discharged from 132 hospitals between December 2002 and January 2003. The objective was to assess the extent to which patient characteristics predict patient ratings (''predictive power'') and the heterogeneity of the characteristics across hospitals. The measures were combined to estimate the impact of each predictor (''impact factor'') high impact variables were selected for adjusting ratings for HCAHPS.

The most important case-mix variables were hospital service (surgery, obstetric, medical), age, race (non-Hispanic black), education, general health status (GHS), speaking Spanish at home, having a circulatory disorder, and interactions of each of these variables with service. Adjustment for GHS and education affected scores in each of the three services, while age and being non-Hispanic black had important impacts for those

receiving surgery or medical services. Circulatory disorder, Spanish language, and Hispanic affected scores for those treated on surgery, obstetrics, and medical services, respectively. Of the 20 medical conditions tested, only circulatory problems had an important impact within any of the services. Results were consistent for the overall ratings of nurse, doctor, and hospital. Although the overall impact of case-mix adjustment is modest, the rankings of some hospitals may be substantially affected. In conclusion, case-mix adjustment had a small impact on hospital ratings, but may lead to important reductions in the bias in comparisons between hospitals.

The stepwise case-mix adjustment analysis below predicts kept items and composites (outcomes) in linear regression at the patient level using hospital intercepts, in addition to case mix adjusters and social risk variables.

Table 1: Stepwise Analysis (All Services)

Predictor Levels	Nurse Rating		Doctor Rating		Hospital Rating	
	Parameter Estimate	Overall p-Value	Parameter Estimate	Overall p-Value	Parameter Estimate	Overall p-Value
In						
Intercept Service	8.20	<.0001	8.74	<.0001	8.28	<.0001
Medical	0.04	<.0001	-0.42***	<.0001	-0.09	<.0001
OB	0.42***		0.03		0.41***	
Health status						
Excellent	0.24***	<.0001	0.28***	<.0001	0.32***	<.0001
Very good						
Good	-0.54***		-0.53***		-0.65***	
Fair	-0.31***		-0.19***		-0.34***	
Poor	-0.96***		-0.86***		-1.14***	
Mental Health						
Excellent	0.17***	<.0001	0.24***	<.0001	0.19***	<.0001
Very good						
Good	-0.19**		-0.24***		-0.05	
Fair	-0.04		-0.09		0.01	
Poor	-0.48**		-0.74***		-0.35*	
Age						
18-24	-0.24***	<.0001	-0.30***	<.0001	-0.22*	<.0001
25-34						
35-44	0.10		0.14*		0.05	
45-54	0.24**		0.27***		0.25**	
55-64	0.53***		0.52***		0.56***	
65-74	0.70***		0.67***		0.76***	
75-79	0.79***		0.69***		0.85***	
80+	0.83***		0.76***		0.99***	
Education						
Eighth grade	0.35***	<.0001	0.30***	<.0001	0.42***	<.0001
Some high school	0.31***		0.26***		0.43***	
High school	0.22***		0.25***		0.29***	
Some college						
College	-0.15*		-0.03		-0.13	
College+	-0.20***		-0.23***		-0.22***	
Race/ethnicity						
Non-Hispanic black	0.41	<.0001	0.29	<.0001	0.43	<.0001
Spanish language	0.71	<.0001	0.61	<.0001	0.88	<.0001
Sex: male	0.24	<.0001			0.19	<.0001
Proxy help	-0.34	<.0001			-0.26	<.0001
Out						
Sex: male				0.348		

Predictor Levels	Nurse Rating		Doctor Rating		Hospital Rating	
	Parameter Estimate	Overall p-Value	Parameter Estimate	Overall p-Value	Parameter Estimate	Overall p-Value
Race/ethnicity						
Hispanic		.1994		.0069		.0785
Asian		.2322		.0763		.0972
Native American		.3018		.0708		.9721
Proxy help				.0124		
Proxy answer		.7628		.7459		.8087

*p<.01,

**p<.001,

***p<.0001. The baseline category of categorical predictors is the level with the blank cell. OB, obstetrics.

Citation for Table 1 shown above:

Case-Mix Adjustment of the CAHPS® Hospital Survey. A. James O'Malley, Alan M. Zaslavsky, Marc N. Elliott, Lawrence Zaborski, and Paul D. Cleary. Health Services Research, 40, part II: 2162-2181. 2005.

Current national HCAHPS top-box patient-mix adjustments for social risk factors are shown below and can be located here: http://hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/january_2019_-mode--patient-mix-adj_pma.pdf

Table 1: HCAHPS Patient-Mix Adjustments of Top-Box for Patients Discharged between Quarter 2, 2017 and Quarter 1, 2018 (April 1, 2017 to March 31, 2018)

Patient-Mix Adjustment (PMA)	Comm with Nurses	Comm with Doctors	Responsiveness of Hosp. Staff	Comm. About Medicines	Cleanliness of Hosp. Env.	Quietness of Hosp. Env.	Discharge Information	Care Transition	Hospital Rating	Recommend the Hospital
Education (per level; 1=8th grade or less and 6=More than 4-year college degree)	1.41%	1.44%	2.27%	2.70%	1.35%	3.38%	0.39%	-0.63%	2.42%	0.92%
Self-Rated Health (per level; 1=Excellent and 5=Poor)	4.82%	4.80%	6.20%	4.94%	4.26%	4.43%	1.08%	6.37%	6.32%	5.54%
Response Percentile (per 1% of response percentile)	0.19%	0.18%	0.24%	0.19%	0.06%	0.04%	0.03%	0.19%	0.17%	0.17%
Spanish	-0.76%	-2.57%	-0.91%	-2.50%	1.27%	-4.88%	-1.72%	0.84%	-11.88%	-9.38%
Chinese	5.65%	4.80%	5.84%	3.55%	3.76%	-2.13%	-1.99%	11.06%	3.56%	2.77%
R/V/P/O (Russian, Vietnamese, Portuguese, Other)	1.15%	0.50%	2.36%	0.51%	3.56%	-6.34%	-0.03%	6.30%	1.56%	-0.34%
English (REFERENCE)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Age 18-24	1.98%	0.70%	3.69%	-9.62%	-0.03%	-4.39%	-3.34%	-8.07%	16.22%	12.55%
Age 25-34	-0.58%	-0.82%	-2.66%	-11.66%	-0.75%	-6.46%	-3.98%	-12.17%	12.20%	7.70%
Age 35-44	-1.57%	-1.69%	-4.08%	-11.75%	0.08%	-5.21%	-4.07%	-11.60%	9.26%	5.44%
Age 45-54	-3.95%	-4.21%	-6.35%	-13.32%	-0.38%	-4.54%	-4.88%	-12.70%	4.09%	1.33%
Age 55-64	-4.98%	-5.01%	-7.06%	-12.07%	-0.24%	-3.13%	-5.08%	-11.95%	0.62%	-0.35%
Age 65-74	-5.53%	-5.70%	-6.38%	-10.22%	0.28%	-2.62%	-4.93%	-10.56%	-2.32%	-1.76%
Age 75-84	-3.37%	-3.50%	-3.75%	-5.47%	1.07%	-1.11%	-2.43%	-4.87%	-2.03%	-1.37%
Age 85+ (REFERENCE)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Male Medical	-1.00%	0.39%	-2.18%	-3.66%	-6.64%	0.42%	-2.57%	-0.96%	-0.47%	-2.05%
Male Surgical	-1.38%	-7.93%	-3.00%	-6.40%	-7.43%	-1.09%	-6.59%	-4.74%	-7.01%	-7.35%
Female Surgical	0.06%	-8.65%	-0.97%	-1.83%	-0.70%	-3.19%	-5.56%	-3.49%	-5.91%	-5.22%
Female Maternity	-6.61%	-12.08%	-14.81%	-12.75%	-0.16%	-12.38%	-6.00%	-4.83%	-14.36%	-14.75%
Female Medical (REFERENCE)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Surgical Line * Age[1]	-0.12%	0.39%	-0.40%	0.03%	-0.28%	0.06%	-0.13%	0.05%	0.36%	0.24%
Maternity Line * Age[1]	0.87%	0.95%	0.74%	1.89%	0.27%	0.64%	0.89%	1.22%	2.52%	2.07%

[1] Age takes on the values of 1 to 8 as follows: (1: 18 to 24); (2: 25 to 34); (3: 35 to 44); (4: 45 to 54); (5: 55 to 64); (6: 65 to 74); (7: 75 to 84); and (8: 85+).

Other analysis examined HCAHPS measure differences by race and ethnicity using survey data collected from 2,684 acute care hospitals for inpatients discharged from October 2006 through June 2007. Analysis compared

the experiences of Hispanic, African American, Asian/Pacific Islander, American Indian/Alaska Native, and multiracial inpatients with those of non-Hispanic White inpatients to understand the roles of between- and within-hospital differences in patients' perspectives of hospital care. The study finds that, on average, non-Hispanic White inpatients receive care at hospitals that provide better experiences for all patients than the hospitals more often used by minority patients. Within hospitals, patient experiences are more similar by race/ethnicity, though some disparities do exist, especially for Asians. This research suggests that targeting hospitals that serve predominantly minority patients, improving the access of minority patients to better hospitals, and targeting the experiences of Asians within hospitals may be promising means of reducing disparities in patient experience.

Racial/Ethnic Differences in Patients' Perceptions of Inpatient Care Using the HCAHPS Survey.

E. Goldstein, M.N. Elliott, W.G. Lehrman, K. Hambarsoomian and L.A. Giordano.

Medical Care Research and Review, 67: 74-92. 2010.

(First published August 3, 2009. Medical Care Research and Review 2009, doi:10.1177/1077558709341066)

Excerpt from Quigley paper (followed by citation):

We used 2014-2015 HCAHPS data to investigate differences in inpatient experiences by preferred language within racial/ethnic groups. HCAHPS is a survey of recently-discharged patients' experiences of hospital care in the United States and includes information on self-reported language preference and race/ethnicity that permits this analysis.32 Specifically, the HCAHPS survey asks which of seven languages the patient primarily speaks at home. Sample sizes used were sufficient to examine all preferred-language groups for which HCAHPS provides translations. As such, this dataset allows one to examine preferred language within racial/ethnic groups among even smaller groups, such as Portuguese-preferring Hispanics and Vietnamese-preferring Asian/Pacific Islanders (API).

HCAHPS measures experiences of inpatients of all payer types (Medicaid, Medicare, and all others) who are 18 years or older at admission, stay overnight in the hospital with a principal diagnosis for medical, surgical or maternity care, and are discharged alive.34 Our analysis included the 5,480,308 completed surveys from all 4,517 hospitals in the 50 states and DC that submitted HCAHPS data to the Centers for Medicare & Medicaid Services (CMS) during the eight quarters of calendar years 2014-2015.

We examined six HCAHPS composite measures: Communication with Doctors, Communication with Nurses, Responsiveness of Hospital Staff, Communication about Medication, Discharge Information, and Care Coordination. Three measures were excluded because they do not rely on conversing in a shared language (Quietness, Cleanliness) or are no longer used for incentive payments (Pain Management). Two global measures (Ratings of Hospital and Recommendation of Hospital) were excluded because prior research suggests that such items may elicit different evaluations of the same care from different racial/ethnic and language groups. The survey items comprising five of the six retained composite measures (all but Discharge Information items, which employ yes/no responses) use a standard set of response options: never, sometimes, e, and always. A description of the composite measures is included in the Appendix in Table A.1.

HCAHPS respondents are asked to self-report whether they are of Spanish, Hispanic or Latino origin or descent. They are then asked to select at least one race, with response options of White, Black or African American, American Indian or Alaska Native (AI/AN), Asian, Native Hawaiian or other Pacific Islander, and some other race. Six mutually-exclusive racial/ethnic categories were created using these two items: (1) Hispanic; and non-Hispanic (2) White, (3) Black, (4) API, (5) AI/AN, and (6) multiracial. Following the Office of Management and Budget approach, we classified any patient as Hispanic who endorsed Hispanic ethnicity. Non-Hispanic patients who endorsed exactly one race were classified as that race; those who endorsed Asian plus Native Hawaiian or other Pacific Islander were classified as API; the remaining non-Hispanic patients who endorsed two or more races were classified as multiracial. Our analysis excluded data from multiracial patients (3%), a heterogeneous and difficult-to-interpret group, and patients who did not answer the race item (7%).

Because several languages measured by the survey are associated almost exclusively with a single racial/ethnic group, language was considered within racial/ethnic categories. The HCAHPS survey asks, "What language do you mainly speak at home?" with response options of English, Spanish, Chinese, Russian, Vietnamese, Portuguese, and "some other language." We included all combinations of preferred language (that is, language spoken at home) and race/ethnicity among our seven languages and five racial/ethnic groups for which at least 400 completed surveys were available nationally: Hispanics (languages included Spanish, English, Portuguese, Other), API (English, Chinese, Vietnamese, Other), Blacks (English, Spanish, Other), AI/AN (English, Other), and Whites (English, Russian, Spanish, Portuguese, Other).

To analyze the types of hospitals utilized by language-within-race/ethnicity groups, we examined key hospital characteristics of bed size (200 or more beds), rural location, profit status (for profit, not-for profit, governmental status), and service line composition (percent medical, surgical, maternity). We also calculated by preferred language within racial/ethnic groups: the average hospital-level proportion of non-English language-preferring patients, the average hospital-level proportion of the matching racial/ethnic group, and the average hospital-level proportion of their same racial/ethnic and language group. Linear regression compared overall, within-hospital and between-hospital patient experiences by preferred language within racial/ethnic groups using standard patient-mix adjustors. Following the CMS approach, we used patient-mix adjusted top-box-scored measures for all composite measures, scoring the most positive response option as 100 and all other responses as 0 prior to averaging non-missing items to create composite scores. The top-box response is "always" for four HCAHPS composites (Communication with Nurses, Communication with Doctors, Responsiveness of Hospital Staff, and Communication about Medication), "yes" for the Discharge Information composite, and "strongly agree" for the Care Coordination composite. To illustrate, the score for a respondent who answered "always," "always," "never," and "sometimes" to four items within a composite would be (100+100+0+0=200)/4 = 50. "Patient mix" refers to patient characteristics not under the control of the hospital that may affect scores of patient experience measures. Patient-mix adjustment accounts for between-hospital differences in the patient population to estimate the scores each hospital would have received if all had treated the same patients. Standard HCAHPS patient-mix adjustors are patient age; service line (maternity, surgical, and medical [reference category]); self-reported education; self-reported overall patient health; response percentile (a rank-based measure of the latency between discharge date and survey completion that addresses the tendency of later responders to indicate worse care experiences); interactions of maternity and surgical service line with linearly-scored patient age; and preferred language spoken at home. Here, preferred language spoken at home was treated as the primary independent variable, rather than as a patient-mix adjustor.

In Table 3, differences in patient experiences for six HCAHPS measures are shown by language within racial/ethnic group. Generally, non-English-preferring Black, Hispanic, API and AI/AN patients reported worse experiences than their English-preferring counterparts, except for Russian-preferring White patients. Differences between English and non-English preferring patients within the same racial/ethnic group were largest and most consistent (i.e., the findings were both statistically significant and had the same sign) for Care Coordination and smallest and least consistent for Discharge Information and Communication about Medication.

The experiences of White patients were not consistent across measures and language preference. Spanish-preferring and other-language-preferring Whites reported less positive experiences than English-preferring Whites, except for Doctor Communication where experiences were similar for English-preferring and other other-language-preferring Whites. Russian-preferring Whites reported the best experiences among Whites, except for Care Coordination.

Non-English-preferring Black patients reported consistently worse experiences than their English-preferring counterparts, with all differences at least moderate in magnitude (3+ points).

Among Hispanics, Spanish-preferring and other-language-preferring patients reported worse experiences than English-preferring Hispanics except for similar experiences for Communication about Medication for English-preferring and other-language-preferring patients. Portuguese-preferring Hispanics reported worse

experiences for only Doctor Communication and Care Coordination. Generally, differences between English-preferring and non-English preferring Hispanics were small (<3 points), except for Care Coordination.

Among API patients, each non-English-preferring group (Chinese-, Vietnamese-, and other-language-preferring) reported worse experiences than English-preferring API, except for Discharge Information. Within API, negative differences compared to English-preferring API tended to be largest for Chinese-preferring API.

Within AI/AN, other-language-preferring AI/AN reported worse care experiences than English-preferring AI/AN; differences were moderate or larger, except for Communication About Medicines.

Table 3. Differences in HCAHPS top-box scored composite measures by preferred language within racial/ethnic group, adjusted for patient-mix and survey year (N = 5,480,308)

Racial/Ethnic Group	Preferred language	Nurse communication		Doctor communication		Hospital staff responsiveness		Communication about medication		Discharge information		Care coordination	
		Est	P	Est	P	Est	P	Est	P	Est	P	Est	P
White	English	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA
White	Spanish	-3.2	***	-1.5	***	-2.4	***	-1.7	**	**-3.3**	***	**-5.9**	***
White	Russian	*2.0*	***	**4.8**	***	**5.6**	***	*2.4*	***	*0.9*	**	**-5.8**	***
White	Portuguese	-0.4		*2.1*	**	-1.7		1.0		0.8		-2.1	*
White	Other language	-0.6	**	-0.2		-0.9	***	-1.7	***	-1.2	***	**-7.1**	***
Joint test of language within r/e group		NA	***		***		***		***		***		***
Black	English	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA
Black	Spanish	**-5.4**	***	**-5.1**	***	**-6.6**	***	**-10.1**	***	**-3.9**	**	**-8.2**	***
Black	Other language	**-4.1**	***	**-4.2**	***	**-4.0**	***	**-4.2**	***	**-4.9**	***	**-6.8**	***
Joint test of language within r/e group		NA	***		***		***		***		***		***
Hispanic	English	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA
Hispanic	Spanish	-2.2	***	-0.4	***	-1.1	***	-0.9	***	-0.5	***	**-3.7**	***
Hispanic	Portuguese	0.5		-1.8	*	-0.3		-2.4		0.1		**-4.6**	***
Hispanic	Other language	-1.4	***	-1.5	***	-2.9	***	-0.2		-0.8	*	**-8.4**	***
Joint test of language within r/e group		NA	***		***		***		***		***		***
API	English	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA
API	Chinese	**-8.8**	***	**-7.6**	***	**-7.6**	***	**-6.0**	***	**3.2**	***	**-10.6**	***
API	Vietnamese	**-5.9**	***	**-5.7**	***	**-4.4**	***	-1.5	*	**3.3**	***	-2.8	***
API	Other language	-1.5	***	-1.6	***	**-3.1**	***	-1.1	***	*1.6*	***	**-4.6**	***
Joint test of language within r/e group		NA	***		***		***		***		***		***
AI/AN	English	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA	0.0	NA
AI/AN	Other language	**-6.0**	***	**-5.6**	***	**-4.2**	***	-2.3		**-3.9**	***	**-6.8**	***

Note: Bold font is used to highlight statistically significant positive coefficients greater than 3. Italics is used to highlight statistically significant positive coefficients less than 3. Underlining is used to highlight statistically significant negative coefficients less than 3. Bold font and underlining is used to highlight statistically significant negative coefficients greater than 3.

Est, estimate.

*P < 0.5. **P < 0.01. ***P < 0.001.

Quigley DD, Elliott MN, Hambarsoomian K, et al. Inpatient care experiences differ by preferred language within racial/ethnic groups. Health Serv Res. 2019;00:1–12. https://doi.org/10.1111/1475-6773.13105

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of*

unique variation in the outcome, assessment of between-unit effects and within-unit effects.) **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

We followed the CAHPS approach for case-mix adjustment (which we refer to as patient-mix adjustment in the HCAHPS program). The selection of factors in the HCAHPS patient-mix adjustment model is based on evidence that these factors met CAHPS case-mix adjustment criteria: exogeneity, significant variation between hospitals, significant association with CAHPS outcomes within hospitals. All patient-mix adjustment factors, including the social risk factors, are measured at the patient level, are exogenous and not caused by the hospital; varied to a substantial and statistically significant extent within hospitals; and are associated with patient experience of care outcomes after controlling for other factors in the patient-mix adjustment model. Providers who were at the extremes of social risk factors that were associated with less positive response tendency (e.g. hospitals with high proportions of patients whose preferred language is Chinese) received substantial positive adjustments.

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

For a description of the development and testing of the patient-mix adjustment and survey mode adjustment of the HCAHPS Survey, please see:

Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS)

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/final-draft-description-of-hcahps-mode-and-pma-with-bottom-box-modedoc-april-30-2008.pdf

Also: "Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS® Hospital Survey Scores." M.N. Elliott, A.M. Zaslavsky, E. Goldstein, W. Lehrman, K. Hambarsoomian, M.K. Beckett and L. Giordano. Health Services Research, 44 (2): 501-518. 2009.

Current survey mode adjustments can be found at:

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/april_2017_mode_patient-mix_adjustment_hcahps_survey_mode_adjustment.pdf

HCAHPS Survey Mode Adjustments of Top Box and Bottom Box Percentages (after PMA) to Adjust Other Modes to a Reference of Mail

Derived from 2016 Mode Experiment; for HCAHPS score adjustments beginning with April 2017 discharges

	Bottom Box			Top Box		
	Phone Only	**Mixed**	**IVR**	**Phone Only**	**Mixed**	**IVR**
HCAHPS Composite Measures						
Communication with Nurses	0.1%	1.3%	-1.8%	-4.2%	-3.6%	-2.3%
Communication with Doctors	-0.6%	-0.9%	-2.2%	-2.8%	-1.8%	0.3%
Responsiveness of Hospital Staff	0.5%	1.9%	-0.9%	-0.8%	-3.4%	2.0%
Pain Management	-1.7%	-0.7%	-4.3%	-3.7%	-2.3%	0.1%
Communication about Medicines	-1.5%	-1.1%	-2.3%	-1.7%	-0.9%	-0.1%
Discharge Information	1.7%	1.2%	1.6%	-1.7%	-1.2%	-1.6%
Care Transition	1.4%	0.9%	-0.5%	-0.6%	-1.3%	-0.1%
HCAHPS Individual items						
Cleanliness of Hospital Environment	-0.8%	0.6%	-1.9%	-2.8%	-3.8%	-0.5%
Quietness of Hospital Environment	1.6%	2.5%	-0.1%	-8.6%	-5.6%	-6.4%
HCAHPS Global Items						
Overall Hospital Rating	1.6%	1.3%	-0.5%	-2.0%	-3.0%	4.0%
Recommend the Hospital	0.6%	0.9%	-1.8%	-3.5%	-2.1%	0.1%

Current patient-mix adjustments shown below can be found at https://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/january_2019_-mode--patient-mix-adj_pma.pdf

Table 1: HCAHPS Patient-Mix Adjustments of Top-Box for Patients Discharged between Quarter 2, 2017 and Quarter 1, 2018 (April 1, 2017 to March 31, 2018)

Patient-Mix Adjustment (PMA)	Comm with Nurses	Comm with Doctors	Responsiveness of Hosp. Staff	Comm. About Medicines	Cleanliness of Hosp. Env.	Quietness of Hosp. Env.	Discharge Information	Care Transition	Hospital Rating	Recommend the Hospital
Education (per level; 1=8th grade or less and 6=More than 4-year college degree)	1.41%	1.44%	2.27%	2.70%	1.35%	3.38%	0.39%	-0.63%	2.42%	0.92%
Self-Rated Health (per level; 1=Excellent and 5=Poor)	4.82%	4.80%	6.20%	4.94%	4.26%	4.43%	1.08%	6.37%	6.32%	5.54%
Response Percentile (per 1% of response percentile)	0.19%	0.18%	0.24%	0.19%	0.06%	0.04%	0.03%	0.19%	0.17%	0.17%
Spanish	-0.76%	-2.57%	-0.91%	-2.50%	1.27%	-4.88%	-1.72%	0.84%	-11.88%	-9.38%
Chinese	5.65%	4.80%	5.84%	3.55%	3.76%	-2.13%	-1.99%	11.06%	3.56%	2.77%
R/V/P/O (Russian, Vietnamese, Portuguese, Other)	1.15%	0.50%	2.36%	0.51%	3.56%	-6.34%	-0.03%	6.30%	1.56%	-0.34%
English (REFERENCE)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Age 18-24	1.98%	0.70%	3.69%	-9.62%	-0.03%	-4.39%	-3.34%	-8.07%	16.22%	12.55%
Age 25-34	-0.58%	-0.82%	-2.66%	-11.66%	-0.75%	-6.46%	-3.98%	-12.17%	12.20%	7.70%
Age 35-44	-1.57%	-1.69%	-4.08%	-11.75%	0.08%	-5.21%	-4.07%	-11.60%	9.26%	5.44%
Age 45-54	-3.95%	-4.21%	-6.35%	-13.32%	-0.38%	-4.54%	-4.88%	-12.70%	4.09%	1.33%
Age 55-64	-4.98%	-5.01%	-7.06%	-12.07%	-0.24%	-3.13%	-5.08%	-11.95%	0.62%	-0.35%
Age 65-74	-5.53%	-5.70%	-6.38%	-10.22%	0.28%	-2.62%	-4.93%	-10.56%	-2.32%	-1.76%
Age 75-84	-3.37%	-3.50%	-3.75%	-5.47%	1.07%	-1.11%	-2.43%	-4.87%	-2.03%	-1.37%
Age 85+ (REFERENCE)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Male Medical	-1.00%	0.39%	-2.18%	-3.66%	-6.64%	0.42%	-2.57%	-0.96%	-0.47%	-2.05%
Male Surgical	-1.38%	-7.93%	-3.00%	-6.40%	-7.43%	-1.09%	-6.59%	-4.74%	-7.01%	-7.35%
Female Surgical	0.06%	-8.65%	-0.97%	-1.83%	-0.70%	-3.19%	-5.56%	-3.49%	-5.91%	-5.22%
Female Maternity	-6.61%	-12.08%	-14.81%	-12.75%	-0.16%	-12.38%	-6.00%	-4.83%	-14.36%	-14.75%
Female Medical (REFERENCE)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Surgical Line * Age[1]	-0.12%	0.39%	-0.40%	0.03%	-0.28%	0.06%	-0.13%	0.05%	0.36%	0.24%
Maternity Line * Age[1]	0.87%	0.95%	0.74%	1.89%	0.27%	0.64%	0.89%	1.22%	2.52%	2.07%

[1] Age takes on the values of 1 to 8 as follows: (1: 18 to 24); (2: 25 to 34); (3: 35 to 44); (4: 45 to 54); (5: 55 to 64); (6: 65 to 74); (7: 75 to 84); and (8: 85+).

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Please see: Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS)

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/final-draft-description-of-hcahps-mode-and-pma-with-bottom-box-modedoc-april-30-2008.pdf

From cited publication below:

We use linear rather than logistic regression models because they are almost identical when sample sizes are large and outcomes are predominantly between 20 and 80 percent (as is the case here) and because linear regression supports simple linear adjustments and variance decompositions. For composite scores, we report the average of the coefficients in models for the constituent report items, which is consistent with the method used to adjust these composites for public reporting. To characterize the importance of patient characteristic adjustments, we standardize the coefficients in units of hospital-level SDs for each outcome. Thus, these coefficients can be interpreted as the change in hospital ranking (in SD) on a given outcome attributable to the change in hospital ranking associated with a one-unit deviation from the overall hospital mean in a single hospital's mean value for a patient-mix variable, holding other patient-mix variables constant. Explanatory power (Zaslavsky 1998) was used to assess the relative importance of individual PMA variables to hospital-level adjustment. Explanatory power is the product of two components: (1) the individual predictive power of a PMA variable (as measured by the improvement in R2 attributable to a candidate predictor) and (2) the hospital-level heterogeneity of a PMA variable.

Explanatory power was greatest for self-reported health status, followed by education, service line, age, emergency room admission, and response order percentile (results not shown). The SD of total adjustment from PMA in terms of hospital-level SDs calculated. These range from 0.19 to 0.50, indicating small to moderate typical adjustments.

"Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS® Hospital Survey Scores." M.N. Elliott, A.M. Zaslavsky, E. Goldstein, W. Lehrman, K. Hambarsoomian, M.K. Beckett and L. Giordano. Health Services Research, 44 (2): 501-518. 2009.

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Please see: Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS)

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/final-draft-description-of-hcahps-mode-and-pma-with-bottom-box-modedoc-april-30-2008.pdf

Please note: For continuous survey-based measures these calibration statistics are not applicable.

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Please see: Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS)

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/final-draft-description-of-hcahps-mode-and-pma-with-bottom-box-modedoc-april-30-2008.pdf

Please note: For continuous survey-based measures these calibration statistics are not applicable.

2b3.9. Results of Risk Stratification Analysis:

Please note: For continuous survey-based measures these calibration statistics are not applicable.

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

The PMA employed for HCAHPS result in data that are more useful to all concerned with improving hospital quality, including the hospitals themselves because changes in survey vendors, survey modes, or patient populations will not disrupt or distort the continuity of valid, comparable scores over time. As previously discussed, explanatory power was greatest for self-reported health status, followed by education, service line,

age, emergency room admission, and response order percentile (results not shown). The SD of total adjustment from PMA in terms of hospital-level SDs calculated. These range from 0.19 to 0.50, indicating small to moderate typical adjustments.

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

For a description of the development and testing of the patient-mix adjustment and survey mode adjustment of the HCAHPS Survey, please see: Mode and Patient-mix Adjustment of the CAHPS® Hospital Survey (HCAHPS)

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/final-draft-description-of-hcahps-mode-and-pma-with-bottom-box-modedoc-april-30-2008.pdf

Also: "Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS® Hospital Survey Scores." M.N. Elliott, A.M. Zaslavsky, E. Goldstein, W. Lehrman, K. Hambarsoomian, M.K. Beckett and L. Giordano. Health Services Research, 44 (2): 501-518. 2009.

Current survey mode adjustments can be found at:

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/april_2017_mode_patient-mix_adjustment_hcahps_survey_mode_adjustment.pdf

Current patient-mix adjustments can be found at:

http://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/2017-12_mode-patient-adj_pma.pdf

---

b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

There are heuristics for noting meaningful differences in CAHPS scores in general (1=small, 3=medium. 5=large on a 0-100 scale). These can be used as an aid to the interpretation of HCAHPS scores. Statistical significance is based on the coefficients for hospital intercepts using a criterion of $p<0.05$ from two-sided hypothesis tests of the hospital mean against the overall mean.

Quigley DD, Elliott MN, Setodji CM, Hays RD (2018) "Quantifying Magnitude of Differences in Patient Experiences with Healthcare Measures" Health Services Research 53(4): 3027-3051 DOI: 10.1111/1475-6773.12828

When a new quarter of HCAHPS data is rolled into public reporting, which occurs four times per year, we publicly report the percentile distribution of top-box and bottom-box results for each HCAHPS measure on our official HCAHPS On-Line Web site, www.HCAHPSonline.org.

Key percentiles are displayed in the **HCAHPS Percentiles Table** for each of the ten publicly reported HCAHPS measures. Both "top box" (most positive) and "bottom box" (least positive) values are shown at the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles. Using the HCAHPS Percentiles Table, one can easily see where a hospital's "top-box" and "bottom-box" score place it relative to other hospitals on each HCAHPS measure.

The "top-box" score indicates how often patients selected the most positive response category when asked about their hospital experience. The higher a hospital's "top-box" score, the higher it ranks among participating hospitals.

The "bottom-box" score, on the other hand, reflects how frequently patients chose negative responses when asked about their hospital experience. Low "bottom-box" scores, and thus a low percentile rank, indicate that a relatively small percentage of a hospital's patients responded negatively about their hospital experience. Conversely, a hospital with a high "bottom-box" score, and thus a high percentile rank, had a relatively large percentage of patients who were critical of their hospital experience.

Please note that while a high "top-box" score is usually associated with a low "bottom-box" score, this is not always the case for a specific hospital. This is because for all HCAHPS measures except Discharge Information, there is also a "middle-box" category, such as *"Usually,"* that varies in size. More information about HCAHPS "box" scores can be found at http://www.hcahpsonline.org/en/summary-analyses/ .

In addition to the percentiles table for the current public reporting period, we provide an archive of results from each public reporting period, beginning with January-December 2008 discharges, at https://www.hcahpsonline.org/en/summary-analyses/.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

Mean, Standard Deviation and Percentiles of HCAHPS Measures, 2016

(Data: 1Q16-4Q16 discharges, ~4,300 hospitals, ~3.1 million completed surveys)

	Mean	Std Dev	Percentile											Inter-Q Range
			100%	99%	95%	90%	75%	50%	25%	10%	5%	1%	0%	
Communication with Nurses	80.34	5.90	100	96	90	88	84	80	77	74	71	65	16	7
Communication with Doctors	82.05	5.80	100	98	92	90	85	82	78	76	74	69	29	7
Responsiveness of Hospital Staff	68.82	9.71	100	95	87	82	75	68	62	58	55	49	20	13
Communication about Medicines	65.32	7.80	100	90	79	75	69	64	61	58	55	48	1	8
Cleanliness	74.50	8.46	100	96	89	85	80	74	69	65	62	56	7	11
Quietness	62.79	10.72	100	90	82	77	69	62	56	50	46	41	3	13
Discharge Information	87.25	4.74	100	97	93	92	90	88	85	82	79	72	6	5
Overall Rating	72.77	9.01	100	93	87	84	78	73	67	62	58	49	24	11
Recommend Hospital	72.14	9.95	100	94	88	84	79	73	66	60	56	48	9	13
Care Transition	52.36	7.79	100	74	65	62	56	52	48	44	40	34	0	8

Please note that the high hospital-level HCAHPS reliabilities (see section 2.a2.3 ensure very good ability to distinguish hospital performance, including establishing differences from the overall average.

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

There are statistically significant and meaningful differences among hospitals on each of the HCAHPS measures. At the recommended sample size of 300 completed surveys, the reliability of all HCAHPS measures is high, indicating good ability to distinguish the performance of a given hospital at the overall national mean. Additional research establishes differences of 5 points on a 0-100 scale as being large differences for HCAHPS.

We observe a very large range of performance across all measures. Some hospitals get nearly perfect scores, median performance is generally between 52% and 88% of the maximum scores, with some hospitals scoring below 30% on each of the measures.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Not Applicable.

Note: *This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

Nonresponse Findings

Logistic regression was used to model the probability of response by eligible patients as a function of available administrative variables (age, gender, service line, emergency room admission, and discharge status). Predictors also included dummies for hospitals and survey modes. A second model added interactions between survey mode and other administrative predictors to test the possibility that patterns of nonresponse differed by mode. Nonresponse weights were defined as the inverse predicted probabilities of response under this model. In order to assess the extent to which the nonresponse weights might correct bias in hospital-level means, for each of the six composites and two global items we assessed the correlation between nonresponse weights and patient-level residuals from the two primary sets of outcome models (with and without PMA).

Nonresponse patterns are summarized in Table 1 and are similar to those observed in Elliott et al. (2005). Wald tests of blocks of interactions by survey mode found evidence that patterns of nonresponse for telephone (p5.002) and active IVR (po.0001) differed from the pattern for the reference group of mail, but that patterns for mixed mode did not (p5.2604). Specifically, the tendency for response rates to increase with age was not as strong with telephone and active IVR as with mail, and the mail tendency for higher response rates for maternity than medical service line was not evident in telephone and active IVR modes (results not shown).

To assess the extent to which the nonresponse weighting might correct bias in hospital-level means, we examined the correlation between nonresponse weights and patient-level residuals with and without PMA. In each case, the null hypothesis corresponds to no association within hospital between weights and outcomes, which would indicate no evidence that nonresponse weighting could systematically affect estimated means and thereby potentially reduce bias. In the absence of PMA, six of eight outcomes were significantly correlated with nonresponse weights (po.05, results not shown). In all instances, this correlation is negative, indicating higher weights (lower predicted probabilities of response) correspond to lower outcome reports. In other words, as noted previously by Elliott et al. (2005), there is a tendency for those individuals with less positive evaluations to be less likely to respond. In the absence of PMA and nonresponse weighting, this pattern would positively bias the scores of hospitals and the bias might be greater in hospitals with lower response rates. On the other hand, PMA reduced the absolute value of all eight correlations and left only one (communication with doctors) statistically significant. This suggests that the use of key nonresponse variables and response order percentile (lag time) in the PMA model adequately addresses the nonresponse bias that would exist without PMA.

While there was evidence of differential nonresponse overall, and evidence that those with lower response propensity had less positive evaluations of care, there was no evidence that nonresponse weighting based on available data improved the accuracy of hospital scores beyond what could be achieved with PMA.

"Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS® Hospital Survey Scores." M.N. Elliott, A.M. Zaslavsky, E. Goldstein, W. Lehrman, K. Hambarsoomian, M.K. Beckett and L. Giordano. Health Services Research, 44 (2): 501-518. 2009.

Table 1: Characteristics of Sampled Discharges and Respondents

Variable	Value	All Surveyed (n=27,229)		Respondents Only (n=7,555)	
		Frequency	Percent	Frequency	Percent
Administrative variables					
Age	18-24	1951	7	344	5***
	25-34	3753	14	880	12†
	35-44	3127	11	758	10
	45-54	3779	14	1077	14***
	55-64	3903	14	1299	17***
	65-74	3996	15	1398	19***
	75-84	4395	16	1330	18***
	85 or older	2325	9	469	6
Male gender		10461	38	2680	35***
Admitted from emergency		12306	45	2951	39***
Discharge status	Sick	6119	22	1377	18***
	Left against medical advice	142	1	20	< 1*
	All other	20968	77	6158	82†
Service line	Surgical	7978	29	2755	36***
	Maternity	3876	14	924	12
	Medical	15206	56	3837	51†
Survey variables					
Survey mode	Mail	6806	33	2239	30†
	Telephone	6808	27	1607	21***
	Mixed	6808	41	2489	33***
	Active IVR	6807	21	1220	16***‡
Education	8th grade or less	–	–	398	5
	Some HS but did not graduate	–	–	809	11
	HS graduate or GED	–	–	2417	32
	Some college or 2-year degree	–	–	2132	28
	Four-year college graduate	–	–	900	12
	More than 4-year college graduate	–	–	899	12
Self-rated health	Excellent	–	–	972	13
	Very good	–	–	1935	26
	Good	–	–	2392	32
	Fair	–	–	1604	21
	Poor	–	–	652	9
Primary language other than English		–	–	332	4

*p<.05.

***p<.001.

†Reference category for multivariate analyses of nonresponse.

‡69% of mixed mode responses were by mail.

IVR, interactive voice response.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

Missing data varies by survey item (see below) and across hospitals. As can be observed, there are higher missing rates for certain survey questions, specifically items 11, 16, 17, 19 and 20. These items are preceded by a screener item, thus a certain percentage of patients are not eligible for these questions based on their hospital experience, so these questions are not applicable to those patients.

Frequency of Missing Data for HCAHPS Survey Items and Administrative Variables, 2016

(Data: 1Q16-4Q16 discharges, ~4,300 hospitals, ~3.1 million completed surveys)

	Missing Rate
Q1: Nurse Courtesy Respect	0.67%
Q2: Nurse Listen	0.68%
Q3: Nurse Explain	0.56%
Q4: Call Button	1.31%
Q5: Dr Courtesy Respect	0.78%
Q6: Dr Listen	1.17%
Q7: Dr Explain	0.92%
Q8: Cleanliness	1.15%
Q9: Quiet	1.27%
Q11: Bathroom Help	4.70%
Q16: Med Description	7.91%
Q17: Side Effects	8.59%
Q19: Help After Discharge	4.61%
Q20: Symptoms	4.99%
Q21: Overall Rating	1.65%
Q22: Recommend	2.59%
Q23: CT Preferences	4.20%
Q24: CT Understanding	2.58%
Q25: CT Purpose of Medicine	3.32%
Q27: Self-Rated Health	4.40%
Q29: Education	6.80%
Q32: Language Spoken at Home	7.96%
Patient Service Line (administrative)	4.01%
Lag Time (administrative)	0.00%
Patient Age (administrative)	0.03%

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

Missing values for survey questions about patient experience are not imputed or otherwise utilized in HCAHPS score calculations. Missing values for patient demographic or administrative variables are imputed using other patient characteristics and hospital distributions of non-missing values where applicable. The amount of missing data is very small except for items that are not applicable to all HCAHPS eligible patients.

Every quarter the HCAHPS Project Team examines hospitals that have a high rate of missing data for HCAHPS survey items, HCAHPS demographic (patient mix) items, and hospital administrative data. Inquiries are made with the hospital or its survey vendor to understand the reason for the missing data and to discuss mitigation strategies, as necessary.

Although there was evidence of selective nonresponse, the patient-mix adjustment model employed was found to effectively account for any nonresponse bias that could have been addressed through nonresponse weighting. Therefore, no further weighting or adjustment for nonresponse is need

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Other

If other: Patient survey.

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in a combination of electronic sources

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

The HCAHPS Survey is administered in four modes: mail, telephone, mixed (mail with telephone follow-up), or Interactive Voice Response. It is not feasible to convert the HCAHPS Survey into an eMeasure. HCAHPS survey and administration procedures are provided in detail in the HCAHPS Quality Assurance Guidelines, V14.0. See: https://www.hcahpsonline.org/en/quality-assurance/

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For

eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

In terms of survey implementation, data submission and oversight, the HCAHPS Survey is improved on an annual basis. In the annual HCAHPS Update Training and in the annual HCAHPS Quality Assurance Guidelines, as well as in the frequently updated "What´s New" section of the HCAHPS On-Line Web site,

https://www.hcahpsonline.org/en/whats-new/, we provide a running summary of changes to the survey and its administration procedures. For example, see HCAHPS Quality Assurance Guidelines, V14.0, pp. 5-22.

Hospitals that participate in HCAHPS receive, from the CMS Hospital Inpatient Quality Reporting program, an individual Preview Report of their publicly reported scores about two months before these scores are publicly reported on the Hospital Compare Web site.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

There are no fees or charges associated with participation in the HCAHPS Survey, HCAHPS training or oversight, or access to publicly reported HCAHPS scores.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
	Public Reporting Hospital Compare http://www.medicare.gov/hospitalcompare/search.html Hospital Compare http://www.medicare.gov/hospitalcompare/search.html Payment Program Hospital Value-Based Purchasing http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/index.html?redirect=/Hospital-Value-Based-Purchasing/ Hospital Value-Based Purchasing http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/index.html?redirect=/Hospital-Value-Based-Purchasing/

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

Hospital Inpatient Quality Reporting program; Hospital Compare program.
Centers for Medicare & Medicaid Services (CMS)
The Hospital Compare web site contains information about the quality of care at over 4,500 Medicare-certified hospitals. You can use Hospital Compare to find hospitals and compare the quality of their care.
The information on Hospital Compare:
Can help you make decisions about where you get your health care
Encourages hospitals to improve the quality of care they provide
Hospital Value-Base Purchasing Program.
Centers for Medicare & Medicaid Services (CMS)
This website will be CMS´ official source of information about the Hospital Value-based Purchasing (HVBP) Program for hospitals, clinicians, and other stakeholders who share CMS´ commitment to transforming the quality of hospital care by realigning hospitals´ financial incentives to do so.
Over 3,000 IPPS hospitals across the USA participate in the HVBP program.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)
N/A

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

N/A

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

HCAHPS Survey data are publicly reported on Hospital Compare on the Medicare.gov website, as well as in a downloadable database. All of the measures derived from the HCAHPS survey are updated four times per year. The Hospital Compare web site includes additional information about the data and its interpretation. Inpatient Prospective Payment System (IPPS) hospitals are required to contract with an approved HCAHPS Survey vendor and have that vendor administer the HCAHPS Survey and submit data on a quarterly basis to CMS. All hospitals that participate in public reporting on Hospital Compare (~4,500) receive a Preview Report prior to each quarterly public reporting that contains all of their HCAHPS scores and HCAHPS information for them to view prior to the data being publicly reported.

All IPPS hospitals that participate in Hospital Value-Based Purchasing (~3,000) receive additional reports on an annual basis from CMS that contain their scores on the HCAHPS domain used in the Hospital VBP for pay-for-performance program.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

HCAHPS scores for all participating, publicly reported hospitals, which include the top-box, middle-box and bottom-box score and star ratings for each HCAHPS measure; state and national averages; number of completed surveys; and survey response rate, as well explanations of the survey items and measures and any footnotes, are available on the Hospital Compare website at https://www.medicare.gov/hospitalcompare/search.html. HCAHPS scores are updated on a quarterly basis, currently in January, April, July and October. Specific details about the data are provided at https://www.medicare.gov/hospitalcompare/Data/Overview.html.

In addition, Preview Reports are provided to all hospitals prior to their data being publicly reported on Hospital Compare. Preview Reports contain all of the information that will be publicly reported, as well as extensive information on how to interpret CMS quality measures, including HCAHPS measures. Even hospitals whose HCAHPS scores will not be publicly reported, for instance because the number of completed surveys is fewer than 25, will be able to see their HCAHPS scores in the Preview Report.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

Feedback on the HCAHPS Survey is regularly collected from HCAHPS survey vendors and self-administering hospitals via conference calls and site visits. In addition, CMS meets with provider groups, such as hospital associations and patient advocacy groups, to hear their concerns. Feedback is also solicited from anyone through the formal comment and response mechanism of the federal rulemaking process, primarily through the annual Inpatient Prospective Payment System (IPPS) rule, the Outpatient Prospective Payment System (OPPS) rule, and through special rules associated with legislation, such as the Patient Protection and Affordable Care Act of 2010 (PPACA) and the SUPPORT Act of 2018.

In addition, feedback on the HCAHPS survey content and protocols is gathered on an ongoing basis from approved HCAHPS survey vendors and self-administering hospitals during annual or biennial on-site visits or conference calls with the HCAHPS Project Team. Feedback is also received during the annual HCAHPS introductory and update training webinars, and continuously through inquiries to the HCAHPS technical assistance Help Desk.

4a2.2.2. Summarize the feedback obtained from those being measured.

Feedback received reflects the importance of this survey for driving hospital improvement. Recent suggestions for improvement include using a web-based mode to collect survey data and removing the pain items from the HCAHPS Survey. Consequently, the pain items have been removed from the survey and CMS has sought approval from the Office of Management and Budget to test web-based modes for patient surveys.

4a2.2.3. Summarize the feedback obtained from other users

CMS receives feedback on the HCAHPS Survey from many different parties: patients, caregivers, hospitals, hospital associations, physicians, patient advocacy groups, survey vendors, the Office of Management and Budget, researchers and academics. In general, this feedback is supportive of the goals and methods of HCAHPS and attests that HCAHPS both provides important information about hospital quality for consumer choice and motivates hospitals to improve the quality of care they provide to all patients. However, feedback sometimes reflects the contrary goals of different stakeholders, such as shortening the survey (hospitals), or adding items to the survey (patient advocacy groups). CMS monitors and assesses feedback. When the feedback is in accord with the policy objectives of the HCAHPS Survey and practicable, CMS makes changes to the survey or its administration protocols. An example of this responsiveness is the creation of official translations of the survey in other languages, such as Vietnamese, Portuguese and, most recently, German.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

CMS is looking into the possibility of electronic administration of several CAHPS surveys, including Hospital CAHPS. A generic OMB Paperwork Reduction Act package is now going through the approval process to give us the ability to test the web mode for HCAHPS. We modified the survey by removing the pain items that made up the pain measure beginning with patients discharged October 1, 2019.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

Objective. Measure HCAHPS improvement in hospitals participating in the second

and fifth years of HCAHPS public reporting; determine whether change is greater for some hospital types.

Data. Surveys from 4,822,960 adult inpatients discharged July 2007–June 2008 or

July 2010–June 2011 from 3,541 U.S. hospitals.

Study Design. Linear mixed-effect regression models with fixed effects for time,

patient mix, and hospital characteristics (bedsize, ownership, Census division, teaching status, Critical Access status); random effects for hospitals and hospital-time interactions; fixed-effect interactions of hospital characteristics and patient characteristics (gender, health, education) with time predicted HCAHPS measures correcting for regression-to-the-mean biases.

Data Collection Methods. National probability sample of adult inpatients in any of four approved survey modes.

Principal Findings. HCAHPS scores increased by 2.8 percentage points from 2008

to 2011 in the most positive response category. Among the middle 95 percent of hospitals, changes ranged from a 5.1 percent decrease to a 10.2 percent gain overall. The greatest improvement was in for-profit and larger (200 or more beds) hospitals.

Conclusions. Five years after HCAHPS public reporting began, meaningful

improvement of patients' hospital care experiences continues, especially among initially low-scoring hospitals, reducing some gaps among hospitals.

For more details, please see:

"Accelerating Improvement and Narrowing Gaps: Trends in Patients´ Experiences with Hospital Care Reflected in HCAHPS Public Reporting." M.N. Elliott, C.W. Cohea, W.G. Lehrman, E.H. Goldstein, P.D. Cleary, L.A. Giordano, M.K. Beckett and A.M. Zaslavsky. Health Services Research,

50: 1850-1867. 2015. http://onlinelibrary.wiley.com/doi/10.1111/1475-6773.12305/full

For information and findings about the use of HCAHPS scores in the Hospital Value-Based Purchasing program, please see:

"Understanding the Role Played by Medicare's Patient Experience Points System in Hospital Reimbursement." M.N. Elliott, M.K. Beckett, W.G. Lehrman, P.D. Cleary, C.W. Cohea, L.A. Giordano, E.H. Goldstein and C.L. Damberg. Health Affairs, 35 (9): 1673-1680. 2016. Published online, 9-7-16: http://content.healthaffairs.org/content/35/9/1673

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

The most significant unintended consequence for the HCAHPS Survey was the allegation that the three pain management questions originally in the HCAHPS Survey created pressure on hospital physicians to over-prescribe pain relief medications, including opioids, in hopes of obtaining more positive responses to these questions.

CMS and the HCAHPS Project Team never found credible, empirical evidence to substantiate this allegation

- (see, "Measurement of the Patient Experience: Clarifying Facts, Myths, and Approaches." L. Tefera, W.G. Lehrman and P. Conway. Journal of the American Medical Association. 2016. 315: 2167-2168. Published online, 3-10-16. http://jama.jamanetwork.com/article.aspx?articleid=2503222).
- However, from an abundance of caution in the midst of a nationwide opioid over-use crisis, CMS undertook a series of steps in response, which included:
- Removing the Pain Management dimension from the HCAHPS domain in the Hospital VBP program in FY 2017;
- Replacing the original pain management questions with three new questions that focused on hospital staff's communication with patients about pain in January 2018 (see, "A Special Contribution from the Centers for Medicare and Medicaid Services: Valuing Patient Experience While Addressing the Prescription Opioid Epidemic." L. Tefera, W.G. Lehrman, E.G. Goldstein and S. Agrawal. Annals of Emergency Medicine. 2016. Published online, 7-19-16. http://www.annemergmed.com/article/S0196-0644(16)30367-5/fulltext);
- Removing the Pain Management measure from Hospital Compare public reporting in July 2018;
- Finally, in compliance with the national Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act (Pub. L. 115-271) of 2018 (Section 6104), CMS will remove the three Communication About Pain items from the HCAHPS Survey beginning with patients discharged in October 2019. In addition, CMS will not publicly report the Communication About Pain measure.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

There have been a number of benefits associated with the creation and implementation of the HCAHPS Survey, but perhaps the greatest has been the increased attention paid to patient experience as a unique, independent, worthy and actionable element of hospital quality. Evidence of this can be found in an increased interest in ways to improve patient experience of care.

- (see, “CAHPS Surveys: Valid and Valuable Measures of Patient Experience.” W.G. Lehrman and M.W. Friedberg. Hastings Center Report. 45 (6): 3-4. 2015. http://onlinelibrary.wiley.com/doi/10.1002/hast.507/full), In the years since HCAHPS was implemented and publicly reported there has been growth in the number of patient experience (PX) departments and officials in hospitals, PX journals and conferences, published research on patient experience of care, use of HCAHPS-like surveys in other nations, as well as a general improvement in HCAHPS scores. While we do not claim these developments are due to HCAHPS, they do underscore the increased attention to and legitimization of this aspect of hospital quality over the past 15 years.
- - (see, “Hospital Survey Shows Improvements in Patient Experience.” M.N. Elliott, W.G. Lehrman, E.H. Goldstein, L.A. Giordano, M.K. Beckett, C.W. Cohea and P.D. Cleary. Health Affairs, 29 (11): 2061-2067. 2010. http://content.healthaffairs.org/content/29/11/2061.abstract, and
- “Accelerating Improvement and Narrowing Gaps: Trends in Patients´ Experiences with Hospital Care Reflected in HCAHPS Public Reporting.” M.N. Elliott, C.W. Cohea, W.G. Lehrman, E.H. Goldstein, P.D. Cleary, L.A. Giordano, M.K. Beckett and A.M. Zaslavsky. Health Services Research, 50: 1850-1867. 2015. http://onlinelibrary.wiley.com/doi/10.1111/1475-6773.12305/full).

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

2548 : Child Hospital Consumer Assessment of Healthcare Providers and Systems (Child HCAHPS) Survey

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

Child HCAHPS is NQF endorsed: NQF #2548.

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications harmonized to the extent possible?

No

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

We are not aware of other measures that have the same measure focus or target the same population as HCAHPS, NQF 0166.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
N/A

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment **Attachment:** HCAHPS-_NQF_0166-_A1_Supplemental_materials-_4-9-19-636904227582638456.docx

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare & Medicaid Services

Co.2 Point of Contact: Helen, Dollar-Maples, Helen.Dollar-Maples@cms.hhs.gov, 410-786-7214-

Co.3 Measure Developer if different from Measure Steward: AHRQ

Co.4 Point of Contact: Caren, Ginsberg, caren.ginsberg@ahrq.hhs.gov, 301-427-1894-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

The CAHPS® II Investigators and the Agency for Healthcare Research and Quality (AHRQ)

540 Gaither Road

Rockville, Maryland 20850

The Division of Consumer Assessment & Plan Performance of the Centers for Medicare & Medicaid Services (CMS) participated in the development and testing of the HCAHPS Survey, as well as having primary responsibility for its ongoing national implementation, public reporting, oversight, analysis and use in CMS programs.

Members involved in the initial development of the HCAHPS Survey (2002-2005) are listed below.

Chuck Darby, AHRQ

Christine Crofton, AHRQ

Marybeth Farquhar, AHRQ

Liz Goldstein, CMS

William Lehrman, CMS

Steven Garfinkel, AIR

Julie Brown, RAND

Shoshanna Sofaer, AIR

Elizabeth Hoy, Polaris Consulting

Jenny Crabb, Westat

Rebecca Matthew, AIR

Kimberly Hepner, RAND

Paul Cleary, Yale

Carol Edwards, RAND

Katrin Hambarsoomians, RAND

Lawrence Zaborski, Harvard

The following members provided statistical/methodological consultations.

Marc Elliott, RAND

Jack Fowler – University of Massachusetts, Boston

Steven Garfinkel – AIR

Ron Hays – Rand

San Keller – AIR

Roger Levine – AIR

James O'Malley, Harvard

Alan Zaslavsky, Harvard

Currently, the following staff are involved in ongoing HCAHPS updates, testing and statistical analyses.

Marc Elliott, RAND

Laura Giordano, HSAG

Chris Cohea, HSAG

William Lehrman, CMS

Liz Goldstein, CMS

Christine Payne, CMS

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2006

Ad.3 Month and Year of most recent revision: 01, 2018

Ad.4 What is your frequency for review/update of this measure? The HCAHPS is reviewed continually and is updated as needed.

Ad.5 When is the next scheduled review/update for this measure? 10, 2019

Ad.6 Copyright statement: CAHPS® is a registered trademark of the Agency for Healthcare Research and Quality, a U.S. Government agency.

Ad.7 Disclaimers:

Ad.8 Additional Information/Comments: In terms of survey implementation, data submission and oversight, the HCAHPS Survey is improved on an annual basis through annual training and annual updates and revisions to the HCAHPS Quality Assurance Guidelines. All information about the HCAHPS Survey is posted on our official HCAHPS On-Line Web site, www.HCAHPSonline.org. In the near future, CMS will post on HCAHPS On-

Line Web site a matrix of all changes from the previous iteration of the HCAHPS Quality Assurance Guidelines, V13.0, to the current version, HCAHPS Quality Assurance Guidelines, V14.0.