NATIONAL QUALITY FORUM
Driving measurable health improvements together

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through National Quality Forum's (NQF) Consensus Development Process (CDP). The information submitted by the measure developers/stewards is included after the *Brief Measure Information* and *Preliminary Analysis* sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 2651

Measure Title: CAHPS® Hospice Survey, Version 9.0

Measure Steward: Centers for Medicare & Medicaid Services

Brief Description of Measure: The measures submitted here are derived from the CAHPS® Hospice Survey, Version 9.0, a 39-item standardized questionnaire and data collection methodology. The survey is intended to measure the care experiences of hospice decedents and their primary caregivers. Survey respondents are the primary informal caregivers (i.e., family members or friends) of patients who died while receiving hospice care.

The proposed measures include the following six multi-item measures:

- Hospice Team Communication
- Care Preferences
- Getting Timely Care
- Treating Family Member with Respect
- Getting Emotional and Religious Support
- Getting Help for Symptoms

In addition, there are three single-item measures:

- Getting Hospice Training
- Rating of the Hospice
- Willingness to Recommend the Hospice

Following is a list of the survey items included in each measure.

Hospice Team Communication (5 items)

- How often did the hospice team keep you informed about when they would arrive to care for your family member?
- How often did the hospice team explain things in a way that was easy to understand?
- How often did the hospice team listen carefully to you when you talked with them about problems with your family member's hospice care?

- How often did the hospice team keep you informed about your family member's condition?
- While your family member was in hospice care, how often did the hospice team listen carefully to you?

Care Preferences (2 items)

- Did the hospice team make an effort to listen to the things that mattered most to you or your family member?
- Did the hospice team provide care that respected your family member's wishes?

Getting Timely Care (2 items)

- When you or your family member asked for help from the hospice team, how often did you get help as soon as you needed it?
- How often did you get the help you needed from the hospice team during evenings, weekends, or holidays?

Treating Family Member with Respect (2 items)

- How often did the hospice team treat your family member with dignity and respect?
- How often did you feel that the hospice team really cared about your family member?

Getting Emotional and Religious Support (3 items)

- While your family member was in hospice care, how much emotional support did you get from the hospice team?
- In the weeks after your family member died, how much emotional support did you get from the hospice team?
- Support for religious or spiritual beliefs includes talking, praying, quiet time, or other ways of meeting your religious or spiritual needs. While your family member was in hospice care, how much support for your religious and spiritual beliefs did you get from the hospice team?

Getting Help for Symptoms (4 items)

- Did your family member get as much help with pain as he or she needed?
- How often did your family member get the help he or she needed for trouble breathing?
- How often did your family member get the help he or she needed for trouble with constipation?
- How often did your family member get the help he or she needed from the hospice team for feelings of anxiety or sadness?

Getting Hospice Care Training (1 item)

- Hospice teams may teach you how to care for family members who need pain medicine, have trouble breathing, are restless or agitated, or have other care needs. Did the hospice team teach you how to care for your family member?

Rating of Hospice Care (1 item)

- Using any number from 0 to 10, where 0 is the worst hospice care possible and 10 is the best hospice care possible, what number would you use to rate your family member's hospice care?

Willingness to Recommend Hospice (1 item)

- Would you recommend this hospice to your friends and family?

A complete list of proposed CAHPS Hospice Survey measures, including response options for each item, is available in Appendix B.

Developer Rationale: A growing number of Medicare decedents use hospice care: in 2019, more than 1.6 million Medicare beneficiaries – more than half of decedents – received hospice services, compared to fewer than a quarter of decedents in 2000. Total Medicare hospice expenditures were nearly $21 billion in that year

(Medicare Payment Advisory Commission 2019). The number of hospice providers has also increased dramatically over time, more than doubling from 2,255 in 2000 to 4,840 in 2019.

There is striking variation across hospices with regard to processes of care associated with care quality, including the proportion of patients discharged alive (Prisc, Plotzke et al. 2016), provision of general inpatient care or continuous home care in the last seven days of life (Plotzke, Christian et al., 2014), and visits from professional hospice staff in the last two days of life (Teno, Plotzke et al. 2016). Evidence of variation in hospice care processes points to the need for performance measures that can be systematically implemented to monitor hospice performance and make fair comparisons across hospices.

Patient- and family-centeredness of care is a central goal of hospice care and can be directly measured through surveys of family caregivers of hospice patients (Anhang Price and Elliott 2018). CMS created the CAHPS® Hospice Survey, a component of the Hospice Quality Reporting Program (HQRP), to ensure that an assessment of the patient- and family-centeredness of care would be included to monitor hospice performance, promote quality improvement, and inform consumer decision making in the selection of a hospice via public reporting of results. To date, scores on the CAHPS Hospice Survey measures have shown more variation and potential for improvement than other measures within the HQRP.

The CAHPS Hospice Survey is the first survey to use standardized content and implementation and analytic protocols to allow for fair comparisons between hospices. CMS continuously reviews these protocols to ensure that the survey is as feasible and fair as possible.

Citations:

1. Anhang Price R, Elliott MN. (2018). Measuring Patient-Centeredness of Care for Seriously Ill Individuals: Challenges and Opportunities for Accountability Initiatives. J Pall Med. 21(S2):S28-S35.
2. Medicare Payment Advisory Commission. (March 2021). Chapter 11: Hospice Services. In: Report to the Congress: Medicare Payment Policy.
3. Prsic E, Plotzke M, Christian TJ, Gozalo P, Teno JM. (2016). A National Study of Live Hospice Discharges Between 2000 and 2012. J Pall Med. 19(9):987-90.
4. Teno JM, Plotzke M, Christian T, Gozalo P. (2016). Examining Variation in Hospice Visits by Professional Staff in the Last 2 Days of Life. JAMA Intern Med. 176(3):364-370.

Numerator Statement: CMS calculates CAHPS Hospice Survey measure scores using top-, middle- and bottom-box scoring. The top-box score refers to the percentage of caregiver respondents that give the most positive response(s). The bottom box score refers to the percentage of caregiver respondents that give the least positive response(s). The middle box is the proportion remaining after the top and bottom boxes have been calculated; see below for details. Details regarding the definition of most and least positive response(s) are noted in Section SP.14 below.

Denominator Statement: In national implementation and public reporting, CAHPS® Hospice Survey measure scores are calculated only for hospices that had at least 30 completed questionnaires over the most recent eight quarters of data collection.

The target population for the survey are the adult primary caregivers of hospice decedents. Respondent eligibility and exclusions are defined in detail in the sections that follow. A survey is defined as completed when at least 50 percent of the questions applicable to all decedents/caregivers are answered. The survey uses screener questions to identify respondents eligible to respond to subsequent items. Therefore, denominators vary by survey item (and corresponding multi-item measures, if applicable) according to the eligibility of respondents for each item. In addition, for the Getting Hospice Care Training measure, scores are calculated only among those respondents who indicate that their family member received hospice care at home or in an assisted living facility.

Denominator Exclusions: The exclusions noted here are those who are ineligible to participate in the survey. The one exception is caregivers who report on the survey that they “never” oversaw or took part in the decedent’s care; these respondents are instructed to complete the “About You” and “About Your Family Member” sections of the survey only.

Cases are excluded from the survey target population if:

- The hospice patient is still alive
- The decedent’s age at death was less than 18
- The decedent died within 48 hours of his/her last admission to hospice care
- The decedent had no caregiver of record
- The decedent had a caregiver of record, but the caregiver does not have a U.S. or U.S. Territory home address
- The decedent had no caregiver other than a nonfamilial legal guardian
- The decedent or caregiver requested that they not be contacted (i.e., by signing a no publicity request while under the care of hospice or otherwise directly requesting not to be contacted)
- The caregiver is institutionalized, has mental/physical incapacity, has a language barrier, or is deceased
- The caregiver reports on the survey that he or she “never” oversaw or took part in decedent’s hospice care

Measure Type: Outcome: PRO-PM

Data Source: Instrument-Based Data

Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: 10/26/2016 08:28 AM

Most Recent Endorsement Date: 11/20/2020 12:23:38 PM

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement, endorsed measures are evaluated periodically to ensure that the measure 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 a 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 are not available, data demonstrating wide variation in performance can be used, assuming the data are from a robust number of providers and the results are not subject to systematic bias. For measures derived from a patient report, the evidence also should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

The developer provides the following description for this patient reported outcome maintenance measure:

- Since the prior review in 2019, updates were made in response to stakeholder requests for the survey instrument to be shortened and simplified. New content related to Care Preferences was added in response to feedback received. All proposed changes were tested with caregivers of hospice decedents via cognitive interviews. CMS tested the revised survey instrument among caregivers from 56 hospices during a 2021 mode experiment.
- The developer provides a logic model that describes updates from the 2019 submission. The developer stated that the conceptual model described in the previous submission is still applicable to the revised CAHPS Hospice Survey measures under review. Minor updates were made to the model to reflect the important role that hospice has in explaining care options, formulating goals of care that reflect patient and family preferences, and then creating a plan of care that aims to achieve those goals. The key process of hospice care is assessed by the new CAHPS Hospice Survey Care Preferences measure.

This submission includes measures derived from the CAHPS® Hospice Survey, Version 9.0, a 39-item standardized questionnaire and data collection methodology. The survey is intended to measure the care experiences of hospice decedents and their primary caregivers. Survey respondents are the primary informal caregivers (i.e., family members or friends) of patients who died while receiving hospice care.

The proposed measures include the following six multi-item measures:

- Hospice Team Communication
- Care Preferences
- Getting Timely Care
- Treating Family Member with Respect
- Getting Emotional and Religious Support
- Getting Help for Symptoms

In addition, there are three single-item measures:

- Getting Hospice Training
- Rating of the Hospice
- Willingness to Recommend the Hospice

Following is a list of the survey items included in each measure.

Hospice Team Communication (5 items)

- How often did the hospice team keep you informed about when they would arrive to care for your family member?
- How often did the hospice team explain things in a way that was easy to understand?
- How often did the hospice team listen carefully to you when you talked with them about problems with your family member's hospice care?
- How often did the hospice team keep you informed about your family member's condition?
- While your family member was in hospice care, how often did the hospice team listen carefully to you?

Care Preferences (2 items)

- Did the hospice team make an effort to listen to the things that mattered most to you or your family member?
- Did the hospice team provide care that respected your family member's wishes?

Getting Timely Care (2 items)

- When you or your family member asked for help from the hospice team, how often did you get help as soon as you needed it?

- How often did you get the help you needed from the hospice team during evenings, weekends, or holidays?

Treating Family Member with Respect (2 items)

- How often did the hospice team treat your family member with dignity and respect?
- How often did you feel that the hospice team really cared about your family member?

Getting Emotional and Religious Support (3 items)

- While your family member was in hospice care, how much emotional support did you get from the hospice team?
- In the weeks after your family member died, how much emotional support did you get from the hospice team?
- Support for religious or spiritual beliefs includes talking, praying, quiet time, or other ways of meeting your religious or spiritual needs. While your family member was in hospice care, how much support for your religious and spiritual beliefs did you get from the hospice team?

Getting Help for Symptoms (4 items)

- Did your family member get as much help with pain as he or she needed?
- How often did your family member get the help he or she needed for trouble breathing?
- How often did your family member get the help he or she needed for trouble with constipation?
- How often did your family member get the help he or she needed from the hospice team for feelings of anxiety or sadness?

Getting Hospice Care Training (1 item)

- Hospice teams may teach you how to care for family members who need pain medicine, have trouble breathing, are restless or agitated, or have other care needs. Did the hospice team teach you how to care for your family member?

Rating of Hospice Care (1 item)

- Using any number from 0 to 10, where 0 is the worst hospice care possible and 10 is the best hospice care possible, what number would you use to rate your family member's hospice care?

Willingness to Recommend Hospice (1 item)

- Would you recommend this hospice to your friends and family?

A complete list of proposed CAHPS Hospice Survey measures, including response options for each item, is available in Appendix B.

Summary of prior review in 2019

- The developer identified several structures and processes of care that can impact the first six measures. Although not stated explicitly, these activities likely also would affect overall ratings of the care provided and willingness to recommend the hospice.
- To assess topics of interest important to caregivers of hospice patients, the developer conducted focus group and individual interviews with family members of hospice decedents. Results from these activities suggested that caregivers find communication, information, and respect to be important facets of high-quality hospice care.
- The Committee agreed the evidence presented met NQF's requirements for patient-reported outcome measures and passed all eight measures on the evidence criterion.

Changes to evidence from the 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:

Question for the Standing Committee:

- *The developer attests the underlying evidence for the measure has not changed since the last NQF endorsement review. Does the Committee agree the evidence basis for the measure has not changed and there is no need for repeat discussion and vote on Evidence?*

Guidance From the Evidence Algorithm

Pro-based measure (Box 1) A Relationship between the outcome and at least one healthcare action is identified and supported by the rationale (Box 2) PASS

Preliminary rating for evidence: ☒ Pass ☐ No Pass

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

Maintenance measures – increased emphasis on gap and variation

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

- Performance data were calculated from the 54 hospices participating in the 2021 mode experiment that had 30 or more completed surveys. Measures with the highest top-box scores were Getting Emotional and Religious Support (top-box mean=92.4; IQR: 4.4), Treating Family Member with Respect (top-box mean=91.3; IQR: 3.3), and Care Preferences (top-box mean=90.2; IQR =5.7). The measures with the lowest top-box scores were Getting Timely Care (top-box mean=77.5; IQR: 9.2) and Getting Help for Symptoms (top-box mean=78.8; IQR: 6.4).

Table 1. Distribution of Top-Box Scores for CAHPS Hospice Survey Measures, 2021 Mode Experiment

*	Mean	Standard Deviation	25th Percentile	50th Percentile	75th Percentile
Multi-item Measures	*	*	*	*	*
Hospice Team Communication	81.5	5.6	79.8	82.5	84.9
Getting Timely Care	77.5	7.7	73.9	78.3	83.1
Treating Family Member with Respect	91.3	3.7	90.3	92.0	93.6
Getting Emotional and Religious Support	92.4	3.5	90.2	92.7	94.6
Getting Help for Symptoms	78.8	5.6	76.2	79.9	82.6
Care Preferences	90.2	4.4	87.5	91.4	93.2
Single-Item Measures	*	*	*	*	*
Getting Hospice Care Training	82.4	7.6	77.9	84.2	86.8
Rating of Hospice	84.8	7.2	81.7	86.3	89.6
Willingness to Recommend	86.0	6.9	82.3	87.9	90.8

* Cell intentionally left empty

Disparities

- The developer summarized analyses of 2021 CAHPS Hospice Survey mode experiment data, including 4,749 caregiver respondents from 56 hospices noting potential disparities in the experience of care measures. Findings include: caregivers of male decedents report poorer care experiences than caregivers of female decedents by 1.5 to 3.7 percentage points for all measures; caregivers of patients age 65 to 74 report better care experiences than caregivers of patients age 90 or older by 0.2 to 4.4 percentage points for all measures; caregivers of decedents with both Medicare and Medicaid insurance report poorer care experiences than caregivers of those with only Medicare insurance by 4.0 to 13.9 percentage points for eight of nine measures; and caregivers of decedents who are Black report better experiences than caregivers of White decedents on Hospice Team Communication (4.1 percentage points higher; $p<0.05$) but similar or worse experiences for all other measures, and caregivers of decedents who are Hispanic report worse experiences than caregivers of White decedents with regard to Emotional and Religious Support (3.4 percentage points lower; $p<0.05$), but similar care experiences for other measures.

Questions for the Standing Committee:

- *Is there a gap in care that warrants a national performance measure?*

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

Criteria 2: Scientific Acceptability of Measure Properties

Complex measure evaluated by the Scientific Methods Panel (SMP)? ☒ **Yes** ☐ **No**

Evaluators: Staff/ [First Name Last Name]; [First Name Last Name]; [First Name Last Name](Combined Methods Panel Review)

- The SMP passed the measure on Reliability with a score of: H-6; M-3; L-2; I-0.
- The SMP passed the measure on Validity with a score of: H-1; M-5; L-2; I-2.

Measure evaluated by the Technical Expert Panel (TEP)? ☐ **Yes** ☒ **No**

2a. Reliability: Specifications and Testing

For maintenance measures—no change in emphasis—specifications should be evaluated the same as with new measures.

2a1. Specifications require the measure, as specified, to produce consistent (i.e., reliable) and credible (i.e., valid) results about the quality of care when implemented.

For maintenance measures – less emphasis if no new testing data are provided.

2a2. Reliability testing demonstrates whether the measure data elements are repeatable and producing the same results a high proportion of the time when assessed in the same population during the same time period, and/or whether the measure score is precise enough to distinguish differences in performance across providers.

Specifications:

- Have the measure specifications changed since the last review? ☒ Yes ☐ No
    - The survey instrument was revised based on feedback from a developer convened TEP and public comments during the 2019–2020 maintenance cycle to shorten and simplify the instrument as well as to add a new two-item Care Preferences measure.

- Measure specifications include the specific instrument (e.g., patient-reported outcome measure [PROM]); standard methods, modes, and languages of administration; whether proxy responses are allowed; standard sampling procedures; and handling of missing data

Reliability Testing:

- Did the developer conduct new reliability testing? ☐X Yes ☐ No
- Reliability testing conducted at the Patient/Encounter Level:
    - Multi-item measure reliability was assessed using Cronbach's alpha (internal consistency). Cronbach's alpha was >=0.70 for five of six multi-item measures; it was 0.62 for one multiitem measure (Getting Timely Care). Cronbach's alpha when an item was deleted decreased for all but one item.
    - Multi-item measure reliability was also assessed using the person-level Pearson item-total correlation (relation of each item to all other items). Item-total Pearson correlation ranged from 0.45–0.71.
- Reliability testing conducted at the Accountable Entity Level:
    - Inter-unit (hospice-level) reliability was calculated at the mean sample sizes, using intraclass correlation coefficients (ICCs) calculated from the case mix-adjusted 0–100 top-box scores, applying the Spearman Brown prediction formula. Hospice-level reliability at the average number of respondents ranged from 0.70–0.84 on six multi-item measures and 0.70–0.87 on three single-item measures.
    - The developer also cites published research assessing the stability of responses to survey items that assess the overall quality of hospice care and willingness to recommend the hospice with repeated measurement at three, six, and nine months following the death of the patient, with agreement of 86 percent or higher for overall ratings of hospice care quality and 90 percent or higher for willingness to recommend the hospice. Kappa statistics ranged from 0.58 for the willingness to recommend to 0.70 for overall quality in repeated measures.

SMP Summary:

- The SMP did not note any concerns with the reliability of this measure.

Questions for the Standing Committee regarding reliability:

- *Do you have any concerns that the measure cannot be consistently implemented (i.e., are the measure specifications adequate)?*
- *The SMP is satisfied with the reliability testing for the measure. Does the Standing Committee think there is a need to discuss and/or vote on reliability?*
- *The developer attests that the specifications have not changed and that additional reliability testing was not conducted. Does the Standing Committee agree that the measure is still reliable and that there is no need for repeated discussion and a vote on reliability?*

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

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

---

For maintenance measures – less emphasis if no new testing data are provided

2b1. Measure Intent: The measure specifications are consistent with the measure's intent and capture the most inclusive target population.

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.

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

Validity Testing

- Did the developer conduct new validity testing? ☐ X Yes ☐ No
- Validity testing conducted at the Patient/Encounter Level:
    - Exploratory and confirmatory factor analyses (CFA) of newly tested items and items in multiitem measures were conducted using weighted least squares means and variance adjusted (WLSMV) estimation. The assessed overall model fit for the six-factor model using the comparative fit index was 0.997, the root mean square error of approximation was 0.014, and the weighted root mean square residual was 1.068 in CFA. The factor loadings were above 0.70. The overall fit chi-square was 252.83, p<0.001, for a model with 120 degrees of freedom.
    - Construct validity was assessed using Pearson correlations between six multi-item and one single-item measure top-box scores and with two single-item global measures top-box scores. The Pearson correlations ranged from 0.40–0.61 across the measures.
    - Discriminant validity was assessed using Pearson correlations among multi-item measures to evaluate the extent to which they measure different constructs. The Pearson correlations ranged from 0.33–0.64.
- Validity testing conducted at the Accountable-Entity Level:
    - The developer noted that it used individual-level data for updated testing "as estimates of hospice-level associations would be unbiased but imprecise if calculated among the 56 hospices participating in the 2021 mode experiment."
    - Prior testing (the 2019 submission) included both individual- and hospice-level results. The hospice-level Pearson correlations between measures and global rating items ranged from 0.63–0.84. The hospice-level Pearson correlations among multi-item measures ranged from 0.42–0.84.

Exclusions

- Decedents or caregivers who otherwise meet the inclusion criteria are excluded if they have a "no publicity" status. No statistical testing was conducted given the nature of this exclusion.

Risk Adjustment

- The model was developed during the prior maintenance review. It was not retested or updated for this submission.
- There is a risk model with nine risk factors (i.e., response percentile, decedent age, payer, primary diagnosis, length of final episode of hospice care, respondent age, respondent education, relationship of decedent to caregiver, and language).
    - Case-mix coefficients from a linear regression are used to generate case-mix adjustments for each survey question.
    - Publicly reported hospice survey measure scores are adjusted to the overall national mean of case-mix variables across all reporting hospices.
- Published literature and data analyses were used to develop the conceptual model and select the risk adjustment approach.

  - The developer identified characteristics as candidates for adjustment if they were present in the response data and not in the hospice's control. For each adjuster, they examined variation among hospices using ICC, bivariate and multivariate association with selected survey outcomes, and the impact on adjustment and parameterization of adjustors.
  - The following social risk factors were considered: decedent education, primary payer for hospice care, caregiver respondent education, and caregiver respondent language. Primary payer and language are included in the risk adjustment model. The two education variables were associated with the outcomes but moderately correlated with each other;therefore, caregiver education was retainedwhile decedent education was not.
- No discrimination or calibration statistics were provided.

Meaningful Differences

- The developer calculated number and percentages of hospices significantly above or below the mode experiment hospice average for each measure; scores were adjusted for mode and case mix. Between 13 and 26 percent of hospices were statistically different (above or below) the mode experiment hospice average.

Missing Data

- Survey response rates ranged from 31–45 percent across modes. Item-level missing data due to inappropriate skips ranged from 0.5–5.0 percent.
- The developer cites prior research that indicates that nonresponse weighting to account for potential bias is not needed after case-mix adjustment.

Comparability

- Linear regression was used to evaluate the effects of different survey modes on survey outcomes. The model included case-mix adjusters, hospice indicators, and the month of death.
- There were significant effects of survey mode on several survey outcomes. Consequently, the survey scores should be adjusted for the mode of administration.

SMP Summary:

- One member noted that at the hospice level the submission refers to above or below the average of experiment's participating hospices and it is whether the developers mean top box scores or actual mean scores. The same member noted concerns with case-mix adjustment particularly around language and mode of administration, sampling due to possible bias introduced from the poor response rate, and consistency for vendors with multiple hospices. 11
- Regarding patient/encounter level reliability, one member stated that the submission notes differences in rating within each domain based on characteristics of decedent, respondent, and mode of survey administration. The member continued that although risk adjustment attempts to correct for these, the methodology ignores the possibility that there are systematic differences in performance associated with decedent characteristics that should not be adjusted away.
- Regarding accountable-entity level reliability, one member noted that per their comment regarding the risk adjustment methodology, the differences in performance across different patients has not been explored. A second member noted that the hospice level reliability (0.03) was poor.
- Regarding the patient/encounter level validity testing, one SMP member noted that within survey validity measures create over endogeneity of results and therefore it is common to use more general/global measures of the quality construct to validate construct-specific measures. Additionally,

it was noted that the scale quality is high enough but there is evidence that hospices may differ along different dimensions of care and because of this a factor analysis could have been useful to present.

- Regarding the accountable-entity level testing and the risk adjustment model, some reviewers noted that the testing was not updated even though the developers updated the survey. Another reviewer noted that the developer did not provide entity level results, rather they relied on respondent-level correlations between measures.
- Regarding missing data, some SMP members noted that there is concern with large nonresponse, however the developer addresses these concerns in the risk adjustment model. Further, an SMP member noted that having characteristics of non-responders would have been helpful to assess whether certain caregivers known to have poorer or better ratings of hospices were more frequently non-responsive.
- The developer provided responses to these concerns, which are available in Appendix B, the SMP Discussion Guide.

Questions for the Standing Committee regarding validity:

- *Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk adjustment approach, etc.)?*
- *The SMP is satisfied with the validity analyses for the measure. Does the Standing Committee think there is a need to discuss and/or vote on validity?*
- *The developer attests that additional validity testing was not conducted. Does the Standing Committee agree that the measure is still valid and that there is no need for repeated discussion and a vote on validity?*

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

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.

- Data for these PRO-PMs are collected via a survey that can be administered via mail, telephone, or web.
- The Hospice CAHPS survey is a survey that is fielded to the caregiver following the death of the hospice patient, and thus the data elements are not routinely generated and used during care delivery.
- The developer indicates that no fees or licensing costs are associated with use of the measures.
- The response rate for the survey is fairly low, although it is comparable to other CAHPS surveys.
- Like many other CAHPS surveys, the CAHPS Hospice Survey relies upon providers to contract with CMS-approved survey vendors. The estimated vendor cost per provider is approximately $4,000 annually.

Questions for the Standing Committee:

- *Is there any evidence that the past few years of implementation have led to unexpectedly burdensome or onerous administration requirements?*

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

Criterion 4: Use and Usability

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 evaluates the extent to which audiences (e.g., consumers, purchasers, providers, and 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 they are not in use at the time of initial endorsement, then a credible plan for implementation within the specified time frames is provided.

Current uses of the measure

Publicly reported? ☒ Yes ☐ No

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

Planned use in an accountability program? ☐ Yes ☐ No ☒ NA

Accountability program details

- Public Reporting: Hospice Quality Reporting Program (HQRP) https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Hospice-Quality-Reporting/index.html

4a.2. Feedback on the measure provided 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; and (3) This feedback has been considered when changes are incorporated into the measure.

Feedback on the measure provided by those being measured or others

- According to the developer, In the 2023 Hospice Payment Rate Update proposed rule, CMS requested comments on potential updates to CAHPS Hospice Survey content and modes of administration. Hospices expressed support for a shorter survey instrument and web-based mode of survey administration.
- During the prior NQF review in 2019, the American Geriatrics Society submitted a public comment suggesting the addition of new survey content related to person-centered care and to assess whether the hospice care team discussed what mattered most to the patient and family.
- To address feedback received, CMS shortened the survey instrument and tested a web-mail mode of survey administration, revised the CAHPS Hospice Survey to be eight questions shorter, added a new Care Preferences measure, and added a web-mail mode of administration.

***Questions for the Standing Committee*:**

- *How have (or can) the performance results be used to further the goal of high quality, efficient healthcare?*
- *How has the measure been vetted in real-world settings by those being measured or others?*

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

RATIONALE: [Rationale for voting low or insufficient]

4b. Usability (4b1. Improvement; 4b2. Benefits of measure)

4b. Usability evaluates the extent to which audiences (e.g., consumers, purchasers, providers, and 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

- Performance data submitted by the developer includes evidence of small improvements in CAHPS Hospice Survey measures, with greatest improvement during the short period corresponding to care provided after public reporting began in February 2018.
- The developer reported that from 2019 to 2021, scores declined slightly for all but one measure, reflecting the responsiveness of the measures to changes in care delivery during the COVID-19 pandemic.

4b2. Benefits versus harms. 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).

Unexpected findings (positive or negative) during implementation

- The developer did not report any unexpected findings.

Potential harms

- The developer did not report any potential harms.

Additional Feedback:

- [Summary of feedback from the Measure Applications Partnership [MAP]]

Questions for the Standing Committee:

- *How can the performance results be used to further the goal of high quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

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

Criterion 5: Related and Competing Measures

Related Measures

- 1623: Bereaved Family Survey
    - The result of the Bereaved Family Survey measure (#1623) is a single score that indicates the family's perceptions of the quality of care that veterans received from the VA during the last month of life; aspects of care included in the measure are communication, emotional and spiritual support, pain management, and personal care needs.

Harmonization

- The measures are harmonized to the extent possible.

Developer Submission

Criteria 1: Importance to Measure and Report

1a. Evidence

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

---

1ma.01. Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.

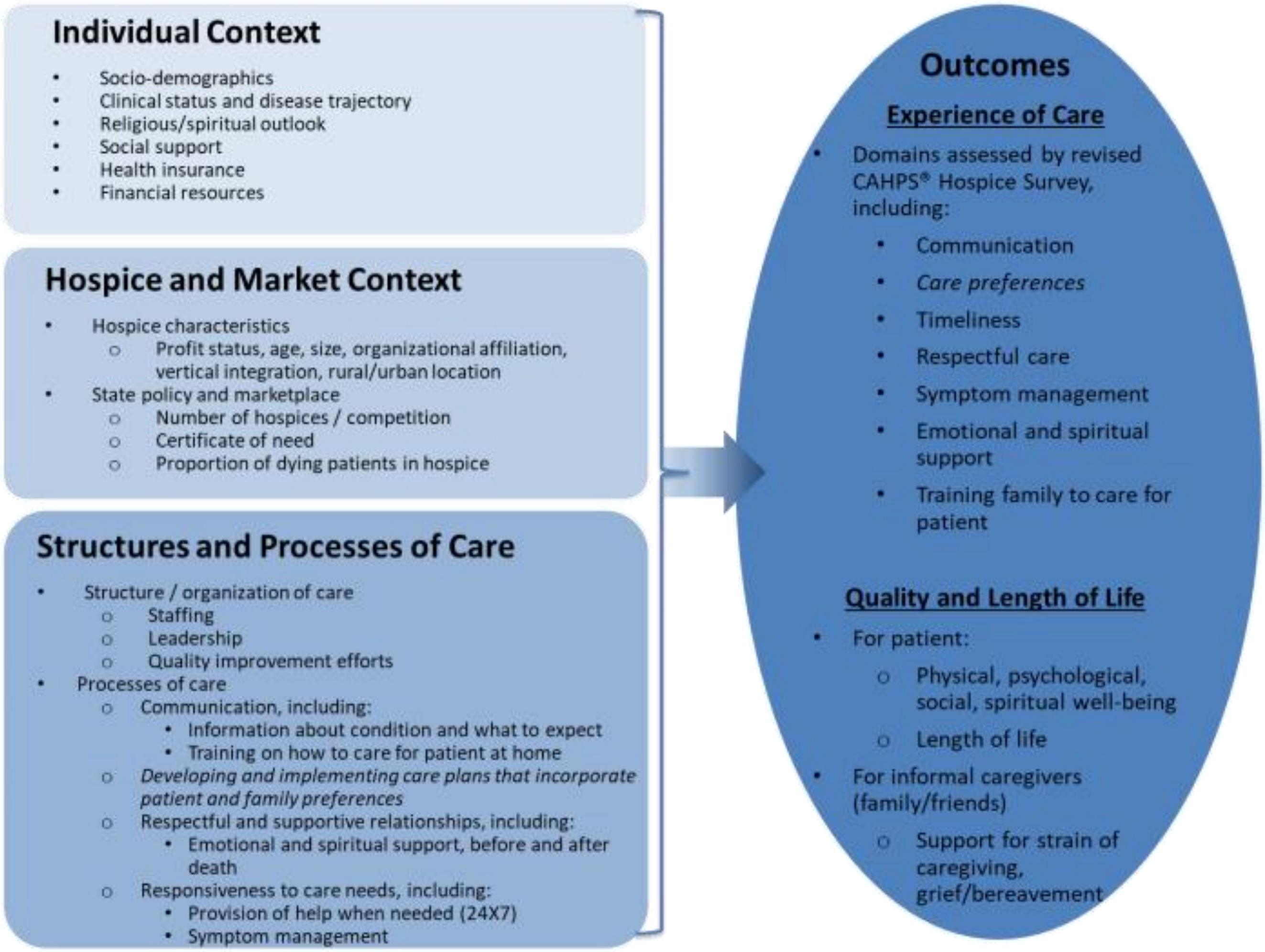
[Response Begins]
No
[Response Ends]

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:
Current Submission:
Updated evidence information here.
Previous (Year) Submission:
Evidence from the previous submission here.

1a.01. Provide a logic model.

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.

[Response Begins]
Current Submission:
The conceptual model described in the previous (2019) submission is still applicable to the revised CAHPS Hospice Survey measures under review. We have made minor updates to the model (shown in *italics*) to reflect the important role that hospice has in explaining care options, formulating goals of care that reflect patient and family preferences, and then creating a plan of care that aims to achieve those goals. This key process of hospice care is assessed by the new CAHPS Hospice Survey Care Preferences measure.
Figure 1. CAHPS Hospice Survey Logic Model, Updated.



Previous (2019) Submission:

The CAHPS® Hospice Survey is grounded in existing guidelines and conceptual models of the quality of hospice care, including NHPCO standards of practice for hospice programs (https://www.nhpco.org/hospice-care-overview/hospice-standards-of-practice/), review of existing guidelines (Teno et al. 2001), a conceptual model of end-of-life care developed by Stewart and colleagues (1999), the National Quality Forum Preferred Practices of Palliative and Hospice Care (2006), the advice of a Technical Expert Panel, and input from qualitative inquiry with caregivers of hospice decedents. Since the development of the survey, the 4th edition of the National Consensus Project Clinical Practice Guidelines for Quality Palliative Care, and a supporting systematic review of 139 studies that synthesized evidence on eight domains of palliative care (Ahluwalia, Chen et al. 2018), have been released. The updated guidelines and supporting evidence continue to emphasize the importance of patient- and family-centeredness of care.

We have updated a conceptual model first presented by Stewart and colleagues in 1999 to reflect the latest guidelines and supporting evidence (Stewart, Teno et al. 1999; Figure 1 below and in the Supplemental Appendix PowerPoint file entitled, "CAHPS Hospice Survey Conceptual Model 2019_10"). For the CAHPS® Hospice survey, the focus is on informal primary caregivers' experiences of hospice care, focusing on the following overarching key processes:

1. **Communication with** the dying person and his or her family. The hospice plays an important role in sensitive **communication** about the patient's condition, educating the family on what to expect, and **training** the family on how to care for the patient. In particular, for patients receiving care within the home, and as patients getting closer to death, the family plays an important role in overseeing care through administration of medications and assistance in the activities of daily living.
2. **Respectful and supportive relationships** with dying patients and their families. At this sentinel period of life, treating the patient and family with **respect**, respecting cultural traditions, and provision of **emotional and spiritual support,** as well as grief services, prior to and after the death of the patient are critically important.
3. **Responsiveness to care needs**. Patient and family questions and concerns arise throughout the day and night, and on every day of the week. Therefore, hospices must respond in a **timely** manner to questions, and provide needed help, even on nights and weekends. Managing pain and other **symptoms**, to the degree desired by patients and families, is a core process of the hospice team. As noted by the systematic review, there is evidence

for the effectiveness of pharmacologic interventions for pain, dyspnea, and constipation, among other symptoms (Ahluwalia, Chen et al. 2018).

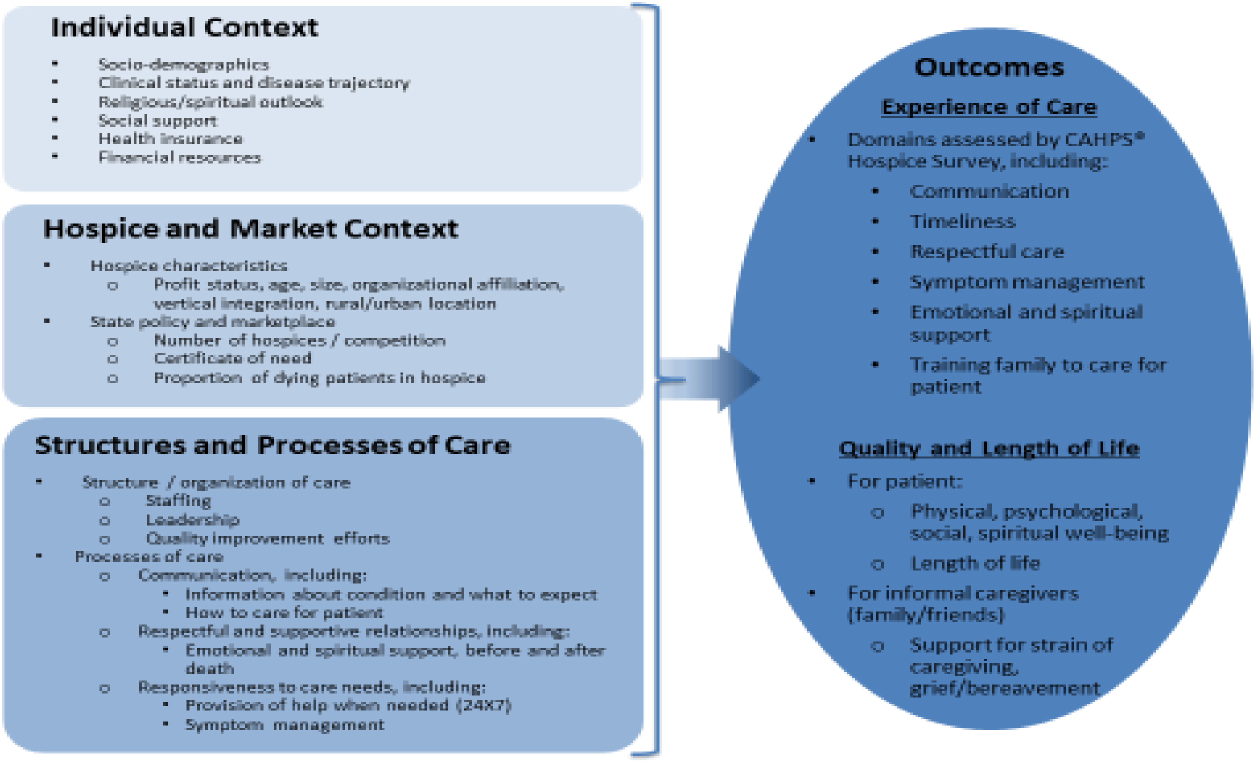
The CAHPS® Hospice Survey asks respondents to assess care experiences that result from these processes in six key domains (addressed by the six multi-item measures):

1. **Hospice Team Communication**: degree to which hospice keeps the hospice primary caregiver informed, listens to their concerns, and explains things in a way they can understand.
2. **Getting Timely Care**: degree to which hospice provides timely help when asked, even on weekends and holidays (i.e., timely assistance 24X7).
3. **Treating Family Member with Respect:** degree to which patient is treated with dignity and respect.
4. **Getting Emotional and Spiritual Support:** degree to which hospice provides the desired amount of emotional and spiritual support prior to and after the patient's death.
5. **Getting Help for Symptoms**: degree to which hospice provides the desired help for management of pain and other symptoms.
6. **Getting Hospice Care Training:** degree to which hospice attends to the needs of the hospice primary caregiver for information and training to safely care for the patient at home.

Citations:

1. Ahluwalia SC, Chen C, Raaen L, Motala A, Walling AM, Chamberlin M O'Hanlon C, Larkin J, Lorenz K, Akinniranye O, Hempel S. (2018). A Systematic Review in Support of the National Consensus Project Clinical Practice Guidelines for Quality Palliative Care, Fourth Edition. *J Pain Symptom Manage*. 56(6): 831-870.
2. National Consensus Project for Quality Palliative Care. (2018). *Clinical Practice Guidelines for Quality Palliative Care, 4th edition*. Richmond, VA: National Coalition for Hospice and Palliative Care; 2018. Available at: https://www.nationalcoalitionhpc.org/ncp/.
3. National Quality Forum. (2006). National Framework and Preferred Practices for Palliative and Hospice Care. Available at: https://www.qualityforum.org/Publications/2006/12/A_National_Framework_and_Preferred_Practices_for_Palliative_and_Hospice_Care_Quality.aspx.
4. Teno JM, Casey VA, Welch L, Edgman-Levitan S. (2001). Patient-Focused, Family-Centered End-of-Life Medical Care: Views of the Guidelines and Bereaved Family Members. *J Pain Symptom Manage-Special Section on Measuring Quality of Care at Life's End II.* 22:738-51.
5. Stewart AL, Teno J, Patrick DL, Lynn J. (1999). The concept of quality of life of dying persons in the context of health care. *J Pain Symptom Manage*. 17:93-108.

Figure 1. CAHPS Hospice Survey Logic Model



[Response Ends]

1a.02. Provide evidence that the target population values the measured outcome, process, or structure and finds it meaningful.

Describe how and from whom input was obtained.

[Response Begins]
Current Submission
In its review of the 2019 submission of CAHPS Hospice Survey measures, NQF's Geriatric and Palliative Care Committee indicated the usefulness of the measures for informing quality improvement and hospice selection. Specifically, "Committee members shared that their organizations carefully reviewed results on this measure and actively worked to improve care to improve measure performance. Committee members also reported using the results on Hospice Compare to find hospice care both for patients and family members. They stated that they found the results of this measure personally useful" (NQF 2020). Revisions to CAHPS Hospice Survey measures were made in response to feedback received from a technical expert panel (CMS 2022), as well as public comments during the 2019-2020 NQF maintenance of endorsement process (NQF 2020). All proposed changes were tested with caregivers of hospice decedents via cognitive interviews. Given that the **revised CAHPS Hospice Survey measures are similar in content to the existing measures, and address content recommended by stakeholders and tested with hospice caregivers, CMS anticipates that the revised measures will continue to be as or more meaningful and useful to hospices.**

Citation:

1. Centers for Medicare & Medicaid Services (CMS). CAHPS Hospice Survey Technical Expert Panel July 2, 2020 1 to 3 pm ET Meeting Summary. Available at: https://hospicecahpssurvey.org/globalassets/hospice-cahps4/whats-new/cahps-hospice-survey-technical-expert-panel-2020_meeting-summary-final.pdf.
2. NQF. (November 17, 2020). Geriatrics and Palliative Care 2019 Cycle Track 2: CDP Report. Available at: https://www.qualityforum.org/Publications/2021/03/Geriatrics_and_Palliative_Care_Final_Report_-_Fall_2019_Cycle.aspx.

Previous (2019) Submission
CAHPS® Hospice Survey measure scores are directly related to caregiver respondents' willingness to recommend the hospice. An analysis of 2015 CAHPS® Hospice Survey data from 2,500 hospices found that compared to caregiver respondents who responded "Always" to all six questions in the Hospice Team Communication composite, respondents who responded "Always" to only three of the six items within the composite had a 39.3% lower chance of definitely recommending the hospice (0.785/2 = 0.393; Anhang Price et al. 2018). These findings in turn suggest that a hospice that scores 10 points lower on Hospice Team Communication can expect definite recommendations 7.9% less often (0.785/10 = 0.079). Differences of 5, 3, and 1 points on a 0–100 scale are sometimes referred to as large, medium, and small for CAHPS® measures; in this analysis, those thresholds for the Hospice Team Communication measure correspond to 4.0%, 2.4%, and 0.8% reductions in the probability of recommending the hospice.
In qualitative interviews with CMS, hospice leaders have indicated that they find CAHPS® Hospice Survey results to be useful in informing quality improvement initiatives.

Citations:

1. Anhang Price R, Stucky B, Parast L, Elliott MN, Haas A, Bradley M, Teno JM. (2018). Development of Valid and Reliable Measures of Patient and Family Experiences of Hospice Care for Public Reporting. *J Pall Med.* 21(7):924-932.
2. Quigley DD, Elliott MN, Setodji CM, Hays RD. (2018). Quantifying Magnitude of Group-Level Differences in Patient Experiences with Health Care. *Hlth Ser Res*. 53(4):3027-3051.

[Response Ends]

1a.03. Provide empirical data demonstrating the relationship between the outcome (or PRO) and at least one healthcare structure, process, intervention, or service.

[Response Begins]
Current Submission

Data from the 2021 CAHPS Hospice Survey mode experiment do not allow for assessment of the relationship between survey measures and structures or processes. However, given that the revised CAHPS Hospice Survey measures are similar in content to the existing measures, CMS anticipates that the revised measures will exhibit similar relationships to those described for the existing measures in the 2019 submission.

Previous (2019) Submission

An analysis of 2016 CAHPS® Hospice Survey data from 2,236 hospices found that hospices' CAHPS® Hospice Survey measure scores were positively associated with provision of professional staff visits to more patients receiving routine hospice care the last two days of life (Teno, Anhang Price et al. 2019). Specifically, for all CAHPS® Hospice Survey measures with the exception of caregiver training, there were significant positive associations between the proportion of patients receiving staff visits and hospices' CAHPS® Hospice Survey measure performance starting at the sixth decile of visits (84.6% and higher). Family caregivers in hospices in the highest decile of professional staff visits (95.1% and higher) rated the hospice 5.2 points higher on timeliness of care, and between 2.3 and 4.5 points higher on other measures, than caregivers in hospices in the lowest decile (67.5% and lower). These findings suggest that the measures are sensitive to differences in best practice clinical processes.

In addition, the following structural characteristics of hospices were shown to be associated with hospices' being in the top quartile of all hospices for which CAHPS® Hospice Survey scores were publicly reported for 2015 and 2016: being in an independent (i.e., non-chain) and non-profit or government hospice and providing care to fewer than 200 patients per year.

Citations:

1. Anhang Price R, Tolpadi A, Schlang D, Bradley MA, Parast L, Teno JM, Elliott MN. (2019) Characteristics of Hospices Providing the Highest Quality Care. Presentation to the AcademyHealth Annual Research Meeting. *(Subsequently published in the Journal of Palliative Medicine; DOI: 10.1089/jpm.2019.0505).*
2. Teno JM, Anhang Price R, Parast L, Haas A, Elliott MN. (2019). More Professional Visits in the Last Days of Life are Associated with Better Hospice Care Experiences. Presentation to the American Academy of Hospice and Palliative Medicine Annual Assembly.

[Response Ends]

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

1b.01. Briefly explain the rationale for this measure.

Explain how the measure will improve the quality of care, and list the benefits or improvements in quality envisioned by use of this measure.

[Response Begins]

A growing number of Medicare decedents use hospice care: in 2019, more than 1.6 million Medicare beneficiaries – more than half of decedents – received hospice services, compared to fewer than a quarter of decedents in 2000. Total Medicare hospice expenditures were nearly $21 billion in that year (Medicare Payment Advisory Commission 2019). The number of hospice providers has also increased dramatically over time, more than doubling from 2,255 in 2000 to 4,840 in 2019.

There is striking variation across hospices with regard to processes of care associated with care quality, including the proportion of patients discharged alive (Prisc, Plotzke et al. 2016), provision of general inpatient care or continuous home care in the last seven days of life (Plotzke, Christian et al., 2014), and visits from professional hospice staff in the last two days of life (Teno, Plotzke et al. 2016). Evidence of variation in hospice care processes points to the need for performance measures that can be systematically implemented to monitor hospice performance and make fair comparisons across hospices.

Patient- and family-centeredness of care is a central goal of hospice care and can be directly measured through surveys of family caregivers of hospice patients (Anhang Price and Elliott 2018). CMS created the CAHPS® Hospice Survey, a component of the Hospice Quality Reporting Program (HQRP), to ensure that an assessment of the patient- and family-centeredness of care would be included to monitor hospice performance, promote quality improvement, and inform consumer decision making in the selection of a hospice via public reporting of results. To date, scores on the CAHPS Hospice Survey measures have shown more variation and potential for improvement than other measures within the HQRP.

The CAHPS Hospice Survey is the first survey to use standardized content and implementation and analytic protocols to allow for fair comparisons between hospices. CMS continuously reviews these protocols to ensure that the survey is as feasible and fair as possible.

Citations:

1. Anhang Price R, Elliott MN. (2018). Measuring Patient-Centeredness of Care for Seriously Ill Individuals: Challenges and Opportunities for Accountability Initiatives. J Pall Med. 21(S2):S28-S35.
2. Medicare Payment Advisory Commission. (March 2021). Chapter 11: Hospice Services. In: Report to the Congress: Medicare Payment Policy.
3. Prsic E, Plotzke M, Christian TJ, Gozalo P, Teno JM. (2016). A National Study of Live Hospice Discharges Between 2000 and 2012. J Pall Med. 19(9):987-90.
4. Teno JM, Plotzke M, Christian T, Gozalo P. (2016). Examining Variation in Hospice Visits by Professional Staff in the Last 2 Days of Life. JAMA Intern Med. 176(3):364-370.

[Response Ends]

1b.02. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.

Include mean, std dev, min, max, interquartile range, and 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 (4b) under Usability and Use.

[Response Begins]

Current Submission

Table 1b.02a displays the mean, standard deviation, 25th, 50th and 75th percentile scores for each of the multi-item and single-item CAHPS Hospice Survey measures. Scores are for the 54 hospices participating in the 2021 mode experiment that had 30 or more completed surveys.

Measures with the highest top-box scores were Getting Emotional and Religious Support (top-box mean=92.4; IQR: 4.4), Treating Family Member with Respect (top-box mean=91.3; IQR: 3.3), and Care Preferences (top-box mean=90.2; IQR =5.7). The measures with the lowest top-box scores were Getting Timely Care (top-box mean=77.5; IQR: 9.2) and Getting Help for Symptoms (top-box mean=78.8; IQR: 6.4).

Table 1b.02a. Distribution of Top-Box Scores for CAHPS Hospice Survey Measures, 2021 Mode Experiment

*	Mean	Standard Deviation	25th Percentile	50th Percentile	75th Percentile
Multi-item Measures	*	*	*	*	*
Hospice Team Communication	81.5	5.6	79.8	82.5	84.9
Getting Timely Care	77.5	7.7	73.9	78.3	83.1
Treating Family Member with Respect	91.3	3.7	90.3	92.0	93.6

*	Mean	Standard Deviation	25th Percentile	50th Percentile	75th Percentile
Getting Emotional and Religious Support	92.4	3.5	90.2	92.7	94.6
Getting Help for Symptoms	78.8	5.6	76.2	79.9	82.6
Care Preferences	90.2	4.4	87.5	91.4	93.2
Single-Item Measures	*	*	*	*	*
Getting Hospice Care Training	82.4	7.6	77.9	84.2	86.8
Rating of Hospice	84.8	7.2	81.7	86.3	89.6
Willingness to Recommend	86.0	6.9	82.3	87.9	90.8

* This cell intentionally left empty.

Additional detail regarding current performance of all items included in the multi-item measures, including minimum, maximum, and decile scores, is attached in the Supplementary Appendix "Section 1b.03 CAHPS Hospice Full Score Distributions 2022_11."

Previous (2019) Submission

Table 1b.2a displays the mean, standard deviation, 25th, 50th and 75th percentile scores for each of the multi-item and global CAHPS Hospice Survey measures. Scores are for the 2,933 hospices that had at least 30 respondents over the eight quarters of Quarter 3 2016 through Quarter 2 2018 (data collected from October 2016 through October 2018 regarding care experiences of patients who died while receiving hospice care from July 2016 through June 2018).

Measures with the highest top-box scores were Treating Family Member with Respect (top-box mean=90.5; interquartile range=88.4,93.2) and Getting Emotional and Religious Support (top-box mean=89.5; interquartile range=87.9, 92.0). Measures with the lowest top-box scores were Getting Help for Symptoms (top-box mean=75.1; interquartile range=71.3,79.1) and Getting Hospice Care Training (top-box mean=75.3; interquartile range=71.3, 80.0).

Table 1b.2a. Distribution of Top-Box Scores for CAHPS Hospice Survey Measures, Quarter 3 2016 to Quarter 2 2018

*	Mean	Standard Deviation	25th Percentile	50th Percentile	75th Percentile
Multi-item Measures	*	*	*	*	*
Hospice Team Communication	80.3	5.3	77.3	80.6	84.0
Getting Timely Care	77.9	7.0	73.6	78.3	82.8
Treating Family Member with Respect	90.5	3.9	88.4	91.0	93.2
Getting Emotional and Religious Support	89.5	3.8	87.9	90.2	92.0
Getting Help for Symptoms	75.1	6.0	71.3	75.2	79.1
Getting Hospice Care Training	75.3	7.0	71.3	75.5	80.0
Global Measures	*	*	*	*	*
Rating of Hospice	80.5	6.6	76.8	81.2	85.0

*	Mean	Standard Deviation	25th Percentile	50th Percentile	75th Percentile
Willingness to Recommend	84.4	7.0	80.4	85.4	89.5

* This cell intentionally left empty.

Additional detail regarding current performance of all items included in the multi-item measures, including minimum, maximum, and decile scores, is attached in the Supplementary Appendix A, "CAHPS Hospice Full Score Distributions of Composites and Items Tables 2019_09_05.xlsx," worksheet entitled "Table 1b.2b."

The PowerPoint file entitled, "Figure 1b.2. Performance scores over time.pptx" in the Supplementary Appendix A, shows trends in performance over time for each of the eight reported measures, for quarters for which data are available from calendar years 2015 through 2018, adjusted for seasonality. These plots are restricted to a common set of hospices: the 1,336 hospices that had scores reported in each quarterly update of Hospice Compare through the most recent time period included in our dataset, and that had at least 30 respondents overall and 11 respondents for each measure within the quarters of available data for the calendar year. Improvements are small but positive for all eight measures. Hospice Team Communication and Rating of Hospice showed the largest improvements, of about 0.6 point (on a 0-100 scale). Notably, for many measures most improvement is concentrated at the end of the time period. Public reporting of CAHPS® Hospice Survey measures began in February 2018, in the last calendar year of the time period analyzed. Evidence from other parallel quality measure initiatives suggests that steeper improvements may be expected following the introduction of public reporting (Elliott, Cohea et al. 2015), as health care providers pursue initiatives to improve their publicly-reported quality performance.

We also calculated trends in performance over time for evaluative items included in the multi-item measures. Improvements were small but positive for all except one of the 22 items tested. The largest improvements (in the range of 0.7 to 0.8 point) were for two items in the Hospice Team Communication composite, "While your family member was in hospice care, how often did the hospice team explain things in a way that was easy to understand?" and "While your family member was in hospice care, how often did the hospice team keep you informed about your family member's condition?" and one item in the Getting Emotional and Religious Support composite, "In the weeks after your family member died, how much emotional support did you get from the hospice team?" For the two items in the Hospice Team Communication measure, about 70% of the overall improvement occurred in 2018, while for the item in the Getting Emotional and Religious Support measure, about 50% of the overall improvement happened in this last year, just as public reporting initiated.

Citation:

Elliott MN, Cohea CW, Lehrman WG, Goldstein EH, Cleary PD, Giordano LA, Beckett MK, Zaslavsky AM. (2015). Accelerating Improvement and Narrowing Gaps: Trends in Patients' Experiences with Hospital Care Reflected in HCAHPS Public Reporting. Health Serv Res (50)6: 1850-1867.

[Response Ends]

1b.03. If no or limited performance data on the measure as specified is reported above, 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. Include citations.

[Response Begins]

The revised CAHPS Hospice Survey measures tested in the 2021 mode experiment have not yet been nationally implemented; therefore, there are no data available to calculate trends in performance over time for the revised measures. Given that the revised CAHPS Hospice Survey measures are similar in content to the existing measures, CMS anticipates that the revised measures will exhibit similar changes over time. To that end, in this section, we describe changes over time in the existing eight CAHPS Hospice Survey measures.

"Section 1b.03. Performance scores over time" in Supplementary Appendix A shows trends in performance over time for each of the eight currently-reported measures, for quarters for which data are available from calendar years 2015 through 2021, adjusted for seasonality. These plots are restricted to a common set of hospices: the 1,418 hospices that had scores publicly reported in each update of Hospice or Care Compare through the most recent time period included in our dataset (Quarter 4, 2021), and that had at least 30 respondents overall and 11 respondents for each measure within the quarters of available data for the calendar year. From 2015 through 2019 (i.e., prior to the COVID-19 pandemic), improvements were small but positive for all eight measures. Hospice Team Communication and Rating of Hospice showed the largest improvements, of about 0.84 and 1.13 points on a 0-100 scale, respectively. From 2019 to 2021, scores declined slightly for all but the Hospice Team Communication measure. The largest declines were in Getting Timely Help and Getting Help for Symptoms (0.84 and 0.64 points on a 0-100 scale, respectively). Changes in scores from 2019 to 2021 reflect the responsiveness of the measures to changes in care delivery during the COVID-19 pandemic.

[Response Ends]

1b.04. 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.

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. 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 (4b) under Usability and Use.

[Response Begins]

Current Submission

Analysis of 2021 CAHPS Hospice Survey mode experiment data, including 4,749 caregiver respondents from 56 hospices, finds that caregivers of male decedents report poorer care experiences than caregivers of female decedents by 1.5 to 3.7 percentage points for all measures (with differences reaching statistical significance for four measures); caregivers of patients age 65 to 74 report better care experiences than caregivers of patients age 90 or older by 0.2 to 4.4 percentage points for all measures (with differences reaching statistical significance for three measures); caregivers of decedents with both Medicare and Medicaid insurance report poorer care experiences than caregivers of those with only Medicare insurance by 4.0 to 13.9 percentage points for eight of nine measures (with differences reaching statistical significance for seven of these measures); and finally, caregivers of decedents who are Black report better experiences than caregivers of White decedents on Hospice Team Communication (4.1 percentage points higher; $p<0.05$) but similar or worse experiences for all other measures, and caregivers of decedents who are Hispanic report worse experiences than caregivers of White decedents with regard to Emotional and Religious Support (3.4 percentage points lower; $p<0.05$), but similar care experiences for other measures.

These results are broadly consistent with patterns observed in data from national implementation of the current CAHPS Hospice Survey, as shown in an Excel file in the Supplemental Appendix entitled, "Section 1b.04. CAHPS Hospice Survey measure scores by subgroup.xls." This file presents scores for the current CAHPS Hospice Survey measures for 2015 through 2020, by decedent age, gender, race, ethnicity, and education, as well as by caregiver respondent language. Given that the revised CAHPS Hospice Survey measures are similar in content to the existing measures, CMS anticipates that the revised measures will exhibit similar patterns in measure scores by subgroup once nationally implemented.

Overall differences between subgroups reflect both differential care experiences within the same hospices and differential concentration of patients between hospices. For example, an analysis of national implementation data for current CAHPS Hospice Survey measures from 464,064 caregivers of Black, Hispanic, and White hospice patients from July 2020—December 2021 found that overall, reported care experiences for Black and Hispanic decedents were similar or better than for White decedents within the same hospices, with the exception of the Emotional and Spiritual Support and

Willing to Recommend the Hospice measures, but that Black and Hispanic decedents more often received care from hospices that provide worse care experiences (Anhang Price, Haas et al., 2022). These findings are consistent with those previously published based on 2015-2016 CAHPS Hospice Survey data (Anhang Price, Parast et al. 2017). Differences in comparison to White decedents have been relatively stable for Black decedents since 2015-2016; however, reported care experiences for Hispanic decedents relative to White decedents worsened slightly since 2015-2016, particularly in 2020-2021, with changes driven largely by worse care within the same hospices rather than concentration of Hispanic decedents in hospices with poorer reported care experiences (Anhang Price, Haas et al., 2022).

Citations:

1. Anhang Price R, Haas A, Teno JM, Elliott MN. (2022). Variation in Caregiver-Reported Hospice Care Experiences by Race and Ethnicity. Presentation at the CAHPS Virtual Research Meeting: Assessing Patient Experience for Insights into Enhancing Equity in Healthcare.
2. Anhang Price R, Parast L, Haas A, Teno JM, Elliott MN. (2017). Black and Hispanic Patients Receive Hospice Care Similar to That of White Patients When in the Same Hospices. Health Affairs. 36(7): 1283-1290.

Previous (2019) Submission

An analysis of 2015-2016 CAHPS Hospice Survey data from nearly 300,000 respondents from 2,497 hospices found that on average, black and Hispanic patients received care from hospices with lower CAHPS® Hospice Survey measure scores (Anhang Price, Parast et a. 2017). Within a given hospice, however, caregivers of black and Hispanic patients reported significantly better hospice care experiences than caregivers of white patients on five of seven CAHPS® Hospice Survey measures studied. Caregivers of black and Hispanic patients reported receiving their desired level of emotional and religious support less often than caregivers of white patients did, and were less likely to recommend the hospice.

An Excel file in the Supplemental Appendix A entitled, "CAHPS Hospice Survey measure scores by subgroup.xls," presents CAHPS® Hospice Survey measure scores for 2015, 2016, and 2017 by decedent age, gender, race, ethnicity, and education, as well as by caregiver respondent language.

Citation:

1. Anhang Price R, Parast L, Haas A, Teno JM, Elliott MN. (2017). Black and Hispanic Patients Receive Hospice Care Similar to That of White Patients When in the Same Hospices. Health Affairs. 36(7): 1283-1290.

[Response Ends]

1b.05. If no or limited data on disparities from the measure as specified is reported above, 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 above.

[Response Begins]

N/A

[Response Ends]

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

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.

---

spma.01. Indicate whether there are changes to the specifications since the last updates/submission. If yes, update the specifications in the Measure Specifications section of the Measure Submission Form, and explain your reasoning for the changes below.

[Response Begins]

Yes

[Yes Please Explain]

Updates were made in response to stakeholder requests for the survey instrument to be shortened and simplified. New content related to Care Preferences was added in response to feedback received from a technical expert panel, as well as public comments during the 2019-2020 NQF maintenance of endorsement process. All proposed changes were tested with caregivers of hospice decedents via cognitive interviews. CMS tested the revised survey instrument among caregivers from 56 hospices during a 2021 mode experiment.

[Response Ends]

spma.02. Briefly describe any important changes to the measure specifications since the last measure update and provide a rationale.

For annual updates, please explain how the change in specifications affects the measure results. If a material change in specification is identified, data from re-testing of the measure with the new specifications is required for early maintenance review.

For example, specifications may have been updated based on suggestions from a previous NQF CDP review.

[Response Begins]

Since the last submission, proposed CAHPS Hospice Survey measures have undergone the following changes:

1. Removal of one survey item regarding confusing or contradictory information from the Hospice Team Communication measure
2. Replacement of the multi-item Getting Hospice Care Training measure with a new, one-item summary measure
3. Addition of a new, two-item Care Preferences measure
4. Simplified wording to component items in the Hospice Team Communication, Getting Timely Care, and Treating Family Member with Respect measures

Updates were made in response to stakeholder requests for the survey instrument to be shortened and simplified. New content related to Care Preferences was added in response to feedback received from a technical expert panel, as well as public comments during the 2019-2020 NQF maintenance of endorsement process. All proposed changes were tested with caregivers of hospice decedents via cognitive interviews. CMS tested the revised survey instrument among caregivers from 56 hospices during a 2021 mode experiment, described in detail in sp.27 and 2a.02 below. Supplementary Appendix A contains a list of survey items in the current and proposed updated measures.

[Response Ends]

sp.01. Provide the measure title.

Measure titles should be concise yet convey who and what is being measured (see What Good Looks Like).

[Response Begins]

CAHPS® Hospice Survey, Version 9.0

[Response Ends]

sp.02. Provide a brief description of the measure.

Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).

[Response Begins]

The measures submitted here are derived from the CAHPS® Hospice Survey, Version 9.0, a 39-item standardized questionnaire and data collection methodology. The survey is intended to measure the care experiences of hospice decedents and their primary caregivers. Survey respondents are the primary informal caregivers (i.e., family members or friends) of patients who died while receiving hospice care.

The proposed measures include the following six multi-item measures:

- Hospice Team Communication
- Care Preferences
- Getting Timely Care
- Treating Family Member with Respect
- Getting Emotional and Religious Support
- Getting Help for Symptoms

In addition, there are three single-item measures:

- Getting Hospice Training
- Rating of the Hospice
- Willingness to Recommend the Hospice

Following is a list of the survey items included in each measure.

Hospice Team Communication (5 items)

- How often did the hospice team keep you informed about when they would arrive to care for your family member?
- How often did the hospice team explain things in a way that was easy to understand?
- How often did the hospice team listen carefully to you when you talked with them about problems with your family member's hospice care?
- How often did the hospice team keep you informed about your family member's condition?
- While your family member was in hospice care, how often did the hospice team listen carefully to you?

Care Preferences (2 items)

- Did the hospice team make an effort to listen to the things that mattered most to you or your family member?
- Did the hospice team provide care that respected your family member's wishes?

Getting Timely Care (2 items)

- When you or your family member asked for help from the hospice team, how often did you get help as soon as you needed it?
- How often did you get the help you needed from the hospice team during evenings, weekends, or holidays?

Treating Family Member with Respect (2 items)

- How often did the hospice team treat your family member with dignity and respect?
- How often did you feel that the hospice team really cared about your family member?

Getting Emotional and Religious Support (3 items)

- While your family member was in hospice care, how much emotional support did you get from the hospice team?
- In the weeks after your family member died, how much emotional support did you get from the hospice team?
- Support for religious or spiritual beliefs includes talking, praying, quiet time, or other ways of meeting your religious or spiritual needs. While your family member was in hospice care, how much support for your religious and spiritual beliefs did you get from the hospice team?

Getting Help for Symptoms (4 items)

- Did your family member get as much help with pain as he or she needed?
- How often did your family member get the help he or she needed for trouble breathing?
- How often did your family member get the help he or she needed for trouble with constipation?
- How often did your family member get the help he or she needed from the hospice team for feelings of anxiety or sadness?

Getting Hospice Care Training (1 item)

- Hospice teams may teach you how to care for family members who need pain medicine, have trouble breathing, are restless or agitated, or have other care needs. Did the hospice team teach you how to care for your family member?

Rating of Hospice Care (1 item)

- Using any number from 0 to 10, where 0 is the worst hospice care possible and 10 is the best hospice care possible, what number would you use to rate your family member's hospice care?

Willingness to Recommend Hospice (1 item)

- Would you recommend this hospice to your friends and family?

A complete list of proposed CAHPS Hospice Survey measures, including response options for each item, is available in Appendix B.

[Response Ends]

sp.04. Check all the clinical condition/topic areas that apply to your measure, below.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Surgery: General*

[Response Begins]

Palliative Care and End-of-Life Care

[Response Ends]

sp.05. Check all the non-condition specific measure domain areas that apply to your measure, below.

[Response Begins]

Person-and Family-Centered Care: Person-and Family-Centered Care

[Response Ends]

sp.06. Select one or more target population categories.

Select only those target populations which can be stratified in the reporting of the measure's result.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]

Elderly (Age >= 65)

Women

[Response Ends]

sp.07. Select the levels of analysis that apply to your measure.

Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

Facility

[Response Ends]

sp.08. Indicate the care settings that apply to your measure.

Check ONLY the settings for which the measure is SPECIFIED and TESTED.

[Response Begins]

Home Care

Inpatient/Hospital

Other

[Response Ends]

sp.09. 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. If no URL is available, indicate "none available".

[Response Begins]

https://hospicecahpssurvey.org/en/public-reporting/scoring-and-analysis/

[Response Ends]

sp.12. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

Attach an excel or csv file; if this poses an issue, contact staff. Provide descriptors for any codes. Use one file with multiple worksheets, if needed.

[Response Begins]

No data dictionary/code table – all information provided in the submission form

[Response Ends]

For the question below: state the outcome being measured. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.13. State the numerator.

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.

[Response Begins]

CMS calculates CAHPS Hospice Survey measure scores using top-, middle- and bottom- box scoring. The top-box score refers to the percentage of caregiver respondents that give the most positive response(s). The bottom box score refers to the percentage of caregiver respondents that give the least positive response(s). The middle box is the proportion remaining after the top and bottom boxes have been calculated; see below for details. Details regarding the definition of most and least positive response(s) are noted in Section SP.14 below.

[Response Ends]

For the question below: describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.14. Provide details needed to calculate the numerator.

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 sp.11.

[Response Begins]

For each survey item, the top and bottom box numerators are the number of respondents who selected the most and least positive response category(ies), respectively, as follows:

For items using a "Never/Sometimes/Usually/Always" response scale, the top box numerator is the number of respondents who answer "Always" and the bottom box numerator is the number of respondents who answer "Never" or "Sometimes."

For items using a "Yes, definitely/Yes, somewhat/No" response scale, the top box numerator is the number of respondents who answer "Yes, definitely" and the bottom box numerator is the number of respondents who answer "No."

For items using a "Too Little/Right Amount/Too Much" response scale, the top box numerator is the number of respondents who answer "Right Amount" and the bottom box numerator is the number of respondents who answer "Too little." "Too much" responses are not included in scoring. (There is no middle box for items using this response scale.)

The top box numerator for the Rating of Hospice item is the number of respondents who answer 9 or 10 for the item (on a scale of 0 to 10, where 10 is the "Best Hospice Care Possible"); the bottom box numerator is the number of respondents who answer 0 to 6.

The top box numerator for the Willingness to Recommend item is the number of respondents who answer "Definitely Yes" (on a scale of "Definitely No/Probably No/Probably Yes/Definitely Yes"); the bottom box numerator is the number of respondents who answer "Probably No" or "Definitely No."

Calculation of hospice-level multi-item measures

0. Score each item using top- box method, possible values of 0 or 100

1. Calculate mode- adjusted scores for each item for each respondent

2. Calculate case-mix adjusted scores for each item for each hospice

3. Take the unweighted means of the mode- and case-mix-adjusted hospice-level items to form multi-item measures

Here is an example of calculations for the measure "Getting Timely Care."

0. Score each item using top box method, possible values of 0 or 100

Both items in "Getting Care Quickly" have four response options: Never, Sometimes, Usually, Always. Recode each item as 100 for "Always" and 0 for "Never", "Sometimes", or "Usually".

Item #1. When you or your family member asked for help from the hospice team, how often did you get help as soon as you needed it?

Item #2. How often did you get the help you needed from the hospice team during evenings, weekends, or holidays?

1. Calculate mode-adjusted scores for each item for each respondent

2. Calculate case-mix adjusted scores for each item for each hospice

Each item is case mix adjusted separately; this step produces case-mix adjusted item-level scores for each hospice.

3. Take the unweighted means of the case-mix adjusted hospice-level items to form multi-item measures.

If the case-mix adjusted scores for a hospice are 95 for item #1 and 90 for item #2, then the hospice-level 'Getting Timely Care' would be calculated as (Item1 + Item2) / 2 = (95 + 90) / 2 = 92.5.

[Response Ends]

For the question below: state the target population for the outcome. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.15. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

In national implementation and public reporting, CAHPS® Hospice Survey measure scores are calculated only for hospices that had at least 30 completed questionnaires over the most recent eight quarters of data collection.

The target population for the survey are the adult primary caregivers of hospice decedents. Respondent eligibility and exclusions are defined in detail in the sections that follow. A survey is defined as completed when at least 50 percent of the questions applicable to all decedents/caregivers are answered. The survey uses screener questions to identify respondents eligible to respond to subsequent items. Therefore, denominators vary by survey item (and corresponding multi-item measures, if applicable) according to the eligibility of respondents for each item. In addition, for the Getting Hospice Care Training measure, scores are calculated only among those respondents who indicate that their family member received hospice care at home or in an assisted living facility.

[Response Ends]

For the question below: describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.16. Provide details needed to calculate the denominator.

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 sp.11.

[Response Begins]

For each item in a measure, the top box denominator is the number of respondents per hospice who answered the item. For each multi-item measure score, the denominator is the number of respondents who answer at least one item within the multi-item measure. Multi-item measure scores are the average proportion of respondents that gave responses in the most positive category across the items in the multi-item measure (as discussed in SP.14).

Survey population: Primary caregivers of patients who died while receiving care from a given hospice in a given month.

Denominator for Multi-Item Measures: The number of respondents who answer at least one item within the multi-item measure.

Denominator for Single-Item Measures: The number of respondents who answered the item. As noted above, for the Getting Hospice Care Training measure, scores are calculated only among those respondents who indicate that their family member received hospice care at home or in an assisted living facility.

[Response Ends]

sp.17. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

The exclusions noted here are those who are ineligible to participate in the survey. The one exception is caregivers who report on the survey that they "never" oversaw or took part in the decedent's care; these respondents are instructed to complete the "About You" and "About Your Family Member" sections of the survey only.

Cases are excluded from the survey target population if:

- The hospice patient is still alive
- The decedent's age at death was less than 18
- The decedent died within 48 hours of his/her last admission to hospice care
- The decedent had no caregiver of record
- The decedent had a caregiver of record, but the caregiver does not have a U.S. or U.S. Territory home address
- The decedent had no caregiver other than a nonfamilial legal guardian
- The decedent or caregiver requested that they not be contacted (i.e., by signing a no publicity request while under the care of hospice or otherwise directly requesting not to be contacted)
- The caregiver is institutionalized, has mental/physical incapacity, has a language barrier, or is deceased
- The caregiver reports on the survey that he or she "never" oversaw or took part in decedent's hospice care

[Response Ends]

sp.18. Provide details needed to calculate the denominator exclusions.

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 sp.11.

[Response Begins]

The CAHPS Hospice Survey Quality Assurance Guidelines (available at: http://www.hospicecahpssurvey.org/en/quality-assurance-guidelines/) contain detailed information regarding how to code decedent/caregiver cases, and how to code appropriately and inappropriately skipped items, as well as items with multiple responses.

[Response Ends]

sp.19. Provide all information required to stratify the measure results, if necessary.

Include 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 in the Data Dictionary field.

[Response Begins]

CAHPS Hospice Survey measure scores are used for reporting at the hospice-level (i.e., not stratified by region or other characteristics).

[Response Ends]

sp.20. Is this measure adjusted for socioeconomic status (SES)?

[Response Begins]

Yes

[Response Ends]

sp.21. Select the risk adjustment type.

Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.

[Response Begins]

Statistical risk model

[Response Ends]

sp.22. Select the most relevant type of score.

Attachment: If available, please provide a sample report.

[Response Begins]

Rate/proportion

[Response Ends]

sp.23. Select the appropriate interpretation of the measure score.

Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score

[Response Begins]

Better quality = Higher score

[Response Ends]

sp.24. Diagram or describe the calculation of the measure score as an ordered sequence of steps.

Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.

[Response Begins]

Top Box Score Calculation:

1. Identify target respondent population (i.e., primary caregivers of hospice patients who died while receiving hospice care from a given hospice in a given month).
2. Identify any exclusions from the respondent population (as described above in S.17).
3. Score each item using top box method, possible values of 0 or 100.
4. Calculate mode adjusted top box scores for each item.
5. Calculate case-mix adjusted top box scores for each item for each hospice; case-mix adjustment is a linear regression based approach that adjusts for all variables listed in S.14. Specifically, a regression model predicting item scores is fit using the case-mix adjustor variables and fixed effects for hospices. Adjusted hospice means are then calculated (e.g., using LSMEANS in SAS).
6. Top-box scores are averaged across the items within each multi-item measure, weighting each item equally. If data are missing for a respondent for an item(s) within a multi-item measure, the respondent's answers to other items within the measure are still used in the calculation of multi-item measure scores.

[Response Ends]

sp.25. Attach a copy of the instrument (e.g. survey, tool, questionnaire, scale) used as a data source for your measure, if available.

[Response Begins]

Copy of instrument is attached.

[Response Ends]

Attachment: 2651_2651_CAHPS Hospice Survey Version 9 2022_8-508.pdf

sp.26. Indicate the responder for your instrument.

[Response Begins]

Family or other caregiver

[Response Ends]

sp.27. If measure testing is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

Examples of samples used for testing:

- *Testing may be conducted on a sample of the accountable entities (e.g., hospital, physician). The analytic unit specified for the particular measure (e.g., physician, hospital, home health agency) determines the sampling strategy for scientific acceptability testing.*

- *The sample should represent the variety of entities whose performance will be measured. The 2010 Measure Testing Task Force recognized that the samples used for reliability and validity testing often have limited generalizability because measured entities volunteer to participate. Ideally, however, all types of entities whose performance will be measured should be included in reliability and validity testing.*
- *The sample should include adequate numbers of units of measurement and adequate numbers of patients to answer the specific reliability or validity question with the chosen statistical method.*
- *When possible, units of measurement and patients within units should be randomly selected.*

[Response Begins]

For the 2021 mode experiment, CMS sampled 15,515 decedents/caregivers from 56 large hospices, recruited for diversity with regard to region, profit status, and prior performance on the CAHPS Hospice Survey. Data collection occurred over six months, reflecting care experiences of those who died while receiving hospice care between March and August 2021. The number of decedents/caregivers sampled from each hospice ranged from 21 to 326, depending on number of survey-eligible decedents in the given month and the number previously sampled in 2021 (so as to allow hospices enough cases to meet their national implementation requirements).

The CAHPS Hospice Survey Quality Assurance Guidelines (available at http://www.hospicecahpssurvey.org/en/quality-assurance-guidelines/) specify guidance for sampling for national implementation. Specifically, hospices with 50 to 699 survey-eligible decedents/caregivers in the prior year are required to survey all cases (conduct a census). Hospices with 700 or more survey-eligible decedents/caregivers in the prior year are required to survey a minimum sample of 700 using an equiprobable approach (simple random sampling) and may conduct a census, if desired. If a hospice chooses to survey more than the required sample of 700 decedents/caregivers, all data collected must be submitted to the CAHPS Hospice Survey Data Warehouse.

[Response Ends]

sp.28. Identify whether and how proxy responses are allowed.

[Response Begins]

The CAHPS Hospice Survey is administered to the primary informal caregiver (i.e., family member or friend) of the decedent, as identified in hospice administrative records. Language in the questionnaire states that the survey should be given to the person in the household who knows the most about the hospice care received by the decedent. If, during a telephone attempt, the sampled caregiver indicates that someone within the household is more knowledgeable about the hospice care that the decedent received, the more knowledgeable person may be a proxy respondent. If a sampled caregiver indicates that he or she never oversaw, was not involved in, or is not knowledgeable about the hospice care provided to the decedent, interviewers may ask if someone else in the household is knowledgeable about the decedent's hospice care. If such a person exists, he or she may be a proxy respondent. Individuals outside of the sampled caregiver's household are not accepted as proxy respondents.

[Response Ends]

sp.29. Survey/Patient-reported data.

Provide instructions for data collection and guidance on minimum response rate. Specify calculation of response rates to be reported with performance measure results.

[Response Begins]

The CAHPS Hospice Survey Quality Assurance Guidelines (available at http:/www.hospicecahpssurvey.org/en/quality-assurance-guidelines/ provides detailed guidance regarding data collection procedures using currently approved modes: Mail Only, Telephone Only, or Mixed Mode (mail plus telephone follow up) administration. Briefly:

Users need to choose a data collection protocol that maximizes the survey response rate at an acceptable cost. Some sponsors, as well as researchers conducting field tests, have found that Mixed Mode (mail with telephone follow-up) yields the highest response rates: results from CAHPS mode experiments and national implementations indicate that Mixed Mode can add 10 to 15 percentage points to the response rate. The 2021 CAHPS Hospice mode experiment found that a Web-Mail mode can yield a similarly large increase in response rates over Mail Only mode among those with available email addresses. The Web-Mail mode tested during the mode experiment consisted of an initial email inviting respondents to complete survey by web, an email reminder to non-respondents two days later, mail survey to non-respondents four days later, and second mail survey to non-respondents 21 days later.

This section provides an overview of the protocol for collecting responses for each of the three modes of data collection currently approved: Mail Only, Telephone Only, and Mixed Mode (mail with telephone follow-up).

Mail Only

- Data collection for sampled decedents/caregivers must be initiated two months following the month of patient death. Survey vendors must send sampled caregivers a first questionnaire with a cover letter within the first seven days of the field period. A second questionnaire with a follow- up cover letter must be sent to all sampled caregivers who did not respond to the first questionnaire, approximately 21 calendar days after the first questionnaire mailing. Data collection must be closed out for a sampled caregiver by six weeks (42 calendar days) following the mailing of the first questionnaire.
- Mailings must include a personalized cover letter, a questionnaire, and a business reply envelope. The cover letters may be sent in both English and one of the official translations. English must be the default language in the continental U.S. and Spanish must be the default language in Puerto Rico. The Mail Only mode of survey administration may be conducted in English, Spanish, Chinese, Russian, Portuguese, Vietnamese, Polish, and Korean.
- Cover letters sent to respondents must be personalized with the decedent's name, the caregiver's name and the hospice's name. The letter must also provide a toll-free number for respondents to call if they have questions. The cover of the questionnaire must include a label indicating the name of the hospice, and if applicable, may include the specific hospice inpatient unit, acute care hospital or nursing home facility in which their family member or friend resided.
- To increase the likelihood that the respondent is the person within the sampled caregiver's household who is most knowledgeable about the decedent's hospice care, language must be included in the questionnaire, and optionally in the cover letter, clearly stating that the survey should be given to the person in the household who knows the most about the hospice care received by the decedent.

Telephone Only

- Data collection may be completed by telephone only. Outbound calling must be scheduled in a manner to ensure all cases have a first attempt within seven days of the start of the field period. Data collection must be closed out for a sampled caregiver by six weeks (42 calendar days) following the first call attempt. If it is known that the caregiver may be available in the latter part of the 42 calendar day data collection time period (e.g., caregiver is on vacation the first two or three weeks of the 42 calendar day field period and there would be an opportunity to reach the caregiver closer to the end of the field period), then survey vendors must use the entire field time period to schedule telephone calls.
- Survey vendors must attempt to reach each and every caregiver in the sample. Telephone call attempts are to be made between the hours of 9 AM and 9 PM respondent time. Repeated attempts must be made until the caregiver is contacted, found ineligible or five attempts have been made. After five attempts to contact the caregiver have been made, no further attempts are to be made.
- Telephone data collection is permitted in English, Spanish and Russian. English must be the default language in the continental U.S. and Spanish must be the default language in Puerto Rico. Survey vendors are provided standardized telephone scripts in both English and Spanish for CAHPS Hospice Survey administration.
- If, during a telephone attempt, the sampled caregiver indicates that someone within the household is more knowledgeable about the hospice care that the decedent received, the more knowledgeable person may be a proxy respondent. If a sampled caregiver indicates that he or she never oversaw, was not involved in, or is not knowledgeable about the hospice care provided to the decedent, interviewers may ask if someone else in the

household is knowledgeable about the decedent's hospice care. If such a person exists, he or she may be a proxy respondent. Interviewers must not accept individuals outside of the sampled caregiver's household as proxy respondents.

- Consistent monitoring of interviewers' work is essential to achieve standardized and accurate results. Properly trained and supervised interviewers ensure that standardized, non-directive interviews are conducted. Interviewers conducting the telephone survey must be trained prior to interviewing. Survey vendors must monitor at least 10 percent of all CAHPS Hospice Survey interviews, dispositions and call attempts in their entirety (English, Spanish and Russian) through silent monitoring of interviewers using the electronic telephone interviewing system software or an alternative system.

Mixed Mode

- Data collection for sampled decedents/caregivers must be initiated two months following the month of patient death within the first seven days of the field period. Survey vendors must send sampled caregivers a questionnaire with a cover letter within the first seven days of the field period, then beginning approximately 21 calendar days after mailing the questionnaire conduct a maximum of five telephone attempts to non-respondents. Data collection must be closed out for a sampled caregiver by six weeks (42 calendar days) following the mailing of the questionnaire.
- Reversing the protocol (telephone attempts followed by mail attempt) is not allowed.
- If a caregiver completes the CAHPS Hospice Survey via the telephone and a questionnaire is subsequently returned by the caregiver, the survey vendor must use the telephone CAHPS Hospice Survey responses since they were completed first. Alternately, if a questionnaire is returned after calling has begun, no further telephone attempts should be made, and the survey vendor must use the responses in the returned mail survey.
- All of the guidelines for both the Mail Only and the Telephone Only survey administration apply to the Mixed Mode survey administration.

The CAHPS Hospice Survey Quality Assurance Guidelines offer the following suggestions for maximizing response rates:

- For Mail Only and Mixed Mode survey administration, vendors must perform address updates for missing or incorrect information, including working with client hospices to obtain the most current caregiver contact information, using the National Change of Address and the United States Postal Service CASS Certified Zip+4 software, and using other means such as commercial software and internet search options to locate current addresses.
- For Telephone Only and Mixed Mode survey administration, vendors must use commercial software or other means to update telephone numbers provided by the hospice for sampled caregivers. This includes running update program software against the sample file just before or after uploading data to survey management systems, utilizing commercial software, Internet directories and/or directory assistance, and contacting the hospice to request updated telephone numbers.
- For Mail Only and Mixed Mode survey administration, send all mailings with first class postage or indicia to ensure delivery in a timely manner and to maximize response rates, as first class mail is more likely to be opened.
- For Telephone Only and Mixed Mode survey administration, make telephone attempts at various times of the day, on different days of the week, and in different weeks to maximize the probability that the survey vendor will contact the caregiver.
- Survey vendors should make every reasonable effort to achieve optimal telephone response rates, such as thoroughly familiarizing interviewers with the study purpose, carefully supervising interviewers, retraining those interviewers having difficulty enlisting cooperation, and re-contacting reluctant respondents at different times until the data collection protocol is completed.

[Response Ends]

sp.30. Select only the data sources for which the measure is specified.

[Response Begins]

Instrument-Based Data

[Response Ends]

sp.31. Identify the specific data source or data collection instrument.

For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.

[Response Begins]

CAHPS Hospice Survey; please see SP.29 for information regarding modes of data collection. The survey instrument is available in English, Spanish, Chinese, Russian, Portuguese, Vietnamese, Polish and Korean.

[Response Ends]

sp.32. Provide the data collection instrument.

[Response Begins]

Available in attached appendix in Question 1 of the Additional Section

[Response Ends]

2ma.01. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

Yes

[Response Ends]

2ma.02. Indicate whether additional empirical validity testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

Yes

[Response Ends]

2ma.03. For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?

[Response Begins]

Yes

[Response Ends]

2ma.04. For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.

Please update the Scientific Acceptability: Validity - Other Threats to Validity section.

Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.

[Response Begins]

No additional risk adjustment analysis included

[Response Ends]

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate fields in the Scientific Acceptability sections of the Measure Submission Form.

- Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.
- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact NQF staff with any questions. Check for resources at the Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the 2021 Measure Evaluation Criteria and Guidance.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a. Reliability testing demonstrates 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. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure;

AND

If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration
- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful 16 differences in performance;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses 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 non-responders) and how the specified handling of missing data minimizes bias.

2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and

2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(if not conducted or results not adequate, justification must be submitted and accepted)

Definitions

Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

Risk factors that influence outcomes should not be specified as exclusions.

With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v.$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous (Year) Submission:

Testing from the previous submission here.

2a.01. Select only the data sources for which the measure is tested.

[Response Begins]

Instrument-Based Data

[Response Ends]

2a.02. 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).

[Response Begins]

Current Submission:

We analyzed survey data collected during a 2021 CAHPS Hospice Survey mode experiment. Sampled decedents/caregivers from each hospice participating in the experiment were assigned to five study arms; in four of these arms, we administered a revised version of the CAHPS Hospice Survey (in Mail Only, Telephone Only, standard Mixed Mode [mail with telephone follow-up] and Web-Mail, respectively). In the fifth arm, we administered the current version of the CAHPS Hospice Survey in Mail Only mode. Data were collected by one survey vendor, the RAND Survey Research Group. Survey eligibility criteria and administration protocols for the three current CMS-approved modes of administration (Mail Only, Telephone Only, Mixed Mode [mail-telephone]) paralleled those in use for national implementation of the current version of the CAHPS Hospice Survey (described in detail in the *CAHPS® Hospice Survey Quality Assurance Guidelines,* available at: https://hospicecahpssurvey.org/en/quality-assurance-guidelines/). The additional, new tested mode, Web-Mail, consisted of an initial email inviting respondents to complete survey by web, an email reminder to non-respondents two days later, mail survey to non-respondents four days later, and second mail survey to non-respondents 21 days later.

Previous (2019) Submission:

We analyzed existing data from the CAHPS® Hospice Survey Data Warehouse. The Warehouse contains survey response data for all hospices eligible to participate in national implementation of the CAHPS® Hospice. Eligible hospices authorize a CMS-approved survey vendor to administer the CAHPS® Hospice Survey and submit data to the Warehouse in accordance with the standard protocols and deadlines outlined in the *CAHPS® Hospice Survey Quality Assurance*

Guidelines (available at http://www.hospicecahpssurvey.org/en/quality-assurance-guidelines). Data submissions to the Warehouse are used to determine hospices' compliance with CMS requirements for national implementation of the CAHPS® Hospice Survey, and to calculate hospices' official scores for the purpose of public reporting on Hospice Compare.

[Response Ends]

2a.03. Provide the dates of the data used in testing.

Use the following format: "MM-DD-YYYY - MM-DD-YYYY"

[Response Begins]

Current Submission:

03-01-2021 – 08-31-2022 (data collected from June 2021 through January 6, 2022, reflecting care experiences of those who died while receiving hospice care between March and August 2021)

Previous (2019) Submission:

07-01-2016 – 06-30-2018 (data collected from October 2016 through October 2018 regarding care experiences of patients who died while receiving hospice care from July 2016 through June 2018)

[Response Ends]

2a.04. Select the levels of analysis for which the measure is tested.

Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

Facility

[Response Ends]

2a.05. List the measured entities 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.

[Response Begins]

Current submission

The measured entity is a hospice agency. Hospice agencies (hereafter, "hospices") are differentiated from one another by their CMS Certification Numbers (CCNs). Hospice-level survey results are calculated across the primary caregiver respondents within a hospice.

Data for these analyses were from primary caregiver respondents whose family member or friend died while receiving hospice care from one of 56 hospices participating in the mode experiment. To ensure that hospices had sufficient sample to promote attainment of the overall targeted sample size for the experiment, we recruited hospices that provided care to 1,000 or more decedents in the most recent year for which data were available. To ensure that participating hospices were diverse with regard to factors associated with response rates and patterns, hospices were proportionately stratified into queues for recruitment using combinations of the following characteristics: region (West, Midwest, South, Northeast), profit status (for-profit versus not-for-profit), and past performance on the CAHPS Hospice Survey (above versus below median performance).

Table 2a.1 below shows the distribution of characteristics for hospices participating in the 2021 CAHPS Hospice Survey mode experiment; characteristics were derived from analyses of the March 2020 Medicare Provider of Services File, Medicare hospice claims data from 2020, and the CMS December 2021 Active Agency List. Thirty four percent of hospices operate in the South, a quarter operate in the Midwest, and approximately one fifth operate in each of the Northeast and West census regions. Eighty-two percent of hospices have been operating for 20 or more years; 55 percent were non-profit.

Table 2a.1. Characteristics of Hospices Participating in the 2021 CAHPS Hospice Survey Mode Experiment, N=56

Hospice Characteristic	n (%)
Census Region: Northeast	11 (19.6%)
Census Region: South	19 (33.9%)
Census Region: Midwest	14 (25.0%)
Census Region: West	12 (21.4%)
Rural	1 (1.8%)
Urban	55 (98.2%)
Size: <1500[a]	16 (28.6%)
Size: 1500-<2000 [a]	15 (26.8%)
Size: 2000-<2500 [a]	10 (17.9%)
Size: 2500+ [a]	15 (26.8%)
Ownership: For-Profit	18 (32.1%)
Ownership: Non-Profit	31 (55.4%)
Ownership: Other	7 (12.5%)
Hospice age: 4-8 years	1 (1.8%)
Hospice age: 9-19 years	9 (16.1%)
Hospice age: 20+ years	46 (82.1%)

Source: CMS March 2020 Provider of Services file, 2020 Medicare hospice claims, and CMS December 2021 Active Agency List.

[a] Hospice size is calculated as the total number of patients, including decedents, live discharges, and patients still under care in 2020 Medicare hospice claims.

[Response Ends]

2a.06. Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.

If there is a minimum case count used for testing, that minimum must be reflected in the specifications.

[Response Begins]

In keeping with CMS requirements for national implementation of the CAHPS Hospice Survey, during the mode experiment, the CAHPS Hospice Survey was completed by the primary informal caregiver (i.e., family member or friend) of the patient who died while receiving hospice care ("decedent"). Primary informal caregivers were identified from hospice administrative records; one primary caregiver was selected to respond to the survey for a given decedent. Decedents/caregivers were eligible for inclusion in the CAHPS® Hospice Survey if:

- Decedents were age 18 or over
- Decedents died at least 48 hours following last admission to hospice care
- Decedents had a caregiver of record
- Decedents had a caregiver with a U.S. or U.S. Territory home address
- Decedents had a caregiver other than a nonfamilial legal guardian.

Decedents or caregivers who requested that they not be contacted (those who signed no publicity requests while under the care of hospice or otherwise directly requested not to be contacted) were also excluded.

Current Submission

Table 2a.2 shows descriptive characteristics of the 5,731 decedents/caregivers for whom survey responses were submitted during the mode experiment. About one-quarter of decedents were age 90 years or older, 76.8 percent had only Medicare as their payer for hospice care, 86.2 percent were White, 5.5 percent were Black, and 5.1 percent were Hispanic. Fifty-four percent of decedents received care only at home, 10.0 percent received care only in a hospice inpatient unit, 8.4 percent received care only in an assisted living facility, and 6.6 percent received care only in a nursing home.

Table 2a.2. Characteristics of Decedents and Caregiver Respondents in the 2021 CAHPS Hospice Survey Mode Experiment, n=5,731

Characteristic	n (%)
Decedent sex: Male	2,683 (46.8%)
Decedent age: 18–54	196 (3.4%)
Decedent age: 55–64	426 (7.4%)
Decedent age: 65–69	435 (7.6%)
Decedent age: 70–74	602 (10.5%)
Decedent age: 75–79	788 (13.7%)
Decedent age: 80–84	848 (14.8%)
Decedent age: 85–89	983 (17.2%)
Decedent age: 90+	1,453 (25.4%)

Characteristic	n (%)
Decedent race/ethnicity: Asian or Pacific Islander	103 (1.8%)
Decedent race/ethnicity: Black	313 (5.5%)
Decedent race/ethnicity: Hispanic	293 (5.1%)
Decedent race/ethnicity: Multiracial or American Indian/Alaska Native	83 (1.4%)
Decedent race/ethnicity: White	4,939 (86.2%)
Primary diagnosis: Alzheimer’s and non-Alzheimer’s dementias	719 (12.5%)
Primary diagnosis: Bladder cancer	68 (1.2%)
Primary diagnosis: Blood and lymphatic cancers	231 (4.0%)
Primary diagnosis: Brain cancer	82 (1.4%)
Primary diagnosis: Breast cancer	133 (2.3%)
Primary diagnosis: Congestive heart failure	402 (7.0%)
Primary diagnosis: Chronic kidney disease	109 (1.9%)
Primary diagnosis: Chronic liver disease	101 (1.8%)
Primary diagnosis: Colorectal cancer	186 (3.2%)
Primary diagnosis: Cerebrovascular accident/Stroke	639 (11.2%)
Primary diagnosis: Liver cancer	119 (2.1%)
Primary diagnosis: Lung and other chest cavity cancer	433 (7.5%)
Primary diagnosis: Non-infectious respiratory	227 (4.0%)
Primary diagnosis: Other heart disease	355 (6.2%)
Primary diagnosis: Pancreatic cancer	192 (3.3%)
Primary diagnosis: Parkinson’s and other degenerative diseases	213 (3.7%)
Primary diagnosis: Pneumonias and other infectious lung diseases	207 (3.6%)
Primary diagnosis: Prostate cancer	125 (2.2%)

Characteristic	n (%)
Primary diagnosis: Other, cancer	555 (9.7%)
Primary diagnosis: Other, non-cancer	635 (11.1%)
Length of final hospice stay: 2–5 days	1,240 (21.6%)
Length of final hospice stay: 6–12 days	1,122 (19.6%)
Length of final hospice stay: 13–29 days	1,055 (18.4%)
Length of final hospice stay: 30–80 days	1,080 (18.8%)
Length of final hospice stay: 81+ days	1,234 (21.5%)
Setting of hospice care: Home only	3,082 (53.8%)
Setting of hospice care: Nursing home only	380 (6.6%)
Setting of hospice care: Hospital only	274 (4.8%)
Setting of hospice care: Hospice inpatient unit only	572 (10.0%)
Setting of hospice care: Assisted living facility only	484 (8.4%)
Setting of hospice care: Other	85 (1.5%)
Setting of hospice care: More than one setting	854 (14.9%)
Caregiver sex: Male	1,649 (28.8%)
Caregiver age: 18–44	353 (6.2%)
Caregiver age: 45–54	657 (11.5%)
Caregiver age: 55–64	1,554 (27.1%)
Caregiver age: 65–74	1,739 (30.4%)
Caregiver age: 75–84	1,034 (18.0%)
Caregiver age: 85+	394 (6.9%)
Caregiver education: No high school degree	200 (3.5%)
Caregiver education: High school graduate or equivalent	1,271 (22.2%)
Caregiver education: Some college	1,827 (31.9%)

Characteristic	n (%)
Caregiver education: 4-year college graduate	1,251 (21.8%)
Caregiver education: More than 4-year college graduate	1,182 (20.6%)
Caregiver relationship (Decedent was the caregiver's ___): Spouse or partner	2,244 (39.2%)
Caregiver relationship (Decedent was the caregiver's ___): Parent	2,638 (46.0%)
Caregiver relationship (Decedent was the caregiver's ___): Mother-in-law or father-in-law	125 (2.2%)
Caregiver relationship (Decedent was the caregiver's ___): Aunt or uncle	90 (1.6%)
Caregiver relationship (Decedent was the caregiver's ___): Sister or brother	273 (4.8%)
Caregiver relationship (Decedent was the caregiver's ___): Child	108 (1.9%)
Caregiver relationship (Decedent was the caregiver's ___): Friend	102 (1.8%)
Caregiver relationship (Decedent was the caregiver's ___): Other	151 (2.6%)

Source: Hospice administrative and survey response data from the 2021 CAHPS Hospice Survey mode experiment, reflecting care experiences of those who died while receiving hospice care between March and August 2021.

Previous (2019) Submission

Table 1.6 shows descriptive characteristics of the 647,694 caregiver respondents for whom survey responses were submitted to the CAHPS® Hospice Survey Data Warehouse from Quarter 3 2016 through Quarter 2 2018 from the hospices described in Table 1.5. The mean age of decedents was 81.8 (Table 1.6); 5.0 percent were black, and 4.8 percent were Hispanic. For more than half of decedents (58.7%), the last setting of hospice care was a home or assisted living facility; last location was a nursing home for 17.6 percent of decedents, a hospice freestanding inpatient unit for 13.7 percent, and an acute care hospital for 3.9 percent. Approximately one-quarter of decedents had a length of stay in hospice of less than one week, while 12.4 percent of decedents had a stay of six months or more.

Table 1.6. Characteristics of Decedents and Caregiver Respondents in National Implementation of the CAHPS® Hospice Survey, Quarter 3 2016 – Quarter 2 2018

Characteristic	n (%)
Decedent Characteristics	*
Sex	*

Characteristic	**n (%)**
Male	294,860 (45.4%)
Missing	8,312 (1.4%)
Age	*
(mean, SD)	81.8 (11.9)
Missing	103 (<0.1%)
Race/Ethnicity	*
White	565,392 (85.9%)
Black	28,596 (5.0%)
Hispanic	27,434 (4.8%)
Asian or Pacific Islander	8,575 (1.5%)
Multiracial or Native American	6,796 (1.1%)
Missing	10,901 (1.7%)
Final Setting of Care	*
Home or Assisted Living Facility	381,655 (58.7%)
Nursing Home	115,982 (17.6%)
Acute Care Hospital	24,062 (3.9%)
Hospice Inpatient Unit	87,227 (13.7%)
Other	3,141 (0.5%)
Missing	35,627 (5.6%)
Length of final episode of hospice care	*
Less than 1 week	156,880 (24.3%)
1 to less than 2 weeks	109,376 (16.9%)
2 to less than 4 weeks	99,279 (15.3%)
1 to less than 2 months	90,543 (13.9%)
2 to less than 4 months	76,017 (11.7%)
4 to less than 6 months	35,762 (5.5%)

Characteristic	n (%)
6 or more months	79,552 (12.4%)
Missing	285 (0.1%)
Primary Diagnosis	*
Dementia	95,004 (14.9%)
Parkinson’s and other degenerative diseases	20,034 (3.1%)
Cancer	214,761 (33.0%)
Non-cancer end organ, except ESRD	147,136 (22.5%)
ESRD	11,588 (1.8%)
Pneumonias and other infectious lung diseases	17,593 (2.7%)
Cerebrovascular accident/Stroke	50,614 (8.1%)
Other	61,091 (9.3%)
Missing	29,873 (4.7%)
Caregiver Characteristics	*
Sex	*
Male (%)	177,189 (27.3%)
Missing	13,493 (2.1%)
Age	*
18 to 24	702 (0.1%)
25 to 34	4,996 (0.8%)
35 to 44	15,716 (2.5%)
45 to 54	65,095 (10.1%)
55 to 64	190,149 (29.4%)
65 to 74	202,970 (31.3%)
75 to 84	116,028 (17.8%)
85 or older	44,637 (6.9%)
Missing	7,401 (1.2%)

Characteristic	n (%)
Relation to decedent (decedent was caregiver's):	*
Spouse or partner	255,086 (39.2%)
Parent	288,635 (44.6%)
Mother-in-law or father-in-law	14,307 (2.2%)
Grandparent	5,340 (0.8%)
Aunt or Uncle	14,514 (2.3%)
Sister or Brother	28,280 (4.5%)
Child	11,212 (1.8%)
Friend	10,712 (1.7%)
Other	13,323 (2.1%)
Missing	6,285 (1.0%)

Source: Hospice administrative and survey response data submitted to the CAHPS® Hospice Survey Data Warehouse, Q3 2016 – Q2 2018.

Note: Counts are unweighted. Percentages and means are weighted to account for probability of sampling and non-response.

* This cell intentionally left empty.

[Response Ends]

2a.07. 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.

[Response Begins]

All analyses with the exception of the development of the risk adjustment model were conducted with the mode experiment dataset described in 2a.02. The risk adjustment model was developed using national implementation data from the previous submission, as described in 2b.20.

[Response Ends]

2a.08. List 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.

[Response Begins]

Current Submission

Social risk factors in the data include (1) decedent education, (2) primary payer for hospice care, (3) caregiver respondent education, and (4) caregiver respondent language (i.e., language that respondents indicate is the one that they speak at home, or in which they complete the survey). Table 2a.3 shows the distribution of these factors for the 5,731 decedents and caregiver respondents for whom survey responses were submitted during the 2021 CAHPS Hospice Survey mode experiment. Nearly half of decedents had a high school education or less; less than 0.5 percent were uninsured, while 6.4 percent were covered by Medicaid alone or in combination with Medicare or private insurance. Approximately one quarter of caregiver respondents had a high school education or less. Spanish was the survey language or language spoken at home for one percent of caregiver respondents.

Table 2a.3. Social Risk Factor Characteristics of Decedents and Caregiver Respondents in the 2021 CAHPS Hospice Survey Mode Experiment, n=5,731

Characteristic	n (%)
Decedent Education: No high school degree	754 (13.2%)
Decedent Education: High school graduate or equivalent	2,018 (35.2%)
Decedent Education: Some college	1,296 (22.6%)
Decedent Education: 4-year college graduate	862 (15.0%)
Decedent Education: More than 4-year college graduate	801 (14.0%)
Payer for hospice care: Medicare only	4,400 (76.8%)
Payer for hospice care: Medicaid only or Medicaid and private insurance	131 (2.3%)
Payer for hospice care: Medicare and Medicaid	235 (4.1%)
Payer for hospice care: Private insurance only	284 (5.0%)
Payer for hospice care: Medicare and private insurance	265 (4.6%)
Payer for hospice care: Uninsured	19 (0.3%)
Payer for hospice care: Other	397 (7.0%)
Caregiver Education: No high school degree	200 (3.5%)
Caregiver Education: High school graduate or equivalent	1,271 (22.2%)
Caregiver Education: Some college	1,827 (31.9%)
Caregiver Education: 4-year college graduate	1,251 (21.8%)
Caregiver Education: More than 4-year college graduate	1,182 (20.6%)
Spanish survey or home language	55 (1.0%)
English or other home language	5,676 (99.0%)

Source: Hospice administrative and survey response data submitted from the 2021 CAHPS Hospice Survey Mode Experiment, reflecting care experiences of those who died while receiving hospice care between March and August 2021.

Previous Submission

Social risk factors in the data include (1) decedent education, (2) primary payer for hospice care, (3) caregiver respondent education, and (4) caregiver respondent language (i.e., language that respondents indicate is the one that they speak at home, or in which they complete the survey). Table 1.8 shows the distribution of these factors for the 647,694 decedents and caregiver respondents for whom survey responses were submitted to the CAHPS® Hospice Survey Data Warehouse from Quarter 3 2016 through Quarter 2 2018 from the hospices described in Table 1.5. More than half of decedents had a high school education or less (Table 1.8); less than one percent were uninsured or had no payer on file with the hospice, while 6.5 percent were covered by Medicaid. Approximately 30 percent of caregiver respondents had a high school education or less. Spanish was the survey language or language spoken at home for 1.5 percent of caregiver respondents.

Table 1.8. Social Risk Factor Characteristics of Decedents and Caregiver Respondents in National Implementation of the CAHPS® Hospice Survey, Quarter 3 2016 – Quarter 2 2018

Characteristic	n (%)
Decedent Characteristics	*
Highest grade or level of school	*
8th grade or less	52,127 (8.1%)
Some high school but did not graduate	57,040 (8.8%)
High school graduate or GED	237,281 (36.2%)
Some college or 2-year degree	137,919 (21.4%)
4-year college graduate	65,511 (10.2%)
More than 4-year college degree	71,062 (11.1%)
Missing	26,754 (4.2%)
Payer for hospice services	*
Medicare only	445,470 (68.6%)
Medicaid only or Medicaid and Private	9,887 (1.6%)
Medicare and Medicaid	30,531 (4.9%)
Private only	25,228 (3.9%)
Private and Medicare	31,292 (5.0%)
Uninsured/no payer	2,392 (0.4%)
Other	73,899 (11.1%)
Missing	28,995 (4.6%)
Caregiver Characteristics	*
Highest grade or level of school	*
8th grade or less	7,375 (1.2%)

Characteristic	n (%)
Some high school but did not graduate	22,941 (3.6%)
High school graduate or GED	165,325 (25.1%)
Some college or 2-year degree	204,696 (31.6%)
4-year college graduate	106,363 (16.6%)
More than 4-year college degree	123,968 (19.4%)
Missing	17,026 (2.7%)
Survey and Home Language	*
Spanish survey or home language	8,703 (1.5%)
Chinese survey or home language	712 (0.1%)
English or other language	638,274 (98.4%)
Missing	5 (<0.1%)

Source: Hospice administrative and survey response data submitted to the CAHPS® Hospice Survey Data Warehouse, Q3 2016 – Q2 2018.

Note: Counts are unweighted. Percentages and means are weighted to account for probability of sampling and non-response.

* This cell intentionally left empty.

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.09 check patient or encounter-level data; in 2a.010 enter "see validity testing section of data elements"; and enter "N/A" for 2a.11 and 2a.12.

2a.09. Select the level of reliability testing conducted.

Choose one or both levels.

[Response Begins]

Patient or Encounter-Level (e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements)

Accountable Entity Level (e.g., signal-to-noise analysis)

[Response Ends]

2a.10. For each level of reliability testing 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.

[Response Begins]

To assess measure reliability, we calculate the following for multi-item measures:

1. The internal consistency of the multi-item measures, estimated by Cronbach's alpha. Cronbach's alpha, a 0 to 1 index of the magnitude of internal consistency for multi-item measures, increases with the number of items in a multi-item measure and their average correlation with each other. Higher values indicate more precise measurement of the multi-item measure. For multi-item measures with more than two items, we also show the effect on Cronbach's alpha of deleting one of the items from the measure.
2. The person-level Pearson item-total correlation, the correlation between a given item and the total multi-item measure score with the given item removed. This metric reflects how related each item is to all other items in the measure.

In addition, we calculate the following for both multi-item and single-item (global) measures:

1. Inter-unit (i.e., hospice-level) reliability, which refers to the degree to which measure scores are able to precisely distinguish between the performances of hospices. We examined reliability for each measure using intra-class correlations (ICCs) computed from the case mix-adjusted 0-100 top-box scores.

Since reliability is designed to be applicable to national implementation of CAHPS Hospice Survey measures, we calculated the ICCs using data from the 2021 mode experiment, and used the most recent mean sample sizes from national implementation data to calculate the reliability of the measure using the Spearman Brown prediction formula (Allen & Yen, 1979). (The most recent national implementation data were from Q3 2019 to Q4 2019 and Q3 2020 to Q4 2021; data from decedents in Q1 and Q2 2020 were not used in public reporting due to the COVID-19 public health emergency.) There is no particular ICC threshold that indicates acceptable reliability; hospice-level reliability is determined by the combination of the ICC with the number of respondents. When entities such as hospices are being compared, multi-item measure reliability greater than 0.70 is commonly considered adequate (Nunnally & Bernstein, 1994).

Given that the primary use of the measures is to distinguish the performance of one hospice from another, hospice-level (rather than person-level) reliability is of primary importance for both the multi-item and single-item CAHPS® Hospice Survey measures. Variance at the hospice level is a sum of person-level and hospice-level variance components; therefore, if hospice-level reliability is adequate for a given measure, person-level reliability will be reliable, as well, since it is not possible for measurement to be reliable at the hospice level by this metric without it being reliable at the person level.

Out of sensitivity to grief and bereavement, CMS does not repeatedly survey the same informal caregivers as would be needed to calculate test-retest reliability of CAHPS® Hospice Survey single-item measures at the person level, even for the purposes of testing. However, prior researchers have established the stability of responses to similar CAHPS® survey items regarding overall ratings of health care by conducting telephone interviews approximately two weeks after survey completion; they report estimated test-retest reliabilities (product-moment correlations) of 0.71 and 0.82 ($P<0.01$) for overall rating of the medical group and overall rating of care, respectively (Solomon, Hays et al. 2005). In the hospice context, DiBiasio and colleagues assessed stability of responses to items assessing the overall quality of hospice care and willingness to recommend the hospice at three, six and nine months following death, and reported agreement of 86% and higher between overall ratings at three and six, three and nine, and six and nine months, and agreement of 90% and higher between reports of willingness to recommend at these time intervals. Corresponding Kappa statistics ranged from 0.58 (for willingness to recommend between three and nine months) to 0.70 (for overall rating between three and six months), suggesting adequate reliability of these single-item measures when repeated over time among the same respondents, even among respondents who demonstrated large changes in grief over time (DiBiasio, Clark et al. 2015).

Citations:

1. Allen M & Yen W. *Introduction to Measurement Theory*. Monterey, CA: Brooks/Cole; 1979.
2. DiBiasio EL, Clark MA, Gozaol PL, Spence C, Casarett DJ, Teno JM. (2015). Timing of Survey Administration After Hospice Patient Death: Stability of Bereaved Respondents. *J Pain Symptom Manage* 50(1): 17-27.
3. Nunnally JC & Bernstein IH. Psychometric Theory. New York: McGraw Hill; 1994.
4. Solomon LS, Hays RD, Zaslavsky AM, Ding L, Cleary PD. (2005). Psychometric Properties of a Group-Level Consumer Assessment of Health Plans Study (CAHPS) Instrument. *Med Care* 43: 53-60.

[Response Ends]

2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?

For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, NQF Measure Evaluation Criteria).

[Response Begins]

Current Submission

Table 2a.4. Cronbach's Alpha Reliability Coefficients (Overall and Dropping Each Item) and Item-total Pearson Correlations for CAHPS Hospice Survey Multi-Item Measures, 2021 Mode Experiment

Multi-Item Measures and items	Item-total Pearson correlation	Cronbach's Alpha with item deleted (95% CI)
Hospice team communication (Cronbach's alpha = 0.84, 95% CI: 0.83, 0.85)	*	*
How often did the hospice team listen carefully to you when you talked with them about problems with your family member's hospice care?	0.69	0.80 (0.78, 0.81)
While your family member was in hospice care, how often did the hospice team listen carefully to you?	0.71	0.79 (0.78, 0.80)
How often did the hospice team explain things in a way that was easy to understand?	0.66	0.80 (0.79, 0.81)
How often did the hospice team keep you informed about your family member's condition?	0.64	0.81 (0.80, 0.82)
How often did the hospice team keep you informed about when they would arrive to care for your family member?	0.52	0.84 (0.83, 0.85)
Care preferences (Cronbach's alpha = 0.75, 95% CI: 0.73, 0.77)	*	*
Did the hospice team make an effort to listen to the things that mattered most to you or your family member?	0.60	N/A
Did the hospice team provide care that respected your family member's wishes?	0.60	N/A
Getting timely care (Cronbach's alpha = 0.62, 95% CI: 0.59, 0.64)	*	*
While your family member was in hospice care, when you or your family member asked for help from the hospice team, how often did you get help as soon as you needed it?	0.45	N/A
How often did you get the help you needed from the hospice team during evenings, weekends, or holidays?	0.45	N/A

Multi-Item Measures and items	**Item-total Pearson correlation**	**Cronbach's Alpha with item deleted (95% CI)**
Treating your family member with respect (Cronbach's alpha = 0.70, 95% CI: 0.68, 0.73)	*	*
While your family member was in hospice care, how often did the hospice team treat your family member with dignity and respect?	0.53	N/A
While your family member was in hospice care, how often did you feel that the hospice team really cared about your family member?	0.53	N/A
Getting emotional and religious support (Cronbach's alpha = 0.71, 95% CI: 0.68, 0.73)	*	*
Support for religious or spiritual beliefs includes talking, praying, quiet time, or other ways of meeting your religious or spiritual needs. While your family member was in hospice care, how much support for your religious and spiritual beliefs did you get from the hospice team?	0.51	0.63 (0.60, 0.66)
While your family member was in hospice care, how much emotional support did you get from the hospice team?	0.58	0.54 (0.50, 0.57)
In the weeks after your family member died, how much emotional support did you get from the hospice team?	0.48	0.67 (0.63, 0.70)
Getting help for symptoms (Cronbach's alpha = 0.73, 95% CI: 0.70, 0.76)	*	*
Did your family member get as much help with pain as he or she needed?	0.46	0.71 (0.68, 0.74)
How often did your family member get the help he or she needed for trouble breathing?	0.50	0.69 (0.65, 0.72)
How often did your family member get the help he or she needed for trouble with constipation?	0.56	0.65 (0.61, 0.69)
How often did your family member get the help he or she needed from the hospice team for feelings of anxiety or sadness?	0.58	0.65 (0.61, 0.69)

Note: N/A indicates that Cronbach's alpha cannot be calculated for a single item. CI = confidence interval.

* This cell intentionally left empty.

Multi-Item and Single-Item Measures

Table 2a.5. Hospice-Level Reliability for CAHPS Hospice Survey Measures, 2021 Mode Experiment

Measure	**Intraclass Correlation Coefficient (ICC) Estimate (95% CI)**	**Average Number of Measure Respondents**	**Reliability at Average Number of Measure Respondents Estimate (95% CI)**
Multi-item measures	*	*	*

Measure	Intraclass Correlation Coefficient (ICC) Estimate (95% CI)	Average Number of Measure Respondents	Reliability at Average Number of Measure Respondents Estimate (95% CI)
Getting Timely Care (two items)	0.023 (0.013, 0.040)	211	0.83 (0.73, 0.90)
Hospice Team Communication (five items)	0.024 (0.013, 0.041)	214	0.84 (0.74, 0.90)
Treating Family Member with Respect (two items)	0.015 (0.008, 0.030)	213	0.77 (0.64, 0.87)
Getting Help for Symptoms (four items)	0.012 (0.006, 0.026)	191	0.70 (0.53, 0.84)
Getting Emotional and Religious Support (three items)	0.018 (0.010, 0.035)	209	0.79 (0.69, 0.88)
Care Preferences (two items)[a]	0.016 (0009, 0.032)	213	0.78 (0.65, 0.87)
Single-item measures	*	*	*
Getting Hospice Care Training [a, b]	0.017 (0.007, 0.039)	130	0.70 (0.49, 0.84)
Overall Rating	0.028 (0.017, 0.047)	211	0.86 (0.78, 0.91)
Willingness to Recommend	0.030 (0.018, 0.050)	211	0.87 (0.80, 0.92)

CI = confidence interval.

[a] For the Care Preferences and single-item Getting Hospice Training measures, which are not on the survey instrument currently used in national implementation, the average number of completes per hospice was extrapolated based on the percent of respondents completing those measures in the 2021 mode experiment.

[b] ICCs and reliability for the Getting Hospice Care Training measure are calculated only for those decedents who received hospice care at home or in an assisted living facility.

* This cell intentionally left empty.

Previous (2019) Submission

Table 2a2.3a shows the Cronbach's alpha for each multi-item measure, the Cronbach's alpha for each multi-item measure with each of its component items deleted, as well as the item-total Pearson correlations (i.e., the correlation between each item and the multi-item measure with the given item removed).

Table 2a2.3a. Cronbach's Alpha Reliability Coefficients (Overall and Dropping Each Item) and Item-total Pearson Correlations for CAHPS® Hospice Survey Measures, Quarter 3 2016 – Quarter 2 2018

Multi-Item Measures and items	**Item-total Pearson correlation**	**Cronbach's Alpha with item deleted**
Hospice team communication (Cronbach's alpha = 0.82)	*	*
How often did the hospice team listen carefully to you when you talked with them about problems with your family member's hospice care?	0.66	0.76
While your family member was in hospice care, how often did the hospice team listen carefully to you?	0.61	0.78
While your family member was in hospice care, how often did the hospice team explain things in a way that was easy to understand?	0.62	0.78
While your family member was in hospice care, how often did the hospice team keep you informed about your family member's condition?	0.61	0.79
While your family member was in hospice care, how often did the hospice team keep you informed about when they would arrive to care for your family member?	0.48	0.83
While your family member was in hospice care, how often did anyone from the hospice team give you confusing or contradictory information about your family member's condition or care?	0.39	0.82
Getting timely care (Cronbach's alpha = 0.61)	*	*
While your family member was in hospice care, when you or your family member asked for help from the hospice team, how often did you get help as soon as you needed it?	0.43	N/A
How often did you get the help you needed from the hospice team during evenings, weekends, or holidays?	0.43	N/A
Treating your family member with respect (Cronbach's alpha = 0.61)	*	*
While your family member was in hospice care, how often did the hospice team treat your family member with dignity and respect?	0.48	N/A
While your family member was in hospice care, how often did you feel that the hospice team really cared about your family member?	0.48	N/A
Getting emotional and religious support (Cronbach's alpha = 0.66)	*	*
In the weeks after your family member died, how much emotional support did you get from the hospice team?	0.46	0.62
While your family member was in hospice care, how much emotional support did you get from the hospice team?	0.54	0.49
Support for religious or spiritual beliefs includes talking, praying, quiet time, or other ways of meeting your religious or spiritual needs. While your family member was in hospice care, how much support for your religious and spiritual beliefs did you get from the hospice team?	0.46	0.59

Multi-Item Measures and items	**Item-total Pearson correlation**	**Cronbach's Alpha with item deleted**
Getting help for symptoms (Cronbach's alpha = 0.74)	*	*
How often did your family member receive the help he or she needed from the hospice team for feelings of anxiety or sadness?	0.50	0.65
How often did your family member get the help he or she needed for trouble with constipation?	0.51	0.67
How often did your family member get the help he or she needed for trouble breathing?	0.46	0.69
Did your family member get as much help with pain as he or she needed?	0.41	0.72
Getting hospice care training (Cronbach's alpha = 0.84)	*	*
Did the hospice team give you enough training about how to help your family member if he or she had trouble breathing?	0.63	0.81
Did the hospice team give you enough training about what side effects to watch for from pain medicine?	0.70	0.79
Did the hospice team give you enough training about what to do if your family member became restless or agitated?	0.62	0.81
Did the hospice team give you enough training about if and when to give more pain medicine to your family member?	0.61	0.81
Side effects of pain medicine include things like sleepiness. Did any member of the hospice team discuss side effects of pain medicine with you or your family member?	0.62	0.81

Note: N/A indicates that Cronbach's alpha cannot be calculated for a single item.

* This cell intentionally left empty.

Multi-Item and Single-Item Measures

Table 2a2.3b shows the hospice-level reliability for each measure (including both multi-item and single-item measures) at the average number of completed surveys per hospice between Q3 2016 and Q2 2018.

Table 2a2.3b. Hospice-Level Reliability for CAHPS® Hospice Survey Measures, Quarter 3 2016 – Quarter 2 2018

Multi-item or single-item measure	**Intraclass Correlation Coefficient (ICC)**	**Expected average number of completed surveys per hospice**	**Hospice reliability @ expected average number of completes**
Multi-Item Measures	*	*	*
Hospice Team Communication (6-items)	0.023	221	0.84
Getting Timely Care (2-items)	0.019	218	0.81

Multi-item or single-item measure	**Intraclass Correlation Coefficient (ICC)**	**Expected average number of completed surveys per hospice**	**Hospice reliability @ expected average number of completes**
Treating Family Member with Respect (2-items)	0.013	220	0.75
Getting Emotional and Religious Support (3-items)	0.016	217	0.78
Getting Help for Symptoms (4-items)	0.012	197	0.71
Getting Hospice Care Training (5-items)	0.022	145	0.76
Global Measures	*	*	*
Rating of Hospice	0.018	218	0.80
Willingness to Recommend	0.025	218	0.85

* This cell intentionally left empty.

[Response Ends]

2a.12. Interpret the results, in terms of how they demonstrate reliability.

(In other words, what do the results mean and what are the norms for the test conducted?)

[Response Begins]

Current Submission

Multi-Item Measures

Multi-item measures with the highest Cronbach's alpha reliability (internal consistency) coefficients were *Hospice Team Communication* (alpha = 0.84), and *Care Preferences* (alpha = 0.75; Table 2a.4). Cronbach's alphas of 0.70 or higher are considered adequate for group comparisons (Nunnally & Bernstein 1994).

As shown in Table 2a.4, the Cronbach's alpha with item deleted was smaller than the Cronbach's alpha for multi-item measures for all but "How often did the hospice team keep you informed about when they would arrive to care for your family member?" This item is maintained within the measure due to its importance to patients and families and its high feasibility as a target for quality improvement efforts. Item-total Pearson correlations are generally moderate to high, indicating that items are related to all other items within their multi-item measures.

Multi-Item and Single-Item Measures

All nine measures exhibit acceptable hospice-level reliability of 0.70 or greater at the expected average number of completed surveys per hospice (Table 2a.5*)*. As noted above in 2a,10, adequate hospice-level reliability is also an indication of adequate person-level reliability for multi-item and single-item measures.

Previous (2019) Submission

Multi-Item Measures

Multi-item measures with the highest Cronbach's alpha reliability (internal consistency) coefficients were *Getting Hospice Care Training* (alpha = 0.84)*, Hospice Team Communication* (alpha = 0.82), and *Getting Help for Symptoms* (alpha = 0.74; Table 2a2.3a). Cronbach's alphas of 0.70 or higher are considered adequate for group comparisons (Nunnally & Bernstein 1994).

As shown in Table 2a2.3a, the Cronbach's alpha with item deleted was smaller than the Cronbach's alpha for multi-item measures for all but two items. These items, about informing family about when the hospice would arrive and provision of confusing or contradictory information, are maintained within the measure due to their importance to patients and families and their high feasibility as targets for quality improvement efforts. Item-total Pearson correlations are generally moderate to high, indicating that items are related to all other items within their multi-item measures.

Multi-Item and Single-Item Measures

All eight measures exhibit acceptable hospice-level reliability of 0.70 or greater at the average number of completed surveys per hospice (Table 2a2.3b*)*. As noted above in 2a2.2, adequate hospice-level reliability is also an indication of adequate person-level reliability for multi-item and single-item measures.

Citation:

1. Nunnally JC & Bernstein IH. Psychometric Theory. New York: McGraw Hill; 1994.

[Response Ends]

2b. Validity

2b.01. Select the level of validity testing that was conducted.

[Response Begins]

Patient or Encounter-Level (data element validity must address ALL critical data elements)

Accountable Entity Level (e.g. hospitals, clinicians)

Empirical validity testing

[Response Ends]

2b.02. 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.

[Response Begins]

In keeping with the development of the original CAHPS Hospice Survey measures (Anhang Price, Quigley et al. 2014), updated CAHPS Hospice Survey measures were developed following best practices, including a request for public comment, a review of the literature and environmental scan of existing tools for measuring experiences with end-of-life care (to identify new content in topic areas suggested by stakeholders), qualitative inquiry with primary caregivers of hospice patients, input and feedback from technical experts, cognitive testing, and field-testing of a draft survey instrument (via the 2021 mode experiment).

To refine and update CAHPS Hospice Survey measures using data from the 2021 mode experiment, and evaluate the data element validity (i.e., the degree to which survey items included in each measure are consistent with their theoretical construct), we conducted exploratory factor analysis and confirmatory factor analysis of newly tested items and the items in the current multi-item measures. We used means and variance adjusted weighted least squares to account for the dichotomous nature of top-box item scores (Flora and Curran, 2004). We used a criterion of factor loadings greater than or equal to 0.40 for inclusion within the composite (Brown, 2015) and assessed overall model fit using the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) in EFA or weighted root mean square residual (WRMR) in CFA. Prior research indicates that a model with a good fit typically has a CFI > 0.95, RMSEA < 0.05, SRMR < 0.08, and WRMR < 1.0, with WRMR being less critical (DiStefano et al., 2018; Hu and Bentler, 1999). The model chi-square statistic and standard error of model estimates were adjusted to account for the clustering of patients within hospices (Asparouhov, 2006; Wu and Kwok, 2012). All models were estimated in Mplus V8 (Muthén and Muthén, 2017). For exploratory factor analysis, we obtained factor loadings and inter-factor correlations using the Geomin oblique rotation method.

We used data from the 2021 mode experiment to examine the relationships between each measure's top box score and the top box score for the global measures of "Using any number from 0 to 10, where 0 is the worst hospice care possible and 10 is the best hospice care possible, what number would you use to rate your family member's hospice care?" and "Would you recommend this hospice to your friends and family?" We calculated Pearson correlations between each of the two global measures and all other measures, adjusting for case mix and mode of survey administration.

We also examined Pearson correlations among the multi-item measures to assess the magnitude of association between the measures, as a proxy for understanding the extent to which they measure different constructs. Moderate intercorrelations are to be expected, given that all measures assess aspects of care experiences. Very high intercorrelations may indicate that measures should be combined, while low intercorrelations may be seen as evidence of discriminant validity (i.e., that measure content that we would expect not to be substantially related is in fact not related very highly).

For the current submission, construct and discriminant validity analyses were conducted using individual-level data, as estimates of hospice-level associations would be unbiased but imprecise if calculated among the 56 hospices participating in the 2021 mode experiment. For the previous (2019) submission, results are shown at both the individual- and hospice-level, as national implementation data reflects thousands of hospices, allowing for more precise estimates.

An analysis of 2016 CAHPS® Hospice Survey data from 2,236 hospices examined criterion validity at the quality measure level by assessing the association between hospices' CAHPS® Hospice Survey measure scores and provision of professional staff visits to patients receiving routine hospice care the last two days of life (Teno, Anhang Price et al. 2019), a best practice clinical process.

Citations:

1. Anhang Price R, Quigley DD, Bradley MA, Teno JM, Parast L, Elliott MN, Haas AC, Stucky BD, Mingura BE, & Lorenz K. (2014). Hospice Experience of Care Survey: Development and Field Test. RAND Corporation, RR-657-CMS. Available at: https://hospicecahpssurvey.org/globalassets/hospice-cahps4/home-page/hospice_field_test_report_2014.pdf.
2. Asparouhov, Tihomir, "General Multi-Level Modeling With Sampling Weights," *Communications in Statistics: Theory and Methods*, Vol. 35, No. 3, 2006, pp. 439–460.
3. Brown, Timothy A., *Confirmatory Factor Analysis for Applied Research*, 2nd ed., New York: Guilford Press, 2015.
4. DiStefano, Christine, Jin Liu, Ning Jiang, and Dexin Shi, "Examination of the Weighted Root Mean Square Residual: Evidence for Trustworthiness?" *Structural Equation Modeling*, Vol. 25, No. 3, 2018, pp. 453–466.
5. Flora, David B., and Patrick J. Curran, "An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis with Ordinal Data," *Psychological Methods*, Vol. 9, No. 4, December 2004, pp. 466–491.
6. Hu, Li-tze, and Peter M. Bentler, "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling*, Vol. 6, No. 1, 1999, pp. 1–55.
7. Teno JM, Anhang Price R, Parast L, Haas A, Elliott MN. (2019). More Professional Visits in the Last Days of Life are Associated with Better Hospice Care Experiences. Presentation to the American Academy of Hospice and Palliative Medicine Annual Assembly.
8. Wu, Jiun-Yu, and Oi-man Kwok, "Using SEM to Analyze Complex Survey Data: A Comparison Between Design-Based Single-Level and Model-Based Multilevel Approaches," *Structural Equation Modeling*, Vol. 19, No. 1, 2012, pp. 16–35.

[Response Ends]

2b.03. Provide the statistical results from validity testing.

Examples may include correlations or t-test results.

[Response Begins]

Current Submission

The six-factor model (Hospice Team Communication, Getting Timely Care, Treating Family Member with Respect, Getting Help for Symptoms, Getting Emotional and Religious Support, Care Preferences) provides an excellent fit to the data, $\chi2(120) = 252.83$, $p < .001$; comparative fit index (CFI) = 0.997; root mean square error of approximation (RMSEA) = 0.014; weighted root mean square residual (WRMR) = 1.068. Factor loadings are all above 0.70, suggesting these data elements are strong indicators of the corresponding construct.

Table 2b.1. Individual-level Pearson Correlations Between Other CAHPS Hospice Survey Measures and Global Ratings, 2021 Mode Experiment

Measure	Correlation with Overall Rating of Hospice Care	Correlation with Willingness to Recommend Hospice
Hospice Team Communication	0.61*	0.60*
Care Preferences	0.59*	0.61*
Treating Family Member with Respect	0.53*	0.54*
Getting Timely Care	0.49*	0.47*
Getting Help for Symptoms	0.48*	0.44*
Getting Emotional and Religious Support	0.46*	0.43*
Getting Hospice Care Training [a]	0.43*	0.40*

* $p < 0.001$.

[a] Correlations with the Getting Hospice Care Training measure are calculated only for those decedents who received hospice care at home or in an assisted living facility.

Table 2b.2. Individual-level Pearson Correlations between CAHPS® Hospice Survey Measures, 2021 Mode Experiment

Measure	Hospice Team Communication	Care Preferences	Treating Family Member with Respect	Getting Timely Care	Getting Help for Symptoms	Getting Hospice Care Training
Hospice Team Communication	1	*	*	*	*	*
Care Preferences	0.64[b]	1	*	*	*	*
Treating Family Member with Respect	0.64[b]	0.61[b]	1	*	*	*
Getting Timely Care	0.61[b]	0.47[b]	0.48[b]	1	*	*

Measure	Hospice Team Communication	Care Preferences	Treating Family Member with Respect	Getting Timely Care	Getting Help for Symptoms	Getting Hospice Care Training
Getting Help for Symptoms	0.56[b]	0.47[b]	0.45[b]	0.48[b]	1	*
Getting Hospice Care Training [a]	0.54[b]	0.51[b]	0.43[b]	0.39[b]	0.45[b]	1
Getting Emotional and Religious Support	0.43[b]	0.49[b]	0.44[b]	0.33[b]	0.35[b]	0.37[b]

[a] Correlations with the Getting Hospice Care Training measure are calculated only for those decedents who received hospice care at home or in an assisted living facility.

[b] p<.001

* This cell intentionally left empty.

The analysis of 2016 CAHPS Hospice Survey data found that for all CAHPS Hospice Survey measures with the exception of caregiver training, there were significant positive associations between the proportion of patients receiving staff visits in the last two days of life and hospices' CAHPS Hospice Survey measure performance starting at the sixth decile of visits (84.6% and higher). Family caregivers in hospices in the highest decile of professional staff visits (95.1% and higher) rated the hospice 5.2 points higher on timeliness of care, and between 2.3 and 4.5 points higher on other measures, than caregivers in hospices in the lowest decile (67.5% and lower). These findings suggest that the measures are sensitive to differences in best practice clinical processes, thereby demonstrating their criterion validity.

Previous (2019) Submission

Table 2b1.3a. Individual- and Hospice-level Pearson Correlations Between Multi-Item Measures and Rating of Hospice, Quarter 3 2016 – Quarter 2 2018

Measure	**Rating of Hospice: Individual-Level**	**Rating of Hospice: Hospice-Level**
Hospice Team Communication	0.61*	0.84*
Treating Family Member with Respect	0.54*	0.80*
Getting Timely Care	0.48*	0.78*
Getting Help for Symptoms	0.46*	0.73*
Getting Hospice Care Training	0.45*	0.63*
Getting Emotional and Religious Support	0.42*	0.63*

* p<.001

Table 2b1.3b. Individual- and Hospice-level Pearson Correlations Between Multi-Item Measures and Willingness to Recommend Hospice, Quarter 3 2016 – Quarter 2 2018

Measure	Willingness to Recommend Hospice: Individual-Level	Willingness to Recommend Hospice: Hospice-Level
Hospice Team Communication	0.58*	0.81*
Treating Family Member with Respect	0.53*	0.79*
Getting Timely Care	0.46*	0.75*
Getting Help for Symptoms	0.43*	0.69*
Getting Hospice Care Training	0.43*	0.60*
Getting Emotional and Religious Support	0.41*	0.63*

*p<.001

Table 2b1.3c. Individual-level Pearson Correlations between CAHPS® Hospice Survey Multi-Item Measures, Quarter 3 2016 – Quarter 2 2018

Measure	Hospice Team Communication	Treating Family Member with Respect	Getting Timely Care	Getting Help for Symptoms	Getting Hospice Care Training	Getting Emotional and Religious Support
Hospice Team Communication	1	*	*	*	*	*
Treating Family Member with Respect	0.63[a]	1	*	*	*	*
Getting Timely Care	0.58[a]	0.46[a]	1	*	*	*
Getting Help for Symptoms	0.55[a]	0.43[a]	0.44[a]	1	*	*
Getting Hospice Care Training	0.57[a]	0.41[a]	0.39[a]	0.52[a]	1	*
Getting Emotional and Religious Support	0.43[a]	0.41[a]	0.31[a]	0.33[a]	0.34[a]	1

[a] p<.001

* This cell intentionally left empty.

Table 2b1.3d. Hospice-level Pearson Correlations between CAHPS® Hospice Survey Multi-Item Measures, Quarter 3 2016 – Quarter 2 2018

Measure	Hospice Team Communication	Treating Family Member with Respect	Getting Timely Care	Getting Help for Symptoms	Getting Hospice Care Training	Getting Emotional and Religious Support
Hospice Team Communication	1	*	*	*	*	*
Treating Family Member with Respect	0.84[a]	1	*	*	*	*

Measure	Hospice Team Communication	Treating Family Member with Respect	Getting Timely Care	Getting Help for Symptoms	Getting Hospice Care Training	Getting Emotional and Religious Support
Getting Timely Care	0.82[a]	0.75[a]	1	*	*	*
Getting Help for Symptoms	0.80[a]	0.72[a]	0.74[a]	1	*	*
Getting Hospice Care Training	0.74[a]	0.61[a]	0.63[a]	0.71[a]	1	*
Getting Emotional and Religious Support	0.58[a]	0.62[a]	0.57[a]	0.52[a]	0.42[a]	1

[a] $p<.001$

* This cell intentionally left empty.

[Response Ends]

2b.04. Provide 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?)

[Response Begins]

Current Submission

Applying Cohen's rule of thumb that a correlation of 0.10-0.29 is small, 0.30-0.49 is medium, and 0.50 or above is large, the results in Table 2b.1 demonstrate that three CAHPS Hospice Survey measures exhibit large associations with respondents' rating of hospice at the individual level (*Hospice Team Communication, Care Preferences* and *Treating Family Member with Respect* with correlations of r=0.61, r=0.59, and r=0.53, respectively), while the remaining multi-item measures exhibit medium associations (range of correlations: r = 0.43 to 0.49). The associations of *Willingness to Recommend* and the multi-item measures are similar to those for *Rating of Hospice*, with *Care Preferences* and *Hospice Team Communication* exhibiting the largest relationship to *Willingness to Recommend,* and *Getting Hospice Care Training* the smallest (Table 2b.1). As is usually observed, the associations between *Willingness to Recommend* and other measures are somewhat smaller than those between *Rating of Hospice* and other measures, since *Willingness to Recommend* may be affected by factors such as geographic location of the respondent. These results demonstrate the construct validity of the proposed measures.

Table 2b.2 indicates that CAHPS Hospice Survey measures generally exhibit medium to large intercorrelations, suggesting that they measure unique but related constructs. Intercorrelations are largest between *Hospice Team Communication and both Care Preferences* and *Treating Family Member with Respect* (intercorrelations of r = 0.64 for both) because these measures assess aspects of communication. The smaller intercorrelations between some of the measures (e.g., individual-level correlation between *Getting Timely Care* and *Getting Emotional and Religious Support*, r = 0.33) provide some evidence of discriminant validity (i.e., that measures that we would not expect to be substantially related are in fact not correlated very highly).

Previous (2019) Submission

Applying Cohen's rule of thumb that a correlation of 0.10-0.29 is small, 0.30-0.49 is medium, and 0.50 or above is large, the results in Table 2b1.3a demonstrate that two CAHPS® Hospice Survey multi-item measures exhibit large associations with respondents' rating of hospice at the individual level (*Hospice Team Communication* and *Treating Family Member with Respect* with individual-level correlations of r=0.61 and r=0.54, respectively), while the remaining multi-item measures exhibit medium associations at the individual level (range of individual-level correlations: r = 0.42 to 0.48); at the hospice level, all CAHPS® Hospice Survey multi-item measures exhibit large associations to *Rating of Hospice* (range of hospice-level correlations: r = 0.63 to 0.84). The associations of *Willingness to Recommend* and the multi-item measures are similar to those for *Rating of Hospice*, with *Hospice Team Communication* exhibiting the largest relationship to

Willingness to Recommend, and *Getting Help for Symptoms, Getting Hospice Care Training* and *Getting Emotional and Spiritual Support* the smallest (Table 2b1.3b). As is usually observed, the associations between composites and *Willingness to Recommend* are somewhat smaller than those between composites and *Rating of Hospice*, since *Willingness to Recommend* may be affected by factors such as geographic location of the respondent. These results demonstrate the construct validity of the proposed measures.

Table 2b1.3c indicates that CAHPS® Hospice Survey measures generally exhibit medium to large intercorrelations, suggesting that they measure unique but related constructs. Intercorrelations are largest between *Hospice Team Communication* and *Treating Family Member with Respect* (individual-level correlation of r = 0.63) because these two measures assess forms of communication. The smaller intercorrelations between some of the multi-item measures (e.g., individual-level correlation between *Getting Timely Care* and *Getting Emotional and Religious Support*, r = 0.31) provide some evidence of discriminant validity (i.e., that measures that we would not expect to be substantially related are in fact not correlated very highly).

[Response Ends]

2b.05. 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 in Importance to Measure and Report: Gap in Care/Disparities.

[Response Begins]

To examine the ability of CAHPS® Hospice Survey measures to identify high or low performing hospices, we calculated the number and percentage of hospices that were significantly above or below the mode experiment hospice average for each measure. (In the previous [2019] submission, we pursued the same approach in comparison to the national hospice average.) All scores were adjusted for mode and case mix and were scored using top-box scoring. A two-sided alpha=0.05 level test was used to test for significance.

[Response Ends]

2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.

Examples may include 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.

[Response Begins]

Current Submission

Table 2b.1. Number of Hospices Significantly Above or Below the 2021 Mode Experiment Hospice Average

Measures	Count of Hospices Significantly Above	Count of Hospices Significantly Below	% Statistically Different from Mode Experiment Hospice Average*
Hospice Team Communication	5	7	22%

Measures	Count of Hospices Significantly Above	Count of Hospices Significantly Below	% Statistically Different from Mode Experiment Hospice Average*
Care Preferences	3	6	17%
Getting Timely Care	7	6	24%
Treating Your Family Member with Respect	1	6	13%
Getting Help with Symptoms	5	6	20%
Getting Emotional and Religious Support	6	6	22%
Getting Hospice Care Training	5	4	17%
Rating of Hospice	3	6	17%
Willingness to Recommend	6	8	26%

* Restricted to hospices that received a score for the respective measure.

Previous (2019) Submission

Table 2b4.2. Number of Hospices Significantly Above or Below the National Hospice Average, Q3 2016 – Q2 2018

Hospice Multi-Item Measure/Items	Count of Hospices Significantly Above	Count of Hospices Significantly Below	% Statistically Different from Hospice National Average*
Hospice Team Communication	523	547	36%
Getting Timely Care	486	516	34%
Treating Your Family Member with Respect	325	412	25%
Getting Help with Symptoms	409	482	30%
Getting Emotional and Religious Support	394	386	27%
Getting Hospice Care Training	350	404	26%
Rating of Hospice	465	452	31%
Willingness to Recommend	574	513	37%

* Restricted to hospices that received a score for the respective measure.

[Response Ends]

2b.07. Provide 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.

In other words, what do the results mean in terms of statistical and meaningful differences?

[Response Begins]

Current Submission

Table 2b.1 indicates that CAHPS Hospice Survey measures can discriminate between hospices' performance as compared to an average of participants in an experiment. Across measures, between 13 percent and 26 percent of hospices participating in the mode experiment scored either significantly above or below the average of the experiment's participating hospices. Among hospice scores that were significantly above or below the average, the mean absolute difference between the hospices' scores and the mode experiment average hospice score for a given measure was 10.1 points on a 0 to 100 scale. The range of the mean absolute differences was 5.8 to 13.8 across the nine measures. Prior studies have suggested that differences of 1, 3, and 5 points on patient experience measures may be interpreted as small, medium, and large, respectively (Quigley et al., 2018). Thus, the mean absolute differences reported in Table 2b.1 correspond to large differences from mean hospice performance.

Citation:

1. Quigley DD, Elliott MN, Setodji CM, Hays RD. Quantifying Magnitude of Group-Level Differences in Patient Experiences with Health Care. *Health Serv Res* 2018;**53**(4):3027-3051.

Previous (2019) Submission

Table 2b4.2 indicates that CAHPS® Hospice Survey measures can discriminate between hospices' performance as compared to a national average. For each of multi-item and global measures, one quarter or more of all hospices score either significantly above or below the national average. Among hospice scores that were significantly above or below the average, the mean absolute difference between the hospices' scores and the national average hospice score for a given measure was 8.0 points on a 0 to 100 scale. The range of the absolute differences was 5.3 to 10.1 across the eight measures. For reference, analyses of the initial two quarters of national CAHPS® Hospice Survey data indicate that a hospice that scores 10 points lower Hospice Team Communication can expect to be definitely recommended eight percent less often (Anhang Price et al. 2018).

Citation:

1. Anhang Price R, Stucky B, Parast L, Elliott MN, Haas A, Bradley M, Teno JM. (2018). Development of Valid and Reliable Measures of Patient and Family Experiences of Hospice Care for Public Reporting. *Journal of Palliative Medicine*.

[Response Ends]

2b.08. Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.

Describe the steps—do not just name a method; what statistical analysis was used.

[Response Begins]

Survey Non-Response

The *CAHPS® Hospice Survey Quality Assurance Guidelines* outlines the survey response rate calculation, as follows:

Response Rate = (Total Number of Completed Surveys) / (Total Number of Surveys Fielded – Total Number of Ineligible Surveys)

The Total Number of Completed Surveys is the total number of surveys for which the caregiver respondent answers at least 50 percent of the questions applicable to all decedents/caregivers. The Total Number of Surveys Fielded is the total sample, and Total Number of Ineligible Surveys is the total number of surveys for which it is determined that the decedent/caregiver did not meet the eligibility criteria outlined above in Section 1.6; the respondent has a language barrier, mental/physical incapacity, or is institutionalized or deceased; or the respondent indicates that they were never involved in decedent care.

The following are **not** removed from the denominator of the response rate calculation: break-off surveys, refusals, non-response after maximum attempts at data collection, bad or no address or telephone number, incomplete caregiver or decedent name, and respondents' disavowal that the decedent received care from any hospice or the named hospice.

The *CAHPS® Hospice Survey Quality Assurance Guidelines* provide advice for maximizing survey response rates, including:

- For Mail Only and Mixed Mode survey administration, vendors must perform address updates for missing or incorrect information, including working with client hospices to obtain the most current caregiver contact information, using the National Change of Address and the United States Postal Service CASS Certified Zip+4 software, and using other means such as commercial software and internet search options to locate current addresses.
- For Telephone Only and Mixed Mode survey administration, vendors must use commercial software or other means to update telephone numbers provided by the hospice for sampled caregivers. This includes running update program software against the sample file just before or after uploading data to survey management systems, utilizing commercial software, Internet directories and/or directory assistance, and contacting the hospice to request updated telephone numbers.
- For Mail Only and Mixed Mode survey administration, send all mailings with first class postage or indicia to ensure delivery in a timely manner and to maximize response rates, as first class mail is more likely to be opened.
- For Telephone Only and Mixed Mode survey administration, make telephone attempts at various times of the day, on different days of the week, and in different weeks to maximize the probability that the survey vendor will contact the caregiver.
- Survey vendors should make every reasonable effort to achieve optimal telephone response rates, such as thoroughly familiarizing interviewers with the study purpose, carefully supervising interviewers, retraining those interviewers having difficulty enlisting cooperation, and re- contacting reluctant respondents at different times until the data collection protocol is completed.

We assessed the association between survey nonresponse and several caregiver and decedent characteristics, including relationship of caregiver to the decedent, and decedent age at death, sex, race/ethnicity, payer for hospice care, final setting of care, length of final episode of hospice care, and primary diagnosis.

Item Non-Response

In 2b.09, we present nonresponse to evaluative items among unit respondents. Specifically, we report the proportion of respondents that skipped each item *appropriately* (i.e., dictated by the survey's skip logic instructions), *inappropriately* (i.e., *not* dictated by the survey's skip logic instructions), as well as the total proportion of missing data for each evaluative item on the survey.

[Response Ends]

2b.09. Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.

For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).

[Response Begins]

Current submission

The estimated response rate for the revised CAHPS Hospice Survey, administered during the 2021 mode experiment, was: 31.5 percent in Telephone Only mode, 35.1 percent in Mail Only mode, 39.7 percent in Web-Mail mode, and 45.3 percent in Mixed Mode (mail with telephone follow-up). (Note that across the 56 participating hospices over the six months of the mode experiment, 31.4 percent of caregivers had available email addresses. The response rate to Web-Mail mode among those *without* an available email address was 35.2 percent; the response rate to that mode among those *with* an available email address was 49.6 percent.)

In addition to survey mode, the factors most strongly associated with the odds of response were the caregiver's relationship to the decedent; decedent age, with caregivers of older decedents more likely to respond than younger decedents; and race/ethnicity, with caregivers of black and Hispanic decedents less likely to respond than caregivers of white decedents. In addition, caregivers of decedents whose final setting of care was a nursing home or acute care hospital were less likely to respond than those whose final setting was at home. In addition, caregivers of decedents with shorter final episodes were less likely to respond than those with longer episodes.

Table 2b.2. CAHPS® Hospice Survey Missing Data, Mode Experiment 2021

CAHPS® Hospice Survey Evaluative Item	**% Missing due to Appropriate Skip**	**% Missing due to Inappropriate Skip**	**% Missing (Total)**
Hospice Team Communication	*	*	*
How often did the hospice team let you know when they would arrive to care for your family member? (Q6)	0.0	3.0	3.0
How often did the hospice team explain things in a way that was easy to understand? (Q8)	0.0	0.5	0.5
How often did the hospice team listen carefully to you when you talked with them about problems with your family member's hospice care? (Q15)	54.4	2.3	56.7
How often did the hospice team keep you informed about your family member's condition? (Q9)	0.0	0.8	0.8
While your family member was in hospice care, how often did the hospice team listen carefully to you? (Q25)	0.0	0.8	0.8
Getting Timely Care	*	*	*
When you or your family member asked for help from the hospice team, how often did you get help as soon as you needed it? (Q7)	0.0	1.8	1.8
How often did you get the help you needed from the hospice team during evenings, weekends, or holidays? (Q5)	32.1	1.6	33.7
Treating Family Member with Respect	*	*	*
How often did the hospice team treat your family member with dignity and respect? (Q10)	0.0	0.6	0.6
How often did you feel that the hospice team really cared about your family member? (Q11)	0.0	0.5	0.5

CAHPS® Hospice Survey Evaluative Item	% Missing due to Appropriate Skip	% Missing due to Inappropriate Skip	% Missing (Total)
Support for Emotional and Religious Beliefs	*	*	*
While your family member was in hospice care, how much emotional support did you get from the hospice team? (Q28)	0.0	1.7	1.7
In the weeks after your family member died, how much emotional support did you get from the hospice team? (Q29)	0.0	3.0	3.0
Support for religious or spiritual beliefs includes talking, praying, quiet time, or other ways of meeting your religious or spiritual needs. While your family member was in hospice care, how much support for your religious and spiritual beliefs did you get from the hospice team? (Q27)	0.0	5.0	5.0
Getting Help for Symptoms	*	*	*
Did your family member get as much help with pain as he or she needed? (Q17)	27.5	2.2	29.7
How often did your family member get the help he or she needed for trouble breathing? (Q19)	49.0	2.2	51.2
How often did your family member get the help he or she needed for trouble with constipation? (Q21)	56.7	4.1	60.9
How often did your family member get the help he or she needed from the hospice team for feelings of anxiety or sadness? (Q23)	41.4	3.2	44.7
Getting Hospice Care Training	*	*	*
Hospice teams may teach you how to care for family members who need pain medicine, have trouble breathing, are restless or agitated, or have other care needs. Did the hospice team teach you how to care for your family member? (Q24)[a]	17.8	1.1	18.8
Care preferences	*	*	*
Did the hospice team make an effort to listen to the things that mattered most to you or your family member? (Q13)	0.0	0.8	0.8
Did the hospice team provide care that respected your family member's wishes? (Q12)	0.0	1.1	1.1
Rating of Hospice	*	*	*

CAHPS® Hospice Survey Evaluative Item	% Missing due to Appropriate Skip	% Missing due to Inappropriate Skip	% Missing (Total)
Please answer the following questions about the hospice named on the survey cover. Do not include care from other hospices in your answers. Using any number from 0 to 10, where 0 is the worst hospice care possible and 10 is the best hospice care possible, what number would you use to rate your family member's hospice care? (Q30)	0.0	1.1	1.1
Willingness to Recommend	*	*	*
Would you recommend this hospice to your friends and family? (Q31)	0.0	0.9	0.9

[a] Rates calculated restricting the denominator to those respondents who reported that the decedent received care at home or in an assisted living facility.

* This cell intentionally left empty.

Previous (2019) submission

The overall CAHPS® Hospice Survey response rate in Quarter 3 2016 through Quarter 2 2018 was 32.63 percent. The vast majority of surveys (95.5 percent) were administered via Mail Only administration, with a response rate of 32.6 percent. Response rates were 26.3 percent for Telephone Only Mode and 40.9 percent for Mixed Mode (mail with telephone follow-up). In addition to survey mode, the factors most strongly associated with the odds of response were the caregiver's relationship to the decedent; decedent age, with caregivers of older decedents more likely to respond than younger decedents; and race/ethnicity, with caregivers of white decedents having twice the odds of responding as caregivers of black and Hispanic decedents. In addition, caregivers of decedents with Medicaid were substantially less likely to respond than those with Medicare, and caregivers of decedents whose final setting of care was a nursing home, acute care hospital or other setting were less likely to respond than those whose final setting was at home. In addition, caregivers of decedents with shorter final episodes were less likely to respond than those with longer episodes.

Table 2b6.2a. CAHPS® Hospice Survey Missing Data, Quarter 3 2016 – Quarter 2 2018

CAHPS® Hospice Survey Evaluative Item	% Missing due to Appropriate Skip	% Missing due to Inappropriate Skip	% Missing (Total)
Hospice Team Communication	*	*	*
While your family member was in hospice care, how often did the hospice team keep you informed about when they would arrive to care for your family member? (Q6)	0.0	1.7	1.7
While your family member was in hospice care, how often did the hospice team explain things in a way that was easy to understand? (Q8)	0.0	0.7	0.7
How often did the hospice team listen carefully to you when you talked with them about problems with your family member's hospice care? (Q14)	58.2	2.5	60.7

CAHPS® Hospice Survey Evaluative Item	% Missing due to Appropriate Skip	% Missing due to Inappropriate Skip	% Missing (Total)
While your family member was in hospice care, how often did the hospice team keep you informed about your family member’s condition? (Q9)	0.0	0.9	0.9
While your family member was in hospice care, how often did the hospice team listen carefully to you? (Q35)	0.0	1.9	1.9
While your family member was in hospice care, how often did anyone from the hospice team give you confusing or contradictory information about your family member’s condition or care? (Q10)	0.0	1.0	1.0
Getting Timely Care	*	*	*
While your family member was in hospice care, when you or your family member asked for help from the hospice team, how often did you get help as soon as you needed it? (Q7)	0.0	1.7	1.7
How often did you get the help you needed from the hospice team during evenings, weekends, or holidays? (Q5)	32.7	2.0	34.7
Treating Family Member with Respect	*	*	*
While your family member was in hospice care, how often did the hospice team treat your family member with dignity and respect? (Q11)	0.0	0.5	0.5
While your family member was in hospice care, how often did you feel that the hospice team really cared about your family member? (Q12)	0.0	0.6	0.6
Support for Emotional and Religious Beliefs	*	*	*
While your family member was in hospice care, how much emotional support did you get from the hospice team? (Q37)	0.0	2.6	2.6
In the weeks after your family member died, how much emotional support did you get from the hospice team? (Q38)	0.0	4.6	4.6
Support for religious or spiritual beliefs includes talking, praying, quiet time, or other ways of meeting your religious or spiritual needs. While your family member was in hospice care, how much support for your religious and spiritual beliefs did you get from the hospice team? (Q36)	0.0	4.8	4.8
Getting Help for Symptoms	*	*	*

CAHPS® Hospice Survey Evaluative Item	% Missing due to Appropriate Skip	% Missing due to Inappropriate Skip	% Missing (Total)
Did your family member get as much help with pain as he or she needed? (Q16)	27.1	2.1	29.2
How often did your family member get the help he or she needed for trouble breathing? (Q22)	47.3	2.7	50.0
How often did your family member get the help he or she needed for trouble with constipation? (Q25)	56.4	3.5	59.9
How often did your family member receive the help he or she needed from the hospice team for feelings of anxiety or sadness? (Q27)	42.6	3.5	46.1
Getting Hospice Care Training	*	*	*
Did the hospice team give you enough training about what side effects to watch for from pain medicine? (Q19)	9.4	2.5	12.0
Did the hospice team give you the training you needed about if and when to give more pain medicine to your family member? (Q20)	21.6	2.4	24.0
Did the hospice team give you the training you needed about how to help your family member if he or she had trouble breathing? (Q23)	52.4	2.3	54.7
Did the hospice team give you the training you needed about what to do if your family member became restless or agitated? (Q29)	35.1	2.6	37.7
Side effects of pain medicine include things like sleepiness. Did any member of the hospice team discuss side effects of pain medicine with you or your family member? (Q18)	9.4	2.1	11.5
Rating of Hospice	*	*	*
Using any number from 0 to 10, where 0 is the worst hospice care possible and 10 is the best hospice care possible, what number would you use to rate your family member's hospice care? (Q39)	0.0	1.4	1.4
Willingness to Recommend	*	*	*
Would you recommend this hospice to your friends and family? (Q40)	0.0	1.3	1.3

[a] Rates calculated restricting the denominator to those respondents who reported that the decedent received care at home or in an assisted living facility.

* This cell intentionally left empty.

[Response Ends]

2b.10. Provide 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 non-responders), and how the specified handling of missing data minimizes bias.

In other words, 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 was conducted, justify the selected approach for missing data.

[Response Begins]

Current submission

Unit response rates observed during the CAHPS Hospice Survey 2021 mode experiment are better than those observed in other facility-level experience of care surveys; for reference, national mean hospital-level response rate for HCAHPS was 25 percent during a similar time period. As noted in the previous (2019) submission, although our analyses indicate that response propensity varies by certain caregiver and decedent characteristics, previous work in other CAHPS settings has demonstrated that nonresponse weighting to account for potential bias is not needed after case-mix adjustment (see, for example, Elliott, Edwards et al. 2005 and Elliott, Zaslavsky et al. 2009). When case-mix adjustment suffices to address nonresponse bias, it generally does so with greater statistical efficiency than nonresponse weighting, resulting in estimates of equal reliability and precision with smaller sample sizes than would be required with nonresponse weighting.

Across evaluative items, 5 percent or fewer respondents inappropriately skipped items. Item missingness tended to be higher for respondents for whom the survey was longer (i.e., respondents who were eligible for more survey items due to skip logic; data not shown). This finding suggests that it is unlikely that CAHPS® Hospice Survey item results are biased due to systematic skipping of items by respondents. The CAHPS® Hospice Survey employs skip logic to promote appropriate skipping among respondents who are not qualified to answer an item (Rodriguez et al., 2009).

Previous (2019) submission

Unit response rates for the CAHPS® Hospice Survey are comparable to those observed in other facility-level experience of care surveys; for reference, national mean hospital-level response rate for HCAHPS was 27 percent during a similar time period (HCAHPS Online). Although our analyses indicate that response propensity varies by certain caregiver and decedent characteristics, previous work in other CAHPS settings has demonstrated that nonresponse weighting to account for potential bias is not needed after case-mix adjustment (see, for example, Elliott, Edwards et al. 2005 and Elliott, Zaslavsky et al. 2009). When case-mix adjustment suffices to address nonresponse bias, it generally does so with greater statistical efficiency than nonresponse weighting, resulting in estimates of equal reliability and precision with smaller sample sizes than would be required with nonresponse weighting.

Across evaluative items, less than 5 percent of respondents inappropriately skipped items. Item missingness tended to be higher for respondents for whom the survey was longer (i.e., respondents who were eligible for more survey items due to skip logic; data not shown). This finding suggests that it is unlikely that CAHPS® Hospice Survey item results are biased due to systematic skipping of items by respondents. The CAHPS® Hospice Survey employs skip logic to promote appropriate skipping among respondents who are not qualified to answer an item (Rodriguez et al., 2009).

Citations:

1. Centers for Medicare & Medicaid Services. Summary of HCAHPS Survey Results, July 2017 to June 2018 Discharges. Online at: https://www.hcahpsonline.org/en/summary-analyses/.
2. Elliott MN, Edwards C, Angeles J, Hays RD (2005). "Patterns of unit and item non-response in the CAHPS® Hospital Survey." *Hlth Serv Res* 40(6): 2096-2119.
3. Elliott MN, Zaslavsky AM, Goldstein E, Lehrman W, Hambarsoomian K, Beckett MK, Giordano L (2009). "Effects of survey mode, patient mix, and nonresponse on CAHPS Hospital Survey scores." *Hlth Serv Res* 44(2): 501-508.
4. Rodriguez HP, Glahn TV, Li A, Rogers WH, Safran DG. The effect of item screeners on the quality of patient survey data: a randomized experiment of ambulatory care experience measures. *Patient*. 2009 Jun 1; 2(2):135-41.

[Response Ends]

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 eCQMs). 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.

2b.11. Indicate whether there is more than one set of specifications for this measure.

[Response Begins]

Yes, there is more than one set of specifications for this measure

[Response Ends]

2b.12. 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. Indicate what statistical analysis was used.

[Response Begins]

Current submission

Currently, hospices participating in national implementation of the CAHPS® Hospice Survey may choose from one of three modes of survey administration: Mail Only, Telephone Only, or Mixed Mode (mail with telephone follow-up). To assess the effect of these modes, as well as a Web-Mail mode, on response rates and response patterns, and determine whether survey mode adjustments are needed to fairly compare CAHPS® Hospice Survey scores across hospices using different modes of administration, CMS conducted the 2021 mode experiment, a randomized experiment in which 15,515 decedent/caregiver pairs were sampled from 56 large hospice programs over six months. Each hospice provided a monthly sample frame of their decedents/caregivers, and we took a random sample of decedents/caregivers who met the eligibility criteria for national implementation of the CAHPS® Hospice Survey. Within each hospice, we randomly assigned cases to receive a revised version of the survey in each of four modes of data collection: Mail Only, Telephone Only, Mixed Mode (mail with telephone follow-up), and Web-Mail; an additional group received the current version of the survey in the Mail Only mode. One vendor collected survey data for all hospices using a standardized data collection protocol. To evaluate possible survey mode effects, we conducted linear regression analysis predicting each of the CAHPS Hospice Survey outcomes (on a 0-100 scale) from survey mode (Mail Only, Revised Survey as reference category), case-mix adjustors, hospice indicators, and month of death, scoring the outcomes using the top box approach. Because survey mode is selected at the hospice level, we translated these effect sizes to hospice-level standard deviations.

Previous (2019) submission

Hospices participating in national implementation of the CAHPS® Hospice Survey may choose from one of three modes of survey administration: Mail Only, Telephone Only, or Mixed Mode (mail with telephone follow-up). To assess the effect of mode on response rates and response patterns, and determine whether survey mode adjustments were needed to fairly compare CAHPS® Hospice Survey scores across hospices using different modes of administration, we conducted a randomized mode experiment, sampling 17,121 decedent/caregiver pairs from 59 large hospice programs in Quarter 1 and Quarter 2 2015 (Parast et al. 2018). Each hospice provided a sample of decedents/caregivers who met the eligibility criteria for national implementation of the CAHPS® Hospice Survey. Within each hospice, we randomly assigned onethird of cases to each of the three modes of data collection. One vendor collected survey data for all hospices using a standardized data collection protocol. To evaluate possible survey mode effects, we conducted linear regression analysis predicting each of the CAHPS® Hospice Survey outcomes (on a 0-100 scale) from survey mode (Mail Only as reference category), case-mix adjustors and hospice indicators, scoring the outcomes using the top box approach.

Citation:

1. Parast L, Elliott MN, Hambarsoomian K, Teno JM, Anhang Price R (2018). Effects of Survey Mode on Consumer Assessment of Healthcare Providers and Systems (CAHPS) Hospice Survey Scores. *J Am Geriatr Soc.* 66(3):546-552.

[Response Ends]

2b.13. Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.

Examples may include correlation, and/or rank order.

[Response Begins]

Current submission

We found significant effects of survey mode on responses to several survey outcomes. Telephone Only respondents were significantly more likely to report worse experiences than Mail Only respondents for four items and were significantly more likely to report better experiences for two items. Mixed Mode (Mail-Telephone) and Web-Mail respondents each reported significantly worse experiences than Mail Only respondents for one item. Mode effects were sometimes large. For example, in terms of hospice-level standard deviations (SDs), the six significant Telephone Only mode effects ranged from 3.4 SDs lower to 1.2 SDs higher than scores from Mail Only.

Previous (2019) submission

We found significant effects of survey mode on responses to several survey outcomes. Telephone Only respondents reported significantly worse experiences of care than Mail Only respondents for 11 of the 24 evaluative items composing our measures ($p<0.05$); there were no outcomes for which the reverse was true. For example, Telephone Only respondents tended to give lower hospice ratings (regression coefficient for Telephone Only = -3.93 compared to Mail Only, $p<0.01$). Telephone Only respondents were also significantly more likely to report worse experiences than Mixed Mode respondents for 6 of these 11 items.

Citation:

1. Parast L, Elliott MN, Hambarsoomian K, Teno JM, Anhang Price R (2018). Effects of Survey Mode on Consumer Assessment of Healthcare Providers and Systems (CAHPS) Hospice Survey Scores. *J Am Geriatr Soc.* 66(3):546-552.

[Response Ends]

2b.14. Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.

In other words, what do the results mean and what are the norms for the test conducted.

[Response Begins]

Current and Previous (2019) Submission

To ensure fair comparisons across hospices, CAHPS® Hospice Survey scores must be adjusted for mode of survey administration, which affects scores but is not related to quality of hospice care.

[Response Ends]

2b.15. Indicate whether the measure uses exclusions.

[Response Begins]

Yes, the measure uses exclusions.

[Response Ends]

2b.16. Describe the method of testing exclusions and what was tested.

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?

[Response Begins]

Current and Previous (2019) submission

The CAHPS Hospice Survey is intended to assess the care experiences of patients (and their family caregivers) of all payer types who die while under the care of hospice. In keeping with CMS requirements for national implementation of the CAHPS Hospice Survey, decedents/caregivers were eligible for inclusion in the CAHPS Hospice Survey 2021 mode experiment if:

- Decedents were age 18 or over (to ensure that results reflect adults only, as pediatric patients may have different care needs)
- Decedents died at least 48 hours following last admission to hospice care (to ensure that the respondent had sufficient time to observe hospice care, and that the hospice had sufficient time to provide required services)
- Decedents had a caregiver of record (to allow for identification of a survey respondent)
- Decedents had a caregiver with a U.S. or U.S. Territory home address (to allow for survey administration within the U.S. or U.S. Territory)
- Decedents had a caregiver other than a nonfamilial legal guardian (to allow for identification of an informal caregiver respondent).

Decedents or caregivers who are otherwise eligible for the CAHPS® Hospice Survey are excluded from the sample frame if they have “no publicity” status. “No publicity” decedents/caregivers are those who initiate or voluntarily request at any time during their stay that the hospice: 1) not reveal the patient’s identity; and/or 2) not survey him or her. Hospices must retain documentation of the “no publicity” request for a minimum of three years.

CAHPS® Hospice Survey eligibility and exclusion criteria are designed to honor the direct requests of individuals who do not wish to be contacted, and to ensure identification of informal caregiver respondents with sufficient experience with hospice care to assess its quality. They are not based on statistical testing.

[Response Ends]

2b.17. Provide 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.

[Response Begins]

N/A

[Response Ends]

2b.18. Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.

In other words, 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.

[Response Begins]

N/A

[Response Ends]

2b.19. Check all methods used to address risk factors.

[Response Begins]

Statistical risk model with risk factors (specify number of risk factors)

[Statistical risk model with risk factors (specify number of risk factors) Please Explain]

9. (Please see 2b.20 for a complete list of the 9 risk factors included in the model.)

[Response Ends]

2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.

[Response Begins]

Although there are many advantages to mode experiment data – among others, randomization is needed to generate unbiased estimates of mode effects – mode experiment data reflect experiences with a smaller number of hospices, contain too few respondents to evaluate case-mix adjustment, and may exhibit smaller variation in decedent and caregiver characteristics between hospices than observed in national implementation data. Therefore, 2021 mode experiment data were case-mix adjusted using the approach developed from and applied to national implementation data, as described in the previous (2019) submission.

To ensure that comparisons between hospices reflect differences in performance rather than differences in patient and/or caregiver characteristics, CAHPS Hospice Survey responses are adjusted for “case mix” (i.e., variations of such characteristics across hospices). The case-mix adjustment model includes the following variables:

- response percentile (calculated by ranking lag time—that is, days between death and survey response—among respondents for each hospice in each quarter, then dividing by total sample size)
- decedent age (categories: 18-54; 55-64; 65-69; 70-74; 75-79; 80-84; 85-89; 90+)
- payer for hospice care (categories: Medicare only; Medicaid only or Medicaid and private insurance; Medicare and Medicaid; Private insurance only; Medicare and private insurance; Other)
- primary diagnosis (categories: Alzheimer’s and non-Alzheimer’s dementias; Bladder cancer; Blood and lymphatic cancers; Brain cancer; Breast cancer; Congestive heart failure, etc.)
- length of final episode of hospice care (categories: 2-5 days; 6-12 days; 13-29 days; 30-80 days; 81+ days)
- respondent age (categories: 18-44; 45-54; 55-64; 65-74; 75-84; 85+)
- respondent education (categories: 8th grade or less; Some high school; High school graduate or GED; Some college; 4-year college graduate; More than 4-year college graduate)
- relationship of decedent to caregiver (categories: Spouse or partner; Parent; Mother-in-law or father-in-law; Aunt or uncle; Sister or brother; Child; Friend; Other)
- language (categories: Survey language or home language was Spanish; All others)

The CAHPS® Hospice Survey website (http://www.hospicecahpssurvey.org/en/scoring-and-analysis) presents more information regarding case-mix adjustment, including the case-mix adjustment factors for each item in CAHPS® Hospice Survey measures for each quarter.

Case-mix adjustment is performed within each quarter of CAHPS® Hospice Survey data after data cleaning and adjustment for mode of survey administration. *Coefficients* obtained in linear regression models estimate the tendency of caregivers to respond more positively or negatively. The *adjustments* needed to counter that tendency are obtained by multiplying the case-mix coefficients by (-1.0). Adjustments are updated and posted quarterly on the CAHPS® Hospice Survey website (http://www.hospicecahpssurvey.org/en/scoring-and-analysis/) for each question composing a publicly reported CAHPS® Hospice Survey measure score.

Publicly reported CAHPS® Hospice Survey measure scores are adjusted to the overall national mean of case-mix variables across all reporting hospices (overall national means are also updated and posted quarterly on the CAHPS® Hospice Survey website). Thus, whether the scores of a given hospice are adjusted upward or downward for a given measure depends not only on these case-mix adjustments, but also on the case mix of that hospice relative to the national average of these case-mix characteristics. Specifically, the total case mix-adjustment for a given hospice is the sum of a series of products, where each product multiplies the adjustments by the difference between the hospice's mean on the corresponding case-mix variable and the national mean on that case-mix variable.

Four sets of numbers are needed to calculate final case-mix adjusted top- or bottom-box scores for a given hospice for a given quarter: (1) Mean top- or bottom-box scores of the items that compose each CAHPS® Hospice Survey measure for the hospice in question that have been adjusted for survey mode; (2) item-level case-mix adjustments; (3) that hospice's means on case-mix variables; and (4) national means on case-mix variables.

The formula for applying case-mix adjustment is as follows:

- Let y be the mode-adjusted hospice mean of an item that composes a CAHPS® Hospice Survey measure
- Let a1-a54 be the corresponding individual-level adjustments (note that though there are 9 risk factors, most risk factors involve multiple categories e.g. decedent age is 18-54, 55-64, 65-69 etc. thus resulting in 54 non-reference categories coded as indicators in the models)
- Let m1-m54 be the national means for the CMA variables
- Let h1-h54 be the CMA means for the hospice in question
- Then $y'=y+a1(h1-m1)+a2(h2-m2)+\ldots+a54(h54-m54)$ is the case-mix and mode-adjusted hospice score for that item

[Response Ends]

2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.

[Response Begins]

[Response Ends]

2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.

[Response Begins]

Published literature

Internal data analysis

[Response Ends]

2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.

Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10 or other statistical tests; correlation of x or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).

[Response Begins]

Previous research, on both CAHPS surveys and other types of surveys, has identified respondent characteristics that are not under the control of the entities being assessed but tend to be related to survey responses. For example, individuals who are older, those with less education, and those in better overall and mental health generally tend to give more positive ratings and reports of care in Medicare CAHPS (Elliott, Swartz et al. 2001; Zaslavsky, Zaborski et al. 2001; Elliott, Zaslavsky et al. 2009). Hence, entities with disproportionate numbers of respondents with such characteristics (favorable case mix) are advantaged relative to those with a less favorable case mix. To ensure that comparisons between hospices reflect differences in performance rather than differences in patient and/or caregiver characteristics, hospice scores should be adjusted for variations of such characteristics across hospices.

Details regarding the analyses that underpin the CAHPS Hospice Survey case-mix adjustment model were included in the measures' first NQF endorsement in 2016 and described by Parast et al. (2018). Briefly, we identified patient and caregiver characteristics as candidates for case-mix adjustment if they were available in hospice administrative or survey response data and were not within the hospice's control. For each potential case-mix adjustor, we examined (a) variation among hospices using intraclass correlation coefficients (ICCs), (b) bivariate and multivariate association with selected CAHPS Hospice Survey outcomes, (c) impact on adjustment, and (d) appropriate parameterization of adjustors. We selected outcomes for assessment that had substantial variation across hospices, different response scales, and addressed conceptually distinct aspects of hospice care experiences.

There is no ordering of risk factor inclusion in the linear regression model used to create case mix adjustments for the CAHPS® Hospice Survey items.

Citations:

1. Elliott MN, Swartz R, Adams J, Spritzer KL and Hays R (2001). "Case-mix adjustment of the National CAHPS® Benchmarking Data 1.0: A violation of model assumptions?" *Hlth Serv Res* 36(3): 555-574.
2. Elliott MN, Zaslavsky AM, Goldstein E, Lehrman W, Hambarsoomian K, Beckett MK and Giordano L (2009). "Effects of survey mode, patient mix, and nonresponse on CAHPS Hospital Survey scores." *Hlth Serv Res* 44(2): 501-508.
3. Parast L, Haas A, Tolpadi A, Elliott MN, Teno JM, Zaslavsky AM, Anhang Price R (2018). Effects of caregiver and decedent characteristics on CAHPS® Hospice Survey scores. *J Pain Symptom Manage,* 56(4):519-529.
4. Zaslavsky AM, Zaborski LB, Ding L, Shaul JA, Cioffi MJ, Cleary PD (2001). Adjusting performance measures to ensure equitable plan comparisons. *Health Care Fin Rev*; 22(3): 109-126.

[Response Ends]

2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.

[Response Begins]

Details regarding the statistical results of our analyses to select risk factors (i.e., select the CAHPS Hospice Survey case-mix adjustment model) were included in the measures' first NQF endorsement in 2016 and described by Parast et al. (2018). We briefly describe those results here. All tested characteristics varied moderately or substantially across hospices as measured by ICC> 0.01 except for decedent gender, caregiver gender, and caregiver age. When examining bivariate associations, all examined characteristics, except for caregiver gender and decedent gender, were significantly (p<0.001) and strongly (regression coefficient greater than 5 points on a 0-100 scale) associated with respondent assessments of at least one outcome. Both caregiver and decedent education were significantly and strongly predictive of the outcomes. These two variables were moderately correlated with one another (R=0.38); the coefficients for caregiver education were generally larger, indicating a stronger adjusted relationship with the outcomes. Because of the need to limit survey length in current and future versions of the survey, only caregiver education was retained in the model. Thus, multivariate models included all characteristics except for caregiver gender, decedent gender, and decedent education. In these multivariate models, each characteristic was significantly associated with at least one of the outcomes (p<0.001). For example, caregivers of decedents that had Medicaid as the payer for hospice care tended to respond more negatively than those of decedents with Medicare only and caregivers who were more highly educated tended to respond more negatively than caregivers with only a high school degree for most measures. In terms of impact, several characteristics had a notable impact on at least one of the outcomes, as measured by a relative variance greater than 0.1 SD: payer for hospice care, caregiver education, and Spanish survey or home language. All other characteristics in the multivariate model had a moderate impact of at least 0.01 SD on at least one outcome. These results supported a CMA model employing the following variables: decedent age, payer for hospice care, primary diagnosis, length of final episode of hospice care, caregiver age, caregiver education, relationship to decedent, language, and response percentile.

Citation:

1. Parast L, Haas A, Tolpadi A, Elliott MN, Teno JM, Zaslavsky AM, Anhang Price R (2018). Effects of caregiver and decedent characteristics on CAHPS® Hospice Survey scores. *J Pain Symptom Manage,* 56(4):519-529.

[Response Ends]

2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.

Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.

[Response Begins]

All four social risk factors [(1) decedent education, (2) primary payer for hospice care, (3) caregiver respondent education and (4) caregiver respondent language] were considered in the analyses to select risk factors. The variables capturing the primary payer for hospice care and caregiver respondent language (factors 2 and 4 above) are included in our risk adjustment model. Both decedent education and caregiver respondent education were significantly and strongly associated with the outcomes; however, they were moderately correlated with each other. Therefore, only caregiver education was retained in the risk adjustment model.

On average, the impact of adjustment is small, as described below in section 2b.26. However, the impact is larger for hospices that provide care to those at the high or low extremes of risk. Parast et al. (2018) provide an example of the impact of adjustment on a hospice with a large proportion of responding caregivers whose decedents had Medicare and Medicaid as the payer for hospice care (51%) and a large proportion of highly educated responding caregivers (53% with more education than 4-year college); this hospice would have a Hospice Team Communication score at the 25th percentile of national hospice scores without case mix adjustment, but a score at the 59th percentile if this case mix was accounted for in adjustment.

Citation:

1. Parast L, Haas A, Tolpadi A, Elliott MN, Teno JM, Zaslavsky AM, Anhang Price R (2018). Effects of caregiver and decedent characteristics on CAHPS® Hospice Survey scores. *J Pain Symptom Manage,* 56(4):519-529.

[Response Ends]

2b.26. 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 control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter "N/A" for questions about the statistical risk model discrimination and calibration statistics.

Validation testing should be conducted in a data set that is separate from the one used to develop the model.

[Response Begins]

To investigate CMA's overall effect on each CAHPS Hospice Survey measure, we compared hospice-level estimates without adjustment versus hospice-level estimates after adjusting for case mix using the multivariate model. We calculated Kendall's tau, a measure of rank correlation, which expresses the proportion of hospice pairs whose relative rankings were reversed by adjustment, scaled from 1 for no changes to −1 for a complete reversal of rankings. A tau value near 0 would indicate very little correlation between the unadjusted and adjusted scores and a large tau value near 1 would indicate almost perfect correlation between the scores. A tau estimate equal to 1 would indicate that case-mix adjustment has no effect on the hospice-level scores, which would be concerning since case-mix adjustment is expected to have some effect. A tau estimate very close to -1 would indicate that case-mix adjustment almost completely re-ranked all hospices, which would also be concerning since case-mix adjustment would not be expected to have such a dramatic effect. Based on prior CAHPS work, tau for these types of measures is expected to be between 0.80 and 0.95 (Kim, Zaslavsky et al. 2005; Zaslavsky, Zaborski et al. 2001). In addition, Kendall's tau is directly related to the proportion of pairs of hospices that would switch ordering as a consequence of case-mix adjustment (Zaslavsky 1998). If Kendall's tau is denoted as K, a value of K would indicate that [(1-K)/2] % of hospice pairs switched rankings due to case–mix adjustment.

Table 2b3.5. Summary of Impact of Case-Mix Adjustment Variables on CAHPS® Hospice Survey Measure Scores, Q3 2016 – Q2 2018

*	Kendall's Tau Comparing Unadjusted and Case-Mix Adjusted Hospice-level Means
Multi-Item Measures	*
Hospice Team Communication	0.8836
Getting Timely Care	0.9121
Treating Family Member with Respect	0.9183
Support for Emotional and Religious Beliefs	0.9123
Getting Help for Symptoms	0.8904
Getting Hospice Care Training	0.8862
Global Measures	*
Rating of Hospice	0.9213
Willingness to Recommend	0.9441

* This cell intentionally left empty.

Citations:

1. Kim M, Zaslavsky AM & Cleary PD (2005). Adjusting pediatric Consumer Assessment of Health Plans Study (CAHPS) scores to ensure fair comparison of health plan performances. *Med Care*, *43*(1), 44-52.
2. Zaslavsky AM (1998). Issues in case-mix adjustment of measures of the quality of health plans. Proceedings, Government and Social Statistics Sections.
3. Zaslavsky AM, Zaborski LB, Ding L, Shaul JA, Cioffi MJ, Cleary PD. (2001). Adjusting performance measures to ensure equitable plan comparisons. *Health Care Financing Review*, *22*(3), 109.

[Response Ends]

2b.27. Provide risk model discrimination statistics.

For example, provide c-statistics or R-squared values.

[Response Begins]

N/A

[Response Ends]

2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).

[Response Begins]

N/A

[Response Ends]

2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.

The preferred file format is .png, but most image formats are acceptable.

[Response Begins]

N/A

[Response Ends]

2b.30. Provide the results of the risk stratification analysis.

[Response Begins]

N/A

[Response Ends]

2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).

In other words, what do the results mean and what are the norms for the test conducted?

[Response Begins]

These findings support the use of case-mix adjustment. They suggest that case-mix adjustment has a small effect, but one that is likely to be important for hospices with unusual case mix.

Results of the comparison between adjustments from the multivariate model versus the null model are shown in Table 2b3.5. Kendall's tau comparing scores between null and multivariate model adjustments for each measure are in the expected range, from 0.88 for *Hospice Team Communication* to 0.94 for *Willingness to Recommend.* This means that applying the full set of recommended adjustors, six percent of hospice pairs would switch in terms of relative rankings for *Hospice Team Communication* and three percent of hospice pairs would switch in terms of relative rankings for *Willingness to Recommend*.

[Response Ends]

2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.

Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.

[Response Begins]

None

[Response Ends]

---

3. Feasibilitiy

3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.

[Response Begins]

Other (Please describe)

[Other (Please describe) Please Explain]

Survey of informal caregiver (i.e., family member or friend) of hospice decedent

[Response Ends]

3.02. Detail to what extent the specified data elements are available electronically in defined fields.

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.

[Response Begins]

Patient/family reported information (may be electronic or paper)

[Response Ends]

3.03. 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 data elements not from electronic sources.

[Response Begins]

The CAHPS Hospice Survey assesses experiences of hospice care from the perspective of an informal caregiver (i.e., family member or friend) of the hospice decedent. Survey responses are therefore collected directly from the informal caregiver via a mail, telephone, or web-based survey.

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]

N/A

[Response Ends]

3.06. 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.

[Response Begins]

Hospices have been collecting CAHPS® Hospice Survey data since 2015. In 2021, approximately 64 percent of all active hospice agencies submitted one or more months of CAHPS data. Ninety-nine percent of the hospices that did not submit CAHPS data were exempt from CAHPS reporting requirements because they met CMS's exemption for size in the prior year.

Data submission and oversight of the CAHPS Hospice Survey is refined on an annual basis, incorporating feedback from vendors and hospices. Updates are conveyed to vendors and hospices during the yearly vendor training and the annually-updated Quality Assurance Guidelines, which is posted on the project website: https://www.hospicecahpssurvey.org/

The overall response rate to the CAHPS Hospice Survey is similar to other CAHPS surveys at approximately 30%. The survey is offered in three modes to allow hospices to select a mode that meets their preferences and budgets. The vast majority of hospices use the mail-only mode, for which the response rate is 30%. Sampling is conducted by the vendor based upon administrative data submitted by the hospices. Most hospices use a census sample.

Like many other CAHPS surveys, the CAHPS® Hospice Survey relies upon providers to contract with CMS-approved survey vendors from which providers can choose. We estimate the vendor cost per hospice is approximately $4,000 annually on average. We believe the questionnaire takes the typical respondent no more than 15 minutes to answer. Although the survey is currently offered in the three modes described above (mail, telephone only, and mail with telephone follow up), the 2021 mode experiment found promising evidence for a web-based mode of administration.

The mail version of the survey is offered in eight languages: English, Spanish, Chinese (simplified and traditional), Russian, Portuguese, Vietnamese, Polish and Korean. The telephone version of the survey is offered in English, Spanish, and Russian. Currently, 99% of the surveys are completed in English.

We do not offer survey respondents confidentiality. The respondents are family members or friends of the deceased patient; they are not themselves patients of the hospice and thus are at less risk if their names are revealed. In addition, since many hospices have a small sample of decedents/caregivers, hospices may be able to identify respondents by inference. We allow hospices to obtain detailed information about survey responses for quality improvement activities. We do not, however, allow hospices to contact respondents to ask about their survey responses.

[Response Ends]

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

3.07. Detail 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),

Attach the fee schedule here, if applicable.

[Response Begins]

There are no fees or licensing for use of the CAHPS Hospice Survey, training or oversight activities, or for accessing publicly reported CAHPS Hospice Survey measure scores.

[Response Ends]

Criterion 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.

3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.

[Response Begins]

Other (Please describe)

[Other (Please describe) Please Explain]

Survey of informal caregiver (i.e., family member or friend) of hospice decedent

[Response Ends]

3.02. Detail to what extent the specified data elements are available electronically in defined fields.

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.

[Response Begins]

Patient/family reported information (may be electronic or paper)

[Response Ends]

3.03. 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 data elements not from electronic sources.

[Response Begins]

The CAHPS Hospice Survey assesses experiences of hospice care from the perspective of an informal caregiver (i.e., family member or friend) of the hospice decedent. Survey responses are therefore collected directly from the informal caregiver via a mail, telephone, or web-based survey.

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]

N/A

[Response Ends]

3.06. 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.

[Response Begins]

Hospices have been collecting CAHPS® Hospice Survey data since 2015. In 2021, approximately 64 percent of all active hospice agencies submitted one or more months of CAHPS data. Ninety-nine percent of the hospices that did not submit CAHPS data were exempt from CAHPS reporting requirements because they met CMS's exemption for size in the prior year.

Data submission and oversight of the CAHPS Hospice Survey is refined on an annual basis, incorporating feedback from vendors and hospices. Updates are conveyed to vendors and hospices during the yearly vendor training and the annually-updated Quality Assurance Guidelines, which is posted on the project website: https://www.hospicecahpssurvey.org/

The overall response rate to the CAHPS Hospice Survey is similar to other CAHPS surveys at approximately 30%. The survey is offered in three modes to allow hospices to select a mode that meets their preferences and budgets. The vast majority of hospices use the mail-only mode, for which the response rate is 30%. Sampling is conducted by the vendor based upon administrative data submitted by the hospices. Most hospices use a census sample.

Like many other CAHPS surveys, the CAHPS® Hospice Survey relies upon providers to contract with CMS-approved survey vendors from which providers can choose. We estimate the vendor cost per hospice is approximately $4,000 annually on average. We believe the questionnaire takes the typical respondent no more than 15 minutes to answer. Although the survey is currently offered in the three modes described above (mail, telephone only, and mail with telephone follow up), the 2021 mode experiment found promising evidence for a web-based mode of administration.

The mail version of the survey is offered in eight languages: English, Spanish, Chinese (simplified and traditional), Russian, Portuguese, Vietnamese, Polish and Korean. The telephone version of the survey is offered in English, Spanish, and Russian. Currently, 99% of the surveys are completed in English.

We do not offer survey respondents confidentiality. The respondents are family members or friends of the deceased patient; they are not themselves patients of the hospice and thus are at less risk if their names are revealed. In addition, since many hospices have a small sample of decedents/caregivers, hospices may be able to identify respondents by inference. We allow hospices to obtain detailed information about survey responses for quality improvement activities. We do not, however, allow hospices to contact respondents to ask about their survey responses.

[Response Ends]

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

3.07. Detail 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),

Attach the fee schedule here, if applicable.

[Response Begins]

There are no fees or licensing for use of the CAHPS Hospice Survey, training or oversight activities, or for accessing publicly reported CAHPS Hospice Survey measure scores.

[Response Ends]

Criterion 4: Use and Usability

4a. 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.

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.

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 demonstrating performance improvement.

4a.01. Check all current uses. For each current use checked, please provide:

- **Name of program and sponsor**
- **URL**
- **Purpose**
- **Geographic area and number and percentage of accountable entities and patients included**
- **Level of measurement and setting**

[Response Begins]

Public Reporting

[Public Reporting Please Explain]

Name of program: Hospice Quality Reporting Program (HQRP)

Sponsor: Centers for Medicare & Medicaid Services (CMS)

Purpose: The HQRP, mandated by Section 3004 of the Affordable Care Act, directs the Secretary to establish quality reporting requirements for hospice programs. The HQRP aims to promote quality improvement and improve transparency through public reporting.

Geographic area and number and percentage of accountable entities and patients included: In the most recent public reporting period, CAHPS Hospice Survey measure scores were publicly reported for 2,996 (50%) of the 5,996 active Medicare-certified hospices. These hospices provided care to 96 percent of all hospice decedents.

Level of measurement: Hospice program

Setting: The survey assesses hospice care received in patients' homes, in assisted living and nursing home facilities, in hospice inpatient units, and in acute care hospitals.

Payment Program

[Payment Program Please Explain]

Failure of hospices to comply with quality data reporting requirements of the Hospice Quality Reporting Program results in a percentage-point reduction to the APU for the corresponding fiscal year (FY). Effective with the FY 2022 Final Rule, beginning with the FY 2024 APU and for each subsequent year, the APU penalty is 4% for hospices that do not comply with the HQRP for that FY.

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Public reporting

Payment Program

[Response Ends]

4a.03. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?

[Response Begins]

N/A

[Response Ends]

4a.04. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

A 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.

[Response Begins]

N/A

[Response Ends]

4a.05. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

Detail 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.

[Response Begins]

Since February 2018, scores for the eight current CAHPS Hospice Survey measures have been publicly reported on Hospice Compare and its successor, Care Compare, web pages of www.Medicare.gov. Scores for each CAHPS Hospice Survey measure are calculated using the most recent eight quarters of data. The number of responses varies from quarter to quarter. For the Q3 2019 through Q4 2019 and Q3 2020 through Q4 2021 reporting period (reflecting care experiences of patients who died while receiving hospice care from July 2019 through December 2019 and July 2020 through December 2021, and excluding Q1 and Q2 2020, for which hospices were exempt from data collection due to the public health emergency), 663,597 CAHPS Hospice Survey responses were submitted to CMS on behalf of 3,978 hospices. CMS publicly reported scores for the 3,017 of these hospices that had at least 30 completed surveys over the reporting period. The overall response rate was 30%.

[Response Ends]

4a.06. Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

[Response Begins]

CAHPS Hospice Survey results are updated quarterly in February, May, August and November. Each display of results is calculated using the most recent eight quarters of data. Details regarding how scores are calculated, including adjustment for mode of survey administration and case mix, are available on the CAHPS Hospice Survey website (https://hospicecahpssurvey.org/en/public-reporting/scoring-and-analysis/). Prior to each quarterly release of data on Care Compare, hospice providers are given the opportunity to review their CAHPS Hospice Survey measure scores during a 30-day preview period via a Provider Preview Report. The purpose of these reports is to give providers the opportunity to preview their CAHPS Hospice Survey results on each measure prior to public display on Hospice Compare. Hospices access their CAHPS Hospice Provider Preview Reports via the Certification and Survey Provider Enhanced Reports (CASPER) application.

[Response Ends]

4a.07. Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.

[Response Begins]

In the FY 2023 Hospice Payment Rate Update proposed rule, CMS requested comments on potential updates to CAHPS Hospice Survey content and modes of administration.

[Response Ends]

4a.08. Summarize the feedback obtained from those being measured.

[Response Begins]

In response to the FY 2023 proposed rule, hospices expressed support for a shorter survey instrument and web-based mode of survey administration.

[Response Ends]

4a.09. Summarize the feedback obtained from other users.

[Response Begins]

During the 2019-2020 NQF maintenance of endorsement process for the CAHPS Hospice Survey measures in 2019, the American Geriatrics Society submitted a public comment suggesting the addition of new survey content related to person-centered care and to assess whether the hospice care team discussed what mattered most to the patient and family.

[Response Ends]

4a.10. Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

[Response Begins]

As requested by hospices, CMS shortened the survey instrument and tested a web-mail mode of survey administration; in addition, CMS developed and tested new survey items to address the content areas suggested by public comment. Consequently, the revised CAHPS Hospice Survey is eight questions shorter than the current version of the survey, a new Care Preferences measure has been added, and a web-mail mode is being considered as an approved mode of administration.

[Response Ends]

4b. Usability

4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, 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, provide an explanation. 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.

[Response Begins]

As noted above in Section 1b.03, from 2015 through 2019 (i.e., prior to the COVID-19 pandemic), there was evidence of small improvements in CAHPS Hospice Survey measures, with greatest improvement during the short period corresponding to care provided after public reporting began in February 2018. From 2019 to 2021, scores declined slightly for all but one measure, reflecting the responsiveness of the measures to changes in care delivery during the COVID-19 pandemic.

[Response Ends]

4b.02. Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.

[Response Begins]

No unexpected findings have been observed.

[Response Ends]

4b.03. Explain any unexpected benefits realized from implementation of this measure.

[Response Begins]

The majority of the nine hospice leaders interviewed by CMS in summer 2018 about the content, administration, and uses of the CAHPS Hospice Survey reported that they regularly review their hospice's results, and direct quality improvement efforts in response to their measure performance and open-ended comments provided by respondents.

[Response Ends]

Criterion 5: Related and 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.

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02, if the measures are NQF endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

(Can search and select measures.)

[Response Begins]

1623: Bereaved Family Survey

3665: Ambulatory Palliative Care Patients' Experience of Feeling Heard and Understood

3666: Ambulatory Palliative Care Patients' Experience of Receiving Desired Help for Pain

3726: Serious Illness Survey for Home-Based Programs

0208: Family Evaluation of Hospice Care

[Response Ends]

5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus or target population).

(Can search and select measures.)

[Response Begins]

[Response Ends]

5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

[Response Begins]

N/A

[Response Ends]

5.04. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

[Response Begins]

Yes

[Response Ends]

5.05. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

[Response Begins]

N/A

[Response Ends]

5.06. Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.

Provide analyses when possible.

[Response Begins]

0208 Family Evaluation of Hospice Care.

The Family Evaluation of Hospice Care Survey (FEHC), developed more than 20 years ago, assesses hospice care experiences from the perspective of bereaved family members. The CAHPS Hospice Survey covers similar domains, but includes important methodological improvements in the response task, and is adjusted for case mix and mode. Additionally, more stringent survey administration guidelines are in place to permit public reporting of the survey results and valid comparison across hospice programs. FEHC measures were maintained by the National Hospice and Palliative Care Organization (NHPCO), which operated a voluntary repository that provided hospice programs with national benchmarks for FEHC measures. With the national implementation of the CAHPS Hospice Survey, NHPCO shut down the voluntary repository. NQF endorsement of FEHC measures was removed in January 2018.

1623 Bereaved Family Survey.

The Department of Veterans Affairs Bereaved Family Survey assesses experiences of veterans' health care in the last month of life from the perspective of bereaved family members. Importantly, the Bereaved Family Survey assesses care for those who die in inpatient settings, regardless of whether they have received hospice care; this is distinct from respondents to the CAHPS Hospice Survey, who include informal caregivers of decedents who received hospice care across a range of care settings (including both inpatient and other settings).

3665 Ambulatory Palliative Care Patients' Experience of Feeling Heard and Understood and 3666 Ambulatory Palliative Care Patients' Experience of Receiving Desired Help for Pain

These American Academy of Hospice and Palliative Medicine measures assess care experiences of living patients who have received specialty palliative care in an outpatient setting. This is distinct from respondents to the CAHPS Hospice Survey, who are informal caregivers of those who died while receiving hospice care across a range of care settings.

3726 Serious Illness Survey for Home-based Programs

The Serious Illness Survey for Home-Based Programs assesses care experiences of living patients who have received care from a special program that provides care to seriously ill patients in their homes. This is distinct from respondents to the CAHPS Hospice Survey, who are informal caregivers of those who died while receiving hospice care across a range of care settings.

[Response Ends]