NATIONAL QUALITY FORUM

MEASURE WORKSHEET

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

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

Purple text represents the responses from measure developers.

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

Brief Measure Information

NQF #: 2614

Corresponding Measures: N/A

De.2. Measure Title: CoreQ: Short Stay Discharge Measure

Co.1.1. Measure Steward: AHCA/NCAL

De.3. Brief Description of Measure: The measure calculates the percentage of individuals discharged in a six month time period from a SNF, within 100 days of admission, who are satisfied (see: S.5 for details of the time-frame). This patient reported outcome measure is based on the CoreQ: Short Stay Discharge questionnaire that utilizes four items.

1b.1. Developer Rationale: Collecting satisfaction information from skilled nursing facility (SNF) patients is more important now than ever. We have seen a philosophical change in healthcare that now includes the patient and their preferences as an integral part of the system of care. The Institute of Medicine (IOM) endorses this change by putting the patient as central to the care system (IOM, 2001). For this philosophical change to person-centered care to succeed, we have to be able to measure patient satisfaction for these three reasons:

(1) Measuring satisfaction is necessary to understand patient preferences.

(2) Measuring and reporting satisfaction with care helps patients and their families choose and trust a health care facility.

(3) Satisfaction information can help facilities improve the quality of care they provide.

The implementation of person-centered care in SNFs has already begun, but there is still room for improvement. The Centers for Medicare and Medicaid Services (CMS) demonstrated interest in consumers' perspective on quality of care by supporting the development of the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey for patients in nursing facilities (Sangl et al., 2007).

Further supporting person-centered care and resident satisfaction are ongoing organizational change initiatives. These include: the Advancing Excellence in America's Nursing Homes campaign (2006), which lists person-centered care as one of its goals; Action Pact, Inc., which provides workshops and consultations with nursing facilities on how to be more person-centered through their physical environment and organizational structure; and Eden Alternative, which uses education, consultation, and outreach to further person-centered care in nursing facilities. All of these initiatives have identified the measurement of resident satisfaction as an essential part in making, evaluating, and sustaining effective clinical and organizational changes that ultimately result in a person-centered philosophy of care.

The importance of measuring resident satisfaction as part of quality improvement cannot be stressed enough. Quality improvement initiatives, such as total quality management (TQM) and continuous quality

improvement (CQI), emphasize meeting or exceeding "customer" expectations. William Deming, one of the first proponents of quality improvement, noted that "one of the five hallmarks of a quality organization is knowing your customer's needs and expectations and working to meet or exceed them" (Deming, 1986). Measuring resident satisfaction can help organizations identify deficiencies that other quality metrics may struggle to identify, such as communication between a patient and the provider.

As part of the U.S. Department of Commerce renowned Baldrige Criteria for organizational excellence, applicants are assessed on their ability to describe the links between their mission, key customers, and strategic position. Applicants are also required to show evidence of successful improvements resulting from their performance improvement system. An essential component of this process is the measurement of customer, or resident, satisfaction (Shook & Chenoweth, 2012).

The CoreQ: Short Stay Discharge questionnaire can strategically help nursing facilities achieve organizational excellence and provide high quality care by being a tool that targets a unique and growing patient population. Over the past several decades, care in nursing facilities has changed substantially. Statistics show that more than half of all elders cared for in nursing homes are now discharged home (approximately 1.6 million residents; CMS, 2009). Moreover, when satisfaction information from current residents (i.e., long stay residents) is compared with those of elders discharged home, substantial differences exist (Castle, 2007). This indicates that long stay and short stay residents are different populations with different needs in the nursing facilities. Thus, the CoreQ: Short Stay Discharge questionnaire measure is needed to improve the care for short stay SNF patients.

Furthermore, improving the care for short stay nursing home patients is tenable. A review of the literature on satisfaction surveys in nursing facilities (Castle, 2007) concluded that substantial improvements in resident satisfaction could be made in many nursing facilities by improving care (i.e., changing either structural or process aspects of care). This was based on satisfaction scores ranging from 60 to 80% on average.

It is worth noting, few other generalizations could be made because existing instruments used to collect satisfaction information are not standardized. Thus, benchmarking scores and comparison scores (i.e., best in class) were difficult to establish. The CoreQ: Short Stay Discharge measure has considerable relevance in establishing benchmarking scores and comparison scores.

This measure's relevance is furthered by recent federal legislative actions. The Affordable Care Act of 2010 requires the Secretary of Health and Human Services (HHS) to implement a Quality Assurance & Performance Improvement Program (QAPI) within nursing facilities. This means all nursing facilities have increased accountability for continuous quality improvement efforts. In CMS's "QAPI at a Glance" document there are references to customer-satisfaction surveys and organizations utilizing them to identify opportunities for improvement. Lastly, the new "Medicare and Medicaid Programs; Reform of Requirements for Long-Term Care Facilities" proposed rule includes language purporting the importance of satisfaction and measuring satisfaction. CMS states "CMS is committed to strengthening and modernizing the nation's health care system to provide access to high quality care and improved health at lower cost. This includes improving the patient experience of care, both quality and satisfaction, improving the health of populations, and reducing the per capita cost of health care." There are also other references in the proposed rule speaking to improving resident satisfaction and increasing person-centered care (Medicare and Medicaid Programs; Reform of Requirements for Long-Term Care Facilities, 2015). The CoreQ: Short Stay Discharge measure has considerable applicability to both of these initiatives.

Castle, N.G. (2007). A literature review of satisfaction instruments used in long-term care settings. Journal of Aging and Social Policy, 19(2), 9-42.

CMS (2009). Skilled Nursing Facilities Non Swing Bed - Medicare National Summary. http://www.cms.hhs.gov/MedicareFeeforSvcPartsAB/Downloads/NationalSum2007.pdf

CMS, University of Minnesota, and Stratis Health. QAPI at a Glance: A step by step guide to implementing quality assurance and performance improvement (QAPI) in your nursing home. https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/QAPI/Downloads/QAPIAtaGlance.pdf.

Deming, W.E. (1986). Out of the crisis. Cambridge, MA. Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Institute of Medicine (2001). Improving the Quality of Long Term Care, National Academy Press, Washington, D.C., 2001.

Medicare and Medicaid Programs; Reform of Requirements for Long-Term Care Facilities; Department of Health and Human Services. 80 Fed. Reg. 136 (July 16, 2015) (to be codified at 42 CFR Parts 405, 431, 447, et al.).

MedPAC. (2015). Report to the Congress: Medicare Payment Policy. http://www.medpac.gov/documents/reports/mar2015_entirereport_revised.pdf?sfvrsn=0.

Sangl, J., Bernard, S., Buchanan, J., Keller, S., Mitchell, N., Castle, N.G., Cosenza, C., Brown, J., Sekscenski, E., and Larwood, D. (2007). The development of a CAHPS instrument for nursing home residents. Journal of Aging and Social Policy, 19(2), 63-82.

Shook, J., & Chenoweth, J. (2012, October). 100 Top Hospitals CEO Insights: Adoption Rates of Select Baldrige Award Practices and Processes. Truven Health Analytics. http://www.nist.gov/baldrige/upload/100-Top-Hosp-CEO-Insights-RB-final.pdf.

S.4. Numerator Statement: The measure assesses the number of patients who are discharged from a SNF, within 100 days of admission, who are satisfied. The numerator is the sum of the individuals in the facility that have an average satisfaction score of =>3 for the four questions on the CoreQ: Short Stay Discharge questionnaire.

S.6. Denominator Statement: The denominator includes all of the patients that are admitted to the SNF, regardless of payor source, for post-acute care, that are discharged within 100 days; who receive the survey (e.g. people meeting exclusions do not receive a questionnaire) and who respond to the CoreQ: Short Stay Discharge questionnaire within the time window.

S.8. Denominator Exclusions: Exclusions used are made at the time of sample selection and include:

(1) Patients who died during their SNF stay;

(2) Patients discharged to a hospital, another SNF, psychiatric facility, inpatient rehabilitation facility or long term care hospital;

(3) Patients with court appointed legal guardian for all decisions;

(4) Patients discharged on hospice;

(5) Patients who left the nursing facility against medical advice (AMA);

(6) Patients who have dementia impairing their ability to answer the questionnaire defined as having a BIMS score on the MDS as 7 or lower. [Note: we understand that some SNCCs may not have information on cognitive function available to help with sample selection. In that case, we suggest administering the survey to all residents and assume that those with cognitive impairment will not complete the survey or have someone else complete on their behalf which in either case will exclude them from the analysis.]

(7) Patients who responded after the two month response period; and

(8) Patients whose responses were filled out by someone else.

De.1. Measure Type: Outcome: PRO-PM

S.17. Data Source: Instrument-Based Data

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Oct 25, 2016 **Most Recent Endorsement Date:** Oct 25, 2016

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not Applicable

Preliminary Analysis: Maintenance of Endorsement

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

Criteria 1: Importance to Measure and Report

1a. Evidence

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

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

Summary of prior review in 2016

- This patient-reported outcome performance measure calculates the percentage of individuals discharged in a six month time period from a SNF, within 100 days of admission, who are satisfied (see: S.5 for details of the time-frame). This measure is based on the CoreQ: Short Stay Discharge questionnaire that utilizes four items.
- During the 2016 original endorsement review, Committee members noted that this is a significant measure for those who go into a nursing home or a SNF who will not stay indefinitely or for a long period of time. Measuring patient satisfaction and the rate of discharges back into the community is important to measurement as including the patient and their preferences is becoming an integral part of healthcare's changing landscape. Additionally, measuring and reporting satisfaction with care helps patients and their families choose and trust a healthcare facility and can help facilities improve the quality of the care they provide.
- Overall, committee members liked that there was a logic model at the beginning of the measure submission form that linked the measure with information on additional improvement programs, organizational change initiatives, and policies that are going on both at the federal level and the facility level. This remained the same for this submission.
- In the 2016 submission, the developer noted that "Drivers for high satisfaction rates include competency of staff, care/concern of staff, and responsiveness of management"
- The developer stated "We have seen a philosophical change in healthcare that now includes the patient and their preferences as an integral part of the system of care" and notes that measuring patient satisfaction is required for person-centered care for three reasons:
    - Measuring satisfaction is necessary to understand patient preferences.

- Measuring and reporting satisfaction with care helps patients and their families choose and trust a health care facility.
- Satisfaction information can help facilities improve the quality of care they provide

Changes to evidence from last review

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

☒ The developer provided updated evidence for this measure:

Updates:

- Developer provides evidence of meaningfulness to patients
- In structured interviews with patients in facilities, patients indicated that the questions in the short-stay questionnaire are important to them.

Question for the Committee:

- *Is there at least one healthcare action that the provider can do to achieve a change in the measure results?*
- *This measure is derived from patient report. Does the target population value the measured outcome and find it meaningful?*

Guidance from the Evidence Algorithm

PRO-based measure (Box 1) → Relationship between the outcome and at least one healthcare action is identified and supported by the rationale (Box 2) → PASS (From Algorithm 1, NQF Measure Evaluation Criteria Sept 2019, pg. 15)

Preliminary rating for evidence: ☒ Pass ☐ No Pass

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

Maintenance measures – increased emphasis on gap and variation

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

- Developer provided updated statistics broken out by quarter with each quarter representing a rolling 12-month period of data.
- Long Term Care Trend Tracker
  - Data covers 2016Q1-2019Q4
  - Number of SNFs ranges from 372-1577
  - Mean Satisfaction Rate ranges from 77-85% between quarters, but generally fluctuates between 77-80%
  - SD ranges from 14-19%
- Vendor data from MA, NJ, PA, IL, NY providers
  - Data covers 2019Q1 and 2019Q2
  - Number of SNFs = 831
  - Mean Satisfaction Rate ranges 74-85% between quarters
  - SD ranges from 10-15%

Disparities

- Race – No statistically significant differences

- By race, whites averaged a score of 83.3, Blacks or African-Americans averaged a score of 83.4, and Asians 83.4;
- there were no observations for Native Hawaiians or other Pacific Islanders, American Indian or Alaskan Natives

- Education – No statistically significant differences
  - By highest education level those with those high school but who did not graduate averaged 83.2;
  - high school graduates averaged 83.1;
  - those with some college or a 2-year degree averaged 82.9;
  - 4 year college graduates averaged 83.1;
  - those with more than 4 year college degree averaged 83.8
- Age
  - By age group, residents younger than 65 years old averaged 70.0;
  - those 65-74 averaged 84.8;
  - those 75-84 averaged 84.6;
  - those older than 85 averaged 87.1
- Gender
  - Males averaged a score of 89.2
  - Females averaged a score of 81.4
- Developer states that differences in satisfaction by SDS were not statistically significant. This appears likely for race and education, but there are clear differences in performance by age and gender. Moreover, research over the last 20 years has consistently found poorer care in facilities with high minority populations and that nursing homes remain segregated, with black patients concentrated in poorer-quality homes (as measured by staffing ratios, performance, and financial vulnerability).

Questions for the Committee:

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

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

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

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

- No new studies, target population values outcome.
- Pass; discuss five states involved to date with two quarters - were there limitations set? Question on how "100 day" stay was selected as the time metric -- that does not seem "short stay" to me. Presume this links to a payor criterion?
- The Short Stay measure directly relates to the process being measured through responsiveness of management, care and competency staff. A gap exists for satisfaction levels related to payment.

Bills typically come months later and confusion with payment can impact satisfaction. It is not clear how cost of care is measured as a function of discharge satisfaction. There are no questions about the value of care in the Score of Importance for Questions Included in the CoreQ: Short Stay Questionnaire. Perceived value is a key satisfaction metric. Research studies that may pertain include: Empirical Evaluation of a Conceptual Model for the Perceived Value of Health Services, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6172529/ There are many studies in other industries that tie price to satisfaction. It may be worthwhile to see best practices and determine if transferable to healthcare. The methodology used to eliminate the finance domain, is healthcare based, and not leveraging service industry experience. It is not substantiated that the target population does not value price as a function of satisfaction.

- This is a maintenance measure and new information is included regarding the value of the measure for patients (conducted structure interviews with patients).
- Pass.
- Appropriate evidence.
- Target population values outcome assessed.
- Literature and empirical data are provided to support measure focus.
- Limited significant new information provided.
- The developer provided updates to the evidence based on the results of patient interviews to determine the meaningfulness of the measures of which the majority of the patients interviewed confirmed the measures reflected value and were meaningful to them.

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

- Gap provided.
- High - discuss race data - would like to understand breakdown by area, geography, payor; how does this compare with CAHPS? Data report as mean score; more recent metrics have focused on top box. Is that worthy of a discussion? Was there analysis of "elderly" vs "not elderly?" Some SNFs see higher percentage of younger chronically complex patients and wonder if that influenced findings.
- Opportunity for improvement is demonstrated by data based on variation in performance across providers as shown in the histogram in Section 2b5.2.The distribution of summary scores indicates the scores can be used to differentiate facilities of varying levels of customer satisfaction quality. Bias from imputation was minimal due to the rate of the number of missing survey questions is considered low. Subgroup populations seemed to be measured by medical issues and not demographics. The study states that the correlation with the quality indicators (i.e., restraint use, pressure ulcers, catheter use, antipsychotic use, antidepressant use, antianxiety use, use of hypnotics, and deficiency citations) did not alter the average Summary Scores. A subgroup of people who received a surprise medical bill would be interesting to determine if satisfaction changed. The subgroup based on payment method may demonstrate disparities in the care.
- Data provided to show disparities and population subgroups. Differences are noted by age and gender.
- High.
- Evidence supports some gap.
- Solid performance gap demonstrated; social factor analyses did not show disparity but very few non-white patients in testing sample.
- Satisfaction with short stays continues to demonstrate substantial room for improvement.

- Average performance remains >70/100.
- Yes, the literature continues to reflect variation in the quality of care received by select subpopulation groups. While there were no statistically significant differences in satisfaction by patient race and education, there were differences between gender and age. Notably males seem to rate satisfaction higher than females 65 years of age and older. Research also suggests poorer care quality in facilities with larger minority populations and populations in nursing homes remain segregated.
- It can demonstrate a high-level performance gap, although it is less clear exactly how to resolve identified issues. Data on population subgroups is provided and does not appear to identify any obvious disparities; however, it is dependent upon the patient choosing to respond and mail back the survey, so may not be fully representative of the overall population of patients.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

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

2c. For composite measures: empirical analysis support composite approach

Reliability

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

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

Validity

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

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

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

Evaluators: NQF Patient Experience and Function Staff

Full NQF Staff Evaluation

Staff Evaluation Summary:

Reliability

- Developer used the same testing from the 2016 submission
- Measure developer performed both data element level and score level reliability testing
- Data element reliability testing included test-retest analysis on a convenience sample of 100 patients

- Developer calculated the distribution of responses by question in the original round of surveys, and then again in the follow-up surveys (they should be distributed similarly);
- Developer subsequently calculated the correlations between the original and follow-up responses by question (they should be highly correlated).

- The stability of the facility-level score was tested using bootstrapping with 10,000 repetitions of the facility score calculation.
  - Developer presented the percent of facility resamples where the facility score is within 1 percentage point, 3 percentage points, 5 percentage points, and 10 percentage points of the original score.
- Data element testing showed very high levels of agreement and no statistically significant difference in the responses to each question between the original and re-test results. Average Percent Agreement between 1st and 2nd Administered Surveys:

Questionnaire Item	**Percent Agreement**
1. In recommending this facility to your friends and family, how would you rate it overall?	96.8%
2. Overall, how would you rate the staff?	97.8%
3. How would you rate the care you receive?	98.2%
4. How would you rate the discharge process?	98.2%

- Person/questionnaire level agreement showed very high levels of agreement and no statistically significant difference in the responses to each question

		Re- administered Response	
		Poor (1) or Average (2)	Good (3), Very Good (4), or Excellent (5)
Pilot Response	Poor (1) or Average (2)	98.5%	98%
	Good (3), Very Good (4), or Excellent (5)	98.5%	99%

- Measure level testing also demonstrated agreement:
  - 17.82% of bootstrap repetition scores were within 1 percentage point of the score under the original pilot sample
  - 38.14% were within 3 percentage points
  - 61.05% were within 5 percentage points
  - 87.05% were within 10 percentage points

Validity

- Developer resubmitted validity testing from the previous submission in 2016.
- Validity testing of the CoreQ: Short Stay Discharge questionnaire included both data element level and score level testing.
- Data element level

- Exploratory factor analysis (EFA) were used to further refine the pilot instrument. This was an iterative process that included using Eigenvalues from the principal factors (unrotated) and correlation analysis of the individual items.
- Correlation analysis and a factor analysis conducted
- Face validity evaluated via literature review and review of 12 commonly used satisfaction surveys; also examined face validity of domains and the response scale, using 40 patients in 5 nursing homes. The Flesch-Kinkaid scale was used to determine if patients understood the questions.
- Also examined correlation between the four items in the measure and all of the items on the pilot instrument.

- Measure score level
  - Convergent validity testing was performed. Developers examined correlations between the CoreQ: Short Stay Discharge measure scores and i) measures of regulatory compliance and other quality metrics from the Certification and Survey Provider Enhanced Reporting (CASPER) data, ii) several other quality metrics from Nursing Home Compare, iii) risk adjusted discharge to community measure and iv) risk adjusted PointRight® Pro30™ Rehospitalizations
- Data element level results
  - Testing the Items for the CoreQ: Short Stay Discharge Questionnaire
    - The percent of missing responses for the items is very low. The distribution of the summary score is wide. This is important for quality improvement purposes, as nursing facilities can use benchmarks.
    - EFA shows that one factor explains the common variance of the items. A single factor can be interpreted as the only "concept" being measured by those variables. This means that the instrument measures the global concept of satisfaction and not multiple areas of satisfaction. This supports the validity of the CoreQ instrument as measuring a single concept of "customer satisfaction". This testing indicates a high degree of criterion validity.
  - Determine if a Sub-Set of Items Could Reliably be Used to Produce an Overall Indicator of Satisfaction (The Core Q: Short Stay Discharge Measure).
    - Using the correlation information of the Core Q: Short Stay Discharge questionnaire (22 items) and the 4 items representing the CoreQ: Short Stay Discharge questionnaire a high degree of correlation was identified. This testing indicates a high degree of criterion validity. The correlation of the 4 item CoreQ: Short Stay Discharge measure summary score with the overall satisfaction score from all 22 items gave a value of 0.94.
    - EFA shows that one factor explains the common variance of the items. A single factor can be interpreted as the only "concept" being measured by those variables. This means that the instrument measures the global concept of satisfaction and not multiple areas of satisfaction. This supports the validity of the CoreQ instrument as measuring a single concept of "customer satisfaction". This testing indicates a high degree of criterion validity.
  - Developer states that the face validity testing shows the following:
    - The literature review shows that domains used in the Pilot CoreQ: Short Stay Discharge items have a high degree of both face validity and content validity.
    - Patients overall rankings show the general "domain" areas used indicates a high degree of both face validity and content validity. The results show that 100% of residents are able to complete the response format used.

        - The Flesch-Kinkaid scale score achieved for all questions indicates that respondents have a high degree of understanding of the items.

- Score level results
    - Convergent Validity
        - The 8 CASPER quality indicators had a low to moderate level of negative correlation with the CoreQ: Short Stay Discharge measure.
        - The Nursing Home Compare (NHC) Quality Indicators, Five Star ratings, and staffing levels all had moderately high levels of correlation and in the direction predicted with the CoreQ: Short-Stay Discharge measure. These correlations range from ± 0.120 to 0.330.
        - The risk-adjusted Discharge to community measure was negatively correlated to the CoreQ: Short Stay Discharge measure. The correlations were small ranging from -0.05 to -0.16. This was not as hypothesized which may be related to some SNFs that specialize in long stay, have very low discharge to community rates as admissions do not have a plan to go home.
        - The risk-adjusted PointRight® Pro 30™ Rehospitalizations was negatively correlated to the CoreQ: Short Stay Discharge measure. The correlations were modest ranging from -0.22 to -0.31, and all of them were statistically significant at the p-value of 0.05. This is expected because lower rehospitalization rates (an indicator of high quality) are associated with higher satisfaction. This was as hypothesized. This testing indicates a reasonable degree of construct validity and convergent validity.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- Do you agree with the NQF staff assessment of the reliability testing provided by the developer?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- Do you agree with the NQF staff assessment of the validity testing provided by the developer?

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

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

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

- Concern re: implementation in non-English speakers.
- High rating - very comprehensive; no questions or concerns.
- Table 1.6 Patient Demographics omits a data element that may directly correlate to satisfaction. It does not ask payment source. It would be good to know if private insurance, Medicare or Medicaid is the source of payment. This data element is not clearly defined and is not provided. In terms of reliability for implementation, the process is valid.
- No concerns.

- High.
- No Concerns.
- Measure result reliability solid.
- Test-retest reliability strong.
- No change from prior approved measure. Several exclusions effectively limit measure to successful discharges among cognitively intact and excludes proxies.
- The same reliability testing was used from the 2016 submission, including data element and score level testing.
- No concerns about the ability to consistently implement it as specified.

2a.2. Reliability - Testing: Do you have any concerns about the reliability of the measure?

- No.
- No concerns, but one question and not certain what category to place. The rating scale - how does that compare to other metrics? (e.g. Average is 2 of 5?)
- Based on the process used and the Source 2 analysis, I do not have concerns with the reliability of the measure. However, I believe the cross tabulation by payment method can yield important information about satisfaction. The process to re-administer the questionnaire to residents 1 month after their completion of the first survey may show reliability in the results because typically bills have not been processed yet. So, the second survey demonstrates reliability but may miss a key factor for satisfaction: the billing/payment process.
- No concerns.
- No.
- No.
- No.
- No.
- No change from endorsed measure.
- No concerns- reliability seems to be high.
- No.

2b1. Validity -Testing: Do you have any concerns with the testing results?

- No.
- No concerns.
- As stated, "the intent of the Pilot instrument was to have items that represented the most important areas of satisfaction (as identified above) in a parsimonious manner." The questions developed were based on the highest ranking from Table 2b2.3.a: Survey Domain Score out of 12. According to the table Finances scored 1 out of 12. This domain ranking validated the elimination of the question in the survey. Literature reviews were used for validation. For example, the domain of clinical care 10 out of the 12 surveys identified in the literature. Using healthcare literature to validate domains in satisfaction may be misleading. Healthcare is behind other service industries with satisfaction measurements. Therefore, literature available is minimal in this arena.
- No concerns.
- High.
- No.
- In addition to the absence of any addressing or mention of response bias, I am somewhat concerned about some measure exclusions - in particular, exclusion of patients readmitted to acute care facilities (a group they specifically leverage in their validity testing) - makes me concerned the

measure is missing key populations; I am also concerned about the rarity of non-white patients and question the existence of disparities based on this limited population.

- No - data provided for face validity and construct validity.
- No change from endorsed measure.
- No concerns with validity testing, which seems high.
- No concerns about the results.

2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment)2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure?2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk- adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?

- Wonder if survey could include family for patients unable to complete survey? Was that ever discussed? Appreciated the linkage to rehospitalization from an outcome perspective. Would appreciate more info on the risk-adjustment model.
- It is unclear how responses are cross tabulated based on payment method. The results do not indicate if patient groups are from the same payee population. It would be helpful to know segmentation based on private pay, Medicare or Medicaid. There are two levels of exclusions not tracked in the sample that may have impact on satisfaction. These are patients with durable power of attorney for all decisions and patients who left against medical advice. “The exclusions of the patients that had left against medical advice or had a durable power of attorney were not tracked in this sample.” It is unclear beyond perceived “bias” why a POA cannot fill out the survey on behalf of a patient. There were no literature citations to justify this elimination. Also, patients that leave against medical advice, clearly have satisfaction issues with the facility. The justification of these patients having a “bias” and “likely distortion of the results” pg. 23 is not a valid reason to not exclude. This is a survey for patient satisfaction and the survey cannot exclude patients that are known to change results.
- No concerns.
- No.
- Acceptable.
- See above re exclusions; measure not risk adjusted for social risk and does not account for response bias.
- Exclusions vetted with expert panel; no risk adjustment.
- Exclusion of persons discharge AMA (provider label), discharged to another facility or hospital, or needing proxy response likely removes persons with more negative experiences from the measure.
- Exclusions and social risk factors seem appropriate for the measure. The results from the analysis appear to be appropriate for the measure. The measure is not risk adjusted.
- This measure cannot capture the views of those whose social risk factors may make it more difficult for them to respond.

2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5.Comparability of performance scores: If multiple sets of specifications: Do analyses indicate

they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?

- No.
- Agree with NQF staff report (very detailed and thorough).
- No concerns noted.
- No.
- No.
- I think non-response bias is a huge potential threat to validity, not missing data among those who respond.
- No.
- Mail response rates are consistent with other mailed surveys.
- No threats to validity determined.
- Of those that respond, it does appear that this can indicate meaningful differences in care from the patient perspective.

Criterion 3. Feasibility

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

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

- The collection instrument is the CoreQ: Short Stay Discharge questionnaire and Resident Assessment Instrument Minimum Data Set (MDS) version 3.0.
- This is a patient satisfaction survey conducted via mailed survey.
- No fees required to use the measure; the developer did not indicate if there are fees associated with the use of the survey.

Questions for the Committee:

- How burdensome is the implementation of the measure to providers? To patients?
- Is the data collection strategy ready to be put into operational use?

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

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

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

- In use.
- Would appreciate understanding the fees for this survey. SNF industry under financial constraints and would look for any efficiencies to minimize expense (e.g. link with Nursing Home Compare or are these satisfaction vs experience measures?). Do not understand why sampling is not used with this being a mail survey to minimize costs.
- It is feasible to ask POA for their input to surveys. There is no literature review to justify the removal of their input for the survey. Also, people who ask to be discharged against medical advice need to

have an opportunity to be heard. Not collecting these two types of input is acknowledged in this write-up to eliminate bias. This is a concern for data collection.

- Survey is conducted via mail based on demographic information from the facility (which is available electronically).
- Moderate.
- Agree with measure worksheet.
- Agree with rating of moderate - only 4 questions which is pretty short for PRO.
- Moderate.
- No change from endorsed measure.
- This reviewer does not have concerns about the feasibility of the measure. It is a patient survey that is mailed. There are no fees for using the measure. Clarification is needed to determine whether there is a fee to use the survey.
- It would be better if there were options in addition to mail to capture these data. Ideally immediately after a patient is discharged in real time, if possible.

Criterion 4: Usability and Use

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

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

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

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

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

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

OR

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

Accountability program details

- Professional Certification or Recognition Program
  - AHCA Quality Initiative
  - AHCA Quality Awards
- Quality Improvement (external benchmarking to organizations)
  - AHCA NCAL Long Term Care Trend Tracker
- Developer notes that a number of states are implementing the CoreQ survey inside of state incentive programs, including NJ, MA, TN, GA and others.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the

measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- Assistance with data and generally understanding the measure is provided through the open-source measure website where the public can find the manual (containing the measure algorithm) as well as the participating vendors (with their direct contacts).
- Feedback on the performance results and data is provided via a quarterly push report, called Top-Line, sent to all members who have access to the Long-Term Care Trend Tracker (LTCTT).
- CoreQ report which is updated quarterly directly being fed data from the API that vendors use to upload CoreQ measure scores. Therefore, those that are choosing to participate in this, will automatically see their results on the benchmarking tool.
- All those that enter data or have vendors enter their data, obtain this feedback and resources.

Additional Feedback: None

Questions for the Committee:

- How have the performance results been 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**

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

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

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

Improvement results

- Developer provided year-over-year performance data in the Performance Gap section, but this did not show clear improvement in the mean performance with the exception of the final quarter reported where there was a large jump in mean performance.
- Developer states that they have been actively monitoring improvement of AHCA membership as part of the Quality Initiative
    - Members need to improve the measures by 10% from their baseline of 2017Q1 score or achieve a score of more than 90% (satisfaction rate) by March 2021.
    - The current iteration of the quality initiative kicked off in 2018. On average, approximately 10-12% of membership submits CoreQ data every quarter. Thus far, on average, 18% of the providers who nationally submitted data have met this goal for at least one of the CoreQ measures (SS discharge, or LS residents, or LS family). This represents facilities across all 50 states plus the District of Columbia.
- With regards to CoreQ short stay discharge, 8% of the total facilities in the nation that have submitted CoreQ data have met the goal of either having >90% satisfaction rate or at least a 10% improvement in the satisfaction rate since 2017Q1.
    - The states of IA, NH, NJ, NM, and RD have at least 20% of the SNFs that submitted data meet the quality initiative goal.
    - MD, CO, MI, MT, and WV have at least 15% of the SNFs that submitted data meet the quality initiative goal.

- All but two states had at least a facility meet the quality initiative goal.

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

Unexpected findings (positive or negative) during implementation N/A

Potential harms N/A

Additional Feedback: None

Questions for the 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 and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

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

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

- In appropriate use.
- Need more info here please - did not see details on feedback (sought or given from SNFs or patient focus groups); mention of developer volunteer to assist - would think this is important to factor into an endorsement to proceed; Baldrige feedback provided (would be interested in recommendations).
- Nursing home compare is the publicly reported system.
- Measure is publicly reported and used for several accountability programs. Performance results are shared publicly and with the facility.
- Pass.
- Nothing to add.
- Measure is in programmatic use.
- Publicly reported.
- No change from endorsed measure.
- Various certification, recognition programs (e.g., ACHA Initiative and Awards) and quality improvement programs (e.g., ACHA Long- Term Acre Trends Tracker) along with several states are using the measures. Feedback on data and results is provided via quarterly reports.

4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations?4b2. Usability –

Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.

- Translate and validate in other languages.
- Improvement gains achieved. No harms described.
- The elimination of patients POA and patient who left against medical advice for survey completion may be an unintended consequence of missing key satisfaction data that could identify real issues in a facility. POA by definition are responsible for the oversight of the patient. Their input may be a useful and valid metric. It is unclear or supported by research that a POA has less ability than an elderly patient for accurate satisfaction input. Also, patients leaving against medical advice may be because of satisfaction. The unintended consequence may be missing an opportunity to measure medical harm.
- Data can be used to improved low scoring areas and compare with other facilities. No harm is noted.
- Moderate.
- Data can be used to improved low scoring areas and compare with other facilities. No harm is noted.
- I am concerned about costs - direct and indirect - of implementing a new survey.
- Minimal.
- No change from endorsed measure.
- Benchmarking reports can be useful in comparing patient satisfaction across facilities as well as help identify deficiencies or areas in need of improvement. Patients' willingness to refer others to facilities is an indication of a patient's level of confidence in the quality of care and the staff's responsiveness in meeting patient needs. The benefits of the measures (i.e., use of quality improvement purposes) are beneficial to the providers and provide patient feedback on their experience. No harms are foreseen unless the patient feels pressured to respond to the survey in general or to answer the questions more positively than they feel.
- While there seem to be no harms to this measure, it is unclear what specific changes can be made at the facility to address responses that indicate dissatisfaction. The questions are very high level. I suppose one of the main benefits for a patient would be to see the data when publicly reported if he/she has a choice of facilities. However, this is not entirely fair to the facility, as they cannot glean specific actions to take based on the results--of course, if they see consistent dissatisfaction with staff or discharge, then there are perhaps some more clear reforms that could take place in those areas.

Criterion 5: Related and Competing Measures

Related or competing measures

- Developer did not identify any related or competing measures
- Staff did not identify any either
- Developer notes that "The CoreQ: Short Stay Discharge measure does not conceptually address either the same measure focus or the same target population as any other NQF-endorsed measures."

Harmonization

N/A

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

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: Month/Day/Year

- **Of the XXX NQF members who have submitted a support/non-support choice:**
  - XX support the measure
  - YY do not support the measure

NQF Staff Scientific Acceptability Evaluation

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 2614

Measure Title: CoreQ: Short Stay Discharge Measure

Type of measure:

☐ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**
☐ **Outcome** ☒ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☒ **Instrument-Based Data** ☐ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**

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

Measure is:

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

RELIABILITY: SPECIFICATIONS

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

   Submission document: Developer submission, items S.1-S.22

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

2. **Briefly summarize any concerns about the measure specifications.**
   - None identified

RELIABILITY: TESTING

Submission document: Specifications, testing attachment questions 1.1-1.4 and section 2a2

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

   ☐ **Yes** ☐ **No**

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

   Submission document: Testing attachment, section 2a2.2

   - Developer used the same testing from the 2016 submission
   - Measure developer performed both data element level and score level reliability testing
   - Data element reliability testing included test-retest analysis on a convenience sample of 100 patients
     - Developer calculated the distribution of responses by question in the original round of surveys, and then again in the follow-up surveys (they should be distributed similarly);
     - Developer subsequently calculated the correlations between the original and follow-up responses by question (they should be highly correlated).
   - The stability of the facility-level score was tested using bootstrapping with 10,000 repetitions of the facility score calculation.
     - Developer presented the percent of facility resamples where the facility score is within 1 percentage point, 3 percentage points, 5 percentage points, and 10 percentage points of the original score.

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

   Submission document: Testing attachment, section 2a2.3

   - Data element testing showed very high levels of agreement and no statistically significant difference in the responses to each question between the original and re-test results. Average Percent Agreement between 1st and 2nd Administered Surveys:

Questionnaire Item	Percent Agreement

5. In recommending this facility to your friends and family, how would you rate it overall?	96.8%
6. Overall, how would you rate the staff?	97.8%
7. How would you rate the care you receive?	98.2%
8. How would you rate the discharge process?	98.2%

- Person/questionnaire level agreement showed very high levels of agreement and no statistically significant difference in the responses to each question

		Re- administered Response	
		Poor (1) or Average (2)	Good (3), Very Good (4), or Excellent (5)
Pilot Response	Poor (1) or Average (2)	98.5%	98%
	Good (3), Very Good (4), or Excellent (5)	98.5%	99%

- Measure level testing also demonstrated agreement:
  - 17.82% of bootstrap repetition scores were within 1 percentage point of the score under the original pilot sample
  - 38.14% were within 3 percentage points
  - 61.05% were within 5 percentage points
  - 87.05% were within 10 percentage points

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

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

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

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

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

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

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

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

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

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

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

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**
    - Clear specifications appropriately tested with strong results.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    Submission document: Testing attachment, section 2b2.

    - No concerns identified by staff; exclusions appear appropriate
    - Developer was advised by an expert panel to exclude patients who
        - Died
        - Discharged to a hospital
        - Durable power of attorney for all decisions
        - Hospice
        - Low BIMS scores
        - Left against medical advice
    - Developer noted that these exclusions are often used with satisfaction surveys. Developer was not able to calculate the mean CoreQ: Short Stay Discharge scores with and without the exclusions.
    - The first exclusion analysis included responses from 10,319 patients (described elsewhere).
        - The exclusions were tracked and included 1,970 patients (19.1%) discharged to the hospital; 5 (0.05%) discharged to hospice; and, 10 (0.09%) expired.
        - The exclusions of the patients that had left against medical advice or had a durable power of attorney were not tracked in this sample.
    - The second exclusion analysis included 100 nursing homes and data from the first 1000 patients that were included in this initiative:
        - 791 patients (7.9%) were discharged to the hospital; 48 (0.48%) were discharged to hospice; 41 (0.41%) expired; 23 (0.23%) left against medical advice; and 46 (0.46%) had a durable power of attorney.

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

    Submission document: Testing attachment, section 2b4.

    - Developer provided a histogram of performance by providers, demonstrating a normal distribution of performance and a moderate IQR:

	min	p25	p50	p75	max
Summary Score	25.0	75.0	82.5	88.6	100.0

    - The distribution of summary scores is quite wide, indicating the scores can be used to differentiate facilities of varying levels of customer satisfaction quality.
    - No concerns from staff.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: Testing attachment, section 2b5.

- N/A

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

Submission document: Testing attachment, section 2b6.

- Developer describes the following approach to assessing missing data:
    - In calculating the CoreQ: Short Stay Discharge measure if 1 item of 4 is missing then imputation is used, and if 2 (or more) of the 4 items is missing, the respondent is excluded. The imputation method consists of using the average score from the items answered. The testing to identify the extent and distribution of missing data included examining the frequency of missing responses for each of the 4 CoreQ: Short Stay Discharges questionnaire items and the extent and distribution of missing data for more than one missing response for the items.
    - The method of testing to identify if the performance results were biased included examining the correlation with the quality indicators (described above) when imputation was and was not used.
- From the testing of 10,319 residents (described elsewhere) we found:
    - In recommending this facility to your friends and family, how would you rate it overall? That missing responses occurred in 3.71% (n=383) cases.
    - Overall, how would you rate the staff? Missing responses occurred in 3.54% (n=365) cases.
    - How would you rate the care you receive? Missing responses occurred in 3.9% (n=402) cases.
    - How would you rate how well your discharge needs were met? Missing responses occurred in 5.21% (n=538) cases.
    - Two (or more) missing responses occurred in 347 cases. Thus, the degree of missing data was very small (=2.4%). Imputation was used in 1341 cases or 12.9% of respondents.
- Using the cases with 1 missing value (i.e., those with imputation) the correlation with the quality indicators described above (i.e., restraint use, pressure ulcers, catheter use, antipsychotic use, antidepressant use, antianxiety use, use of hypnotics, and deficiency citations) was unchanged compared to those with no imputation.
- No concerns from NQF staff.

16. Risk Adjustment

16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

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

☒ Yes ☐ No ☐ Not applicable

16c. **Social risk adjustment:**

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

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

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

16d. **Risk adjustment summary:**

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

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

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

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

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

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

- Developer performed analyses that demonstrated that the educational makeup of the respondents or the racial makeup of the respondents does not influence the measure.

VALIDITY: TESTING

17. Validity testing level: ☐ Measure score ☐ Data element ☒ Both

18. Method of establishing validity of the measure score:

☒ Face validity

☒ Empirical validity testing of the measure score

☐ N/A (score-level testing not conducted)

19. Assess the method(s) for establishing validity

Submission document: Testing attachment, section 2b2.2

- Developer resubmitted validity testing from the previous submission in 2016.
- Validity testing of the CoreQ: Short Stay Discharge questionnaire included both data element level and score level testing.
- Data element level
    - Exploratory factor analysis (EFA) were used to further refine the pilot instrument. This was an iterative process that included using Eigenvalues from the principal factors (unrotated) and correlation analysis of the individual items.
    - Correlation analysis and a factor analysis conducted
    - Face validity evaluated via literature review and review of 12 commonly used satisfaction surveys; also examined face validity of domains and the response scale, using 40 patients in 5 nursing homes. The Flesch-Kinkaid scale was used to determine if patients understood the questions.
    - Also examined correlation between the four items in the measure and all of the items on the pilot instrument.
- Measure score level
    - Convergent validity testing was performed. Developers examined correlations between the CoreQ: Short Stay Discharge measure scores and i) measures of regulatory compliance and other quality metrics from the Certification and Survey Provider Enhanced Reporting (CASPER) data, ii) several other quality metrics from Nursing Home Compare, iii) risk adjusted discharge to community measure and iv) risk adjusted PointRight® Pro30™ Rehospitalizations

20. Assess the results(s) for establishing validity

Submission document: Testing attachment, section 2b2.3

- Data element level results
    - Testing the Items for the CoreQ: Short Stay Discharge Questionnaire
        - The percent of missing responses for the items is very low. The distribution of the summary score is wide. This is important for quality improvement purposes, as nursing facilities can use benchmarks.
        - EFA shows that one factor explains the common variance of the items. A single factor can be interpreted as the only "concept" being measured by those variables. This means that the instrument measures the global concept of satisfaction and not multiple areas of satisfaction. This supports the validity of the CoreQ instrument as

measuring a single concept of "customer satisfaction". This testing indicates a high degree of criterion validity.

- Determine if a Sub-Set of Items Could Reliably be Used to Produce an Overall Indicator of Satisfaction (The Core Q: Short Stay Discharge Measure).
  - Using the correlation information of the Core Q: Short Stay Discharge questionnaire (22 items) and the 4 items representing the CoreQ: Short Stay Discharge questionnaire a high degree of correlation was identified. This testing indicates a high degree of criterion validity. The correlation of the 4 item CoreQ: Short Stay Discharge measure summary score with the overall satisfaction score from all 22 items gave a value of 0.94.
  - EFA shows that one factor explains the common variance of the items. A single factor can be interpreted as the only "concept" being measured by those variables. This means that the instrument measures the global concept of satisfaction and not multiple areas of satisfaction. This supports the validity of the CoreQ instrument as measuring a single concept of "customer satisfaction". This testing indicates a high degree of criterion validity.
- Developer states that the face validity testing shows the following:
  - The literature review shows that domains used in the Pilot CoreQ: Short Stay Discharge items have a high degree of both face validity and content validity.
  - Patients overall rankings show the general "domain" areas used indicates a high degree of both face validity and content validity. The results show that 100% of residents are able to complete the response format used.
  - The Flesch-Kinkaid scale score achieved for all questions indicates that respondents have a high degree of understanding of the items.

- Score level results
  - Convergent Validity
    - The 8 CASPER quality indicators had a low to moderate level of negative correlation with the CoreQ: Short Stay Discharge measure.
    - The Nursing Home Compare (NHC) Quality Indicators, Five Star ratings, and staffing levels all had moderately high levels of correlation and in the direction predicted with the CoreQ: Short-Stay Discharge measure. These correlations range from ± 0.120 to 0.330.
    - The risk-adjusted Discharge to community measure was negatively correlated to the CoreQ: Short Stay Discharge measure. The correlations were small ranging from -0.05 to -0.16. This was not as hypothesized which may be related to some SNFs that specialize in long stay, have very low discharge to community rates as admissions do not have a plan to go home.
    - The risk-adjusted PointRight® Pro 30™ Rehospitalizations was negatively correlated to the CoreQ: Short Stay Discharge measure. The correlations were modest ranging from -0.22 to -0.31, and all of them were statistically significant at the p-value of 0.05. This is expected because lower rehospitalization rates (an indicator of high quality) are associated with higher satisfaction. This was as hypothesized. This testing indicates a reasonable degree of construct validity and convergent validity.

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

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

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

22. Was the method described and appropriate for assessing the accuracy of ALL critical data elements?

NOTE that data element validation from the literature is acceptable.

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

☒ **Yes**

☐ **No**

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

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

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

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

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

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

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

ADDITIONAL RECOMMENDATIONS

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

- No concerns from staff

Developer Submission

Additional evaluations and submission materials attachments…

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

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

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

CoreQ_Short_Stay_Evidence_Final-635949676534319959.docx

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

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

No

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 2614

Measure Title: CoreQ: Short Stay Discharge Measure

Date of Submission: 4/9/2020

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

Outcome

☐ Outcome: Click here to name the health outcome

☒ Patient-reported outcome (PRO): Customer Satisfaction

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

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☐ Process: Click here to name what is being measured

☐ Appropriate use measure: Click here to name what is being measured

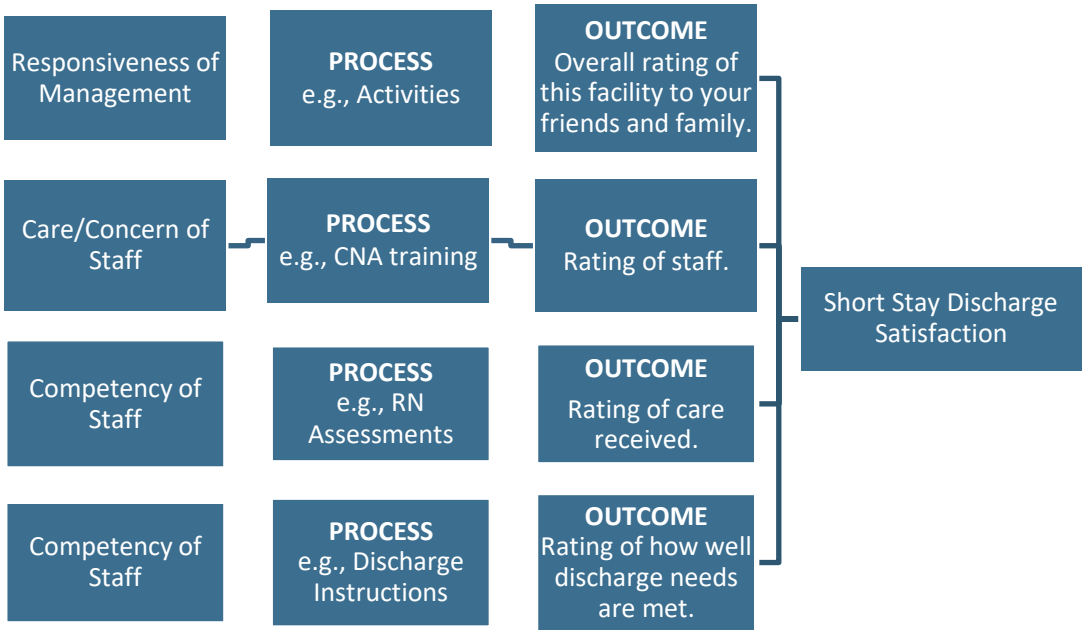
☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

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

Short stay discharge satisfaction can be looked at as the outcome (encompassing the four outcomes shown in the diagram) for a number of structures and processes within skilled nursing care centers. Drivers for

satisfaction include competency of staff, care/concern of staff, and responsiveness of management (National Research Corporation, 2014).



Castle, N.G. (2007). A literature review of satisfaction instruments used in long-term care settings. *Journal of Aging and Social Policy*, 19(2), 9-42.

Donabedian, A. (1985). Twenty years of research on the quality of medical care: 1964-1984. *Evaluation and the Health Professions,* 8, 243-65.

Donabedian, A. (1988). The quality of care. *Journal of the American Medical Association,* 260, 1743-1748.

Donabedian, A. (1996). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly*, 44(1), 166-203.

Glass, A. (1991). Nursing home quality: A framework for analysis. *Journal of Applied Gerontology*, 10(1), 5-18.

National Research Corporation. (2014). 2014 National Research Report Empowering Customer-Centric Healthcare Across the Continuum.

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

The meaningfulness of the measure was determined using residents (n=40) in five nursing facilities in the Pittsburgh region. All short-stay residents were cognitively intact. Permission to approach residents was given by facility management. Most residents (40 of 50) agreed to be interviewed. The interview was conducted within three days of discharge. An informed consent was signed by each resident. Apart from informed consent, the interviews were anonymous. The interviews were not recorded, but notes were taken by the interviewer, Dr. Nicholas Castle. The interviews were conducted at sites that ensured confidentiality (e.g., resident rooms or private areas) and no staff were in the vicinity. Interviews were conducted in a standardized format with the same script for

each resident. The interviews were used to assess the importance of domains used in the satisfaction measure. . The items assessing overall satisfaction were shown to be extremely important using a scale from 1 as most important to 22 as least important. Respondents could pick a maximum of 5 questions as most important. The CoreQ questions were ranked as follows:

1. In recommending this facility to your friends and family, how would you rate it overall? Of the 40 respondents, 39 ranked this question as most important.
2. Overall, how would you rate the staff? Of the 40 respondents, 36 ranked this question as most important.
3. How would you rate the care you received? Of the 40 respondents, 36 ranked this question as most important.
4. How would you rate how well your discharge needs were met? Of the 40 respondents, 37 ranked this question as most important.

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

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

In a review of the satisfaction literature, Castle (2007) noted that the structure, process, outcome model was most commonly used to identify the factors that influence satisfaction. The table below provides the structure and process drivers that are associated with our stated outcome of customer satisfaction.

Table 1a.2.1: The structure and process drivers associated with short stay discharge satisfaction.

Authors	***Structure or Process and Driver of Short Stay Discharge Satisfaction***	***Summary Statement showing structures, processes, interventions and services and influence short-stay discharge satisfaction.***	***Citation***
Reinhardt, et al., 2014	Process Responsiveness of management and care/concern of staff	Conversations regarding end-of-life care options with family members show higher overall satisfaction with care and more use of advance directives.	Reinhardt, J.P., Chichin, E., Posner, L., & Kassabian, S. (2014). Vital conversations with family in the nursing home: preparation for end-stage dementia care. *Journal Of Social Work In End-Of-Life & Palliative Care*. 10(2):112-26.

Lin et al., 2014.	Process Competency of Staff	Significant difference for overall resident satisfaction with higher perceived service quality.	Lin, J., Hsiao, C.T., Glen, R., Pai, J.Y., & Zeng, S.H. (2014). Perceived service quality, perceived value, overall satisfaction and happiness of outlook for long-term care institution residents. *Health Expectations*. 17(3):311-20.
Van Uden et al. (2013).	Process Competency of Staff	For nursing home residents with dementia improved symptom management is associated with higher satisfaction with care.	Van Uden, N., Van den Block, L., van der Steen, J.T., Onwuteaka-Philipsen, B.D., Vandervoort, A., Vander Stichele, R., & Deliens, L. (2013). Quality of dying of nursing home residents with dementia as judged by relatives. *International Psychogeriatrics.* 25(10):1697-707.
Li et al. (2013).	Structure Competency of Staff	Higher overall nursing home satisfaction scores were associated with higher nursing staffing levels and fewer deficiency citations.	Li, Y., Cai, X., Ye, Z., Glance, L.G., Harrington, C., & Mukamel, D.B. (2013). Satisfaction with Massachusetts nursing home care was generally high during 2005-09, with some variability across facilities. *Health Affairs*. 32(8):1416-25.
Authors	***Structure or Process***	***Summary Statement showing structures, processes, interventions and services and influence short-stay discharge satisfaction.***	***Citation***
Brownie & Nancarrow (2013).	Structure & Process Responsiveness of management and Care/concern of staff	Implementation of person-centered care is associated with higher levels of satisfaction.	Brownie, S. & Nancarrow, S. (2013). Effects of person-centered care on residents and staff in aged-care facilities: a systematic review. *Clinical Interventions In Aging*. 8:1-10.
Kleijer et al., 2014	Process Competency of staff	Residents perceive a low level of quality of care in centers where there is a high level of antipsychotic use.	Kleijer, B., Van Marum, R., Frijeters, D., Jansen, P., Ribbe, M., Egberts, A., & Heerdink, E. (2014). Variability between nursing homes in prevalence of antipsychotic use in patients with dementia. *International Psychogeriatrics,* 26(3), 363-371.
Bishop et al., 2008	Structure	CNA's that receive a good supervision are more committed to staying in their jobs. This	Bishop, C., Weinberg, D., Leutz, W., Dossa, A., Pfefferle, S., & Zincavage, R. (2008). Nursing assistants' job commitment: Effect of nursing

	Care/concern of staff	commitment in turn leads to positive relationships with resident and higher resident satisfaction.	home organizational factors and impact on resident well-being. *The Gerontologist*, 48(1), 36-45.
Kayser-Jones et al., 1999	Structure Responsiveness of management and competency of staff	Higher levels of RN and LPN staffing have been associated with better quality outcomes such as ADL maintenance and hydration. Centers that have a family council in addition to the required resident council have higher resident satisfaction.	Kayser-Jones, J., Schell, E.S., Poter, C., Barbaccia, J.C., & Shaw, H. (1999). Factors contributing to dehydration in nursing homes: Inadequate staffing and lack of professional supervision. *Journal of the American Geriatrics Society*, 47(10), 1187-1194.

Castle, N.G. (2007). A literature review of satisfaction instruments used in long-term care settings. *Journal of Aging and Social Policy*, 19(2), 9-42.

Donabedian, A. (1985). Twenty years of research on the quality of medical care: 1964-1984. *Evaluation and the Health Professions,* 8, 243-65.

Donabedian, A. (1988). The quality of care. *Journal of the American Medical Association,* 260, 1743-1748.

Donabedian, A. (1996). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly*, 44(1), 166-203.

Glass, A. (1991). Nursing home quality: A framework for analysis. *Journal of Applied Gerontology*, 10(1), 5-18.

Kleijer, B., Van Marum, R., Frijeters, D., Jansen, P., Ribbe, M., Egberts, A., & Heerdink, E. (2014). Variability between nursing homes in prevalence of antipsychotic use in patients with dementia. *International Psychogeriatrics,* 26(3), 363-371.

Bishop, C., Weinberg, D., Leutz, W., Dossa, A., Pfefferle, S., & Zincavage, R. (2008). Nursing assistants' job commitment: Effect of nursing home organizational factors and impact on resident well-being. *The Gerontologist*, 48(1), 36-45.

Lucas, J.A., Lowe, T.J., Robertson, B., Akincigil, A., Sambamoorthi, Q., Bilder, S., Paek, E.K., & Crystal, S. (2007). The relationship between organizational factors and resident satisfaction with nursing home care and life. *Journal of Aging & Social Policy,* 19(2), 125-151.

Kayser-Jones, J., Schell, E.S., Poter, C., Barbaccia, J.C., & Shaw, H. (1999). Factors contributing to dehydration in nursing homes: Inadequate staffing and lack of professional supervision. *Journal of the American Geriatrics Society*, 47(10), 1187-1194.

Kane, R.L., & Kane, R.A. (2001). What older people want from long-term care, and how can they get it. *Health Affairs*, 20(6), 114-127.

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

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

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

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

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies?	

• Quality – what type of studies?	
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

1a.4 OTHER SOURCE OF EVIDENCE

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

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

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

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

1b. Performance Gap

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

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

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

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

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

Collecting satisfaction information from skilled nursing facility (SNF) patients is more important now than ever. We have seen a philosophical change in healthcare that now includes the patient and their preferences as an integral part of the system of care. The Institute of Medicine (IOM) endorses this change by putting the patient as central to the care system (IOM, 2001). For this philosophical change to person-centered care to succeed, we have to be able to measure patient satisfaction for these three reasons:

(1) Measuring satisfaction is necessary to understand patient preferences.

(2) Measuring and reporting satisfaction with care helps patients and their families choose and trust a health care facility.

(3) Satisfaction information can help facilities improve the quality of care they provide.

The implementation of person-centered care in SNFs has already begun, but there is still room for improvement. The Centers for Medicare and Medicaid Services (CMS) demonstrated interest in consumers'

perspective on quality of care by supporting the development of the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey for patients in nursing facilities (Sangl et al., 2007).

Further supporting person-centered care and resident satisfaction are ongoing organizational change initiatives. These include: the Advancing Excellence in America's Nursing Homes campaign (2006), which lists person-centered care as one of its goals; Action Pact, Inc., which provides workshops and consultations with nursing facilities on how to be more person-centered through their physical environment and organizational structure; and Eden Alternative, which uses education, consultation, and outreach to further person-centered care in nursing facilities. All of these initiatives have identified the measurement of resident satisfaction as an essential part in making, evaluating, and sustaining effective clinical and organizational changes that ultimately result in a person-centered philosophy of care.

The importance of measuring resident satisfaction as part of quality improvement cannot be stressed enough. Quality improvement initiatives, such as total quality management (TQM) and continuous quality improvement (CQI), emphasize meeting or exceeding "customer" expectations. William Deming, one of the first proponents of quality improvement, noted that "one of the five hallmarks of a quality organization is knowing your customer's needs and expectations and working to meet or exceed them" (Deming, 1986). Measuring resident satisfaction can help organizations identify deficiencies that other quality metrics may struggle to identify, such as communication between a patient and the provider.

As part of the U.S. Department of Commerce renowned Baldrige Criteria for organizational excellence, applicants are assessed on their ability to describe the links between their mission, key customers, and strategic position. Applicants are also required to show evidence of successful improvements resulting from their performance improvement system. An essential component of this process is the measurement of customer, or resident, satisfaction (Shook & Chenoweth, 2012).

The CoreQ: Short Stay Discharge questionnaire can strategically help nursing facilities achieve organizational excellence and provide high quality care by being a tool that targets a unique and growing patient population. Over the past several decades, care in nursing facilities has changed substantially. Statistics show that more than half of all elders cared for in nursing homes are now discharged home (approximately 1.6 million residents; CMS, 2009). Moreover, when satisfaction information from current residents (i.e., long stay residents) is compared with those of elders discharged home, substantial differences exist (Castle, 2007). This indicates that long stay and short stay residents are different populations with different needs in the nursing facilities. Thus, the CoreQ: Short Stay Discharge questionnaire measure is needed to improve the care for short stay SNF patients.

Furthermore, improving the care for short stay nursing home patients is tenable. A review of the literature on satisfaction surveys in nursing facilities (Castle, 2007) concluded that substantial improvements in resident satisfaction could be made in many nursing facilities by improving care (i.e., changing either structural or process aspects of care). This was based on satisfaction scores ranging from 60 to 80% on average.

It is worth noting, few other generalizations could be made because existing instruments used to collect satisfaction information are not standardized. Thus, benchmarking scores and comparison scores (i.e., best in class) were difficult to establish. The CoreQ: Short Stay Discharge measure has considerable relevance in establishing benchmarking scores and comparison scores.

This measure's relevance is furthered by recent federal legislative actions. The Affordable Care Act of 2010 requires the Secretary of Health and Human Services (HHS) to implement a Quality Assurance & Performance Improvement Program (QAPI) within nursing facilities. This means all nursing facilities have increased accountability for continuous quality improvement efforts. In CMS's "QAPI at a Glance" document there are references to customer-satisfaction surveys and organizations utilizing them to identify opportunities for improvement. Lastly, the new "Medicare and Medicaid Programs; Reform of Requirements for Long-Term Care Facilities" proposed rule includes language purporting the importance of satisfaction and measuring satisfaction. CMS states "CMS is committed to strengthening and modernizing the nation's health care system to provide access to high quality care and improved health at lower cost. This includes improving the patient

experience of care, both quality and satisfaction, improving the health of populations, and reducing the per capita cost of health care." There are also other references in the proposed rule speaking to improving resident satisfaction and increasing person-centered care (Medicare and Medicaid Programs; Reform of Requirements for Long-Term Care Facilities, 2015). The CoreQ: Short Stay Discharge measure has considerable applicability to both of these initiatives.

Castle, N.G. (2007). A literature review of satisfaction instruments used in long-term care settings. Journal of Aging and Social Policy, 19(2), 9-42.

CMS (2009). Skilled Nursing Facilities Non Swing Bed - Medicare National Summary. http://www.cms.hhs.gov/MedicareFeeforSvcPartsAB/Downloads/NationalSum2007.pdf

CMS, University of Minnesota, and Stratis Health. QAPI at a Glance: A step by step guide to implementing quality assurance and performance improvement (QAPI) in your nursing home. https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/QAPI/Downloads/QAPIAtaGlance.pdf.

Deming, W.E. (1986). Out of the crisis. Cambridge, MA. Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Institute of Medicine (2001). Improving the Quality of Long Term Care, National Academy Press, Washington, D.C., 2001.

Medicare and Medicaid Programs; Reform of Requirements for Long-Term Care Facilities; Department of Health and Human Services. 80 Fed. Reg. 136 (July 16, 2015) (to be codified at 42 CFR Parts 405, 431, 447, et al.).

MedPAC. (2015). Report to the Congress: Medicare Payment Policy. http://www.medpac.gov/documents/reports/mar2015_entirereport_revised.pdf?sfvrsn=0.

Sangl, J., Bernard, S., Buchanan, J., Keller, S., Mitchell, N., Castle, N.G., Cosenza, C., Brown, J., Sekscenski, E., and Larwood, D. (2007). The development of a CAHPS instrument for nursing home residents. Journal of Aging and Social Policy, 19(2), 63-82.

Shook, J., & Chenoweth, J. (2012, October). 100 Top Hospitals CEO Insights: Adoption Rates of Select Baldrige Award Practices and Processes. Truven Health Analytics. http://www.nist.gov/baldrige/upload/100-Top-Hosp-CEO-Insights-RB-final.pdf.

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

(Updated for Maintenance of Endorsement)

Below are the statistics requested in this question broken out but quarter, each quarter representing a rolling 12-month of data, akin to measures in the public domain. Section 1 contains data from Long Term Care Trend Tracker, whereas Section 2 contains data from a vendor on facilities in MA, NJ, PA, IL, NY (not included in Section 1).

For a more user-friendly view of these stats, please see appendix Table 1b.2.e and 1.b.2f (section 1 and 2, respectively).

Section 1: Data from Long Term Care Trend Tracker where all vendors upload CoreQ data

Survey dates in this dataset containing 16 quarters of data ranged from August 2016 to December 2019. The data is from Long Term Care Trend Tracker (https://www.ahcancal.org/research_data/trendtracker/Pages/default.aspx) where vendors and member user can upload data, and vendors can also upload non-member data.

2016Q1

Nr_SNFs:372 Mean Satisfaction Rate: 79.09% STD:15.23% Min:0.00% Max:100.00% Q1:70.40% Q3:89.48% IQR:19.08% Total Nr. of Respondents:13988 Decile 1: 62.20% Decile 2:68.10% Decile 3:72.70% Decile 4:76.40% Decile 5: 79.80% Decile 6:83.30% Decile 7:87.50% Decile 8:91.87% Decile 9: 100.00% Decile 10:100.00%

2016Q2

Nr_SNFs:393 Mean Satisfaction Rate: 79.82% STD:14.39% Min:25.00% Max:100.00% Q1:71.40% Q3:90.28% IQR:18.88% Total Nr. of Respondents:16285 Decile 1: 62.50% Decile 2:69.20% Decile 3:73.20% Decile 4:77.20% Decile 5:80.60% Decile 6:83.78% Decile 7:88.34% Decile 8:92.66% Decile 9: 100.00% Decile 10:100.00%

2016Q3

Nr_SNFs:473 Mean Satisfaction Rate: 80.64% STD:16.45% Min:0.00% Max:100.00% Q1:71.70% Q3:93.20% IQR:21.50% Total Nr. of Respondents:18757 Decile 1:61.50% Decile 2:69.20% Decile 3:73.70% Decile 4:78.60% Decile 5:82.44% Decile 6:86.09% Decile 7:90.60% Decile 8:96.88% Decile 9:100.00% Decile 10:100.00%

2016Q4

Nr_SNFs:963 Mean Satisfaction Rate: 79.09% STD:17.95% Min:0.00% Max:100.00% Q1:70.40% Q3:92.00% IQR:21.60% Total Nr. of Respondents:29593 Decile 1:59.26%Decile 2:66.70% Decile 3:72.70% Decile 4:76.90% Decile 5:80.60% Decile 6:85.00% Decile 7:89.78%Decile 8:96.10% Decile 9:100.00% Decile 10:100.00%

2017Q1

Nr_SNFs:977 Mean Satisfaction Rate: 80.28% STD:17.99% Min:0.00% Max:100.00% Q1:71.40% Q3:94.10% IQR:22.70% Total Nr. of Respondents:24903 Decile 1: 58.30% Decile 2:68.00% Decile 3:75.00% Decile 4:78.60% Decile 5:82.50% Decile 6:87.50% Decile 7:91.25% Decile 8:98.00% Decile 9:100.00% Decile 10:100.00%

2017Q2

Nr_SNFs:998 Mean Satisfaction Rate: 80.15% STD:15.99% Min:0.00% Max:100.00% Q1:71.00% Q3:92.00% IQR:21.00% Total Nr. of Respondents:35560 Decile 1: 60.00% Decile 2:68.30%Decile 3:74.70%Decile 4:78.30%Decile 5:82.58%Decile 6:86.33%Decile 7:90.00% Decile 8:94.44%Decile 9:100.00% Decile 10:100.00%

2017Q3

Nr_SNFs:1056 Mean Satisfaction Rate: 80.67% STD:15.94% Min:0.00% Max:100.00% Q1:71.85% Q3:92.85% IQR:21.00% Total Nr. of Respondents:38372 Decile 1:60.00%Decile 2:68.80%Decile 3:75.00% Decile 4:79.15%Decile 5:83.30%Decile 6:87.00%Decile 7:90.42%Decile 8:95.00%Decile 9:100.00% Decile 10:100.00%

2017Q4

Nr_SNFs:1247 Mean Satisfaction Rate: 79.44% STD:16.96% Min:0.00% Max:100.00% Q1:70.65% Q3:92.00% IQR:21.35% Total Nr. of Respondents:47785 Decile 1:59.15%Decile 2:66.70% Decile 3:73.30%Decile 4:77.80%Decile 5:81.40%Decile 6:85.70%Decile 7:89.80%Decile 8:94.18% Decile 9:100.00% Decile 10:100.00%

2018Q1

Nr_SNFs:1132 Mean Satisfaction Rate: 79.01% STD:16.72% Min: 0.00% Max:100.00% Q1:69.50% Q3:91.73% IQR:22.23% Total Nr. of Respondents:46096 Decile 1:58.60%Decile 2:66.66%Decile 3:72.45% Decile 4:76.88%Decile 5:80.87%Decile 6:85.03%Decile 7:88.75%Decile 8:94.00%Decile 9:100.00% Decile 10:100.00%

2018Q2

Nr_SNFs:1220 Mean Satisfaction Rate: 77.44% STD:17.06% Min:0.00% Max:100.00%

Q1:67.08% Q3:90.13% IQR:23.05% Total Nr. of Respondents:52910 Decile 1:56.63%Decile 2:65.50% Decile 3:70.00%Decile 4:75.00%Decile 5:79.00% Decile 6:83.30%Decile 7:87.50%Decile 8:92.90% Decile 9:100.00% Decile 10:100.00%

2018Q3

Nr_SNFs:1241 Mean Satisfaction Rate: 77.10% STD:16.83% Min:0.00% Max:100.00% Q1:66.70% Q3:88.95% IQR:22.25% Total Nr. of Respondents:59681 Decile 1:56.30%Decile 2:65.00%Decile 3:69.35% Decile 4:75.00%Decile 5:78.87%Decile 6:82.60%Decile 7:86.70%Decile 8:91.70%Decile 9:100.00% Decile 10:100.00%

2018Q4

Nr_SNFs:1145 Mean Satisfaction Rate: 77.56% STD:17.63% Min:0.00% Max:100.00% Q1:66.70% Q3:91.00% IQR:24.30% Total Nr. of Respondents:50577 Decile 1:55.60%Decile 2:64.62%Decile 3:69.60% Decile 4:75.00% Decile 5:80.00%Decile 6:84.00%Decile 7:88.10%Decile 8:93.80% Decile 9:100.00% Decile 10:100.00%

2019Q1

Nr_SNFs:1395 Mean Satisfaction Rate: 77.79% STD:18.21% Min:0.00% Max:100.00% Q1:66.70% Q3:92.45% IQR:25.75% Total Nr. of Respondents:61909 Decile 1:54.55% Decile 2:64.30% Decile 3:69.13%Decile 4:75.00%Decile 5:80.00%Decile 6:84.20%Decile 7:89.50%Decile 8:95.78% Decile 9:100.00% Decile 10:100.00%

2019Q2

Nr_SNFs:1170 Mean Satisfaction Rate: 78.51% STD:18.90% Min:0.00% Max:100.00% Q1:67.93% Q3:94.34% IQR:26.41% Total Nr. of Respondents:58560 Decile 1:55.00%Decile 2:65.41%Decile 3:70.89% Decile 4:76.50%Decile 5:81.00%Decile 6:85.70%Decile 7:91.34%Decile 8:96.67%Decile 9:100.00% Decile 10:100.00%

2019Q3

Nr_SNFs:1577 Mean Satisfaction Rate: 78.36% STD:17.66% Min:0.00% Max:100.00% Q1:69.44% Q3:91.40% IQR:21.96% Total Nr. of Respondents:73091 Decile 1:56.65%Decile 2:66.70%Decile 3:72.50% Decile 4:76.20%Decile 5:80.00%Decile 6:84.70%Decile 7:88.90%Decile 8:93.94%Decile 9:100.00% Decile 10:100.00%

2019Q4

Nr_SNFs:419 Mean Satisfaction Rate: 85.83% STD:14.61% Min:16.65% Max:100.00% Q1:78.71% Q3:97.42% IQR:18.71% Total Nr. of Respondents:24943 Decile 1:69.20%Decile 2:75.91% Decile 3:81.88%Decile 4:85.50%Decile 5:89.70%Decile 6:92.19%Decile 7:95.93%Decile 8:98.77% Decile 9:100.00% Decile 10:100.00%

Section 2: Data from one of the vendors (non-Long Term Trend Tracker data) representing facilities in MA, NJ, PA, IL, NY

A. 2019Q1 Score (%):

a.

mean 84.98

min 40.00

max 100.00

Sdv 10.33

Q1 81.00

Q3 93.00

IQR 12.00
p10 76.00
p20 80.00
p30 82.00
p40 84.00
p50 86.00
p60 88.00
p70 91.00
p80 93.00
p90 96.00
p100 100.00
N of SNFs 831
B. 2019Q1 Response Rate:
a.
mean 78.25%
min 22.73%
max 97.37%
Sdv 13.84%
Q1 73.53%
Q3 87.88%
IQR 14.35%
p10 56.52%
p20 70.51%
p30 75.00%
p40 79.55%
p50 82.65%
p60 84.75%
p70 86.57%
p80 88.51%
p90 90.91%
p100 97.37%
N of SNFs 831
C. 2019Q2 Score (%):
a.
74.15 mean
40.00 min
100.00 max
14.60 Sdv
60.00 Q1

88.00 Q3

28.00 IQR

58.00 p10

60.00 p20

63.00 p30

70.00 p40

77.00 p50

81.00 p60

85.00 p70

88.00 p80

92.00 p90

100.00 p100

831 N of SNFs

D. 2019Q2 Response Rate:

a. N

mean 71.97%

min 27.03%

max 100.00%

Sdv 16.38%

Q1 60.59%

Q3 84.65%

IQR 24.06%

p10 48.89%

p20 57.00%

p30 63.04%

p40 67.74%

p50 71.91%

p60 76.92%

p70 82.11%

p80 87.62%

p90 92.86%

p100 100.00%

N of SNFs 831

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

Not Applicable

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities;*

number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.

We did not risk adjust the measure by sociodemographic status due to no statistically significant differences (at the 5% level) in the scores between the SDS categories. See Table 2b4.4b.b in the Testing section. By race, whites averaged a score of 83.3, Blacks or African-Americans averaged a score of 83.4, and Asians 83.4; there were no observations for Native Hawaiians or other Pacific Islanders, American Indian or Alaskan Natives (Table 2b4.4b.c in the Testing section). By highest education level those with those high school but who did not graduate averaged 83.2, high school graduates averaged 83.1, those with some college or a 2-year degree averaged 82.9, 4 year college graduates averaged 83.1, and those with more than 4 year college degree averaged 83.8 (Table 2b4.4b.c in the Testing section). By age group, residents younger than 65 years old averaged 70.0, those 65-74 averaged 84.8, those 75-84 averaged 84.6, and those older than 85 averaged 87.1 (Table 1b.4.a in the Appendix). Furthermore, by gender, males averaged a score of 89.2 and females averaged a score of 81.4 (Table 1b.4.b in the Appendix).

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

Multiple studies in the past twenty years have examined racial disparities in the care of nursing facility residents and have consistently found poorer care in facilities with high minority populations (Fennell et al., 2000; Mor et al., 2004; Smith et al., 2007). Work on racial disparities in nursing facilities' quality of care between elderly white and black residents within nursing facility has shown clearly that nursing homes remain relatively segregated and that specifically nursing home care can be described as a tiered system in which blacks are concentrated in marginal-quality homes (Li, Ye, Glance & Temkin-Greener, 2014; Fennell, Feng, Clark & Mor, 2010; Li, Yin, Cai, Temkin-Greener, Mukamel, 2011; Chisholm, Weech-Maldonado, Laberge, Lin, & Hyer, 2013; Mor et al., 2004; Smith et al., 2007). Such homes tend to have serious deficiencies in staffing ratios, performance, and are more financially vulnerable (Smith et al, 2007; Chisholm et al., 2013). Based on a review of the nursing facility disparities literature, Konetzka and Werner concluded that disparities in care are likely related to this racial and socioeconomic segregation as opposed to within-provider discrimination (Konetzka and Werner 2009). This conclusion is supported, for example, by Grunier and colleagues who found that as the proportion of black residents in the nursing home increased the risk of hospitalization among all residents, regardless of race, also increased (Grunier et al., 2008). Thus, adjusting for racial status has the unintended effect of adjusting for poor quality providers not to differences due to racial status and not within-provider discrimination.

Therefore, lower satisfaction scores for both Caucasian and Blacks and other ethnicities are likely to increase as the proportion of black residents increases in a SNF, indicating that the best measure of racial disparities in satisfaction rates is one that measures scores at the facility level. That is, ethnic and social economic status differences are related to inter-facility differences not to intra-facility differences in care. Therefore, the literature suggests that racial status should not be risk adjusted otherwise one is adjusting for the poor quality of the SNFs rather than differences due to racial status.

Chisholm L, Weech-Maldonado R, Laberge A, Lin FC, Hyer K. (2013). Nursing home quality and financial performance: does the racial composition of residents matter? Health Serv Res;48(6 Pt 1):2060–2080.

Fennell ML, Feng Z, Clark MA, Mor V. (2010). Elderly Hispanics more likely to reside in poor-quality nursing homes. Health Aff (Millwood);29(1):65–73.

Grabowski, D.C. (2004). The admission of Blacks to high-deficiency nursing homes. Medical Care 42(5): 456-464.

Gruneir, A., Miller, S. C., Feng, Z., Intrator, O., & Mor, V. (2008). Relationship between state Medicaid policies, nursing home racial composition, and the risk of hospitalization for black and white residents. Health Services Research, 43(3), 869-881.

Konetzka, R. T., & Werner, R. M. (2009). Review: Disparities in long-term care building equity into market-based reforms. Medical Care Research and Review, 66(5), 491-521.

Li Y, Yin J, Cai X, Temkin-Greener J, Mukamel DB. (2011). Association of race and sites of care with pressure ulcers in high-risk nursing home residents. JAMA;306(2):179–186.

Li Y, Ye Zhiqiu, Glance, Laurent & Temkin-Greener, Helena. (2014). Trends in family rating experience with care and racial disparities among Maryland nursing homes. Med Care, 52(7): 641-648.

Mor, V., Zinn, J., Angelelli, J., Teno, J. M., & Miller, S. C. (2004). Driven to tiers: socioeconomic and racial disparities in the quality of nursing home care. Milbank Quarterly, 82(2), 227-256.

Smith, D. B., Feng, Z., Fennell, M. L., Zinn, J. S., & Mor, V. (2007). Separate and unequal: racial segregation and disparities in quality across US nursing homes. Health Affairs, 26(5): 1448-1458.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

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

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

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

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

Person-and Family-Centered Care

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

Elderly

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

http://www.coreq.org/

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

This is not an eMeasure **Attachment:**

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

No data dictionary **Attachment:**

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

Attachment: CoreQ_Short_Stay_Instrument.docx

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

Patient

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

No

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

No changes to the measure specifications since the last measure update. Since the last update, we have created a website (as specified in section S.1.) with information on all CoreQ measures.

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

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

The measure assesses the number of patients who are discharged from a SNF, within 100 days of admission, who are satisfied. The numerator is the sum of the individuals in the facility that have an average satisfaction score of =>3 for the four questions on the CoreQ: Short Stay Discharge questionnaire.

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

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

The numerator includes all of the patients who were discharged within 100 days of admission and had an average response =>3 on the CoreQ: Short Stay Discharge questionnaire.

The calculation of the individual patient's average satisfaction score is done in the following manner:

-A numeric score is associated with each response scale option on the CoreQ: Short Stay Discharge questionnaire (that is, Poor=1, Average=2, Good=3, Very Good=4, and Excellent=5).

-The following formula is utilized to calculate the individual's average satisfaction score: [Numeric Score Question 1 + Numeric Score Question 2 + Numeric Score Question 3 + Numeric Score Question 4]/4

-The number of respondents whose average satisfaction score >=3 are summed together and function as the numerator.

For patients with one missing data point (from the four items included in the questionnaire) imputation is used (representing the average value from the other three available responses). Patients with more than one missing data point, are excluded from the analyses (i.e., no imputation will be used for these patients). Imputation details are described further below (S.22).

No risk-adjustment is used (See S.18).

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

The denominator includes all of the patients that are admitted to the SNF, regardless of payor source, for post-acute care, that are discharged within 100 days; who receive the survey (e.g. people meeting exclusions do not receive a questionnaire) and who respond to the CoreQ: Short Stay Discharge questionnaire within the time window.

S.7. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection*

items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)

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

The target population includes all of the individuals who respond to the CoreQ: Short Stay Discharge questionnaire within the time window (See: S.5).

The data is collected over a maximum 6 month time window. A shorter period can be used if the sample size (125) meets the specifications described below. The questionnaire is administered to discharged patients within 2 weeks of their discharge date. The discharge date is identified from nursing facility records (e.g., MDS, wherein a discharge MDS record is created that includes a discharge date). Note, the questionnaire must be administered after the patient is discharged and not on the day of the discharge. Patients must respond to the CoreQ: Short Stay Discharge questionnaire within 2 months of receiving the questionnaire.

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

Exclusions used are made at the time of sample selection and include:

(1) Patients who died during their SNF stay;

(2) Patients discharged to a hospital, another SNF, psychiatric facility, inpatient rehabilitation facility or long term care hospital;

(3) Patients with court appointed legal guardian for all decisions;

(4) Patients discharged on hospice;

(5) Patients who left the nursing facility against medical advice (AMA);

(6) Patients who have dementia impairing their ability to answer the questionnaire defined as having a BIMS score on the MDS as 7 or lower. [Note: we understand that some SNCCs may not have information on cognitive function available to help with sample selection. In that case, we suggest administering the survey to all residents and assume that those with cognitive impairment will not complete the survey or have someone else complete on their behalf which in either case will exclude them from the analysis.]

(7) Patients who responded after the two month response period; and

(8) Patients whose responses were filled out by someone else.

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

Individuals are excluded based on information from the admission Minimum Data Set (MDS) 3.0 assessment.

(1) Patients who die: This is recorded in the MDS as Deceased (A2100 = 08).

(2) Patients who were discharged to a hospital, another SNCC, psychiatric facility, Inpatient Rehabilitation Facilities (IRF), or MR/DD facility: This is recorded in the MDS as Discharge to hospital (A2100 = 03); another SNCC (A2100 = 02); psychiatric facility (A2100 = 04); Inpatient Rehabilitation Facilities (A2100 = 05); ID/DD facility (A2100 = 06).

(3) Patients with Court appointed legal guardian for all decisions as identified from the nursing facility health information system.

(4) Patients on hospice: This is recorded in the MDS as Hospice O0100K1 = 1 (“the patient was on hospice in the last 14 days while not a resident”), O0100K2 = 1 (“the patient was on hospice in the last 14 days while a resident”), A1800=07 (“entered from hospice”), or A2100=07 (“discharged to hospice”).

(5) Patients who left the nursing facility against medical advice (AMA) as identified from nursing facility health information systems.

(6) Patients with a BIMS score on the MDS as 7 or lower. This is recorded in the MDS as C0500 <= 7.

(7) Patients who respond after the two month response period.

(8) Patients whose responses were filled out by somebody other than him/herself, as identified by the additional questions on the questionnaire.

Surveys returned as undeliverable are also excluded from the denominator.

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

No stratification is used.

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

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Other (specify):

If other: Non-weighted score. Score is a percentage.

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

Better quality = Higher score

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

1.Identify SNF patients that are discharged within 100 days after admission

a.Calculate the duration of the SNF stay [MDS discharge date (A2000) - MDS admission date (A1900)] to determine if it is = 100 days.

2.Take the patients that have a SNF stay of = 100 days and exclude the following:

a.Patients who died; patients discharged to a hospital; patients with Court appointed legal guardian for all decisions; patients with hospice; patients who left the nursing facility against medical advice (AMA), and patients with a BIMS score of less than 7 do not receive that survey as a result of the exclusions (described in detail above).

i.Patients who die: This is recorded in the MDS as Die during stay (A2100 =08)

ii.Patients who were discharged to a hospital, another SNCC, psychiatric facility, Inpatient Rehabilitation Facility, or MR/DD facility (A2100 = 06): This is recorded in the MDS as Discharge to hospital (A2100 = 03); another SNCC (A2100 = 02); psychiatric facility (A2100 = 04); Inpatient Rehabilitation Facility (A2100 = 05); MR/DD facility (A2100 = 06).

iii.Patients with Court appointed legal guardian for all decisions will be identified from nursing facility health information system.

iv.Patients on hospice: This is recorded in the MDS as Hospice O0100K1 = 1 (“the patient was on hospice in the last 14 days while not a resident”), O0100K2 = 1 (“the patient was on hospice in the last 14 days while a resident”), A1800=07 (“entered from hospice”), or A2100=07 (“discharged to hospice”).

v.Patients who left the nursing facility against medical advice (AMA) will be identified from nursing facility health information systems.

vi.Patients with a BIMS score of 7 or less. This is recorded in the MDS as C0500 <= 7.

3.Administer the CoreQ: Short Stay Discharge questionnaire (See S.25) to these individuals. The questionnaire should be administered to patients discharged within 2 weeks of discharge. Provide individuals 2 months to respond to the survey.

a.Create a tracking sheet with the following columns:

i. Data Administered

ii.Data Response Received

iii.Time to Receive Response ([Date Response Received – Date Administered])

b. Exclude any surveys where Time to Receive Response >2 Months

4.Collect data over a maximum 6 month time window or until 125 consecutive usable surveys are received (See S.21).

5.Exclude responses not completed by the intended recipient (e.g. questions were answered by a friend or family members. It is important to note that cases in which the residents had help with reading the questions, or writing down their responses, are included in the measure, because in these cases the residents answer the questions themselves).

6.Exclude surveys that are returned after two months

7.Combine the CoreQ: Short Stay Discharge questionnaire items to calculate a patient level score. Responses for each item should be given the following scores:

a.Poor = 1,

b.Average = 2,

c.Good = 3,

d.Very good =4 and

e.Excellent = 5.

8.Impute missing data if only one of the four questions are missing data by taking the average of the other questions responses.

9.Exclude any survey with 2 or more survey questions that have missing data.

10.Calculated patient score from usable surveys.

Patient score= (Score for Item 1 + Score for Item 2 + Score for Item 3 + Score for Item 4) / 4.

a.For example, a patient rates their satisfaction on the CoreQ questions as excellent = 5, very good = 4, very good = 4, and good = 3. The resident's total score will be 5 + 4 + 4 + 3 for a total of 16. The patient's total score (16) will then be divided by the number of questions (4), which equals 4. Thus the patients average satisfaction rating is 4.0. This individual would be counted in the numerator since their average score is >3.0.

11.Flag those patients with an average score equal to or greater than 3.0

12.Calculate the CoreQ: Short Stay Discharge measure which represents the percent of patients with average scores of 3.0 or above.

CoreQ: Short Stay Measure= ([number of valid responses with an average score of =3.0] / [total number of valid responses])*100

13.No risk-adjustment is used.

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

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

No sampling is used. No proxy responses are allowed.

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

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

1.Administer that CoreQ: Short Stay Discharge questionnaire to SNF patients discharged within 100 days of admission and who do not fall into one of the exclusions noted below.

a.Identify that SNF patient is discharged within 100 days of admission

i.Calculate the duration of the SNF stay [MDS discharge date (A2000) - MDS admission date (A1900)] to determine if it is = 100 days.

b.Remove individuals with the following exclusions from the sample:

i.Patients who die: This is recorded in the MDS as Die during stay (A2100 =08)

ii.Patients who were discharged to a hospital, another SNCC, psychiatric facility, Inpatient Rehabilitation Facility, or MR/DD facility (A2100 = 06). This is recorded in the MDS as Discharge to hospital (A2100 = 03); another SNCC (A2100 = 02); psychiatric facility (A2100 = 04); Inpatient Rehabilitation Facility (A2100 = 05); MR/DD facility (A2100 = 06).

iii.Patients with Court appointed legal guardian for all decisions will be identified from nursing facility health information system.

iv.Patients on hospice: This is recorded in the MDS as Hospice O0100K1 = 1 ("the patient was on hospice in the last 14 days while not a resident"), O0100K2 = 1 ("the patient was on hospice in the last 14 days while a resident"), A1800=07 ("entered from hospice"), or A2100=07 ("discharged to hospice").

v.Patients who left the nursing facility against medical advice (AMA) will be identified from nursing facility health information system.

vi.Patients with a BIMS score of 7 or lower. This is recorded in the MDS as C0500 <=7.

2.Administer the CoreQ: Short Stay Discharge questionnaire to patients discharged, within two weeks of discharge (ideally, within one week). The questionnaire should be administered after discharge, not the day of discharge. Optional but not required, reminders or duplicate questionnaires can be administered to patients to help increase response rate.

3.Instruct individuals that they must respond to the survey within two months.

4.Collect the responses continuously for all eligible discharges. The maximum time period for data collection is 6 months. However, a SNF may optionally stop data collection if they consecutively receive =125 usable surveys and calculate the measure.

5.A minimum response rate of 30% needs to be achieved for results to be reported for a SNF.

a.The response rate is calculated as the number of valid returned questionnaires divided by the number of questionnaires administered. Those returned as undeliverable are excluded as well as those completed by another person on behalf of the patient and those with missing data on 2 or more of the 4 questions.

6.Regardless of response rate, SNFs must also achieve a minimum number of 20 usable questionnaires (e.g. denominator). If after 6 month, less than 20 usable questionnaires are received than a facility level satisfaction measure cannot be reported.

7.All the questionnaires that are received (other than those with more than one missing value; or those returned as undeliverable; or those returned after two months; or those completed by another person) must be used in the calculations.

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

If other, please describe in S.18.

Instrument-Based Data

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

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

The collection instrument is the CoreQ: Short Stay Discharge questionnaire and Resident Assessment Minimum Data Set (MDS) version 3.0.

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

Available in attached appendix at A.1

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

Facility

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

Post-Acute Care

If other:

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

Not Applicable

2. Validity – See attached Measure Testing Submission Form

CoreQ_Short_Stay_Testing_Final_-1-.docx,CoreQ_Short_Stay_Testing_Final_v7.1-637202268917480393.docx

2.1 For maintenance of endorsement

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

No

2.2 For maintenance of endorsement

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

No

2.3 For maintenance of endorsement

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

No - This measure is not risk-adjusted

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 2614
Measure Title: CoreQ: Short Stay Discharge Measure
Date of Submission: 1/28/2020

Type of Measure:

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

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

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

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

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	Measure Tested with Data From:
☐ abstracted from paper record	☐ abstracted from paper record
☐ claims	☐ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☒ other: CoreQ: Short Stay Discharge questionnaire	☒ other: CoreQ: Short Stay Discharge questionnaire, Pilot CoreQ: Short Stay Discharge questionnaire, Nursing Home Compare, and CASPER

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

Data utilized for testing came from CoreQ: Short Stay Discharge questionnaire. To validate the measure; we also utilized CASPER Quality Indicators and data form Nursing Home Compare. Additionally, Pilot CoreQ: Short Stay Discharge questionnaire containing an extended list of questions included on the CoreQ: Short Stay Discharge questionnaire was utilized for reliability and validity testing.

1.3. What are the dates of the data used in testing? June 2014 - September 2014

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

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

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

The testing and analysis included four data sources (Table 1.5 below):

1. Reliability and validity testing of the Pilot CoreQ: Short Stay Discharge questionnaire was examined using responses from 853 patients from a national sample of facilities.
2. Validity testing of the Pilot CoreQ: Short Stay Discharge questionnaire was examined using responses from 100 patients from the Pittsburgh area.
3. CoreQ: Short Stay Discharge measure was examined using 282 facilities and included responses from 10,319 patients. These facilities were located across multiple states.
4. In addition, patient-level sociodemographic (SDS) variables were examined using a sample of 1012 patients in nursing facilities in Massachusetts. This included 121 facilities.

Table 1.5: Demographics of Data Sources

Data Source	Average Number of Licensed Beds	Average Daily Census	Average Monthly Number of New Patients	Sample Size of Patients (N)
Source 1	122	112	37	853
Source 2	202	188	49	100
Source 3	135	108	34	10 319
Source 4	140	133	29	1,012

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

Patient Level of Analysis

Data was used from the CoreQ: Short Stay Discharge questionnaire. The questionnaire was mailed to all patients discharged within 2 weeks of their discharge date (with the exclusions described in the Specification section). The testing and analysis included:

1. The Pilot CoreQ: Short Stay Discharge questionnaire was examined using responses from 853 patients from a national sample of facilities.
2. Validity testing of the Pilot CoreQ: Short Stay Discharge questionnaire was examined using responses from 100 patients from the Pittsburgh area.
3. CoreQ: Short Stay Discharge measure was examined using 282 facilities and included responses from 10,319 patients. These facilities were located across multiple states.
4. In addition, patient-level sociodemographic (SDS) variables were examined using a sample of 1012 patients in nursing facilities in Massachusetts. This included 121 facilities.

The descriptive characteristics of the residents are given in the following table that includes information from all of the data used (the education level and race information comes only from the sample described above with 1012 respondents, as this data was not collected for the other samples).

Table 1.6: Descriptive Characteristics of Patients Included in the Analysis (all samples pooled)

DEMOGRAPHICS		Percent
How long were you a resident at this facility?	<1 Month	60.88%
	1-3Months	34.59%
	3-6Months	2.89%
Are you male or female?	Male	39%
	Female	61%
What year were you born?	Average	1936
What is the highest grade or level of school that you have completed?	Some HS	15%
	HS or GED	41%
	Some College/ 2yr Degree	23%
	4yr College Degree	11%
	>4yr College Degree	10%
Are you of Hispanic or Latino origin or descent?	Yes	2%
	No	98%
What is your race?	White	86%
	Black	13%
	Asian	1%
	Native Hawaiian	0%
	American Indian	0%

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

We conducted two levels of testing in the development of the CoreQ: Short Stay Discharge measure. The first focused on testing (e.g., reliability, validity, exclusions) of the CoreQ: Short Stay Discharge questionnaire. The first source of data (pilot data) was utilized in developing and choosing the items to be included in the CoreQ: Short Stay Discharge questionnaire. This included using a questionnaire with 22 items. Below we call this the Pilot CoreQ: Short Stay Discharge questionnaire.

Once the CoreQ: Short Stay Discharge questionnaire was developed, a second source of data was used to test the validity of the CoreQ: Short Stay Discharge measure (i.e., facility and summary score validity).

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

The following patient-level sociodemographic variables were available for analysis. For the distribution of these categories, see Table 1.6 above.

- Age
  - Exact date of birth
- Sex
  - Male
  - Female
- Highest level of education
  - Some high school, but did not graduate
  - High school graduate or GED
  - Some college or 2 year degree
  - 4 year college graduate
  - More than 4 year college degree
- Hispanic Descent
  - Yes
  - No
- Race
  - White
  - Black or African American
  - Asian
  - Native Hawaiian or other Pacific Islander
  - American Indian or Alaskan Native.

2a2. RELIABILITY TESTING

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

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)
☒ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)
☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

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

We measured reliability at the: (1) data element level; (2) the person/questionnaire level; and, (3) at the measure (i.e., facility) level. More detail of each analysis follows.

(1) DATA ELEMENT LEVEL

To determine if the CoreQ: Short Stay Discharge questionnaire data elements were repeatable (i.e. producing the same results a high proportion of the time when assessed in the same population in the same time period) we re-administered the questionnaire to patients 1 month after the submission of their first survey. The Pilot CoreQ: Short Stay Discharge questionnaire had responses from 853 patients; we re-administered the survey to 100 patients. The re-administered sample was a sample of convenience as they represented patients from the Pittsburgh area (the location of the team testing the questionnaire). To measure the agreement, we calculated first the distribution of responses by question in the original round of surveys, and then again in the follow-up surveys (they should be distributed similarly); and second, calculated the correlations between the original and follow-up responses by question (they should be highly correlated).

(2) PERSON/QUESTIONNAIRE LEVEL

Having tested whether the *data elements* matched between the pilot responses and the re- administered responses, we then examined whether the *person-level* results matched between the Pilot CoreQ: Short Stay Discharge questionnaire responses and their corresponding re- administered responses. In particular, we calculated the percent of time that there was agreement between whether or not the pilot response was poor, average, good, very good or excellent, and whether or not the re- administered response was poor, average, good, very good or excellent.

(3) MEASURE (FACILITY) LEVEL

Last, we measured stability of the facility-level measure when the facility's score is calculated using multiple "draws" from the same population. This measures how stable the facility's score would be if the underlying patients are from the same population but are subject to the kind of natural sample variation that occurs over time. We did this by bootstrap with 10,000 repetitions of the facility score calculation, and present the percent of facility resamples where the facility score is within 1 percentage point, 3 percentage points, 5 percentage points, and 10 percentage points of the original score calculated on the Pilot CoreQ: Short Stay Discharge questionnaire sample.

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

(1) DATA ELEMENT LEVEL

Table 2a2.3.a shows the four CoreQ: Short Stay Discharge questionnaire items, and the response per item for both the pilot survey of 853 patients and the re-administered survey of 100 patients. The responses in the pilot survey are not statistically significant from the re-administered survey. This shows that the data elements were highly repeatable and produced the same results a high proportion of the time when assessing the same population in the same time period.

Table 2a2.3.a: CoreQ: Short Stay Discharge Questionnaire Responses from the Pilot and Re-administered Surveys

Questionnaire Item	Response	Pilot Survey (N=853)	Readministered Survey Percent (N=100)
1. In recommending this facility to your friends and family, how would you rate it overall?	Poor	10%	11%
	Average	10%	9%
	Good	15%	13%
	Very Good	33%	35%
	Excellent	33%	33%
2. Overall, how would you rate the staff?	Poor	4%	4%
	Average	10%	10%
	Good	17%	16%
	Very Good	40%	42%
	Excellent	30%	29%
3. How would you rate the care you received?	Poor	5%	5%
	Average	12%	13%
	Good	18%	18%
	Very Good	37%	36%
	Excellent	28%	27%
4. How would you rate the discharge process?	Poor	8%	8%
	Average	12%	13%
	Good	20%	20%
	Very Good	34%	33%
	Excellent	26%	25%

Table 2a2.3.b shows the average of the percent agreement from the first survey score to the second survey score for each item in the CoreQ: Short Stay Discharge questionnaire. This shows very high levels of agreement.

Table 2a2.3.b: Average Percent Agreement Between 1st and 2nd Administered Surveys

Questionnaire Item	Percent Agreement
9. In recommending this facility to your friends and family, how would you rate it overall?	96.8%
10. Overall, how would you rate the staff?	97.8%
11. How would you rate the care you receive?	98.2%
12. How would you rate the discharge process?	98.2%

(2) PERSON/QUESTIONNAIRE LEVEL

Table 2a2.3.c shows the CoreQ: Short Stay Discharge questionnaire items, and the agreement in response per item and responses for both the pilot survey of 853 patients compared with the re- administered survey of 100 patients. The person-level responses in the pilot survey are not statistically significant from the re-administered survey. This shows that a high percent of time there was agreement between whether or not the pilot response was poor, average, good, very good or excellent, and whether or not the re- administered response was poor, average, good, very good or excellent. Table 2a2.3.d shows the average percent agreement between the pilot and re- administered responses In summary, 98% or more of the re-administered responses agreed with their corresponding pilot responses, in terms of whether or not they were rated in the categories of poor or average or good, very good or excellent.

Table 2a2.3.c: Average Percent Agreement between Responses per Item for the Pilot Survey and Re-administered Survey

Questionnaire Item	Response	Percent Person-Level Agreement in Response for the Pilot Survey (N=853) vs. Re-administered Survey (N=100)
1. In recommending this facility to your friends and family, how would you rate it overall?	Poor	96%
	Average	96%
	Good	95%
	Very Good	98%
	Excellent	99%
2. Overall, how would you rate the staff?	Poor	99%
	Average	98%
	Good	98%
	Very Good	96%
	Excellent	98%
3. How would you rate the care you received?	Poor	99%
	Average	99%
	Good	98%
	Very Good	97%
	Excellent	98%
4. How would you rate the discharge process?	Poor	99%
	Average	97%
	Good	98%
	Very Good	99%
	Excellent	98%

Table 2a2.3.d: Average Percent Agreement between Response Options for the Pilot Survey and Re-administered Survey

	Re- administered Response

		Poor (1) or Average (2)	Good (3), Very Good (4), or Excellent (5)
Pilot Response	Poor (1) or Average (2)	98.5%	98%
	Good (3), Very Good (4), or Excellent (5)	98.5%	99%

(3) MEASURE (FACILITY) LEVEL

After having performed the 10,000-repetition bootstrap, 17.82% of bootstrap repetition scores were within 1 percentage point of the score under the original pilot sample, 38.14% were within 3 percentage points, 61.05% were within 5 percentage points, and 87.05% were within 10 percentage points.

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

In summary, the measure displays a high degree of element-level, questionnaire-level, and measure (facility)-level reliability. First, the CoreQ: Short Stay Discharge questionnaire data elements were highly repeatable, with pilot and re-administered responses agreeing between 94% and 97% of the time, depending on the question. That is, this produced the same results a high proportion of the time when assessed in the same population in the same time period. Second, the questionnaire level scores were also highly repeatable, with pilot and re-administered responses agreeing 98% of the time. Third, a facility drawing patients from the same underlying population only varied modestly. The 10,000-repetition bootstrap results showed that the CoreQ: Short Stay Discharge measure scores from the same facility are very stable, given the minimum sample size of 20 we set for this measure; and the maximum sample size of 196.

2b1. VALIDITY TESTING

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

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

☒ **Performance measure score**

☐ **Empirical validity testing**

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

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

In the development of the CoreQ: Short Stay Discharge questionnaire, four sources of data were used to perform three levels of validity testing. Each is described further below. The first source of data (convenience sampling) was used in developing and choosing the format to be utilized in the CoreQ: Short Stay Discharge questionnaire (i.e., response scale). The second source of data was pilot data collected from 865 patients (described below). This data was used in choosing the items to be used in the CoreQ: Short Stay Discharge

questionnaire. The third source of data (collected from 285 facilities described in Section 1.5) was used to examine the validity of the CoreQ: Short Stay Discharge measure (i.e., facility and summary score validity).

Thus, the following sections describe this validity testing:
1. Validity testing of the questionnaire format used in the CoreQ: Short Stay Discharge questionnaire;
2. Testing the items for the CoreQ: Short Stay Discharge questionnaire;
3. To determine if a sub-set of items could reliably be used to produce an overall indicator of satisfaction (Core Q: Short Stay Discharge measure);
4. Validity Testing for the CoreQ: Short Stay discharge measure.

In summary, the overall intent of these analyses was to determine if a subset of items could reliably be used to produce an overall indicator of satisfaction.

1. Validity Testing for the Questionnaire Format used in the CoreQ: Short Stay Discharge Questionnaire

A. The face validity of the domains used in the CoreQ: Short Stay Discharge questionnaire was evaluated via a literature review. The literature review was conducted to examine important areas of satisfaction for long term care residents. The research team examined 12 commonly used satisfaction surveys and reports to determine the most valued satisfaction domains. These surveys were identified by completing internet searches in PubMed and Google. Key terms that were searched included "resident satisfaction, long-term care satisfaction, and elderly satisfaction".

B. The face validity of the domains was also examined using patients. The overall ranking used was 1=Most important and 22=Least important. The respondents were patients (N=40) in five nursing facilities in the Pittsburgh region.

C. The face validity of the Pilot CoreQ: Short Stay Discharge questionnaire response scale was also examined. The respondents were patients (N=40) in five nursing facilities in the Pittsburgh region. The percent of respondents that stated they "fully understood" how the response scale worked, could complete the scale, AND in cognitive testing understood the scale was used.

D. The Flesch-Kinkaid scale (Streiner & Norman, 1995) was used to determine if respondent correctly understood the questions being asked (Streiner, D. L. & Norman, G.R., 1995).

2. Testing the Items for the CoreQ: Short Stay Discharge Questionnaire

The analyses above were performed to provide validity information on the format in the CoreQ: Short Stay Discharge questionnaire (i.e., domains and format). The second series of validity testing was used to further identify items that should be included in the CoreQ: Short Stay Discharge questionnaire. This analysis was important, as all items in a satisfaction measure should have adequate psychometric properties (such as low basement or ceiling effects). For this testing, a Pilot version of the CoreQ: Short Stay Discharge questionnaire survey was administered consisting of 22 items (N= 853 patients). The testing consisted of:

A. The Pilot CoreQ: Short Stay Discharge questionnaire items performance with respect to the distribution of the response scale and with respect to missing responses.

B. The intent of the pilot instrument was to have items that represented the most important areas of satisfaction (as identified above) and to be parsimonious. Additional analyses were used to eliminate items in the Pilot instrument. More specifically, analyses such as exploratory factor analysis (EFA) were used to further

refine the pilot instrument. This was an iterative process that included using Eigenvalues from the principal factors (unrotated) and correlation analysis of the individual items.

3. Determine if a Sub-Set of Items Could Reliably be used to Produce an Overall Indicator of Satisfaction (The Core Q: Short Stay Discharge measure).

The CoreQ: Short Stay Discharge is meant to represent overall satisfaction with as few items as possible. The testing given below describes how this was achieved.

A. To support the construct validity (i.e. that the CoreQ items measured a single concept of "satisfaction") we performed a correlation analysis using all items in the instrument.

B. In addition, using all items in the instruments a factor analysis was conducted. Using the global items Q1 ("How satisfied are you with the facility?") the Cronbach's Alpha of adding the "best" additional item was explored.

4. *Validity Testing for the Core Q: Short Stay Discharge Measure.*

The overall intent of the analyses described above was to identify if a sub-set of items could reliably be used to produce an overall indicator of satisfaction, the CoreQ: Short Stay Discharge questionnaire. Further testing was conducted to determine if the 4 items in the CoreQ: Short Stay Discharge questionnaire were a reliable indicator of satisfaction.

A. To determine if the 4 items in the CoreQ: Short Stay Discharge questionnaire were a reliable indicator of satisfaction, the correlation between these four items in the CoreQ: Short Stay Discharge Measure and all of the items on the Pilot CoreQ instrument was conducted.

B. We performed additional validity testing of the facility-level CoreQ: Short Stay Discharge measure by measuring the correlations between the CoreQ: Short Stay Discharge measure scores and i) measures of regulatory compliance and other quality metrics from the Certification and Survey Provider Enhanced Reporting (CASPER) data, ii) several other quality metrics from Nursing Home Compare, iii) risk adjusted discharge to community measure and iv) risk adjusted PointRight® Pro 30™ Rehospitalizations. If the CoreQ Short Stay Discharge scores correlate negatively with the measures that decrease as they get better, and positively with the measures that increase as they get better, then this supports the validity of the CoreQ Short Stay Discharge measure.

Streiner, D. L. & Norman, G.R. 1995. Health measurement scales: A practical guide to their development and use. 2nd ed. New York: Oxford.

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

1. Validity Testing for the Questionnaire Format used in the CoreQ: Short Stay Discharge questionnaire A. The face validity of the Domains used in the CoreQ: Short Stay Discharge questionnaire was evaluated via a literature review (described in 2b2.2). Specifically, the research team examined the surveys and reports to identify the different domains that were included. The research team scored the domains by simply counting if an instrument included the domain. Table 2b2.3.a gives the domains that were found throughout the search, as their respective score. An example is the domain food, this was used in 11 out of the 12 surveys. (Note: food was not ultimately included in the final CoreQ Short Stay Discharge because correlation and factor analysis showed that it added little to the survey when the overall question, i.e. CoreQ Question 1 was used). An interpretation of this finding would be that items addressing food are extremely important in satisfaction surveys. These domains were used in developing the pilot CoreQ: Short Stay Discharge questionnaire items.

Table 2b2.3.a: Survey Domain Score out of 12

Domain	Score out of 12	Domain	Score out of 12
Food	11	Spiritual	4
Activities	10	Confidence in Caregivers	3
Administration	10	Language and Communication	3
Clinical Care	10	Personal Suite	3
Staff Interaction	10	Therapy	3
Choice and Decision Making	9	Care Access	2
Facility Environment	9	Case Manager	2
Security and Safety	9	Comfort	2
Overall	8	Maintenance	2
Staff Overall	7	Move In	2
Autonomy and Privacy	6	Non-Clinical Staff Services	2
Housekeeping	6	Transitions	2
Personal Care	6	Transportation	2
Recommend facility	6	Emergency Response	1
Resident to Resident Friendships	5	Finances	1
Family Involvement	4	Time	1
Resident to Staff Friendships	4	Trust	1

B. The face validity of the domains was also examined using patients (described above). The following abbreviated table shows the rank of importance for each group of domains. The overall ranking used was 1=Most important and 22=Least important. The ranking of the 4 areas used in the CoreQ: Short Stay Discharge questionnaire are shown in Table 2b2.3.b.

Table 2b2.3.b: Average Ranking of CoreQ: Average Ranking of CoreQ: Short Stay Discharge Questionnaire Items

Domain (Question)	Average Rank
OVERALL (In recommending this facility to your friends and family, how would you rate it overall?)	2
STAFF (Overall, how would you rate the staff?)	1
CARE (How would you rate the care you received?)	3
DISCHARGE (How would you rate how well your discharge needs were met?)	5

C. The face validity of the pilot CoreQ: Short Stay Discharge questionnaire response scale was also examined (described above). Table 2b2.3.c gives the percent of respondents that stated they fully understood how the response scale worked, could complete the scale, AND in cognitive testing understood the scale.

Table 2b2.3.c: Resident Understanding of Response Scale

Scale Format	Residents
Yes – No	100%
Yes – Somewhat – No	100%
Always – Usually – Sometimes –Never	100%
Very happy – Somewhat happy – Unhappy	100%
Excellent – Good – Fair – Poor	100%
Very Good – Good – Average – Poor – Very Poor	100%
Very Satisfied – Satisfied – Neither Satisfied or Dissatisfied – Dissatisfied – Very Dissatisfied	100%
4 Point Satisfaction Scale (1=Very unsatisfied, 2=Unsatisfied, 3=Neutral, 4=Satisfied)	100%
5 Point Likert Scale (1=Poor, 2=Average, 3=Good, 4=Very Good, 5=Excellent)	100%
Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree	95%
5 Point Importance Scale (1=Very important, 5=Very unimportant)	95%
5 Point Expectancy Scale (1=Not met, 2=Nearly met, 3=Met, 4=Exceeded, 5=Far exceeded expectations)	90%
10 Point Satisfaction Scale (1=Poor, 10=Excellent)	90%
8 Point Satisfaction Scale (1=Very dissatisfied, 2=Dissatisfied, 3=Somewhat dissatisfied, 4=Neither satisfied nor dissatisfied, 5=Somewhat satisfied, 6=Satisfied, 7=Very satisfied, 8=No response)	85%

Note: Highlighted cell represents the scale used in the CoreQ.

D. The CoreQ: Short Stay Discharge questionnaire was purposefully written using simple language. No *a priori* goal for reading level was set, however a Flesch-Kinkaid scale score of six, or lower, is achieved for all questions.

2. Testing the Items for the CoreQ: Short Stay Discharge Questionnaire

A. The pilot CoreQ: Short Stay Discharge questionnaire items are shown below. Table 2b2.3.d in the appendix shows that the items performed well with respect to the distribution of the response scale and with respect to missing responses.

B. Using all items in the instruments (excluding the global item Q1 ("How would you rate the facility?")) exploratory factor analysis (EFA) was used to evaluate the construct validity of the measure. The Eigenvalues from the principal factors (unrotated) are presented. Sensitivity analyses using principal factors and rotating provide highly similar findings.

Table 2b2.3.e: Eigenvalues for Principle Factors

Factor	Eigenvalues
Factor 1	9.61

Factor 2	0.37

3. Determine if a Sub-Set of Items could Reliably be used to Produce an Overall Indicator of Satisfaction (The Core Q: Short Stay Discharge Measure).

A. To support the construct validity that the idea that the CoreQ items measured a single concept of "satisfaction" – we performed a correlation analysis using all items in the instrument. The analysis identifies the pairs of CoreQ items with the highest correlations. The highest correlations are shown in Table 2b2.3.f. Items with the highest correlation are potentially providing similar satisfaction information. Note, the table provides 7 sets of correlations, the analysis was conducted examining all possible correlations between items. Because items with the highest correlation were potentially providing similar satisfaction information they could be eliminated from the instrument.

Table 2b2.3.f: CoreQ: Short Stay Discharge Questionnaire Item Correlations

Correlation Rankings (from high to low)	Question Pairs (Correlation Coefficient)
Highest Correlation	Q8-Q6 (.841)
Next highest Correlation	Q10-Q9 (.842)
Next highest Correlation	Q17-Q20 (.822)
Next highest Correlation	Q6-Q2 (.814)
Next highest Correlation	Q15-Q6 (.804)
Next highest Correlation	Q13-Q10 (.814)
Next highest Correlation	Q9-Q2 (.818)

B. In addition, using all items in the instrument a factor analysis was conducted. Using the global items Q1 ("How satisfied are you with the facility?") the Cronbach's Alpha of adding the "best" additional item is shown in table 2b2.3.g. Chronbach's alpha measures the internal consistency of the values entered into the factor analysis, where a value of 0.7 or higher is generally considered acceptably high. The additional item(s) is considered best in the sense that it is most highly correlated with the existing item, and therefore provides little additional information about the same construct. So this analysis was also used to eliminate items. Note, the table again provides 7 sets of correlations, the analysis was conducted examining all possible correlations between items.

Table 2b2.3.g: Secondary Correlation Analysis of CoreQ: Short Stay Discharge Questionnaire Items

Questions used in analysis	Correlation Coefficients
Q1 + Q10	0.94
Q1 + Q6	0.94
Q1 + Q2	0.93
Q1 + Q2 + Q6	0.93
Q1 + Q10 + Q9	0.93
Q1 + Q9 + Q8	0.92
Q1 + Q10 + Q6	0.94

Q1 + Q9 + Q6	0.93

Thus, using the correlation information and factor analysis 4 items representing the CoreQ: Short Stay Discharge questionnaire were identified.

4. *Validity testing for the Core Q: Short Stay Discharge Measure*

The overall intent of the analyses described above was to identify if a sub-set of items could reliably be used to produce an overall indicator of satisfaction, the CoreQ: Short Stay Discharge questionnaire.

A. The items were all scored according to the rules identified elsewhere. The same scoring was used in creating the 4 item CoreQ: Short Stay Discharge questionnaire summary score and the satisfaction score using the Pilot CoreQ: Short Stay Discharge questionnaire. The correlation was identified as having a value of 0.94.

That is, the correlation score between the final "CoreQ: Short Stay Discharge Measure" and all of the 22 items used in the Pilot instrument indicates that the satisfaction information is approximately the same if we had included either the 4 items or the 22 item Pilot instrument.

B. We performed additional validity testing of the facility-level CoreQ: Short Stay Discharge measure by measuring the correlations between the CoreQ: Short Stay Discharge measure scores and i) measures of regulatory compliance and other quality metrics from the Certification and Survey Provider Enhanced Reporting (CASPER) data, ii) several other quality metrics from Nursing Home Compare, iii) risk-adjusted Discharge to Community Measure [NQF# 2858] and iv) risk-adjusted PointRight® Pro 30™ Rehospitalizations [NQF# 2375].This score should be associated with better quality in the SNF. Therefore, we hypothesize that for each facility in the sample there is a positive correlation with other quality indicators.

(i) Relationship with CASPER Quality Indicators

Certification and Survey Provider Enhanced Reporting (CASPER) contains data collected as part of state/federal nursing home inspections. In short, nursing facilities that accept residents with Medicare and/or Medicaid payments are surveyed; this includes most (i.e., 97% [15,000 facilities]) nursing homes in the U.S. The survey process occurs approximately yearly, and includes the recording of many quality characteristics of the nursing home. The most commonly used CASPER quality indicators are restraint use, pressure ulcers, catheter use, antipsychotic use, antidepressant use, antianxiety use, and, use of hypnotics in SNFs.

In addition, when a SNF is determined not to meet a certification minimum standard a deficiency citation is issued. These deficiency citations are also commonly used in the analyses of the quality of SNFs. Approximately 180 deficiency citations exist and are grouped into 16 categories. These 16 categories group similar areas together. They were developed by CMS and have considerable face validity; although, one limitation of using these categories is that they were not defined using empirical estimation (such as factor analysis).

Table 2b2.3.h: CoreQ: Short Stay Discharge Correlation with Quality Metrics

Quality Indicator	Correlation Coefficients with Satisfaction Summary Score	P-Value
Any Deficiency Citations	-0.11	0.07

Physical Restraint Use	-0.01	0.91
Pressure ulcers	-0.22	<0.01
Catheterized	-0.04	0.56
Antipsychotic medications	-0.06	0.32
Antidepressant medications	0.13	0.03
Antianxiety medications	0.08	0.19
Hypnotic medications	0.04	0.46

(ii) Relationship with Nursing Home Compare (NHC) Quality Indicators, Five Star ratings and staffing levels

Nursing Home Compare (NHC) is a nursing home report card. After several years of pilot testing, the Centers for Medicare and Medicaid Services (CMS) released this report card on the world-wide web in November of 2002. Briefly, Nursing Home Compare provides information for facility location, structural factors (such as ownership), and staffing characteristics (such as registered nurse [RN] staffing levels). Most significantly, standardized quality information is presented in what are called Quality Measures (QMs). These are calculated from MDS information.

At the time period of for this study (i.e., 2014) CMS reported on 19 measures – these are called the core Quality Measures. The Quality Measures address specific areas of resident care, 5 are for short-stay residents and 14 are for long-stay residents. Long-stay measures are for those residents staying at a facility 3 months or more and short-stay measures are for residents staying at a facility less than 3 months. The short-stay measures are most pertinent to the CoreQ: Short Stay Discharge questionnaire; therefore, these were used in the analyses. These are the percent of residents: with delirium; with moderate to severe pain; and, with pressure sores.

Nursing Home Compare also uses a five-star rating for facilities. This is based on information from the health inspection, direct care staffing, and the MDS quality measures. A five star facility is the highest score and a 1 star facility the lowest score. With respect to staffing, two measures are used: 1) RN hours per resident day; and 2) total staffing hours (RN+ LPN+ nurse aide hours) per resident day.

Table 2b2.3.i: CoreQ: Short Stay Discharge Correlation with Short Stay Quality Measures, Five Star ratings, and staffing levels

Quality Indicator	Correlation Coefficients with CoreQ: Short Stay Discharge	P-value
Percent of residents with delirium	-0.12	0.30
Percent of residents with moderate to severe pain	-0.14	0.19
Percent of residents with pressure sores	-0.25	0.08
Five-Star rating	0.33	0.07
RN hours per resident day	0.31	0.11

(iii) Relationship with the risk-adjusted Discharge to Community Measure

The risk adjusted Discharge to Community [NQF# 2858] measure determines the percentage of all new admissions from a hospital who are discharged back to the community within 100 days and remain out of any skilled nursing center for the next 30 days. The measure, referring to a rolling year of MDS entries, is calculated each quarter and includes all new admissions to a SNF regardless of payor source. Unsuccessful discharges will result in the resident becoming a long stay resident, which we hypothesize would increase dissatisfaction in SNFs with poor discharge to community rates.

The results of testing for correlation between risk-adjusted discharge to community measure (from 2015q1) and the CoreQ: Short Stay Discharge measure are provided in the table below.

Table 2b2.3.j: Correlation results between the CoreQ: Short Stay Discharge Measure and Risk-adjusted Discharge to Community Measure

CoreQ: Short Stay Discharge Questions	Correlation Coefficients with Risk-adjusted discharge to community measure	P-Value
Q1: In recommending this facility to your friends and family, how would you rate it overall?	-0.05	0.36
Q2: Overall, how would you rate the staff?	-0.16	0.01
Q3: How would you rate the care you received?	-0.12	0.05
Q4: How would you rate how well your discharge needs were met?	-0.10	0.09
CoreQ: Short Stay Discharge summary score	-0.11	0.06

(iv) Relationship with the risk-adjusted PointRight® Pro 30™ Rehospitalizations

PointRight® Pro 30™ [NQF #2375] is an all-cause, risk adjusted rehospitalization measure. It provides the rate at which all patients (regardless of payer status or diagnosis) who enter skilled nursing facilities (SNFs) from acute hospitals and are subsequently rehospitalized during their SNF stay, within 30 days from their admission to the SNF. Individuals who are rehospitalized after admission are much more likely to become a long stay residents. We hypothesize residents would therefore be more dissatisfied on average in SNFs with high short stay resident rehospitalization rates.

The results of testing for correlation between the risk-adjusted PointRight® Pro 30™ Rehospitalizations measure (from 2015q2) and the CoreQ: Short Stay Discharge measure are provided in the table below.

Table 2b2.3.j: Correlation results between the CoreQ: Short Stay Discharge Measure and Risk-adjusted PointRight® Pro 30™ Rehospitalizations Measure

CoreQ: Short Stay Discharge Questions	Correlation Coefficients with Risk-adjusted PointRight® Pro 30™ Rehospitalizations measure	P-Value
Q1: In recommending this facility to your friends and family, how would you rate it overall?	-0.23	<0.001
Q2: Overall, how would you rate the staff?	-0.28	<0.001
Q3: How would you rate the care you received?	-0.24	<0.001
Q4: How would you rate how well your discharge needs were met?	0.31	<0.001
CoreQ: Short Stay Discharge summary score	-0.28	<0.001

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

1. Validity Testing for the Questionnaire Format used in the CoreQ: Short Stay Discharge Questionnaire

A. The literature review shows that domains used in the Pilot CoreQ: Short Stay Discharge questionnaire items have a high degree of both face validity and content validity.

B. Patients overall rankings, show the general "domain" areas used indicates a high degree of both face validity and content validity.

C. The results show that 100% of residents are able to complete the response format used. This testing indicates a high degree of both face validity and content validity.

D. The Flesch-Kinkaid scale score achieved for all questions indicates that respondents have a high degree of understanding of the items.

2. Testing the Items for the CoreQ: Short Stay Discharge Questionnaire

A. The percent of missing responses for the items is very low. The distribution of the summary score is wide. This is important for quality improvement purposes, as nursing facilities can use benchmarks.

B. EFA shows that one factor explains the common variance of the items. A single factor can be interpreted as the only "concept" being measured by those variables. This means that the instrument measures the global concept of satisfaction and not multiple areas of satisfaction. This supports the validity of the CoreQ instrument as measuring a single concept of "customer satisfaction". This testing indicates a high degree of criterion validity.

3. Determine if a Sub-Set of Items Could Reliably be Used to Produce an Overall Indicator of Satisfaction (The Core Q: Short Stay Discharge Measure).

A. Using the correlation information of t*he Core Q: Short Stay Discharge questionnaire (22 items)* and the 4 items representing the CoreQ: Short Stay Discharge questionnaire a high degree of correlation was identified. This testing indicates a high degree of criterion validity.

B. EFA shows that one factor explains the common variance of the items. A single factor can be interpreted as the only "concept" being measured by those variables. This means that the instrument measures the global concept of satisfaction and not multiple areas of satisfaction. This supports the validity of the CoreQ instrument as measuring a single concept of "customer satisfaction". This testing indicates a high degree of criterion validity.

4. *Validity Testing for the Core Q: Short Stay Discharge Measure.*

A. The correlation of the 4 item CoreQ: Short Stay Discharge measure summary score (identified elsewhere in this document) with the overall satisfaction score (scored using all data and the same scoring metric) gave a value of 0.94.

That is, the correlation score between actual the "CoreQ: Short Stay Discharge Measure" and all of the 22 items used in the Pilot instrument indicates that the satisfaction information is approximately the same if we had included either the 4 items or the 22 item Pilot questions.

This indicates that the CoreQ: Short Stay Discharge instrument summary score adequately represents the overall satisfaction of the facility. This testing indicates a high degree of criterion validity.

B.

(i) Relationship with CASPER Quality Indicators

The 8 CASPER quality indicators had a low to moderate level of negative correlation with the CoreQ: Short Stay Discharge measure. Those that correlate have a clear conceptual link with short stay, and those that do not are more associated with long stay residents or have unclear conceptual links to short stay customer satisfaction. The CASPER quality indicators that correlate with the CoreQ Short Stay Discharge score are any deficiency citations (-0.11; p=0.07), pressure ulcers (-0.22, p<0.01) and antidepressants (+0.13, p=0.03); those that do not correlate are physical restraints (-0.01, p=0.91), catheterization (-0.04, p=0.56), antipsychotic medications (-0.06, p=0.32), antianxiety medications (0.08, p=0.19), and hypnotic medications (0.04, p=0.46). This testing indicates a moderate degree of construct validity and convergent validity.

(ii) Relationship with Nursing Home Compare (NHC) Quality Indicators, Five Star ratings and staffing levels

The Nursing Home Compare (NHC) Quality Indicators, Five Star ratings, and staffing levels all had a moderately high levels of correlation and in the direction predicted with the CoreQ: Short-Stay Discharge measure. These correlations range from ± 0.120 to 0.330. The CoreQ: Short-Stay Discharge measure is associated with these quality indicators, and always in the hypothesized direction (good correlates with good). In particular, as emphasized in the structure-process-outcome framework of the evidence section, the link between staffing and customer satisfaction is particularly high, as confirmed by the correlation coefficients 0.330 for RN hours per resident-day and 0.305 for total staffing hours per resident day. This testing indicates a high degree of construct validity and convergent validity.

(iii) Relationship with the risk-adjusted Discharge to Community Measure

The risk-adjusted Discharge to community measure was negatively correlated to the CoreQ: Short Stay Discharge measure. The correlations were small ranging from -0.05 to -0.16. This was not as hypothesized which may be related to some SNFs that specialize in long stay, have very low discharge to community rates as admissions do not have a plan to go home.

(iv) Relationship with the risk adjusted PointRight® Pro 30™ Rehospitalizations

The risk-adjusted PointRight® Pro 30™ Rehospitalizations was negatively correlated to the CoreQ: Short Stay Discharge measure. The correlations were modest ranging from -0.22 to -0.31, and all of them were statistically significant at the p-value of 0.05. This is expected because lower rehospitalization rates (an indicator of high quality) are associated with higher satisfaction. This was as hypothesized. This testing indicates a reasonable degree of construct validity and convergent validity.

As noted by Mor and associates (2003, p.41) "there is only a low level of correlation among the various measures of quality." Castle and Ferguson (2010) also show the pattern of findings of quality indicators in nursing facilities is consistently moderate with respect to the correlations identified. Thus, it is not surprising that "very high" levels of correlations were not identified. Nevertheless, some correlation was identified.

2b2. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section 2b4*

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

To develop the CoreQ: Short Stay Discharge measure, we convened an expert panel to advise us on aspects such as which exclusions to apply to the measure.

Two sources of data were used to examine the exclusions. The first, included responses from 10,319 patients (Section 1.5). The second exclusion analysis included 100 nursing homes that have used the CoreQ: Short Stay Discharge measure in Massachusetts.

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

The expert panel advised us to exclude patients who died, patients who were discharge to a hospital, patients with durable power of attorney for all decisions, patients on hospice, patients with low BIMS scores, and patients who left against medical advice.

These exclusions are often used with satisfaction surveys. Because the exclusions were made we are not able to confirm if the exclusions actually made a difference to the scores, which is why we cannot calculate the mean CoreQ: Short Stay Discharge scores with and without the exclusions. However, we are able to report descriptive statistics regarding the number of exclusions made.

The first, exclusion analysis included responses from 10,319 patients (described elsewhere). The exclusions were tracked and included 1,970 patients (19.1%) discharged to the hospital; 5 (0.05%) discharged to hospice; and, 10 (0.09%) expired. The exclusions of the patients that had left against medical advice or had a durable power of attorney were not tracked in this sample.

The second exclusion analysis included 100 nursing homes and data from the first 1000 patients that were included in this initiative: 791 patients (7.9%) were discharged to the hospital; 48 (0.48%) were discharged to hospice; 41 (0.41%) expired; 23 (0.23%) left against medical advice; and 46 (0.46%) had a durable power of attorney.

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)
These exclusions were applied because such patients were incapable or unlikely to complete a questionnaire (those who died and those who were discharged to the hospital), patients for whom the burden of completing a questionnaire is potentially unethical (hospice patients who are extremely sick), or patients whose answers we could not be confident were accurate or unbiased (durable power of attorney, left against medical advice). The value of excluding these includes burden on respondents and likely distortion of the results.

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

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

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

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

No risk model used.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.
No research (to date) has risk adjusted or stratified satisfaction information from nursing facilities. Testing on this was conducted as part of the development of the federal initiative to develop a CAHPS®[1] Nursing Home Survey to measure nursing home residents' experience (hereafter referred to as NHCAHPS). No empirical or theoretical or empirical risk adjusted or stratified reporting of satisfaction information was recommended as the evidence showed that no clear relationship existed with respect to resident characteristics and the satisfaction scores.

[1]RTI International, Harvard University, RAND Corporation. *CAHPS Instrument for Persons Residing in Nursing Homes,* Final Report to CMS, CMS Contract No. CMS-01-01176, Sept. 2003.

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

"ordering" of risk factor inclusion; for example, are social risk factors added after all clinical factors?
Not Applicable.

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

- ☒ **Published literature**
- ☐ **Internal data analysis**
- ☐ **Other (please describe)**

2b3.4a. What were the statistical results of the analyses used to select risk factors?
Not Applicable

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

Analyses used to examine social (SDS) factors include: (1) the summary score for each of the 4 CoreQ: Short Stay Discharge questionnaire items; (2) the summary score for the CoreQ: Short Stay Discharge measure; and (3) the summary score from the CoreQ: Short Stay Discharge measure (at the facility level).

(1) Summary Score for each of the 4 CoreQ: Short Stay Discharge Questionnaire Items

The summary score for each of the 4 CoreQ: Short Stay Discharge questionnaire items is calculated in the following way: Respondents answering poor are given a score of 1, average = 2, good =3, very good =4 and excellent =5. Correlation and T-test analyses were used to compare the SDS means with each other (Tables 2b4.4b.a1 - 2b4.4b.a4). These analyses show that the individual item scores used in the CORE Q: Short Stay Discharge measure are not significantly different based on either education level or race. That is, the educational makeup of the respondents or the racial makeup of the respondents does not influence the scores for individual items.

TABLE 2B4.4B.A1: MEAN COREQ: SHORT STAY DISCHARGE ITEM DISTRIBUTION BY EDUCATION

What is the highest grade or level of school that you have completed?	**Respondents**	***Q1 Mean***	***Q2 Mean***	***Q3 Mean***	***Q4 Mean***
Some high school, but did not graduate	***10% (n=103)***	3.99	3.96	4.00	3.93
High school graduate or GED	***36% (n=363)***	3.83	4.03	3.99	3.85
Some college or 2 year degree	***25% (n=256)***	3.83	3.94	3.79	3.80
4 year college graduate	***17% (n=175)***	3.81	3.94	3.93	3.94
More than 4 year college degree	***11% (n=114)***	3.99	3.89	3.94	4.01

TABLE 2B4.4B.A2: MEAN COREQ: SHORT STAY RANK CORRELATION BY EDUCATION

	Question 1	Question 2	Question 3	Question 4
Rank Correlation Coefficients	0.01	0.04	0.04	0.04

Rank Correlation of Items with Education: None significant at p=0.05

TABLE 2B4.4B.A3: MEAN COREQ: SHORT STAY DISCHARGE ITEM DISTRIBUTION BY RACE

What is your race?	Respondents	**Q1 Mean**	**Q2 Mean**	**Q3 Mean**	**Q4 Mean**

White	***95% (n=972)***	3.87	3.99	3.94	3.89
Black or African-American	**3% (n=26)**	3.69	3.79	3.77	3.92
Asian	**2% (n=16)**	4.18	4.06	4.01	4.06
Native Hawaiian or other Pacific Islander	**0% (n=0)**	0	0	0	0
American Indian or Alaskan Native	**0% (n=0)**	0	0	0	0

TABLE 2B4.4B.A4: MEAN COREQ: SHORT STAY DISCHARGE ITEM CORRELATION BY RACE

Comparison Groups	Correlation Q1 Mean	Correlation Q2 Mean	Correlation Q3 Mean	Correlation Q4 Mean
White vs. Black or African American	0.43	0.33	0.88	0.41
White vs. Asian	0.27	0.78	0.54	0.5
Black or African-American vs. Asian	0.15	0.43	0.68	0.33

Note: two sample t-tests, none significant at p=0.05

(2) Summary Score for the CoreQ: Short Stay Discharge Measure

The summary score for each of the 4 CoreQ: Short Stay Discharge questionnaire items is calculated in the following way: Respondents answering poor are given a score of 1, average = 2, good =3, very good =4 and excellent =5. For the 4 questionnaire items the average score for the resident is calculated. Correlation and T-test analyses were used to compare the SDS means with each other (Tables 2b4.4b.b1- 2b4.4b.b3). These analyses show that the CORE Q: Short Stay Discharge measure score is not significantly different based on either education level or race of respondents. That is, the educational makeup of the respondents or the racial makeup of the respondents does not influence the measure score.

TABLE 2B4.4B.B1: MEAN COREQ: SHORT STAY DISCHARGE DISTRIBUTION BY EDUCATION

What is the highest grade or level of school that you have completed?	Respondents	*Mean*
Some high school, but did not graduate	***10% (n=103)***	3.96
High school graduate or GED	***36% (n=363)***	3.93
Some college or 2 year degree	***25% (n=256)***	3.84
4 year college graduate	***17% (n=175)***	3.91
More than 4 year college degree	***11% (n=114)***	3.97

Note: Rank Correlation = 0.01, note not significant p=0.05

TABLE 2B4.4B.B2 MEAN COREQ: SHORT STAY DISCHARGE DISTRIBUTION BY RACE

What is your race?	Respondents	*Mean* CoreQ Score
White	***95% (n=972)***	3.92
Black or African-American	**3% (n=26)**	3.76
Asian	**2% (n=16)**	4.01
Native Hawaiian or other Pacific Islander	**0% (n=0)**	0
American Indian or Alaskan Native	**0% (n=0)**	0

Note: Not statistically significant at p=0.5

Table 2b4.4b.b3 Mean CoreQ: Short Stay Discharge Item Correlation by Race

Comparison Groups	Correlation of Mean CoreQ Score
White vs. Black or African American	0.41*
White vs. Asian	0.50*
Black or African-American vs. Asian	0.33*

Note: *Not statistically significant at p=0.5

***(3)* Summary score from the CoreQ: Short Stay Discharge Measure (at the Facility Level).**

The summary score for each of the 4 CoreQ: Short Stay Discharge questionnaire items is calculated in the following way: Respondents answering poor are given a score of 1, average = 2, good =3, very good =4 and excellent =5. For the 4 questionnaire items the average score for the resident is calculated. The facility score represents the percent of residents with average scores of 3 or above. A t-test a**nalysis was** used to compare the mean scores (Table 2b4.4b.c). This analysis demonstrated the CORE Q: Short Stay Discharge measure is not significantly different based on either education level or race. That is, the educational makeup of the respondents or the racial makeup of the respondents does not influence the measure.

TABLE 2B4.4B.C: COREQ: SHORT STAY DISCHARGE SCORE WITH AND WITHOUT ADJUSTMENT FOR EDUCATION

What is the highest grade or level of school that you have completed?	Respondents	Mean Score with Score with Characteristic	Mean Score without Characteristic	Significance
Some high school, but did not graduate	***10% (n=103)***	83.4	83.2	n.s
High school graduate or GED	***36% (n=363)***	83.4	83.1	n.s
Some college or 2 year degree	***25% (n=256)***	83.4	82.9	n.s
4 year college graduate	***17% (n=175)***	83.4	83.1	n.s
More than 4 year college degree	***11% (n=114)***	83.4	83.8	n.s

N.S. = Not significant at p=0.05

Table 2b4.4b.c CoreQ: Short Stay Discharge Score with and without adjustment for Race

What is your race?	Respondents	Mean Score with Characteristic	Mean Score without Characteristic	Significance
White	***95% (n=972)***	83.4	83.3	n.s
Black or African-American	**3% (n=26)**	83.4	83.4	n.s

Asian	**2% (n=16)**	83.4	83.4	n.s
Native Hawaiian or other Pacific Islander	**0% (n=0)**	0	0	0
American Indian or Alaskan Native	**0% (n=0)**	0	0	0

N.S. = Not significant at p=0.05

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)
Not Applicable.

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**
Not Applicable.

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):
Not Applicable.

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:
Not Applicable.

2b3.9. Results of Risk Stratification Analysis:

Not Applicable.

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)
Not Applicable.

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

Not Applicable.

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*
We performed an analysis to examine whether the CoreQ Short Stay Discharge measure captured clinically/practically meaningful differences between providers. We produced a histogram of the scores for the providers in the Pilot CoreQ: Short Stay Discharge questionnaire sample (figure 1b.2).

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)
See histogram below (figure 1b.2) and table 1b.2.d showcasing the distribution of scores

Figure1b.2: Distribution of CoreQ: Short Stay Discharge Measure

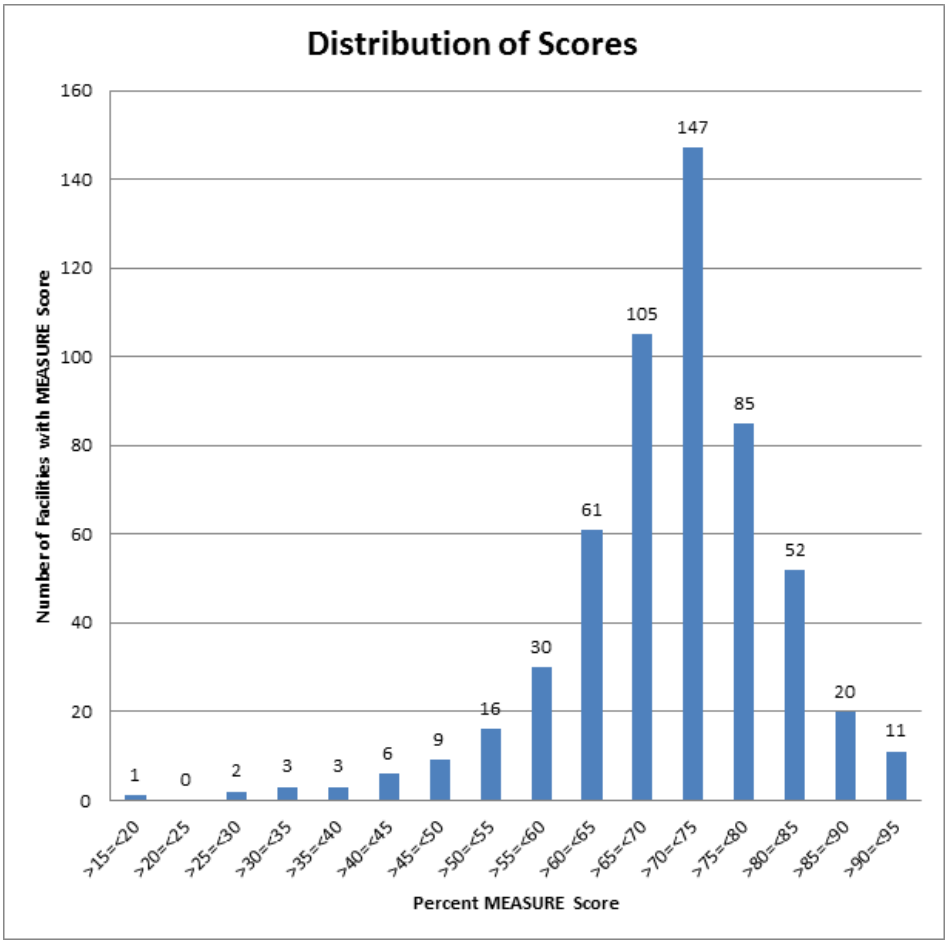


Table 1b.2.d: Overall Descriptive Information for the CoreQ: Short Stay Discharge Measure

Percentile	Measure Score
min	25
p25	75
p50	82.5
p75	88.6
max	100

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i*.e., what do the results mean in terms of statistical and meaningful differences?*)
The CoreQ Short Stay Discharge scores reflect practical and meaningful differences in quality between facilities. The histogram in Section 2b5.2 (figure 1b.2) shows that the distribution of summary scores is quite wide, indicating the scores can be used to differentiate facilities of varying levels of customer satisfaction quality.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*) Not Applicable.

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*) Not Applicable

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*) Not Applicable

---

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)
Four items are used in the CoreQ: Short Stay Discharges questionnaire. In calculating the CoreQ: Short Stay Discharge measure if 1 item of 4 is missing then imputation is used, and if 2 (or more) of the 4 items is missing, the respondent is excluded. The imputation method consists of using the average score from the items answered. The testing to identify the extent and distribution of missing data included examining the frequency of missing responses for each of the 4 CoreQ: Short Stay Discharges questionnaire items and the extent and distribution of missing data for more than one missing response for the items. The method of testing to identify if the performance results were biased included examining the correlation with the quality indicators (described above) when imputation was and was not used.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)
As noted above in section 2b7.1, 4 items are used in the CoreQ: Short Stay Discharges questionnaire. In calculating the CoreQ: Short Stay Discharge measure if 1 item of 4 is missing then imputation is used, and if 2 (or more) of the 4 items is missing, the respondent is excluded. The imputation method consists of using the average score from the items answered. From the testing of 10,319 residents (described elsewhere) we found:

1. In recommending this facility to your friends and family, how would you rate it overall?

- That missing responses occurred in 3.71% (n=383) cases.

2. Overall, how would you rate the staff?
   - Missing responses occurred in 3.54% (n=365) cases.
3. How would you rate the care you receive?
   - Missing responses occurred in 3.9% (n=402) cases.
4. How would you rate how well your discharge needs were met?
   - Missing responses occurred in 5.21% (n=538) cases.

Two (or more) missing responses occurred in 347 cases. Thus, the degree of missing data was very small (=2.4%). Imputation was used in 1341 cases or 12.9% of respondents.

Using the cases with 1 missing value (i.e., those with imputation) the correlation with the quality indicators described above (i.e., restraint use, pressure ulcers, catheter use, antipsychotic use, antidepressant use, antianxiety use, use of hypnotics, and deficiency citations) was unchanged compared to those with no imputation.

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i*.e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

Bias from imputation was minimal. The correlation with the quality indicators described above (i.e., restraint use, pressure ulcers, catheter use, antipsychotic use, antidepressant use, antianxiety use, use of hypnotics, and deficiency citations) was unchanged. When the respondents were removed from the analyses, the average summary scores remained the same.

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

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Other

If other: Satisfaction Survey

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in a combination of electronic sources

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than

electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

Not applicable. In an effort to keep administrative burden low to encourage collection of satisfaction data, which is important in the field, there are no efforts to develop an eCQM.

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

Maintenance of endorsement update:

There have been no reported difficulties. Providers, vendors, patients and family members (or designated party) have liked the fact that it is a short questionnaire. Patients and the family members (or designated party) have expressed appreciation that their satisfaction or lack thereof with the facility is being measures.

From initial endorsement:

Since the CoreQ: Short Stay Discharge measure has been created and utilized in testing and quality improvement, we have modified it in the following ways.

Additionally, we examined how frequently facilities could administer the questionnaire and the impact of waiting longer periods. We recommend that a facility administer the questionnaire weekly (but up to 2 weeks after patient discharge). The facility operating systems are able to generate patient records after these intervals (i.e., 1 week and 2 weeks). Furthermore, it is advantageous if administered weekly as we identified an increase in response rate of approximately 8%. Moreover, this time period is optimal in order to minimize recall bias. Therefore, this recommendation was incorporated into the measure specifications (given above).

We conducted analyses on allowing up to 2 months for a patient to respond. We identified the average (modal) response to occur within 2 weeks. A few responses were still received 6 weeks after administration, however, by 2 months the response was very much lower (<5% of additional returned surveys). Furthermore, in order to ensure that this time frame did not bias the type of responses captured, we analyzed the average score for the surveys returned. We found that the average scores for surveys returned in the first month were almost identical to those returned in the second month. Thus, this recommendation was incorporated into the measure specifications (given above).

We examined the effect of the 6 month survey completion time period on a facility's ability to collect the survey data. Even the largest nursing facilities need an extended period of time to achieve the 20 minimum sample size identified above. We identified that a majority of nursing facilities (i.e., 90%) in our sample could achieve this response rate if given up to 6 months. Therefore, this recommendation was incorporated into the recommendations (given above).

We conducted analyses on collecting data from residents discharged to the hospital. We identified that patients discharged to the hospital did not have high response rates (i.e., 1 out of 25 were returned). Therefore, discharge to an acute care hospital became an exclusion criterion.

Furthermore, we decided that once 125 consecutive responses are received for a particular facility, it is optional to stop the collection prior to the 6 month period and calculate the measure, because past this mark, no additional information effects the SNFs satisfaction score. Moreover, at 125 responses, the confidence interval shrinks, increasing the certainty of the CoreQ: Short Stay Discharge questionnaire as capturing the true population customer satisfaction.

As part of the CoreQ: Short Stay Discharge measure development, existing satisfaction vendors were contacted (including MyInnerView, Symbria, and NRC) for input on the administration and sample selection used. With respect to administration, the 2 month window used for including returned surveys and the 2 week period from discharge to administer the survey were viewed positively and are currently standard time periods used in the industry. With respect to the sample selection, the exclusion criteria (i.e., Patients who die; patients who were discharged to a hospital, another SNCC, psychiatric facility, Inpatient Rehabilitation Facility, or MR/DD facility; patients with Court appointed legal guardian for all decisions; patients on hospice; patients who left the nursing facility against medical advice) were well received by these vendors. In many cases most of these sample selection criteria are already used by the vendors. Also, with respect to the sample selection, the use of the MDS to capture the sample selection criteria (above) were well received by these vendors.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

No fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, and algorithm) exist.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

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

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)

	Professional Certification or Recognition Program AHCA Quality Initiative https://www.ahcancal.org/quality_improvement/qualityinitiative/Pages/default.aspx AHCA Quality Awards https://066b40b5535506586917-68298049b65edbd7ec9f493f0b1c8eb3.ssl.cf2.rackcdn.com/ahca_1ecb9d979e9f049b2382b029da472a1c.pdf Quality Improvement (external benchmarking to organizations) AHCA NCAL Long Term Care Trend Tracker https://www.ahcancal.org/research_data/trendtracker/Pages/default.aspx Quality Improvement (Internal to the specific organization) Large Nursing Home Chain N/A Brookside Inn in CO https://www.youtube.com/watch?v=V5OcpyJDUkQ

4a1.1 For each CURRENT use, checked above (update for <u>maintenance of endorsement</u>), provide:

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

Quality Awards

- Name: National Quality Award Program
- Purpose: This Baldrige-accredited award program seeks to recognize and help skilled nursing facilities and assisted living communities implement the Baldrige framework for performance excellence. The Baldrige framework emphasizes building systems and using data to understand and meet customer and patient needs. As such, silver and gold recipients must submit benchmarked patient satisfaction data in their application, and CoreQ is one of the acceptable measures. Regardless if an applicant receives an award, they receive a feedback report on their application, which discusses their approach, including deployment, learning, and integration of said approach, in understanding and meeting their patient needs. The goal of this feedback report is to help them improve their processes which would eventually lead to higher patient satisfaction and CoreQ scores.
- Geographic Area: There are currently 697 active silver quality award recipients and 44 gold recipients across the country. All recipients are listed online.

LTC Trend Tracker

- Name: LTC Trend Tracker
- Purpose: Provide an online tool for SNF and AL members to trend and benchmark their performance on CoreQ. In addition to have reports where users can login to access, there are PDF publications that are pushed out to users via email with data for their specific facility, including CoreQ. These publications include CoreQ run charts to show their trend over time and whether they have met the Quality Initiative Goal of a 10% improvement since 2017 or achieved a high performance rate of greater than 90%.
- Geographic Area: All of the approximately 10,000 SNF members and 4,000 AL members have access to LTC Trend Tracker and thus these reports and publications.

The CoreQ: Short Stay Discharge measure is currently in use by a large nursing home chain for the purposes of quality improvement.

In addition, Massachusetts Senior Care is currently using the Measure for quality improvement. A total of 150 facilities in Massachusetts are collecting satisfaction data using of the CoreQ: Short Stay Discharge

questionnaire. The CoreQ: Short Stay Discharge measure is calculated and distributed in a report card to each participant.

Furthermore, 27 national satisfaction vendors in the SNF and AL area have agreed to add the CoreQ to their questionnaires and calculate the measure. This is an increase from 10 vendors a few years ago. The following customer satisfaction vendors are using CoreQ (updated list is also found here: http://coreq.org/)

- Align
- A Place For Mom
- Blue Sky Creative
- Brighton Consulting Group
- Care Analytics
- Cortex Health Inc.
- The Doug Williams Group, Inc.
- Healthcare Academy (ReadyQ)
- Holleran
- inQ Experience Surveys
- Lighthouse Care Updates
- Market Research Answers (CareSat)
- Nexus Health Resources, Inc
- NRC Health
- Pinnacle
- Providigm/abaqis
- Qblue Surveys, Inc.
- Qualtrics
- Reputation.com
- Senior Living Alliance
- Sensight Surveys
- Service Trac
- Simplus Surveys
- Sky Care Media
- Sperduto & Associates, Inc.
- Symbria
- Viewpoint 2 Quality

We do not have counts of patients being surveyed and geographical representation from the vendors, however they represent the majority of customer satisfaction vendors currently doing SNF business in the United States.

In 2019, AHCA/NCAL began sharing reports with vendors on the data they have uploaded to LTC Trend Tracker on behalf of their client SNFs and ALs. The purpose of these reports was to show them trends on how many data submissions meet measure requirements, such as sample size and response rate. These reports were in addition to conference calls with the vendors to discuss best practices and potential issues with meeting CoreQ measure requirements.

A letter has been sent to all 10,000 AHCA SNF members indicating which vendors to date have agreed to add the CoreQ to their questionnaire and calculate the measure (see attached letter in appendix, section 4.a.1). A user's manual has been developed and is available on AHCA's website for all satisfaction survey vendors to use, in addition to the measures' specific website: http://coreq.org/.

AHCA and NCAL have also incorporated the CoreQ into their national Quality Initiative goals. AHCA represents nearly 10,000 of the 15,000 SNFs and provides feedback to all of its members on their satisfaction scores using the CoreQ. This has resulted in growing number of members and vendors collecting the data.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

Not applicable, see 4a1.1.

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

AHCA NCAL is in the midst of its the third Quality Initiative, laying out a series of quality improvement and reporting goals for the AHCA membership, which covers approximately 10,000 of all SNFs in the U.S. Among these goals is the improvement of both long-stay measures and the short-stay measure by 10% (baseline rate: 2017Q1), or achieving a rate of >90% by March 2021 (https://www.ahcancal.org/quality_improvement/qualityinitiative/Pages/Customer-Satisfaction.aspx).

Because it has been included in the Quality Initiative 2018-2021, AHCA's machinery for publicizing and encouraging the adoption of the tool has been activated, including AHCA's quality division spending a large number of staff hours working to accomplish this. Part of these initiatives are providing semi-annual quality initiative feedback reports through the LTC Trend TrackerSM to all the providers submitting data. The LTC Trend TrackerSM CoreQ report and upload feature within LTC Trend Tracker includes an API for vendors performing the survey on behalf of SNFs or individual users so that the aggregate CoreQ results will be immediately available to providers as they are collected. Given that LTC Trend TrackerSM is probably the leading method for SNFs to profile their quality and other data, the incorporation of CoreQ into LTC Trend Tracker means it has become the de facto standard for customer satisfaction surveys for the SNF industry. This is evident by having 7 large national satisfaction vendors in the SNF area now use the CoreQ in their questionnaires and calculate the measure.

We also continue working with states who require satisfaction measurement to incorporate the CoreQ into their process. In the state of New Jersey, the Long-stay resident and family measures are part of the Department of Human Services Quality Incentive Payment Program (https://www.spb.nj.gov/humanservices/doas/documents/NF%20Quality%20Incentive%20Payment%20Program%20October%202019%20Final.pdf), but we do not have any data on the count or the geographic distribution of this. The state of Massachusetts has included the CoreQ short stay as part of its current ongoing quality improvement initiatives on measuring satisfaction in SNFs. Tennessee uses the resident and family long-stay measures as part of their Medicaid quality incentive program, to align Medicaid payments with person-centered care initiatives. Similarly, the state of Georgia also uses these measures as part of the SNF Medicaid Quality incentive payment program. Many other states which are in the midst of developing or updating their Medicaid quality incentive payment programs are considering incorporating the long-stay CoreQ measures. AHCA has a presence in each state, and our state affiliates will be promoting the use of the CoreQ in those states that are collecting or considering collecting satisfaction.

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

We provide these via different channels that will be discussed at length in 4a2.1.2. In short, these channels include: the CoreQ website (http://www.coreq.org/), Long-Term Care Trend Tracker (LTCTT, https://www.ahcancal.org/research_data/trendtracker/Pages/default.aspx), Quality Initiative Publications, Top-Line Publications, and various stakeholder meetings and advocacy efforts.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

Assistance with data and generally understanding the measure is provided through the open-source measure website (http://www.coreq.org/) where the public can find the manual (containing the measure algorithm) as well as the participating vendors (with their direct contacts).

Feedback on the performance results and data is provided via a quarterly push report, called Top-Line, sent to all members who have access to the Long Term Care Trend Tracker (LTCTT). They get an email notification when the data is updated. AHCA has held many webinars for membership on how to better understand all the data components. Further, if the providers being measured need assistance, there is a mailbox (LTCTT-specific) included in the push report, where providers may and do directly contact. We have also set up many FAQs and resources in the adult-learning website https://educate.ahcancal.org/

On LTCTT, we also have set up a CoreQ report which is updated quarterly directly being fed data from the API that vendors use to upload CoreQ measure scores. Therefore, those that are choosing to participate in this, will automatically see their results on the benchmarking tool. The tool is also interactive in allowing the user to see scores over time and allow them to benchmark themselves against whatever relevant peer is most pertinent (state, nation, MSA).

Further, because the CoreQ measures are part of the quality Initiative, all members and those who have access to LTCTT receive push semi-annual Quality Initiative Publications. Here too they can see run charts of their performance over time, as well as their state average and state rank for the published time period. Likewise, in this publication, we also provide different resources, two of which are the mailbox for LTCTT and a Quality Initiative-only mailbox, with designated staff continuously monitoring them.

All those that enter data or have vendors enter their data, obtain this feedback and resources.

Finally, in presentations with state affiliates and any other advocacy efforts (such as assistance with developing Medicaid quality incentives), we continue to advocate for the use of CoreQ.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

We are not sure what section 4d.1 refers to as it is not in this page, but in the even that you meant 4b.1, the following is the response:

Because all the data in 4b.1. is from providers who either have individually or through their vendor submitted data that meets the specifications for measurement, they are able to see their own performance (current and over time) via LTCTT (data source in 4b.1.). Additionally, because LTCTT allows for setting peer comparisons for benchmarking, these providers can also see data for whatever region they specify as pertinent (i.e. state, MSA, nation).

4a2.2.2. Summarize the feedback obtained from those being measured.

Because we do not administer the instruments, we have no way of measuring the feedback collected by residents and family members (or designated parties). It is the vendors who informally collect this feedback, but in favor of reducing the burden that patients and family members (or designated parties) may face, they keep the formal solicited CoreQ survey to only the tested questions. Therefore, any such feedback has been informal. With regards to providers, we provide feedback via the channels mentioned in 4a2.1.1. They do not see an administrative burden since most of them have these instruments administered by vendors and calculated by vendors and LTCTT. For those who cannot afford a vendor, Dr. Nick Castle (one of the developers of the measure) has volunteered to administer and collect their satisfaction measure.

4a2.2.3. Summarize the feedback obtained from other users

Other users such as the state agencies involved in the administration of the Medicaid quality incentive program as well as the quality initiative programs have lauded the small number of questions that are included in this measure. This is especially important as more states move to Medicaid value-based care with a person-centered care model.

Additionally, to maintain transparency and increase awareness of the measures, we have presented abstracts and posters on all 5 CoreQ measures (SNF and AL) at various professional conferences: Academy health (2019),

International Association of Gerontology and Geriatrics World Congress (2017), Gerontological Society of America (2019).See below:

Castle, N., Schwartz, L., Gifford, D. (2019, November) Using a Universal Satisfaction Score in Long-Term Care Settings. Paper presented at the annual meeting of the Gerontological Society of America, Austin TX.

Schwartz, L., Castle, N., Domi, M., Gifford, D. (2019, June). CoreQ: Development of a Universal Satisfaction Score for SNF and AL Residents and Families. Poster session presented at the Academy Health Annual Research Meeting, Washington, DC.

Castle, N. & Schwartz, L. (2017, July) Development of a Universal Satisfaction Score for Long-Term Care Facilities. Poster presented at the World Congress of the International Association of Gerontology and Geriatrics.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

Because feedback has been largely positive, we have not had a reason to believe that these measures needed to be modified.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

We have been actively monitoring improvement of membership because it has been part of the Quality Initiative (https://www.ahcancal.org/quality_improvement/qualityinitiative/Pages/default.aspx). Members need to improve the measures by 10% from their baseline of 2017Q1 score, or achieve a score of more than 90% (satisfaction rate) by March 2021. The current iteration of the quality initiative kicked off in 2018. On average, approximately 10-12% of membership submits CoreQ data every quarter. Thus far, on average, 18% of the providers who nationally submitted data have met this goal for at least one of the CoreQ measures (SS discharge, or LS residents, or LS family). This represents facilities across all 50 states plus the District of Columbia.

With regards to CoreQ short stay discharge, 8% of the total facilities in the nation that have submitted CoreQ data have met the goal of either having >90% satisfaction rate or at least a 10% improvement in the satisfaction rate since 2017Q1. The states of IA, NH, NJ, NM, and RD have at least 20% of the SNFs that submitted data meet the quality initiative goal. MD, CO, MI, MT, and WV have at least 15% of the SNFs that submitted data meet the quality initiative goal. All but two states had at least a facility meet the quality initiative goal. We provide all facilities that submit this data and the state affiliates with their progress on semi-annual push reports.

These are promising results, but we believe that these numbers would be higher if the CoreQ was mandated to be collected from federal initiatives. At the very least we would see a higher number of providers submitting data.

4b2. Unintended Consequences

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

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

There were no negative consequences to individuals or populations identified during testing or evidence of unintended negative consequences to individuals or populations reported since the implementation of the CoreQ: Short Stay Discharge questionnaire or the measure that is calculated using this questionnaire.

This is consistent with satisfaction surveys in general in nursing facilities. Many other satisfaction surveys are used in nursing facilities with no reported unintended consequences to patients or their families.

There are no potentially serious physical, psychological, social, legal, or other risks for patients. However, in some cases the satisfaction questionnaire can highlight poor care for some dissatisfied patients, and this may make those patients further dissatisfied.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

A large vendor has reported patients and family members writing in the margins of the survey that they appreciate being asked about their satisfaction in a short questionnaire with the nursing home and the care provided, however, we don't systemic way of capturing this data. Another thing that we are noticing is the peer-effect in nudging the submission of data. Because the data is submitted on LTCTT allowing for providers to set benchmarks and peer comparisons, providers are nudging others in their larger organizations to submit data. Further, they continue to advocate for it through their local channels (state affiliates and conferences), because the more providers submit measurement data, the more robust their peer comparisons will be in LTCTT.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

No

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

5.1b. If related or competing measures are not NQF endorsed, please indicate measure title and steward. 5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications harmonized to the extent possible?

No

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

The CoreQ: Short Stay Discharge measure does not conceptually address either the same measure focus or the same target population as any other NQF-endorsed measures. The CAHPS® Nursing Home Survey of discharged nursing home residents' experiences (NHCAHPS-D) received provisional status by NQF 3 years ago. However, this provisional status has expired. The NHCAHPS-D remains to be fully developed. The NHCAHPS-D is not widely used because little information exists on its content or reliability since there was insufficient sample size of respondents in the initial testing to finalize the instrument.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Not Applicable

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment: CoreQ_Short_Stay_Appendix_Final_updated_Jan2020-637136665931248726.docx

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): AHCA/NCAL

Co.2 Point of Contact: Courtney, Bishnoi, cbishnoi@ahca.org, 202-898-2857-

Co.3 Measure Developer if different from Measure Steward: American Health Care Association

Co.4 Point of Contact: Marsida, Domi, mdomi@ahca.org, 202-898-2807-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

The workgroup gave input, reviewing our suggested administration, required response rate, the manual, and exclusions.

Mary Tess Crotty, Genesis - Also helped provide feedback on the development process and the user manual. Additionally, she reviewed the analyses.

Matt O'Connor HCR Manor Care- Also helped provide feedback on the development process and the user manual. Additionally, he conducted some analyses and reviewed the analyses.

Judy Hoff, Health Care Academy

Rich Kortum, My Innerview/National Research Corporation

Peter Kramer, abaqis/Providigm

Ellen Kuebrich, abaqis/Providigm

Michael Johnson, ServiceTrac

Chris Magelby, Pinnacle

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2015

Ad.3 Month and Year of most recent revision: 10, 2015

Ad.4 What is your frequency for review/update of this measure? Annually

Ad.5 When is the next scheduled review/update for this measure? 03, 2020

Ad.6 Copyright statement: None

Ad.7 Disclaimers: None

Ad.8 Additional Information/Comments: None