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 #: 0674

Corresponding Measures:

De.2. Measure Title: Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay)

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

De.3. Brief Description of Measure: This measure reports the percentage of long-stay residents in a nursing home who have experienced one or more falls resulting in major injury (defined as bone fractures, joint dislocations, closed head injuries with altered consciousness, or subdural hematoma) reported in the look-back period no more than 275 days prior to the target assessment. The long stay nursing home population is defined as residents who have received 101 or more cumulative days of nursing home care by the end of the target assessment period. This measure is based on data obtained through the Minimum Data Set (MDS) 3.0 OBRA, PPS, and/or discharge assessments during the selected quarter(s).

1b.1. Developer Rationale: This outcome-based quality measure reports the percentage of long-stay nursing home residents who have experienced one or more falls resulting in major injury, which include bone fractures, joint dislocations, and closed-head injuries. Injurious falls are important to monitor in the nursing home population because of their impacts on health outcomes, as research has demonstrated injurious falls are the leading causes of disability and death for nursing home residents. Falls with major injury also impact resident quality of life by introducing new functional limitations and psychosocial distress, while potentially influencing providers to increase the use of unwanted physical or chemical restraints. Studies have shown that risk for falling is associated with a variety of resident characteristics, including but not limited to, increasing age, being female, and cognitive decline. Falls are also associated with inappropriate or changing prescriptions. The capacity of nursing homes to provide residents appropriate living accommodations/amenities and sufficient support by appropriately qualified staff around-the-clock to serve the medical needs of residents can mitigate fall risks and prevent or reduce falls.

S.4. Numerator Statement: The numerator is the number of long-stay residents with one or more look-back scan assessments that indicate one or more falls that resulted in major injury.

S.6. Denominator Statement: The denominator consists of all long-stay nursing home residents with one or more look-back scan assessments except those who meet the exclusion criteria.

S.8. Denominator Exclusions: A resident is excluded from the denominator of this quality measure if all look-back scan assessments indicate that data is missing from the data element assessing falls resulting in major injury during the look-back period preceding the target assessment.

De.1. Measure Type: Outcome

S.17. Data Source: Assessment Data

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Mar 03, 2011 **Most Recent Endorsement Date:** Dec 09, 2015

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? This measure is not paired/grouped.

Preliminary Analysis: Maintenance of Endorsement

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

Criteria 1: Importance to Measure and Report

1a. Evidence

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

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

Evidence Summary with Updates:

The developer provides an update to the evidence on this measure, citing additional studies were referencing the following points:

- Injurious falls are important to monitor in the nursing home population because of their impacts on health outcomes. Injurious falls are the leading causes of disability and death for nursing home residents. Falls with major injury also impact resident quality of life by introducing new functional limitations and psychosocial distress, while potentially influencing providers to increase the use of unwanted physical or chemical restraints.
- Some nursing home residents are at higher risk for experiencing falls, as certain resident characteristics and care-related factors influence the rate of falls in a facility.

- Falls are also associated with inappropriate or changing medications. Polypharmacy is a major risk factor for falls in the nursing home population (Ambrose, Paul, & Hausdorff, 2013; Damian et al., 2013; Seppala et al., 2018).
- Several nursing home characteristics may influence the risk for experiencing a fall with major injury, including adequate staffing levels, staff education, and adequate levels of facility equipment, such as accessible computers used to complete assigned falls prevention tasks (Vlaeyen et al., 2017).

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:**

Question for the Committee:

- *Is there at least one thing that the provider can do to achieve a change in the measure results?*

Guidance from the Evidence Algorithm

Box 1 (Health outcome)? – yes -> Box 2 (One or more action?) yes -> PASS

Preliminary rating for evidence: ☒ Pass ☐ No Pass

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

Maintenance measures – increased emphasis on gap and variation

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

Data were drawn from Q2 2019 from a national source, 14,286 facilities (94%) and 1,012,706 (98%) of residents that met inclusion criteria. The facility-level mean score was 3.4% and the median score was 2.9%. The standard deviation was 2.9%, the minimum was 0%, and score at the 90th percentile was 7.1%. The interquartile range for this measure was 3.6%, indicating some room for improvement in this measure. Of the facilities with adequate sample size to report, 19.0% had perfect scores of 0.

Disparities

The data were examined to assess for disparities by age, race, and socioeconomic status. Notably all three factors demonstrated significant relationships in performance at the facility-level.

Age: the developer reports that residents aged 85 years or older are at higher risk for experiencing falls with major injury than residents less than 85 years of age.

Race the developer reports that the White only population (3.73%) is at higher risk for experiencing falls with major injury than the non-White only population (1.8%).

Socioeconomic status the developer reports that that there is a statistically significant relationship between Medicaid enrollment and QM score for this measure.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?
- Based upon the observed differences by age, race, and SES, should this measure be risk-adjusted?

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.

- Evidence appears appropriate
- The evidence supporting falls as an important quality measure is strong, with both evidence of associated harms and potential avoidability based on facility and medical factors. The evidence has been updated by the developers and supports the measure.
- This outcome measure was last endorsed in 2015. In this submission, CMS provides updated evidence demonstrating the need to continue this measure to protect vulnerable elderly population of long-term care residents.
- Injurious falls are important to monitor in the NH population due to impact on health outcomes. Research has demonstrated injurious falls are a major cause of disability and death for NH residents.
- Evidence is Pass.
- pass
- Are the look back scan assessments randomly selected?
- no issues
- Nursing home specific measure for long term stay patients

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?

There is a performance gap. Age, race and SES were examined for disparities. Older residents, white residents and those of lower SES were more at risk

Moderate spread in the data indicating opportunities for improvement, though 1 in 5 facilities had 0 falls. Disparities noted in care that could rise to clinically meaningful differences by age and SES. Risk adjustment should be considered.

The facility data from Q2 2019 show a variation in performance on the measure. The mean facility-level score was 3.4% and the interquartile range was 3.6 percentage points, which indicates a performance gap for improvement. Disparities were examined by age, race and socioeconomic status. It finds residents aged 85 years or older have a higher risk of falls with major injury than residents less than 85 years old; White residents experienced higher risk falls with major injury than the non-White residents population; and there is a statistically significant relationship between Medicaid enrollment and QM score for this measure.

The risk of falls varies among different nursing homes due to resident variations in medical management, staffing, etc.

High - existing performance gap moderate

There are clear disparities. I do not favor risk adjustment.

reasonable gap

moderate evidence for performance gaps - significant disparities in age, race, socioeconomic status

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

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

Reliability

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

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

Validity

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

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

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

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

Evaluators: NQF Scientific Methods Panel Subgroups

Methods Panel Review (Combined)

Methods Panel Evaluation Summary:

This measure was reviewed by the Scientific Methods Panel and discussed on the call. A summary of the measure and the Panel discussion is provided below.

Reliability

Data element reliability was established by assessing agreement between gold-standard nurse abstractor and facility nurse abstractor to ensure critical data element can be uniformly ascertained across facilities. For performance score reliability, the developer calculated facility specific signal-to-noise reliability scores. There were concerns by an SMP member that the data were old for data element reliability testing (15 years old).

Data element: The Kappa for gold-standard to gold-standard on the MDS 3.0 item was 0.967, and the Kappa for gold-standard to facility-nurse agreement on the MDS 3.0 item was 0.945. Ratings of 0.967 and 0.945 are considered "substantial agreement."

Measure-score: The average signal-to-noise reliability score was 0.45, but with 19% of facilities achieving a perfect score of 0.0%, this results in lower variation of scores between facilities. There were concerns by some SMP members that this was low.

Validity

Performance score validity was established by correlations with other measure of nursing home quality. These included related MDS Quality Measures and Facility Five Star Ratings. Variations between states, seasonality, and stability of the measure scores was assessed.

Found low, but positive correlations between facility performance on this measure and other quality measures. Almost all of the correlation values fell below 0.1.

Questions for the Committee regarding reliability:

Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?

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

Questions for the Committee regarding validity:

Do you have any concerns regarding the validity of the measure (e.g., exclusions, lack of risk-adjustment approach, etc.)?

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

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

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

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?

No concerns

Reliability was moderate, but developers noted 1/5 having 0% scores decreased signal-to-noise ratio. I have minimal concerns about this given other measures of spread and reliability present. Because of focus on falls in nursing homes, this is expected. I do wonder if BMI risk adjustment has any concerns, given poor assessment of this in nursing homes (hard to accurately determine height/weight for some residents).

No Concerns.

residents older than 85 are at higher risk of fall with major injury as are white vs. non-white population

Data supporting reliability was 15 years old. Some other concerns as raised by the methods panel, but agree overall reliability is "moderate". One question is how the develop chose the dates used in the definition - i.e. 275 days look back for falls and long term resident = 101 days staying in the facility?

no concerns

Acceptable

reasonable

panel determine moderate reliability

2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure; reliability testing and results for the measure?

no concerns

No

No concerns. While one Scientific Methods Panel (SMP) member expressed concerns over using data that was 15 years old for reliability testing, the SMP is stratified with the reliability testing, and preliminary rating is moderate.

No

No concerns

No concerns

No

No

panel determine moderate reliability

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

No concerns

Minor concerns with low correlations with other MDS measures, especially functional decline.

The SMP passes the validity, ranging from low to high. The preliminary rating is moderate. No concerns.

No

No concerns, agree with "moderate"

No concerns

No

No

Panel determine moderate validity

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 concerns

No

"Overall, 22.7% of facilities scored significantly different than the national mean in FY2019 Quarter 2, indicating that there are meaningful differences in facility-level scores for this measure even though the measure may not differentiate well in the middle of the distribution of facilities." According to the national facility-level scores provided by the developer, the measure provides the opportunity to identify nursing homes that are both particularly good performers, as well as those that are particularly poor performers regarding injurious falls. But, the quality measure scores vary enough from the national mean that there are meaningful differences in facility-level scores for this measure. It also appears that there are not enough missing data to cause concerns to the validity of this measure.

no

No concerns.

No

No

None

panel determine moderate validity

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?

- No risk adjustment
- Exclusions appear appropriate. Risk adjustment was explored, but developers argued that falls as a never event should increase focus on complete prevention.
- Risk adjustment is not applied in this measure, for the developer's risk models do not have sufficient predictive ability. But, more importantly, the developer rightly argues that "a fall resulting in a major injury is considered to be a never event, which is a serious and costly error in the provision of healthcare services that should never happen. Therefore, the importance of this quality measure is in identifying providers with higher frequency of patients experiencing falls with major injury, which contribute to burden on the healthcare system through an increase in negative health outcomes, healthcare utilization, and costs that remain pertinent regardless of provider risk." So, I think no risk adjustment is justified.
- Moderate reliability and validity found by scientific panel
- I share the same question with one of the members from the methods panel. To say that we will not risk adjust for a "never" event makes sense, but how is this measure different than other measures like surgical mortality where we do risk adjustment and seems just as qualified to be a "never" event as a fall. In reality, not all falls can be prevented in the best of circumstances.
- no concerns
- Acceptable results without risk adjustments- my opinion.
- no issues
- Not sure this was risk adjusted

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.

ALL data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS)

The general data collection method for the MDS 3.0 is currently in operational use and mandatory for all Medicare/Medicaid certified nursing facilities.

Questions for the Committee:

Are the required data elements routinely generated and used during care delivery?

Are the required data elements available in electronic form, e.g., EHR or other electronic sources?

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?

- No concerns
- Extractions from MDS and EMR data; no concerns
- No concern and feasibility is rated high.
- this data is routinely generated
- No concerns.
- High feasibility
- None
- Feasible
- High feasibility - built into the MDS system

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 evaluates 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**

Accountability program details

- Public Reporting:
  - Care Compare https://www.medicare.gov/care-compare/
  - Provider Data Catalog https://data.cms.gov/provider-data/
- Quality Improvement (external benchmarking to organizations)
  - Certification and Survey Provider Enhanced Reports (CASPER) https://www.qtso.com/providernh.html

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

The CASPER reports are available to providers on-demand with quality measure data updated monthly. Care Compare reports the rolling average of four quarters for the quality measure, comparing each nursing home's score to both the state and national average; providers can preview this information before it is publicly reported.

Additional Feedback:

Upon review of all inquiries submitted to the quality measure support inbox between 10/2019 and 02/2021, one other user raised a concern regarding the overlap in implementation of NQF 0674 and the SNF QRP application of this measure. There are rare cases where a SNF stay overlaps with a LS Nursing Home stay. In such cases, the resident must be discharged from the short stay as their Medicare PPS Part A stay ends, and after discharge the resident remains in the facility. Under such circumstances, a fall with major injury could be included in both stays (SNF and LS Nursing Home) if the fall with major injury occurred before the Medicare PPS Discharge and within the target assessment period for the long-stay measure. However, NQF 0674 and the SNF QRP application of the measure are not in the same quality reporting program. Only the LS Nursing Home measure is included in the Nursing Home Five-Star Quality Rating System.

***Questions for the Committee**:*

How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?

How has the measure been vetted in real-world settings by those being measured or others?

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

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

4b. Usability evaluates 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

The national facility-level mean and median scores for the Percent of Residents Experiencing One or More Falls with Major Injury demonstrate stability from quarter to quarter (Figure 1 of NQF Testing Form). Overall, the national facility-level mean and median scores have decreased marginally and indicate a slight improvement in performance over time. The mean score for this measure was 3.5% in quarter 1 of 2017 and the median score was 3.0%. In Q2 2019, the mean and median were 3.4% and 2.9%, respectively.

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:

There were no unexpected findings reported by the developer.

Potential harms

None

Additional Feedback:

N/A

***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?

- Publicly reported and used in accountability programs already
- Publicly reported and used in an accountability program. No concerns.
- The measure has been used in both accountability and public transparency programs, such as CMS public reporting for care compare and quality improvement program like CASPER.
- encompass health
- No concerns. Rating is "pass".
- In use
- Data useful to specific facilities wishing to reduce the number of serious falls.
- publicly reported and used
- High use

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.

- no unintended consequences evident
- No reports of unintended consequences.
- The national facility-level mean and median scores have shown a slight but noticeable improvement over time, when comparing mean and median scores between Q1 2017 and Q2 2019. No unintended consequences are identified, and the preliminary rating on usability is moderate.
- publicly reported
- No known unintended consequences.
- benefits > harms
- There is potential for bias against selection of extremely fragile residents more prone to falling with serious outcomes.
- reasonable
- Moderate usability

Criterion 5: Related and Competing Measures

Related or competing measures

0101: Falls: Screening, Risk-Assessment, and Plan of Care to Prevent Future Falls
0141: Patient Fall Rate
0202: Falls with injury

Harmonization

#0202 Falls with Injury - Acute Care Prevention of Falls (rate of inpatient falls with injury per 1,000 patient days): This measure has a similar focus as NQF #0674, but it is different because it focuses on adult acute care inpatient and adult rehabilitation patients and is reported as a rate rather than a percentage. Additionally, this measure includes any injury from minor to major. This is an important distinction. Focusing on falls with minor injury could potentially create inappropriate incentives for nursing homes to reduce resident opportunity for mobility and independence. The selection of the outcome of falls with major injury for NQF #0674 was deliberate to reduce this potential adverse unintended consequence. #0101 Falls Screening, Risk-Assessment, and Plan of Care to Prevent Future Falls: This is a clinical process measure that assesses falls prevention in older adults. The measure has three rates: 1) screening: percentage of patients aged 65 years of age and older who were screened for future fall risk at least once within 12 months; 2) falls risk assessment: percentage of patients aged 65 years of age and older with a history of falls who had a risk assessment for falls completed within 12 months; and 3) plan of care for falls: percentage of patients aged 65 years of age and older with a history of falls who had a plan of care for falls documented within 12 months. This measure is different in that it is a process measure, rather than an outcome measure. #0141 Patient Fall Rate (Total number of patient falls [with or without injury to the patient and whether or not assisted by a staff member] by hospital unit during the calendar month X 1000): This measure has a similar focus as NQF #0674, but it is different because it focuses on the adult acute care inpatient and adult rehabilitation patients and does not discriminate between falls with and without injuries, which is an important distinction. Focusing on falls with minor injury could potentially create inappropriate incentives for nursing homes to reduce resident opportunity for mobility. The selection of the outcome of falls with major injury for NQF #0674 was deliberate to reduce this potential adverse unintended consequence.

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

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

- no competing measures focused on the same population
- Overlaps are appropriate, given different settings and populations for fall prevention
- Three similar measures were identified: NQF #0101, #0141, and 0202. But, they do not involve the same care setting as in #674, do not focus on major and serious injuries that resulted from falls, or are not outcome measure. So, I do not see the need for harmonization.
- n/a
- No concerns.
- Several, listed in document
- no
- Three additional measures that are defined specific to the population or setting. Work could be done to integrate these measures but it really depends on the information system that is capturing and providing the data. The information systems would also have to be redesigned.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 06/03/2021

No NQF Members have submitted support/non-support choices as of this date.

No Public or NQF Member comments submitted as of this date.

Combined Methods Panel Scientific Acceptability Evaluation

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 0674

Measure Title: Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay)

RELIABILITY: SPECIFICATIONS

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

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

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

Briefly summarize any concerns about the measure specifications.

Panel Member 2: The reported average signal-to-noise reliability is 0.45. It will be helpful to understand the distribution of reliability scores among all eligible facilities. It may be useful to establish volume threshold for reporting.

Panel Member 3: No concerns

Panel Member 4: No concerns

Panel Member 5: No Concerns

Panel Member 6: Obviously, the reliability of all elements in the Nursing Home Minimum Data Set is dependent on the training, knowledge and accurate data entry of personnel. This is the source of the falls and injury data.

Panel Member 8: None

RELIABILITY: TESTING

Type of measure:

☒ Outcome (including PRO-PM) ☐ Intermediate Clinical Outcome ☐ Process

☐ Structure ☐ Composite ☐ Cost/Resource Use ☐ Efficiency

Data Source:

☐ Abstracted from Paper Records ☐ Claims ☐ Registry
☐ Abstracted from Electronic Health Record (EHR) ☐ eMeasure (HQMF) implemented in EHRs
☐ Instrument-Based Data ☐ Enrollment Data ☒ Other (please specify)

Panel Member 1: Assessment data
Panel Member 2: Nursing home MDS
Panel Member 3: Percent of Residents Experiencing One or More Falls with Major Injury (Long-Stay)
Panel Member 4: Assessment Data/Nursing Home Minimum Data Set (MDS) 3.0
Panel Member 5: Nursing Home Minimum Data Set (MDS) 3.0

Panel Member 6: The Nursing Home Minimum Data Set
Panel Member 7: Nursing Home MDS
Panel Member 8: MDS

Level of Analysis:

☐ **Individual Clinician** ☐ **Group/Practice** ☒ **Hospital/Facility/Agency** ☐ **Health Plan**
☐ **Population: Regional, State, Community, County or City** ☐ **Accountable Care Organization**
☐ **Integrated Delivery System** ☐ **Other (please specify)**

Measure is:

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

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

Reliability testing level ☒ **Measure score** ☒ **Data element** ☐ **Neither**

Reliability testing was conducted with the data source and level of analysis indicated for this measure ☒ **Yes** ☐ **No**

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**

Assess the method(s) used for reliability testing

Submission document: Testing attachment, section 2a2.2

Panel Member 1: Data element reliability: Assessed inter-rater reliability of the gold-standard nurse to gold-standard nurse & gold-standard nurse to facility nurse and calculated kappas. Measure score reliability: Assessed reliability using both signal-to-noise & split-half. All are appropriate methods for assessing reliability.

Panel Member 2: Data element reliability was established by assessing agreement between gold-standard nurse abstractor and facility nurse abstractor to ensure critical data element can be uniformly ascertained across facilities. For performance score reliability, the developer calculated facility specific signal-to-noise reliability scores. Both approaches were sound. The developer also conducted split-half reliability analysis, however, neither Pear correlation coefficient nor Spearman correlation coefficient is the most appropriate measure. ICC would be more appropriate.

Panel Member 3: No concerns

Panel Member 4: Appropriate for score level and data element level.

Panel Member 5: Both methods appropriate. Facility Nurse agreement using Kappa and Signal to Noise/ Split Half reliability for score level testing.

Panel Member 6: Data element testing was performed on a sample extracted from the entire dataset, including 71 community nursing facilities in 8 states and 19 VA nursing homes (3822 residents). Characteristics were predominantly white, female, and dual-eligible, age 85 or more, diagnoses of arthritis, osteoporosis, and fractures, depression, dementia, and other medical issues common to nursing home residents. A Gold-standard assessor approach measuring the agreement between gold-standard nurses and facility nurses. Inter-rater kappa for gold-standard to gold-standard was 0.967 and the inter-rater kappa for gold-standard to facility nurse was 0.945. Performance score testing included all nationwide nursing home facilities with a minimum denominator of 20 (14,286). Signal to noise ratio for between-provider analysis. As falls are a binary outcome, a beta-binomial Split-half reliability analysis was also utilized for reliability

testing model was used for nursing homes with greater than 40 residents so that each facility would have 20 or more in each of the split halves.

Panel Member 7: Re: Measure --> SNR and split-half

Panel Member 8: Critical data element testing was completed with 15-year-old data. How can we be sure that the data element reliability and validity have been maintained? Comparing community nurses to gold standard nurses seems like a test of validity rather than reliability (and is described as such in the validity testing section). Comparing community to community nurses would be a better test of reliability. The Landis and Koch adjectives pertain to evidence against the null hypothesis of zero agreement. Interpretation of kappa is difficult in the presence of asymmetry https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4236536/ The entity-score reliability methods included SNRs and split-sample reliability, which are strong methods.

Assess the results of reliability testing

Submission document: Testing attachment, section 2a2.3

Panel Member 1: Data element: The Kappa for gold-standard to gold-standard on the MDS 3.0 item was 0.967, and the Kappa for gold-standard to facility-nurse agreement on the MDS 3.0 item was 0.945. Ratings of 0.967 and 0.945 are considered "substantial agreement." Measure-score: The average signal-to-noise reliability score was 0.45, but with 19% of facilities achieving a perfect score of 0.0%, this results in lower variation of scores between facilities.

Panel Member 2: Kappa for critical data element was very high. The developer only reported average signal-to-noise reliability score at 0.45. It is partly due to the nature of the outcome (rare event) and sample size. It would be helpful to report the distribution of reliability scores among all facilities and potentially establish a volume threshold for reporting.

Panel Member 3: Data element reliability was excellent. I agree with the developers that only moderate reliability at the score level was supported, but that this is still sufficient to justify a measure that assesses an event that is expected to have little variance due to many facilities demonstrating no falls within the measurement period.

Panel Member 4: RAND study of the validation of MDS 3.0 national pilot test cited as evidence for Critical Data Element Reliability. Study was conducted 13 years ago. The current MDS 3.0 forms are similar to the ones used for testing in 2008. Performance Measure Score Reliability was assessed via signal-to-noise ratio and split half reliability. Both methods are appropriate.

Panel Member 5: Critical Data Element= facility-nurse agreement on the MDS 3.0 item was 0.945. Kappa is a statistical measure of inter-rater agreement for qualitative data, ranging from 0.0 to 1.0. Ratings of 0.967 and 0.945 are considered "substantial agreements." These results are indicative of data element reliability. Performance Measure Score Reliability Signal-to-noise analysis: The average signal-to-noise reliability score of this quality measure using facility scores based on data from FY2019 Quarter 2 was observed to be 0.45. This suggests that the measure is moderately reliable in separating facility characteristics from variability within facility.

Panel Member 6: Data Element: Kappas for both gold standard nurse to gold standard nurse as well as gold standard nurse to facility nurse indicated substantial agreement of 0.967 and 0.945, respectively. Measure score: Average signal-to-noise reliability on data from FY2019Q2 was 0.45. 19% of the facilities had a perfect score of 0%. Split half reliability analysis had a correlation of 0.18, attributed to a small amount of variation for relatively rare events, again with 19% of the facilities having 0%. The 50th percentile score is 2.9% and the 90th percentile score is 7.1%.

Panel Member 7: The average signal-to-noise reliability score of this quality measure using facility scores based on data from FY2019 Quarter 2 was observed to be 0.45. - no breakdowns. The split-half correlation for this measure was positive ($r = 0.18$, $\rho = 0.18$, $p < .01$), providing limited evidence of internal reliability.

These low correlations were expected due to a small amount of variation in performance among providers due to the rareness of the outcome being measured.

Panel Member 8: Critical data element reliability for gold to community nurses was 0.945. Community to community nurse reliability was not reported. These results are quite old. Median SNR was 0.45. The distribution of SNR was not given. Since half of the entities have reliability <.45, I considered the entity-level reliability to low. The split-sample reliability analysis yielded r = .18 which is very low. We have seen much higher values for measures with similar sample size and skew issues.

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)

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)

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)

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

Panel Member 1: The measure developers demonstrated a high level of inter-rater reliability of the critical data elements. The score-level reliability results were less convincing -- a signal-to-noise statistic of 0.45 is marginal at best, and the split-half analysis demonstrated limited evidence of internal reliability.

Panel Member 2: Data element reliability is very high. Measure score reliability is moderate, but we need to keep in mind that this is an important outcome, and it is somewhat rare.

Panel Member 3: The moderate rating is due to the moderate reliability demonstrated.

Panel Member 4: Based on the results of testing.

Panel Member 5: Critical Data Element Level reliability was high but score level reliability was only moderate on the low side. Authors explained this was due to the low number of events (numerator) and variance in denominators.

Panel Member 6: Under current guidelines the reliability of the measure score is moderate. For future considerations, a higher level of reliability should be expected, even despite the relative infrequency of this event.

Panel Member 7: Rare event. There may not be other options. (Data elements are as they are - not addressed by me)

Panel Member 8: Results of entity-score reliability suggest low reliability for at least half of the facilities. Also, the stability analysis reported as a test of validity shows that in 25% of facilities, Q2Q performance can jump 3 or more deciles.

VALIDITY: TESTING

Validity testing level: ☒ **Measure score** ☐ **Data element** ☒ **Both**

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)

Method of establishing validity of the measure score:

☐ **Face validity**

☒ **Empirical validity testing of the measure score**

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

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)

Assess the method(s) for establishing validity

Submission document: Testing attachment, section 2b2.2

Panel Member 1: Data element: Assessed inter-rater reliability on the collection of data elements (see Reliability testing above) Measure score: Looked at how a facility's performance on this measure is correlated with other related quality measures.

Panel Member 2: The developer correlated this measure with a few other MDS quality measures as well as facility star ratings and claims-based quality measures. The developer also presented the results of variation by state, seasonality, stability analysis, and confidence interval analysis. Correlation analysis is more relevant, while others do not necessarily address the validity issue.

Panel Member 3: No concerns

Panel Member 4: Acceptable.

Panel Member 5: The kappa for gold-standard nurse assessment to facility nurse assessment of falls with major injury was 0.959.

Panel Member 6: Critical data element validity was established by the above-mentioned concurrence between gold standard nurses and gold standards to facility nurses and a feedback survey about appropriateness of the data elements. Performance score validity was established by correlations with other measure of nursing home quality. These included related MDS Quality Measures and Facility Five Star Ratings. Variations between states, seasonality, and stability of the measure scores was assessed.

Panel Member 7: Related MDS Quality Measures (such as percent of long-stay residents who received antipsychotic, antianxiety or hypnotic medication, percentage of long-stay residents who have depressive

symptoms, percent of long-stay residents whose need for help with daily activities has increased, etc. PLUS, by state, seasonality, "stability" and "confidence interval analysis"

Panel Member 8: I appreciate the fact that several tests of validity were performance, some of which were informative.

Assess the results(s) for establishing validity

Submission document: Testing attachment, section 2b2.3

Panel Member 1: Data element: Kappas demonstrated 'substantial agreement' between nurse abstractors Measure score: Found low, but positive correlations between facility performance on this measure and other quality measures. Almost all the correlation values fell below 0.1. The additional analyses appear to address the reliability of the measure more than the validity.

Panel Member 2: The results of correlation analyses are somewhat supportive of validity.

Panel Member 3: Results are satisfactory. Although the empirical validity results demonstrated weak correlations with somewhat related measures, I do not think these attests to the low validity for this measure, which has very strong face validity. However, I appreciate the effort done to try and identify correlations with associated measures. The additional analyses for variations in performance at the state and season are also appreciated.

Panel Member 4: Appropriate.

Panel Member 6: Kappa for gold standard nurse to facility nurse assessment was 0.959. 88% of nurse surveys felt that the fall-related injury definitions were clear and 94% agree that facility falls documentations should include the data elements indicated. Spearman Correlations with items from the MDS Quality Measures are provided and all have significant p values. Correlation with Facility Five-Star Ratings is -0.032 with a significant p value. State-by state analysis is provided, and ANOVA shows the 6.43% of the variation is attributed to the state in which the facility is located. There was no perceptible seasonal variation. Percentile ranking change showed that while 37.8% remained in the same percentile, 20.2% changed by 1 decile, 16,7% by 2 deciles, and 25.3% by 3 deciles. A possible explanation is that since 20% of facilities have 0%, a single event can change the decile a given facility can end up in considerably.

Panel Member 7: Weak correlations (stat sig) with other measures. Other assessments fine (state one is not clear to me).

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

Please describe any concerns you have with measure exclusions.

Submission document: Testing attachment, section 2b2.

Panel Member 1: None. The only exclusion is missing data, for which the measure developer indicated there was minimal missing data.

Panel Member 2: No concern

Panel Member 3: No concerns

Panel Member 4: No concerns

Panel Member 5: No Concerns

Panel Member 6: Missing data information is provided and is insignificant.

Panel Member 8: Critical data element: Validity testing method and results were strong. Correlation with other quality measures: The very small correlations with other measures are hard to interpret in terms of validity evidence. Statistical significance is meaningless in this context with such a large sample. Is there a pattern of results that would have caused concern regarding validity? Variation by state: It unclear if these analyses were done with 3-level hierarchical (patient/facility/state) models or an ANOVA on the facility level

scores. Seasonality: No concerns revealed. Stability analysis: It is concerning that 25% of facilities are jumping 3 or more deciles in performance over a short time interval, likely related to low reliability. The developer not that one fall can shift a facility 3 deciles. These analyses could also have been done in the split samples used for reliability testing. Confidence interval analysis: I think this analysis speaks more to NQF's "gap" criteria more than validity evidence. The stability analysis could have been applied to the 3 categories (Lower than average, etc.).

Risk Adjustment

Submission Document: Testing attachment, section 2b3

Risk-adjustment method ☒ **None** ☐ **Statistical model** ☐ **Stratification**

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

☒ Yes ☒ No ☐ Not applicable

19c. **Social risk adjustment:**

19c.1 Are social risk factors included in risk model? ☐ Yes ☒ No ☒ Not applicable

19c.2 Conceptual rationale for social risk factors included? ☒ Yes ☒ No

19c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☒ Yes ☐ No

19d. **Risk adjustment summary:**

19d.1 All of the risk-adjustment variables present at the start of care? ☒ Yes ☐ No

19d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☒ Yes ☐ No

19d.3 Is the risk adjustment approach appropriately developed and assessed? ☒ Yes ☐ No

19d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☒ Yes ☐ No

19d.5. Appropriate risk-adjustment strategy included in the measure? ☒ Yes ☒ No

19e**. Assess the risk-adjustment approach**

Panel Member 1: As this measure captures a "never event", the measure developers argue that facilities should be held accountable for all of these events, regardless of the risk level of the facility's patients.

Panel Member 2: Certain patients do appear to have higher risks of falls, but the developer argued that providers can still adopt preventive clinical care practices to prevent falls.

Panel Member 3: I don't think that relatively low explanatory power of potential social risk factors is a good reason for not supporting risk adjustment. However, I do agree with the argument for not needing to risk-adjust this measure due to its 'never event' nature and the need to encourage to adopt preventative clinical care practices regardless of the rate of high-risk patients instead of statistical adjustment.

Panel Member 4: Justification provided. No evidence that contradicts the developer's rationale.

Panel Member 5: Authors tested for clinical risk factor and age as a single social risk factor using published literature and internal data analysis. Risk model indicated that none of these factors significantly predicted a fall with major injury, so they were not included in the final analysis.

Panel Member 6: Social risk factors are not explicitly included in the model. Analysis is performed for age, BMI, depression, dementia, weight loss and account for variation in the measure results.

Panel Member 7: There was a limited attempt to try risk adjustment (dichotomization of risk adjustors). With low c-stats and invoking the concept that falls are a never event, risk adjustment was rejected. Some analysis of whether risk adjustment changes deciles/rankings may be worthwhile. Further, falls are never a never event - let's be clear. If they really were a never event, we wouldn't be measuring them. Ref: neutrinos.

Panel Member 8: The decision to not risk adjust this outcome measure is stated: "While some providers do care for a larger proportion of patients or residents at increased risk for falls with major injury, providers with high-risk patient populations should be encouraged to adopt preventative clinical care practices instead of statistical adjustment of these providers' scores. Doing so helps to ensure that providers are held accountable for the safety of all patients and that preventative measures are implemented in a clinically appropriate and equitable manner." Could this decision have an adverse impact on patient selection? The decision to not risk-adjust the measure seems inconsistent with many other CMS outcome measures. How is this outcome different from (for example) surgical mortality or complication measures?

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

Submission document: Testing attachment, section 2b4.

Panel Member 1: None, as we see variation in performance on the measure. The mean facility-level score was 3.4% with a median score of 2.9%. The interquartile range was 3.6 percentage points.

Panel Member 2: The results indicate substantial variation among facilities and there is performance gap among facilities.

Panel Member 3: No concerns

Panel Member 4: No concerns.

Panel Member 5: No Concerns

Panel Member 6: None

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

Submission document: Testing attachment, section 2b5.

Panel Member 1: N/A

Panel Member 2: Same data source

Panel Member 4: No concerns.

Panel Member 6: Not applicable

Panel Member 7: "Overall, 22.7% of facilities scored significantly differently than the national mean in FY2019 Quarter 2, indicating that there are meaningful differences in facility-level scores for this measure even though the measure may not differentiate well in the middle of the distribution of facilities."

Please describe any concerns you have regarding missing data.

Submission document: Testing attachment, section 2b6.

Panel Member 1: None. The measure developer noted that there was negligible missing data.

Panel Member 2: No concern

Panel Member 3: No concerns due to the low exclusion rates

Panel Member 4: No concerns.

Panel Member 5: No concerns, rate of missing data very low

Panel Member 6: None. A very insignificant few had missing data.

For cost/resource use measures ONLY:

Are the specifications in alignment with the stated measure intent?

☐ **Yes** ☐ **Somewhat** ☐ **No (If "Somewhat" or "No", please explain)**

Describe any concerns of threats to validity related to attribution, the costing approach, carve outs, or truncation (approach to outliers):

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.)

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

Panel Member 1: While the measure score validity testing that was provided demonstrated weak correlations with other quality measures, the data element validity testing results were strong (kappas in mid-90s).

Panel Member 2: The results of correlation analyses provide supportive evidence of measure score validity.

Panel Member 4: Based on testing.

Panel Member 5: Differences between highest and lowest performers were demonstrated but very little difference in mid-level performers.

Panel Member 6: Date element validity is rated by expert opinion (gold standard nurses). Performance score validity is rated by correlation with other related quality measures and by variation by state, seasonality, stability analysis, and confidence interval analysis. Decile change was noted but is attributed to the low frequency of events.

Panel Member 7: MOD-LOW

Panel Member 8: Both the results of the stability analysis and lack of a risk adjustment model are concerning.

FOR COMPOSITE MEASURES ONLY: Empirical analyses to support composite construction

What is the level of certainty or confidence that the empirical analysis demonstrates that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct?

☐ **High**

☐ **Moderate**

☐ **Low**

☐ **Insufficient**

Briefly explain rationale for rating of EMPIRICAL ANALYSES TO SUPPORT COMPOSITE CONSTRUCTION

ADDITIONAL RECOMMENDATIONS

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.

Panel Member 1: None.

Panel Member 2: It may be useful to identify a volume threshold for reporting.

Panel Member 5: None

Panel Member 6: none

Developer Submission

NQF #: 0674

Corresponding Measures:

De.2. Measure Title: Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay)

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

De.3. Brief Description of Measure: This measure reports the percentage of long-stay residents in a nursing home who have experienced one or more falls resulting in major injury (defined as bone fractures, joint dislocations, closed head injuries with altered consciousness, or subdural hematoma) reported in the look-back period no more than 275 days prior to the target assessment. The long stay nursing home population is defined as residents who have received 101 or more cumulative days of nursing home care by the end of the target assessment period. This measure is based on data obtained through the Minimum Data Set (MDS) 3.0 OBRA, PPS, and/or discharge assessments during the selected quarter(s).

1b.1. Developer Rationale: This outcome-based quality measure reports the percentage of long-stay nursing home residents who have experienced one or more falls resulting in major injury, which include bone fractures, joint dislocations, and closed-head injuries. Injurious falls are important to monitor in the nursing home population because of their impacts on health outcomes, as research has demonstrated injurious falls are the leading causes of disability and death for nursing home residents. Falls with major injury also impact resident quality of life by introducing new functional limitations and psychosocial distress, while potentially influencing providers to increase the use of unwanted physical or chemical restraints. Studies have shown that risk for falling is associated with a variety of resident characteristics, including but not limited to, increasing age, being female, and cognitive decline. Falls are also associated with inappropriate or changing prescriptions. The capacity of nursing homes to provide residents appropriate living accommodations/amenities and sufficient support by appropriately qualified staff around-the-clock to serve the medical needs of residents can mitigate fall risks and prevent or reduce falls.

S.4. Numerator Statement: The numerator is the number of long-stay residents with one or more look-back scan assessments that indicate one or more falls that resulted in major injury.

S.6. Denominator Statement: The denominator consists of all long-stay nursing home residents with one or more look-back scan assessments except those who meet the exclusion criteria.

S.8. Denominator Exclusions: A resident is excluded from the denominator of this quality measure if all look-back scan assessments indicate that data is missing from the data element assessing falls resulting in major injury during the look-back period preceding the target assessment.

De.1. Measure Type: Outcome

S.17. Data Source: Assessment Data

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Mar 03, 2011 **Most Recent Endorsement Date:** Dec 09, 2015

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? This measure is not paired/grouped.

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

NQF-0674-Evidence-Form-20210408-508.docx

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

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

Yes

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0674

Measure Title: Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay)

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

Date of Submission: **4/2/2021**

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

Outcome

☒ Outcome: **Fall resulting in major injury (defined as bone fractures, joint dislocations, closed head injuries with altered consciousness, or subdural hematoma)**

☐ Patient-reported outcome (PRO):

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

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

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

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

Outcomes and Risk Factors

This outcome-based quality measure reports the percentage of long-stay nursing home residents who have experienced one or more falls resulting in major injury, which include bone fractures, joint

dislocations, and closed-head injuries. Injurious falls are important to monitor in the nursing home population because of their impacts on health outcomes, as injurious falls are the leading causes of disability and death for nursing home residents. Falls with major injury also impact resident quality of life by introducing new functional limitations and psychosocial distress, while potentially influencing providers to increase the use of unwanted physical or chemical restraints.

Some nursing home residents are at higher risk for experiencing falls, as certain resident characteristics and care-related factors influence the rate of falls in a facility. Nursing home residents who have a history of falls, have impaired balance or gait, use walking aids, or have a moderate disability are at high risk for experiencing falls (Deandrea et al., 2013; Roman de Mettelinge & Cambier, 2015; Ambrose, Paul, & Hausdorff, 2013). Other clinical factors that are associated with a high risk for falls include having an overactive bladder, low blood pressure, heart failure, cardiac arrhythmia, obesity, and malnutrition (Szabo et al., 2018; Jansen et al., 2016; Neri et al., 2019; Trevisan et al., 2019). Although residents with obesity are at high risk for experiencing a fall or multiple falls, there is no association between obesity and fall-related injuries (Neri et al., 2019). In addition to these clinical risk factors described above, advancing age, being female, and experiencing cognitive decline all place nursing home residents at high risk for falls (Ambrose, Paul, & Hausdorff, 2013).

Falls are also associated with inappropriate or changing medications. Polypharmacy, or the simultaneous use of multiple drugs by a single patient to treat one or more conditions, is a major risk factor for falls in the nursing home population (Ambrose, Paul, & Hausdorff, 2013; Damian et al., 2013; Seppala et al., 2018). Several classes of drugs are associated with falls, such as loop diuretics, antipsychotics, antidepressants, benzodiazepines, opioids, and anti-epileptics (de Vries et al., 2018; Seppala et al., 2018a; Seppala et al., 2018b). Specifically, the use of loop diuretics, opioids, and anti-epileptics are significantly associated with increased risk of falling.

Evidence for link between structure and quality of care outcomes

Several nursing home characteristics may influence the risk for experiencing a fall with major injury, including adequate staffing levels, staff education, and adequate levels of facility equipment, such as accessible computers used to complete assigned falls prevention tasks (Vlaeyen et al., 2017). One meta-analysis examined multiple studies in the nursing home population to identify facilitators and barriers to fall prevention implementation and found the greatest number of determinants at the social and organizational levels of healthcare (Vlaeyen et al., 2017). Overall, poor communication, such as poor information transfer between shifts among a facility's care providers (i.e., physicians, nurses, aides) and tension between licensed and unlicensed staff, inadequate staffing levels, and limited facility equipment were among the greatest barriers to implementing a falls prevention program in a facility (Vlaeyen et al., 2017). In addition, a facility's lack of leadership accountability, leadership lacking quality improvement skills, and lack of educational structure for staff prevented facilities from adequately implementing and maintaining a falls prevention program (Vlaeyen et al., 2017).

Staffing issues remain at the core of a facility's ability to prevent falls in the nursing home population. High staff turnover rates and lack of adequate staffing levels prevent caretakers from having sufficient time dedicated to a falls prevention program, as they must prioritize other tasks and responsibilities (Vlaeyen et al., 2017). Nursing home staff have cited feeling helpless, frustrated, overwhelmed, and highly concerned about their ability to control fall management (Vlaeyen et al., 2017).

Evidence for link between processes and quality of care outcomes

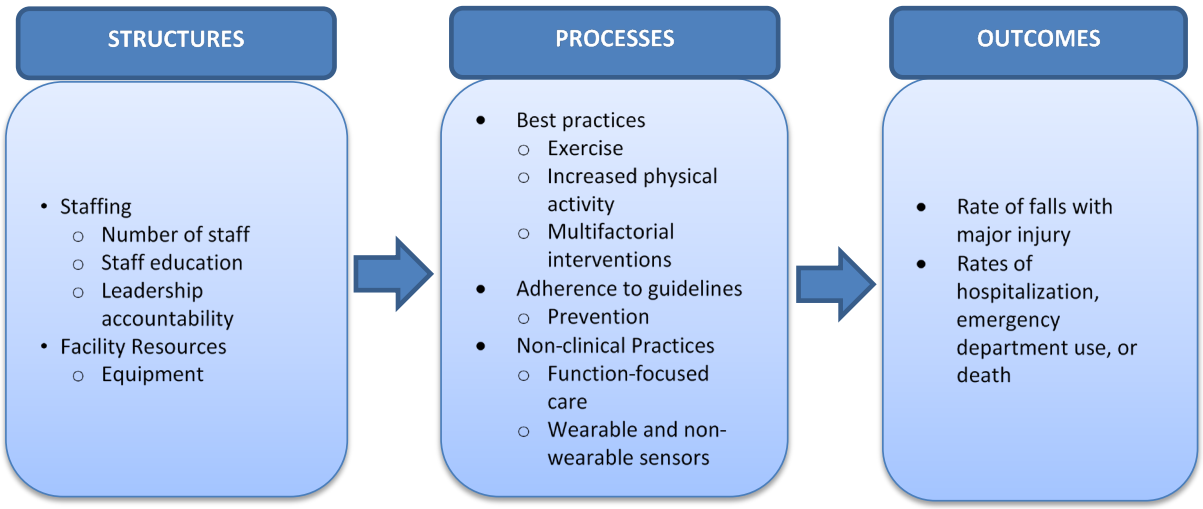
Nursing home characteristics and resource allocation may affect key processes known to influence the rate of falls with major injury within a facility. These key processes include adherence to clinical guidelines and best practices in falls prevention. In addition to these clinical processes, incorporating other supportive practices into resident care, such as the use of sensor technologies aimed at preventing falls, can impact a facility's rate of falls with major injury.

Many clinical practice guidelines for preventing falls and reducing fall-related injuries emphasize the importance of increased physical activity and functional improvement for residents. Function-focused care recognizes an individual's existing functional capacity and helps them to preserve/improve functioning through increased physical activity. Rather than focusing on completing nursing related tasks, such as dressing, bathing or feeding, function-focused care aims to enhance a resident's overall physical activity (Galik et al., 2014). Additionally, it is associated with improved/maintained functionality, it allows patients to maintain a sense of autonomy, it encourages residents to engage in continued activity, and it does not increase risk for falling. (Resnick, 2012; Resnick et al., 2012; Resnick et al., 2011, Resnick et al., 2009). Function-focused care in particular offers several advantages as a care method. The use of function-focused care and effective patient-provider communication as proxies to reducing falls are two examples of adhering to clinical practice guidelines aimed at reducing falls and fall-related injuries (Arling et al., 2014; Galik et al., 2014).

Other supportive practices that can be used to reduce falls in resident care include innovative fall reduction interventions, such as wearable and non-wearable sensor technologies aimed at preventing falls. Evidence on the ability of sensor technologies to reduce or prevent falls, however, is inconsistent with some studies reporting no reductions in fall rates when implementing a sensor system intervention and other studies reporting significant reductions of falls (Kosse et al., 2013). The combination of multifactorial risk factor assessments and interventions such as identification of specific patient characteristics, known risk factors, and available interventions is used to minimize the incidence of falls in nursing homes (Cameron et al., 2012; Davison et al., 2005; Chang et al., 2004).

There are a number of interventions that have been tested to reduce both falls and injurious falls. Evidence-based exercise programs and increased physical activity have demonstrated to be among the best practices for fall prevention (Crandall et al., 2016; Grossman & The USPSTF, 2018; Tricco et al., 2017). Nursing home residents with higher physical activity levels are at lower risk for falls than residents with low physical activity levels (Chang & Do, 2015). One study found that a single intervention of exercise reduced the number of fallers in the nursing home setting by 36%, as well as the number of recurrent fallers by 41% (Gulka et al., 2019). In addition to interventions centered on increased physical activity, one study found that combining cognitive-behavioral therapy and exercise can help residents manage their fears related to falling, as well as their predisposition to depression, while simultaneously increasing their mobility and muscle strength (Huang et al., 2016). As residents became less fearful of falls and increased their mobility and muscle strength, the incidence of falls in the facility decreased significantly (Huang et al., 2016). A number of studies observed multifactorial interventions that included activities such as exercise, medication review, risk assessment, vision assessment, environmental assessment, and staff education can significantly reduce fall rates (Vlaeyen et al., 2015; Gulka et al., 2019; Tricco et al., 2017).

Figure 1: Role of Nursing Home Structures and Processes in Rates of Falls with Major Injury



References:

Ambrose, A. F., Paul, G., & Hausdorff, J. M. (2013). Risk factors for falls among older adults: A review of the literature. Maturitas, 75(1), 51-61. doi:10.1016/j.maturitas.2013.02.009

Arling, P. A., et al. (2014). "Communication and effectiveness in a US nursing home quality-improvement collaborative." Nurs Health Sci 16(3): 291-297.

Cameron ID, Gillespie LD, Robertson MC, Murray GR, Hill KD, Cumming RG, Kerse N. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database of Systematic Reviews 2012, Issue 12. Art. No.: CD005465. DOI: 10.1002/14651858.CD005465.pub3.

Chang JT, Morton SC, Rubenstein LZ, Mojica WA, Maglione M, Suttorp MJ, Roth EA, Shekelle PG. Interventions for the prevention of falls in older adults: systematic review and meta-analysis of randomized clinical trials. BMJ. 2004 Mar 20; 328(7441):680.

Chang, V. C. & Do, M. T. (2015). Risk factors for falls among seniors: Implications of gender. American Journal of Epidemiology, 181(7), 521-31. doi:10.1093/aje/kwu268

Crandall, M., Duncan, T., Mallat, A., Greene, W., Violano, P., & Christmas, B. (2016). Prevention of fall-related injuries in the elderly: An eastern association for the surgery of trauma practice management guideline. Journal of Trauma and Acute Care Surgery, 81(1), 196-206. doi:10.1097/TA.0000000000001025

Damian, J., Pastor-Barriuso, R., Valderrama-Gama, E., & de Pedro-Cuesta, J. (2013). Factors associated with falls among older adults living in institutions. BMC Geriatrics, 13(6), 1-9. doi:10.1186/1471-2318-13-6

Davison J, Bond J, Dawson P et al. Patients with recurrent falls attending Accident & Emergency benefit from multifactorial intervention--a randomised controlled trial. Age Ageing 2005; 34(2):162–168.

de Vries, M., Seppala, L. J., Daams, J. G., van de Glind, E. M. M., Masud, T., & van der Velde, N. (2018). Fall-risk-increasing-drugs: A systematic review and meta-analysis: I. cardiovascular drugs. Journal of the American Medical Directors Association, 19(4), 371e1-371e9. doi:10.1016/j.jamda.2017.12.013

Deandrea, S., Bravi, F., Turati, F., Lucenteforte, E., La Vecchia, C., & Negri, E. (2013). Risk factors for falls in older people in nursing homes and hospitals: A systematic review and meta-analysis. Archives of Gerontology and Geriatrics, 56(3), 407-15. doi:10.1016/j.archger.2012.12.006

Galik, E., et al. (2014). "Optimizing function and physical activity among nursing home residents with dementia: testing the impact of function-focused care." Gerontologist 54(6): 930-943.

Grossman, D. C. & The US Preventative Services Task Force (USPSTF). (2018). Interventions to prevent falls in community-dwelling older adults: US Preventive Services Task Force recommendation statement. Journal of the American Medical Association, 319(16), 1696-1704. doi:10.1001/jama/2018.3097

Gulka, H. J., Patel, V., Arora, T., McArthur, C., & Iaboni, A. (2019). Efficacy and generalizability of falls prevention interventions in nursing homes: A systematic review and meta-analysis. Journal of the American Medical Directors Association,1-12. doi:10.1016/j.jamda.2019.11.012

Huang, T., Chung, M., Chen, F., Chin, Y., & Wang, B. (2016) Evaluation of a combined cognitive-behavioural and exercise intervention to manage fear of falling among elderly residents in nursing homes, Aging & Mental Health, 20:1, 2-12, DOI: 10.1080/13607863.2015.1020411

Jansen, S., Bhangu, J., de Rooij, S., Daams, J., Kenny, R. A., & van der Velde, N. (2016). The association of cardiovascular disorders and falls: A systematic review. Journal of the American Medical Directors Association, 17(3), 193-9. doi:10.1016/j.jamda.2015.08.022

Kosse N.M., Branks K., Bauer J.M., Hortobagui T., Lamoth C.J.C. (2013). Sensor technologies aiming at fall prevention in institutionalized old adults: A synthesis of current knowledge. Journal of Medical Informatics. 82, 743-752.

Neri, S. G. Rr, Oliveira, J. S., Dario, A. B., Lima, R. M., & Tiedemann, A. (2019). Does obesity increase the risk and severity of falls in people aged 60 years and older? A systematic review and meta-analysis of observational studies. Journals of Gerontology: Medical Science, 1-9. doi:10.1093/gerona/glz272

Resnick, B. (2012). Changing the Philosophy of Care — A Function-Focused Care Approach. Aging Well, 5(2), 34.

Resnick B, Galik E, Gruber-Baldini I, Zimmerman S. "Falls and fall-related injuries associated with function-focused care." Clinical nursing research 21.1 (2012): 43-63.

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.)

This is not applicable.

****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.

This is not applicable.

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

Resnick, B., Galik, E., Gruber-Baldini, A.L., Zimmerman, S. (2011). Testing the effect of function-focused care in assisted living. J Am Geriatr Soc, 59(12), 2233-2240.

Source of Systematic Review:

- Title
- Author
- Date
- Citation, including page number
- URL

Title: Preventing Falls and Reducing Injury from Falls

Author: Registered Nurses Association of Ontario (RNAO)

Date: 2017

Citation: Registered Nurses Association of Ontario (RNAO). (2017). Preventing falls and reducing injury from falls. Retrieved from: https://rnao.ca/bpg/guidelines/prevention-falls-and-fall-injuries

URL: https://rnao.ca/bpg/guidelines/prevention-falls-and-fall-injuries

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.

- Recommendation 1.1: Screen all adults to identify those at risk for falls. Conduct screening as part of admission processes, after any significant change in health status, or at least annually. Screening should include the following approaches:
    - identifying a history of previous falls;
    - identifying gait, balance, and/or mobility difficulties; and
    - using clinical judgment.

Recommendation 1.2a: For adults at risk for falls, conduct a comprehensive assessment to identify factors contributing to risk and determine appropriate interventions. Use an approach

- and/or validated tool appropriate to the person and the health-care setting.
- Recommendation 1.2b: Refer adults with recurrent falls, multiple risk factors, or complex needs to the appropriate clinician(s) or to the interprofessional team for further assessment and to identify appropriate interventions.
- Recommendation 2.1: Engage adults at risk for falls and fall injuries using the following actions:

    explore their knowledge and perceptions of risk, and their level of motivation to address risk;

    communicate sensitively about risk and use positive messaging;

    discuss options for interventions and support self-management;

    develop an individualized plan of care in collaboration with the person;

    engage family (as appropriate) and promote social support for interventions; and

    evaluate the plan of care together with the person (and family) and revise as needed.

Recommendation 2.2: Provide education to the person at risk for falls and fall injuries and their family (as appropriate) in conjunction with other falls prevention interventions. This includes providing information about risk for falls, falls prevention, and interventions. Ensure that the information is provided in a variety of formats and in the appropriate language.

Recommendation 2.3: Communicate the person's risk for falls and related plan of care/interventions to the next responsible health-care provider and/or the interprofessional team at all care transitions to ensure continuity of care and to prevent falls or fall injuries.

Recommendation 2.4: Implement a combination of interventions tailored to the person and the healthcare setting to prevent falls or fall injuries.

Recommendation 2.5: Recommend exercise interventions and physical training for adults at risk for falls to improve their strength and balance. Encourage an individualized, multicomponent program/activity that corresponds to the person's current abilities and functioning.

Recommendation 2.6: Collaborate with prescribers and the person at risk for falls to reduce, gradually withdraw, or discontinue medications that are associated with falling, when the person's health condition or change in status allows. This includes the following actions:

- identify polypharmacy and medications that increase risk for falls;
- conduct a medication review, or refer to an appropriate health-care provider and/or the prescriber; and
- monitor for side effects of medications known to contribute to risk for falls.

- Recommendation 2.7: Refer adults at risk for falls or fall injuries to the appropriate health-care provider for advice about vitamin D supplementation.
- Recommendation 2.8: Encourage dietary interventions and other strategies to optimize bone health in adults at risk for falls or fall injuries, particularly those at risk for fracture. Refer to the appropriate health-care provider for advice and individualized interventions.

Recommendation 2.9: Consider hip protectors as an intervention to reduce the risk of hip fracture among adults at risk for falls and hip fracture. Review the evidence, potential benefits, harms, and barriers to use with the person to support individualized decisions.

Recommendation 3.1: After a person falls, provide the following interventions:

- conduct a physical examination to assess for injury and to determine the severity of any fall injuries;
- provide appropriate treatment and care;
- monitor for injuries that may not be immediately apparent;
- conduct a post-fall assessment to determine factors that contributed to the fall;
- collaborate with the person and the interprofessional team to conduct further assessments and determine appropriate interventions; and
- refer the person to the appropriate health-care provider(s) for physical rehabilitation and/or to support psychological well-being (as needed).

Recommendation 4.1: Educational institutions incorporate content on falls prevention and injury reduction into health-care education and training programs.

Recommendation 4.2: Health-care organizations provide ongoing organization-wide education to all staff in conjunction with other activities to help prevent falls and reduce injuries among persons in their care.

Recommendation 5.1: To ensure a safe environment:

- implement universal falls precautions, and
- identify and modify equipment and other factors in the physical/structural environment that contribute to risk for falls and fall injuries.

Recommendation 5.2: Organizational leaders, in collaboration with teams, apply implementation science strategies to enable successful implementation and sustainability of falls prevention/injury reduction initiatives. This includes identifying barriers and establishing formalized supports and structures within the organization.

Recommendation 5.3: Implement rounding as a strategy to proactively meet the person's needs and prevent falls.

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade.

Levels of evidence are assigned to study designs to rank how well each study design can eliminate alternate explanations of the phenomena under study. The higher the level of evidence, the more likely it is that there were fewer potential sources of bias influencing the research findings. However, levels of evidence do not reflect the quality of individual studies or reviews.

In some cases, recommendations in this BPG are assigned more than one level of evidence. This reflects the varied study designs that support the recommendation. For transparency, the level of evidence for each component of the recommendation statement is identified.

LEVEL	SOURCE OF EVIDENCE
Ia	Evidence obtained from meta-analysis or systematic reviews of randomized controlled trials, and/or synthesis of multiple studies primarily of quantitative research.
Ib	Evidence obtained from at least one randomized controlled trial.
IIa	Evidence obtained from at least one well-designed controlled study without randomization.
IIb	Evidence obtained from at least one other type of well-designed quasi-experimental study, without randomization.
III	Synthesis of multiple studies primarily of qualitative research.

LEVEL	SOURCE OF EVIDENCE
IV	Evidence obtained from well-designed non-experimental observational studies, such as analytical studies or descriptive studies, and/or qualitative studies.
V	Evidence obtained from expert opinion or committee reports, and/or clinical experiences of respected authorities.

Provide all other grades and definitions from the evidence grading system.

In addition to the levels of evidence, the quality of each of the reviews cited in the discussion of evidence was appraised and categorized as strong, moderate, or low based on the AMSTAR instrument for reviews. The quality rating is calculated by converting the score on the AMSTAR tool into a percentage. When other guidelines informed the recommendation and discussion of evidence, the AGREE II instrument was used to determine the quality rating. Tables 2 and 3 highlight the quality scores required to achieve a strong, moderate, or low quality rating.

Table 2: Quality Rating for Reviews Using the AMSTAR Tool

QUALITY SCORE ON THE AMSTAR	OVERALL QUALITY RATING
Greater than, or equal to, a converted score of 82.4%	Strong
A converted score of 62.5 – 82.4%	Moderate
Less than, or equal to, a converted score of 62.4%	Low

Table 3: Quality Rating for Guidelines Using the AGREE II Tool

QUALITY SCORE ON THE AGREE II	OVERALL QUALITY RATING
A score of 6 or 7 on the overall guideline quality	Strong
A score of 5 on the overall guideline quality	Moderate
A score of less than 4 on the overall guideline quality	Low (Not used to support recommendations)

Grade assigned to the recommendation with definition of the grade.

This Guideline replaces the RNAO BPG *Prevention of Falls and Fall Injuries in the Older Adult* (2002, 2005) and its supplement (2011).

+ The recommendation and supporting evidence were updated following the systematic review.

NEW A new recommendation was developed following the systematic review.

1.0 Research Question #1: What are the most effective ways to identify adults at risk for falls or for injury due to falls?

PRACTICE RECOMMENDATIONS	LEVEL OF EVIDENCE	STATUS
Recommendation 1.1: Screen all adults to identify those at risk for falls. Conduct screening as part of admission processes, after any significant change in health status, or at least annually. Screening should include the following approaches: • identifying a history of previous falls; • identifying gait, balance, and/or mobility difficulties; and • using clinical judgment.	Ia & V	**NEW**
Recommendation 1.2a: For adults at risk for falls, conduct a comprehensive assessment to identify factors contributing to risk and determine appropriate interventions. Use an approach and/or validated tool appropriate to the person and the health-care setting.	III	+
Recommendation 1.2b: Refer adults with recurrent falls, multiple risk factors, or complex needs to the appropriate clinician(s) or to the interprofessional team for further assessment and to identify appropriate interventions.	V	**NEW**

2.0 Research Question #2: What interventions are effective in preventing falls and reducing the risk for falls or falls-related injury among at-risk adults?

PRACTICE RECOMMENDATIONS	LEVEL OF EVIDENCE	STATUS
Recommendation 2.1: Engage adults at risk for falls and fall injuries using the following actions: • explore their knowledge and perceptions of risk, and their level of motivation to address risk; • communicate sensitively about risk and use positive messaging; • discuss options for interventions and support self-management; • develop an individualized plan of care in collaboration with the person; • engage family (as appropriate) and promote social support for interventions; and • evaluate the plan of care together with the person (and family) and revise as needed.	Ia, III, & V	**NEW**
Recommendation 2.2: Provide education to the person at risk for falls and fall injuries and their family (as appropriate) in conjunction with other falls prevention interventions. This includes providing information about risk for falls, falls prevention, and interventions.	Ia & V	+

PRACTICE RECOMMENDATIONS	LEVEL OF EVIDENCE	STATUS
Ensure that the information is provided in a variety of formats and in the appropriate language.		
Recommendation 2.3: Communicate the person's risk for falls and related plan of care/interventions to the next responsible health-care provider and/or the interprofessional team at all care transitions to ensure continuity of care and to prevent falls or fall injuries.	V	**NEW**
Recommendation 2.4: Implement a combination of interventions tailored to the person and the healthcare setting to prevent falls or fall injuries.	Ia	+
Recommendation 2.5: Recommend exercise interventions and physical training for adults at risk for falls to improve their strength and balance. Encourage an individualized, multicomponent program/ activity that corresponds to the person's current abilities and functioning.	Ia	+

PRACTICE RECOMMENDATIONS	LEVEL OF EVIDENCE	STATUS
Recommendation 2.6: Collaborate with prescribers and the person at risk for falls to reduce, gradually withdraw, or discontinue medications that are associated with falling, when the person's health condition or change in status allows. This includes the following actions: • identify polypharmacy and medications that increase risk for falls; • conduct a medication review, or refer to an appropriate health-care provider and/or the prescriber; and • monitor for side effects of medications known to contribute to risk for falls.	Ia & V	+
Recommendation 2.7: Refer adults at risk for falls or fall injuries to the appropriate health-care provider for advice about vitamin D supplementation.	V	**NEW**
Recommendation 2.8: Encourage dietary interventions and other strategies to optimize bone health in adults at risk for falls or fall injuries, particularly those at risk for fracture. Refer to the appropriate health-care provider for advice and individualized interventions.	V	+
Recommendation 2.9: Consider hip protectors as an intervention to reduce the risk of hip fracture among adults at risk for falls and hip fracture. Review the evidence, potential benefits, harms, and barriers to use with the person to support individualized decisions.	Ia	+

3.0 Research Question #3: What interventions or processes should occur immediately following a fall?

PRACTICE RECOMMENDATIONS	LEVEL OF EVIDENCE	STATUS
Recommendation 3.1: After a person falls, provide the following interventions: • conduct a physical examination to assess for injury and to determine the severity of any fall injuries; • provide appropriate treatment and care; • monitor for injuries that may not be immediately apparent; • conduct a post-fall assessment to determine factors that contributed to the fall; • collaborate with the person and the interprofessional team to conduct further assessments and determine appropriate interventions; and • refer the person to the appropriate health-care provider(s) for physical rehabilitation and/or to support psychological well-being (as needed).	III & V	+

4.0 Research Question #4: What content and educational strategies are necessary to effectively educate nurses and other health-care providers to prevent falls and injury from falls?

EDUCATION RECOMMENDATIONS	LEVEL OF EVIDENCE	STATUS
Recommendation 4.1: Educational institutions incorporate content on falls prevention and injury reduction into health-care education and training programs.	V	+
Recommendation 4.2: Health-care organizations provide ongoing organization-wide education to all staff in conjunction with other activities to help prevent falls and reduce injuries among persons in their care.	Ia	+

5.0 Research Question #5: What organizational policies and system-level supports are required to help prevent falls and injuries from falls among at-risk adults?

EDUCATION RECOMMENDATIONS	LEVEL OF EVIDENCE	STATUS
Recommendation 5.1: To ensure a safe environment: • implement universal falls precautions, and • identify and modify equipment and other factors in the physical/structural environment that contribute to risk for falls and fall injuries.	Ia	+
Recommendation 5.2: Organizational leaders, in collaboration with teams, apply implementation science strategies to enable successful implementation and sustainability of falls prevention/injury reduction initiatives. This includes identifying barriers and establishing formalized supports and structures within the organization.	Ia	NEW
Recommendation 5.3: Implement rounding as a strategy to proactively meet the person's needs and prevent falls.	Ia	NEW

Provide all other grades and definitions from the recommendation grading system.

This is not applicable.

Body of evidence:

Quantity – how many studies?

Quality – what type of studies?

This set of guidelines referenced 179 sources in the review. The authors selected evidence from a variety of study designs, including meta-analyses, systematic reviews, randomized control trials, controlled studies, quasi-experimental studies, observational studies, and qualitative studies. Quality of evidence is appraised and categorized based on the AMSTAR and AGREE II instruments for reviews, which are detailed in Tables 2 and 3 above. The quality rating for reviews is evaluated using the AMSTAR tool, and the quality rating for guidelines is evaluated using the AGREE II tool.

Estimates of benefit and consistency across studies.

The guidelines outlined several benefits identified in the studies. One moderate-rated review demonstrated that cognitive motor interference was shown to be effective for preventing falls among older adults in the short term. There was a strong-rated guideline recommended specific medications for people in long-term care at risk of fracture that should and should not be taken. In a low-rated review it was found that a prompted voiding schedule in long-term care, together with physical activity, reduced falls. In addition to these potential benefits, the guidelines also found that medication management, rounding, and vitamin D supplementation posed potential benefits based on the reviewed studies and guidelines.

What harms were identified?

The guidelines identified potential harms for only one of the recommendations (2.9). These potential harms include a slight increase in the risk of pelvic fractures and skin irritation for the recommendation to use hip protectors as an intervention to reduce the risk of hip fracture among adults at risk for falls and hip fracture.

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?

A 2018 systematic review calculated the population-level impact of evidence-based interventions on reducing falls and lowering direct medical costs among community-dwelling older adults. Vitamin D supplementation was one of the evidence-based interventions that was examined, and researchers observed that correcting vitamin D insufficiency would avert $247 million in direct medical costs. However, another 2018 review of 62 trials on interventions to prevent falls in

community-dwelling older adults conducted by the US Preventative Services Task Force (USPSTF) found that vitamin D supplementation has no benefit in preventing falls in older adults. Therefore, there is insufficient evidence to support the supplementation of vitamin D as a guideline to prevent falls. The RNAO recommends referring adults at risk for falls or fall injuries to the appropriate health-care provider for advice about vitamin D supplementation (recommendation 2.7). However, the 2018 review by the USPSTF mentioned above does not support vitamin D supplementation as an intervention to prevent falls or injurious falls.

References:

Stevens, J. A. & Lee, R. (2018). The potential to reduce falls and avert costs by clinically managing fall risk. American Journal of Preventive Medicine, 55(3), 290-7. doi:10.1016/j.amepre.2018.04.035

Grossman, D. C. & The US Preventative Services Task Force (USPSTF). (2018). Interventions to prevent falls in community-dwelling older adults: US Preventive Services Task Force recommendation statement. Journal of the American Medical Association, 319(16), 1696-1704. doi:10.1001/jama/2018.309

Resnick, B., Gruber-Baldini, A.L., Zimmerman, S., et al. (2009). Nursing home resident outcomes from the Res-Care intervention. J Am Geriatr Soc, 57(7), 1156-1165.

Roman de Mettelinge, T. & Cambier, D. (2015). Understanding the relationship between walking aids and falls in older adults: A prospective cohort study. Journal of Geriatric Physical Therapy, 38(3) 127-32. doi:10.519/JPT.0000000000000031

Seppala, L. J., van de Glind, E. M. M., Daams, J. G., Ploegmakers, K. J., de Vries, M., Wermelink, A. M. A. T., & van der Velde, N. (2018). Fall-risk-increasing drugs: A systematic review and meta-analysis: III. Others. Journal of the American Medical Directors Association, 19(4), 372e1-372e8. doi:10.1016/j.jamda.2017.12.099

Seppala, L. J., Wermelink, A. M. A. T., de Vries, M., Ploegmakers, K. J., van de Glind, E. M. M., Daams, J. G., & van der Velde, N. (2018). Fall-risk-increasing drugs: A systematic review and meta-analysis: II. Psychotropics. Journal of the American Medical Directors Association, 19(4), 371e11-371e17. doi:10.1016/j.jamda.2017.12.098

Szabo, S.M., Gooch, K.L., Walker, D.R., Johnston K.M., & Wagg, A.S. (2018). The association between overactive bladder and falls and fractures: A systematic review. Advances in Therapy, 35(11),1831–41. doi:10.1007/s12325-018-0796-8

Trevisan, C., Crippa, A., Ek, S., Welmer, A.-K., Sergi, G., Maggi, S., Manzato, E., Bea, J. W., Cauley, J. A., Decullier, E., Hirani, V., LaMonte, M. J., Lewis, C. E., Schott, A.-M., Orsini, N., & Rizzuto, D. (2019). Nutritional status, body mass index, and the risk of falls in community-dwelling older adults: A systematic review and meta-analysis. Journal of the American Medical Directors Association,20(5), 569-82. https://doi.org/10.1016/j.jamda.2018.10.027

Tricco, A. C., Thomas, S. M., Veroniki, A. A., Hamid, J. S., Cogo, E., Strifler, L., Khan, P. A., Robson, R., Sibley, K. M., MacDonald, H., Riva, J. J., Thavorn, K., Wilson, C., Holroyd-Leduc, J., Kerr, G. D., Feldman, F., Majumdar, S. R., Jaglal, S. B., Hui, W., & Straus, S. E. (2017). Comparisons of interventions for preventing falls in older adults: A systematic review and meta-analysis. Journal of the American Medical Association 318(17), 1687-1699. doi:10.1001/jama.2017.15006

Vlaeyen, E., Coussement, J., Leysens, G., Van der Elst, E., Delbaere, K., Cambier, D., Denhaerynck, K., Goemaere, S., Wertelaers, A., Dobbels, F., Dejaeger, E., & Milisen, K. (2015). Characteristics and effectiveness of fall prevention

programs in nursing homes: A systematic review and meta-analysis of randomized control trials. Journal of the American Geriatrics Society, 6(3), 211-21. doi:10.1111/jgs.132.54

Vlaeyen, E., Stas, J., Leysens, G., Van der Elst, E., Janssens, E., Dejaeger, E., Dobbels, F., & Milisen, K. (2017). Implementation of fall prevention in residential care facilities: A systematic review of barriers and facilitators. International Journal of Nursing Studies, 70, 110-121. doi:10.1016/j.ijnurstu.2017.02.002

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.

This is not applicable.

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

This is not applicable.

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

This is not applicable.

1b. Performance Gap

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

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

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

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

This outcome-based quality measure reports the percentage of long-stay nursing home residents who have experienced one or more falls resulting in major injury, which include bone fractures, joint dislocations, and closed-head injuries. Injurious falls are important to monitor in the nursing home population because of their impacts on health outcomes, as research has demonstrated injurious falls are the leading causes of disability and death for nursing home residents. Falls with major injury also impact resident quality of life by introducing new functional limitations and psychosocial distress, while potentially influencing providers to increase the use of unwanted physical or chemical restraints. Studies have shown that risk for falling is associated with a variety of resident characteristics, including but not limited to, increasing age, being female, and cognitive decline. Falls are also associated with inappropriate or changing prescriptions. The capacity of nursing homes to provide residents

appropriate living accommodations/amenities and sufficient support by appropriately qualified staff around-the-clock to serve the medical needs of residents can mitigate fall risks and prevent or reduce falls.

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

Current performance: Table 6 of the NQF Testing Form describes the national facility score distribution for Percent of Residents Experiencing One or More Falls with Major Injury. The facility-level mean score for this measure in Quarter 2 (Q2) of 2019 was 3.4% and the median score was 2.9%. The standard deviation was 2.9%, the minimum was 0%, and score at the 90th percentile was 7.1%. The interquartile range for this measure was 3.6%, indicating some room for improvement in this measure. Of the facilities with adequate sample size to report, 19.0% had perfect scores of 0. This analysis is restricted to facilities that had at least 20 residents in the denominator, the minimum denominator threshold for public reporting. In 2019Q2, there were 14,286 facilities (93.9%) and 1,012,706 residents (98.0%) that met the denominator inclusion criteria

n (Facilities): 14,286

k (Residents): 1,012,076

Mean score: 3.4%

Std dev.: 2.9%

10th percentile: 0.0%

25th percentile: 1.3%

50th percentile: 2.9%

75th percentile: 4.9%

90th percentile: 7.1%

Interquartile range: 3.6%.

% of facilities with "perfect scores": 19.0%

Performance Over Time: The national facility-level mean and median scores for the Percent of Residents Experiencing One or More Falls with Major Injury demonstrate stability from quarter to quarter (Figure 1 of NQF Testing Form). Overall, the national facility-level mean and median scores have decreased marginally and indicate a slight improvement in performance over time. The mean score for this measure was 3.5% in quarter 1 of 2017 and the median score was 3.0%. In Q2 2019, the mean and median were 3.4% and 2.9%, respectively. (Data Source: Data are drawn from all United States Nursing Homes with Medicare certified beds and a minimum of 20 long-stay residents in their denominator in each quarter.)

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.

This is not applicable (data are available and described in 1b.2).

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.*

Age

To examine whether facilities with higher percentages of residents aged 85 or older have different performance scores for falls with major injury, analyses were completed comparing the performance of facilities based on their percentage of residents aged 85 or older and residents below the age of 85. First, the percentage of residents experiencing one or more falls with major injury was stratified by age. Residents aged 85 or older represented the highest mean (4.1%) followed by residents below the age of 85 (2.6%). Next, a 2-way chi-squared test for statistical dependence was run that assessed the association between quality measure score and age. The results were significant (p=<.0001) indicating that there is a statistically significant relationship between age and QM score for the measure. The results suggested that residents aged 85 years or older are at higher risk for experiencing falls with major injury than residents less than 85 years of age.

Race

To examine whether facilities with higher percentages of non-White residents have different performance scores for falls with major injury, analyses were completed comparing the performance of facilities based on their percentage of White only and non-White residents. First, the percentage of residents experiencing one or more falls with major injury was stratified by racial identification. American Indian/Alaska Native residents represented the highest mean (4.14%), followed by White residents (3.73%), Hispanic or Latino residents (2.45%), and Black or African American Residents (1.25%). Next a 2-way chi-squared test for statistical dependence was run that assessed the association between quality measure score and race/ethnicity. The results were significant (p=<.0001) indicating that there is a statistically significant relationship between racial composition and QM score for the measure. The results suggested that the White only population (3.73%) is at higher risk for experiencing falls with major injury than the non-White only population (1.8%).

Socioeconomic status

To examine whether facilities with higher percentages of Medicaid-enrolled residents have different performance scores for falls with major injury, analyses were completed comparing the performance of facilities based on their percentage of Medicaid-enrolled residents and residents not enrolled in Medicaid. First, the percentage of residents experiencing one or more falls with major injury was stratified by Medicaid enrollment. Medicaid-enrolled residents represented the highest mean (3.24%), followed by residents not enrolled in Medicaid (3.13%), indicating there are slightly more Medicaid-enrolled residents that experience falls with major injury than residents not enrolled in Medicaid. Next a 2-way chi-squared test for statistical dependence was run that assessed the association between quality measure score and Medicaid enrollment. The results were significant (p=<.0001) indicating that that there is a statistically significant relationship between Medicaid enrollment and QM score for this measure. The results suggested that the non-Medicaid population (4.04%) is at higher risk for experiencing falls with major injury than the Medicaid population (3.07%), indicating there is a relationship between socioeconomic status and falls with major injury among long-stay residents.

SOURCE: Acumen analysis of Q2 2019 MDS 3.0 data

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

This is not applicable.

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):*

Safety, Safety: Complications

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

Elderly, Populations at Risk, Populations at Risk: Individuals with multiple chronic conditions

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.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/NHQIQualityMeasures.html; please see "MDS-3.0-QM-User's-Manual-v14.0.pdf" in the "Users-Manuals-Updated-10-19-2020.zip" zipped folder in the Downloads s

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 **:** MDS-3.0-RAI-Manual-v1.17.1_October_2019.pdf

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

Clinician

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

No

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

There have been no changes to the measure specifications since the last measure update.

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

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

The numerator is the number of long-stay residents with one or more look-back scan assessments that indicate one or more falls that resulted in major injury.

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

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

The numerator is the number of long-stay residents with one or more look-back scan assessments that indicate one or more falls that resulted in major injury (J1900C = [01, 02]). The selection period for the look-back scan consists of all qualifying Reason for Assessments (RFAs) (A0310A = [01, 02, 03, 04, 05, 06] or A0310B = [01] or A0310F = [10, 11]) within the current episode that have target dates no more than 275 days prior to the target assessment. A 275-day time period is used to include up to three quarterly OBRA assessments. The earliest of these assessments would have a look-back period of up to 93 days, which would cover a total of about one year. The look-back scan includes the target assessment and all qualifying earlier assessments in the scan. An earlier assessment should only be included in the scan if it meets all of the following conditions: (a) it is contained within the resident's episode, (b) it has a qualifying RFA, (c) its target date is on or before the target date for the target assessment, and (d) its target date is no more than 275 days prior to the target date of the target assessment. The Centers for Medicare & Medicaid Services (CMS) then scans the target assessment and qualifying earlier assessments to calculate the measure.

Residents are counted in the numerator if they are long-stay residents, defined as residents who have had 101 or more cumulative days of nursing home care by the end of the target period. Residents who return to the nursing home following a hospital discharge will not have their cumulative days in facility reset to zero.

An episode is defined as a period of time spanning one or more stays. An episode begins with an admission and ends with either (a) a discharge, or (b) the end of the target period, whichever comes first. Data are publicly reported on the Nursing Home Compare website and are weighted on an average of four target periods.

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

The denominator consists of all long-stay nursing home residents with one or more look-back scan assessments except those who meet the exclusion criteria.

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).*

Residents are counted in the denominator if they are long-stay residents with one or more look-back scan assessments no more than 275 days prior to the target assessment, except those with exclusions (specified in S.8 and S.9). Long-stay residents are defined as residents who have had 101 or more cumulative days of nursing home care by the end of the target assessment period. Residents who return to the nursing home following a hospital discharge will not have their cumulative days in facility reset to zero. Target assessments may be an OBRA admission, quarterly, annual or significant change/correction assessment (A0310A = [01, 02, 03, 04, 05, 06]); or PPS 5-day assessments (A0310B = [01]); or discharge assessment with or without anticipated return (A0310F = [10, 11]).

A description of the time period for the data included in this measure is provided in S.5 above.

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

A resident is excluded from the denominator of this quality measure if all look-back scan assessments indicate that data is missing from the data element assessing falls resulting in major injury during the look-back period preceding the target assessment.

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.)*

A resident is excluded from the denominator if the following is true for all look-back scan assessments:

1.The number of falls with major injury was not coded (J1900C = [-]).

If the facility sample includes fewer than 20 residents after all other resident-level exclusions are applied, then the facility is suppressed from public reporting because of small sample size.

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.)*

This is not applicable because this measure is not stratified.

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:

Rate/proportion

If other:

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

Better quality = Lower 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.)*

Step 1: Identify the total number of long-stay residents with a qualifying target assessment (OBRA, PPS, or discharge), one or more look-back scan assessments, and who do not meet the exclusion criteria (i.e., if J1900C = [-] on the target assessment or other qualifying assessments).

Step 2: Starting with the set of residents identified in Step 1, determine the total number of long-stay residents with one or more look-back scan assessments that indicate one or more falls that resulted in major injury (J1900C = [1, 2]).

Step 3: Divide the results of step 2 by the results of step 1.

Step 4: Multiply the result of step 3 by 100 to obtain a percent value.

A description of the time period for the data included in this measure is provided in S.5 above.

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.

This is not applicable because the data are not estimated based on samples. Rather, the data include all nursing home residents nationally who do not meet the exclusion criteria.

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.

This is not applicable because this measure is not based on survey/patient-reported data.

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

If other, please describe in S.18.

Assessment 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 data source is the Minimum Data Set (MDS) 3.0, and the collection instrument is the Resident Assessment Instrument (RAI). For MDS 3.0 item sets used to calculate the quality measure, please see: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/NHQIMDS30TechnicalInformation.

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

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

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

Facility

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

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.)*

This is not applicable because this is not a composite performance measure.

2. Validity – See attached Measure Testing Submission Form

NQF-0674-Testing-20210408-508.docx

2.1 For maintenance of endorsement

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

Yes

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include

information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

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

No - This measure is not risk-adjusted

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

Measure Number (*if previously endorsed*)**:** 0674

Measure Title: Percent of Residents Experiencing One or More Falls with Major Injury (Long-Stay)

Date of Submission: 1/5/2021

Type of Measure:

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

*cell intentionally left blank

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
X other: Nursing Home Minimum Data Set (MDS) 3.0	X other: Nursing Home Minimum Data Set (MDS) 3.0

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

The dataset used for testing was the Nursing Home Minimum Data Set (MDS) 3.0, which is one of three components of the Resident Assessment Instrument (RAI). The RAI is a tool used by nursing home staff to collect information on residents' strengths and needs. The MDS contains screening, clinical, and functional status elements, such as definitions and coding categories. These elements form the foundation of the comprehensive RAI for all eligible Medicare and Medicaid beneficiaries who are residents of nursing homes. The MDS items standardize how information about resident status and condition is recorded and shared within the facility, between facilities, and between facilities and outside agencies. Nursing homes are required to complete assessments on a regular basis, and the assessment requirements for the MDS are applicable to all residents in Medicare and/or Medicaid certified long-term care facilities, regardless of payment source or payer source.

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

Critical Data Element Testing

The RAND Development and Validation study from August 2006 to February 2007 on the development and validation of a revised nursing home assessment tool was used for the testing of critical data elements.

Saliba, D., & Buchanan, J. (2008, April). *Development and validation of a revised nursing home assessment tool: MDS 3.0*. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation. Retrieved from

https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf.

Performance Measure Score Testing

MDS 3.0 data from FY2019 Quarter 2 was used to construct this measure and calculate the QM scores. The seasonal trend analysis in **Section 2b1** was conducted using data from FY2017 Quarter 1 to FY2019 Quarter 2. The split-half analysis in **Section 2a2** was conducted using data from FY2019 Quarter 1 to FY2019 Quarter 2. The signal-to-noise analysis in **Section 2a2** and the 95% confidence interval analysis in **Section 2b1** were conducted using data from FY2019 Quarter 2.

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
X hospital/facility/agency	X hospital/facility/agency
☐ health plan	☐ health plan
☐ other:	☐ other:

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

Critical Data Element Testing

The RAND Development and Validation of MDS 3.0 study sample included a representative sample of for-profit and not-for-profit facilities, and hospital-based and freestanding facilities, which were recruited for the study. The sample included 71 community nursing facilities in 8 states and 19 Veterans Affairs (VA) nursing homes (Saliba & Buchanan, 2008).

Saliba, D., & Buchanan, J. (2008, April). *Development and validation of a revised nursing home assessment tool: MDS 3.0*. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation. Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf.

Performance Measure Score Testing

The analysis of MDS 3.0 data included all nationwide nursing home facilities with sufficient denominator size ($n \geq 20$) to publicly report this measure in FY2019 Quarter 2 ($k = 14,286$), unless otherwise noted.

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

Critical Data Element Testing

The RAND Development and Validation of MDS 3.0 study sample included 3,822 residents from community nursing homes and 764 residents from VHA nursing homes (Saliba & Buchanan, 2008).

Saliba, D., & Buchanan, J. (2008, April). *Development and validation of a revised nursing home assessment tool: MDS 3.0*. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation. Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf.

Performance Measure Score Testing

The analysis of MDS 3.0 data included all long-stay residents who met the denominator inclusion criteria for this measure in facilities with sufficient sample size ($n \geq 20$, $k = 14,286$) to report this measure ($n = 1,012,706$) in FY2019 Quarter 2.

Table 1 describes the characteristics of the residents who were counted in the denominator after applying facility sample size restrictions to FY2019 Quarter 2 data ($n = 1,012,706$). The majority of residents who met the denominator criteria were female (64.2%) and white (73.2%), while a smaller proportion of residents were male (35.8%) and Black or African American (16.0%). A majority of residents were also dual-eligible for Medicare and Medicaid (74.8%). More than 40% of residents were over the age of 85, and 25.8% were between the ages of 75-84. The most frequently reported active diagnoses in the past 7 days (as recorded on the target assessment) were Hypertension (77.4%), Depression (53.1%), and Non-Alzheimer's Dementia (49.5%). Other common active diagnoses reported for more than 25% of residents were Arthritis (31.5%), Anemia (30.5%), Diabetes Mellitus (35.2%) and Anxiety Disorders (32.4%). Table 1 also outlines the characteristics of the residents who were counted in the numerator. Compared to the denominator, the numerator population had a higher share of females, residents above the age of 85, and residents diagnosed with osteoporosis and fracture (which is considered a major injury).

Table 1. Characteristics of Long-Stay Residents Included in Analyses, NQF #0674 (FY2019 Quarter 2)

	NQF #0674 Denominator			NQF #0674 Numerator		
Sex	**Frequency (n)**	**Total Observations (N)**	**Percentage (%)**	**Frequency (n)**	**Total Observations (N)**	**Percentage (%)**
Female	650,612	1,012,706	64.2%	24,016	32,521	73.8%
Male	362,094	1,012,706	35.8%	8,505	32,521	26.2%

	NQF #0674 Denominator			NQF #0674 Numerator		
Race/Ethnicity	**Frequency (n)**	**Total Observations (N)**	**Percentage (%)**	**Frequency (n)**	**Total Observations (N)**	**Percentage (%)**
White Only	741,211	1,012,706	73.2%	27,626	32,521	84.9%
Black or African American Only	161,646	1,012,706	16.0%	2,014	32,521	6.2%

Race/Ethnicity	Frequency (n)	Total Observations (N)	Percentage (%)	Frequency (n)	Total Observations (N)	Percentage (%)
Hispanic or Latino Only	59,485	1,012,706	5.9%	1,460	32,521	4.5%
Asian Only	22,461	1,012,706	2.2%	503	32,521	1.5%
American Indian/Alaska Native Only	4,787	1,012,706	0.5%	198	32,521	0.6%
Native Hawaiian or Other Pacific Islander Only	3,699	1,012,706	0.4%	84	32,521	0.3%
Multi-race	3,136	1,012,706	0.3%	70	32,521	0.2%
Missing race	16,281	1,012,706	1.6%	566	32,521	1.7%

	NQF #0674 Denominator			NQF #0674 Numerator		
Medicare-Medicaid Dual Eligibility	Frequency (n)	Total Observations (N)	Percentage (%)	Frequency (n)	Total Observations (N)	Percentage (%)
Dual-Eligible	757,954	1,012,706	74.8%	24,571	32,521	75.6%
Non-Dual	240,825	1,012,706	23.8%	7,533	32,521	23.2%
Missing	13,927	1,012,706	1.4%	417	32,521	1.3%

	NQF #0674 Denominator			NQF #0674 Numerator		
Age	Frequency (n)	Total Observations (N)	Percentage (%)	Frequency (n)	Total Observations (N)	Percentage (%)
<65	155,673	1,012,706	15.4%	2,523	32,521	7.8%
65-74	189,250	1,012,706	18.7%	4,647	32,521	14.3%
75-84	260,783	1,012,706	25.8%	8,672	32,521	26.7%
85+	407,000	1,012,706	40.2%	16,679	32,521	51.3%

	NQF #0674 Denominator			NQF #0674 Numerator		
Diagnoses	Frequency (n)	Total Observations (N)	Percentage (%)	Frequency (n)	Total Observations (N)	Percentage (%)
Arthritis	90,466	286,855	31.5%	3,584	10,397	34.5%
Osteoporosis	37,718	286,864	13.1%	2,070	10,397	19.9%

Diagnoses	Frequency (n)	Total Observations (N)	Percentage (%)	Frequency (n)	Total Observations (N)	Percentage (%)
Hip Fracture	17,172	926,551	1.9%	5,547	28,395	19.5%
Other Fracture	29,252	926,553	3.2%	6,687	28,397	23.5%
Depression	491,690	926,495	53.1%	17,077	28,400	60.1%
Stroke	116,855	926,537	12.6%	3,021	28,401	10.6%
Alzheimer's Disease	146,634	926,526	15.8%	5,515	28,397	19.4%
Non-Alzheimer's Dementia	458,179	926,486	49.5%	16,506	28,398	58.1%
Malnutrition or at risk for malnutrition	54,154	1,012,601	5.3%	2,007	32,518	6.2%
Cancer	19,727	286,815	6.9%	777	10,392	7.5%
Anemia	282,673	926,430	30.5%	9,535	28,399	33.6%
Heart Failure	197,036	926,532	21.3%	6,323	28,400	22.3%
Hypertension	716,697	926,492	77.4%	22,198	28,400	78.2%
Diabetes Mellitus	356,196	1,012,598	35.2%	10,263	32,519	31.6%
Anxiety Disorder	327,717	1,012,525	32.4%	12,585	32,517	38.7%
Asthma, Chronic Obstructive Pulmonary Disease, or Chronic Lung Disease	217,589	926,538	23.5%	7,019	28,406	24.7%

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.

Data for Critical Data Elements

RAND reliability analysis of data elements used the same sample as described in **Sections 1.5** and **1.6** (Saliba & Buchanan, 2008).

Saliba, D., & Buchanan, J. (2008, April). Development and validation of a revised nursing home assessment tool: MDS 3.0. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation. Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf.

Data for Measure Performance Score Testing

All analyses used the same data as described above in **Sections 1.2, 1.3, 1.5, and 1.6**.

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.

Resident-level social risk factor variables related to falls with major injury that were available in the MDS 3.0 dataset were selected, including age, race, Medicaid status, and gender. The descriptive statistics for all of these characteristics are listed in Table 1 under item 1.6 in response to NQF prompting for sample resident characteristics.

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*)

Critical Data Element Reliability

The national test of MDS 3.0 items examined the agreement between assessors (reliability). Quality Improvement Organizations were employed to identify gold-standard (research) nurses and recruit community nursing facilities to participate in the national evaluation (Saliba & Buchanan, 2008). The gold-standard nurses were trained in the MDS 3.0 instrument, and they, in turn, trained a facility nurse from each participating nursing facility in their home states. Residents participating in the test were selected to capture a representative sample of short- and long-stay residents. In this national test of the Falls with Major Injury item, the gold-standard nurse to gold-standard nurse comparison (measures instrument performance with highly trained nurses using research protocols) and gold-standard to facility-nurse comparison (measures performance in more operational environment in which one assessor has ongoing facility responsibilities) were examined. Saliba and Buchanan (2008) present Falls with Major Injury rates using the MDS 3.0 items at the resident-level, as well as Cohen's kappas, which were calculated to assess item reliability. Kappa is a statistical measure of inter-rater agreement for qualitative data, ranging from 0.0 to 1.0, where a rating of greater than 0.60 is considered substantial agreement (Landis & Koch, 1977).

Landis, JR, Koch, GG. The measurement of observer agreement for categorical data. *Biometrics 33*(1), p 159-174, 1977.

Saliba, D., & Buchanan, J. (2008, April). *Development and validation of a revised nursing home assessment tool: MDS 3.0*. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation. Retrieved from

https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf.

Performance Measure Score Reliability

a. Signal-to-noise analysis: The signal-to-noise ratio gives the proportion of variability in measure performance that can be explained by between-provider differences in provider performance rather than variability within a provider (e.g., measurement or sampling error). Since a fall with major injury is a binary outcome, the reliability was estimated using a beta-binomial model. The beta-binomial model assumes that the provider QM score for falls with major injury is a binomial random variable, conditional on the provider's true value that comes from a beta distribution. Data from FY2019 Quarter 2 were used to conduct this analysis by fitting the beta binomial model to the data. The estimated alpha and beta parameters from the model were used to calculate the provider-to-provider variance:

$$\sigma^2_{\text{provider-to-provider}} = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

The provider-specific error was calculated using the following formula, where "p" is each facility's QM score and "n" is the number of residents in each facility:

$$\sigma^2_{\text{provider-specific-error}} = \frac{p\,(1-p)}{n}$$

The reliability score for each facility was then calculated using the following formula:

$$\text{reliability} = \frac{\sigma^2_{\text{provider-to-provider}}}{\sigma^2_{\text{provider-to-provider}} + \sigma^2_{\text{provider-specific-error}}}$$

A reliability score closer to 1 implies that most of the variability is attributable to between-provider differences in performance, and a score closer to 0 implies that most of the variability in the measure is attributable to variation within providers.

b. Split-half reliability analysis: Split-half reliability assesses the internal consistency of a quality measure by randomly dividing the residents within each nursing facility into two halves and calculating the correlation between the nursing facility's quality measure scores based on the two randomly divided halves. When a nursing facility's residents, randomly divided, have similar scores to one another, the quality measure score is more likely to reflect systematic differences in nursing home-level quality rather than random variation. A split-half reliability analysis was conducted on all facilities with 40 or more residents counted in the measure denominator, with each facility having 20 or more residents in both of its random samples. Data from FY2019 Quarter 1 to FY2019 Quarter 2 were used to calculate the Pearson Product-Moment Correlation (r) and Spearman Rank Correlation (ρ) to measure the internal reliability.

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*)

Critical Data Element Reliability

a. To assess reliability of the MDS 3.0, RAND calculated the gold-standard nurse to gold-standard nurse comparison (measured instrument performance with highly trained nurses using research protocols)

and gold-standard to facility-nurse comparison (measured performance in more operational environment in which one assessor had ongoing facility responsibilities) for Falls with Major Injury (Saliba & Buchanan, 2008). The Kappa for gold-standard to gold-standard on the MDS 3.0 item was 0.967, and the Kappa for gold-standard to facility-nurse agreement on the MDS 3.0 item was 0.945. Kappa is a statistical measure of inter-rater agreement for qualitative data, ranging from 0.0 to 1.0. Ratings of 0.967 and 0.945 are considered "substantial agreements." These results are indicative of data element reliability.

Saliba, D., & Buchanan, J. (2008, April). *Development and validation of a revised nursing home assessment tool: MDS 3.0*. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation. Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf.

Performance Measure Score Reliability

a. Signal-to-noise analysis: The average signal-to-noise reliability score of this quality measure using facility scores based on data from FY2019 Quarter 2 was observed to be 0.45. This suggests that the measure is moderately reliable in separating facility characteristics from variability within facility. This moderate variability is expected for this outcome measure because falls with major injury is considered to be a "never event" or an adverse event that should never occur. With approximately 19% of facilities achieving a perfect score of 0.0% in the national level data, this results in a low variation of scores between facilities. Hence, a reliability score of 0.45 would imply that we would be able to detect enough differences in provider performance.

b. Split-half reliability analysis: The split-half correlation for this measure was positive ($r = 0.18$, $\rho = 0.18$, $p < .01$), providing limited evidence of internal reliability. These low correlations were expected due to a small amount of variation in performance among providers due to the rareness of the outcome being measured. The national-level distribution indicates that 19% of all providers achieved a perfect score of 0.0% (lower scores demonstrate higher quality), which gives rise to a positively skewed distribution of provider performance. Taking this finding into account, intra-facility splits of the data would result in pairs of zero values (0.0%) being compared for 19% of all facilities. Since correlations are calculated using the covariance of the data and the individual variances to naturalize the covariance to report a value range between -1 and 1, the small amount of variance in performance was expected to yield low correlation coefficients. Additionally, given the overall high-performance scores among non-perfect providers and the variation in denominator sizes, it is likely that splitting the data would also result in scores of 0% being compared to non-zero values. This would also result in low correlations because the data may not vary in a systematic or predictable way across the split samples. As described below, **Table 6** in **Section 2b4.2** demonstrates that the variation in scores is still sufficient to distinguish poor performers. The 50th percentile score is 2.9% and the 90th percentile score is 7.1%.

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?*)

Critical Data Element Reliability

The RAND Development and Validation of MDS 3.0 national pilot test study demonstrated excellent reliability for MDS 3.0 items used to calculate this measure. Although the RAND testing was conducted 13 years ago, the MDS 3.0 forms used in the RAND study are similar to the latest MDS 3.0 forms used in the testing of this measure. The MDS 3.0 item set has remained stable since RAND created the recommended MDS 3.0 form in 2008, except for select changes in item specifications and the addition of some new

items. In particular, the Falls with Major Injury item has the same look-back period and the same item wording in the latest MDS 3.0 form and the 2008 recommended form.

Performance Measure Score Reliability

These analyses demonstrate that the falls with major injury measure shows moderate evidence of internal reliability. The average signal-to-noise ratio across all providers is 0.45, which suggests that the measure is moderately reliable in separating provider characteristics from variability within provider. Additionally, as explained above, there is a strong explanation (i.e., positively skewed distribution of provider performance) of why the split-half reliability analysis would yield a more limited correlation.

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)*

Critical Data Elements

The RAND validation of MDS 3.0 study tested the criterion validity of the items by comparing how different nurses assessed the same residents using MDS 3.0. They compared gold-standard research nurses to gold-standard nurses, and they compare gold-standard nurses to staff nurses trained by the gold-standard nurses. Kappa statistic was calculated. Nurses who participated in RAND's national study on the development and validation of MDS 3.0 also completed a feedback survey at the end of the study to provide feedback on the falls items.

References:

1. Saliba, D., & Buchanan, J. (2008, April). *Development and validation of a revised nursing home assessment tool: MDS 3.0*. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation. Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf.

Performance Measure Score Validity

a. Correlation with related quality measures: Groups of quality measures that reflect similar care processes or outcomes were examined with the hypothesis that a facility's percentile ranking (compared to all facilities reporting the measure) may be somewhat consistent among related quality measures. Related MDS Quality Measures (such as percent of long-stay residents who received antipsychotic, antianxiety or hypnotic medication, percentage of long-stay residents who have depressive symptoms, percent of long-stay residents whose need for help with daily activities has

increased, etc.) that are associated with drug-related fall risk, Facility Five-Star Ratings, and claims-based quality measures were examined for this purpose (de Jong, Van der Elst, & Hartholt, 2013). Public reporting data was used to calculate these correlations between NQF #0674 (Percent of Residents Experiencing One or More Falls with Major Injury (long stay)) and related quality measures. Since the related quality measures do not reflect very closely tied care processes to NQF #0674 and instead capture other dimensions of quality, we expect the correlations to be small if present. We also anticipate that the direction of any correlation between Falls with Major Injury and Percent of Long-Stay Residents Who Were Physically Restrained and Overall Facility Five-Star Ratings will be negative, while the direction of any correlation between Falls with Major Injury and the following MDS Quality Measures to be positive: Percent of Long-Stay Residents Who Received Antipsychotic, Antianxiety or Hypnotic Medication, Percent of Long-Stay Residents Who Have Depressive Symptoms, Percent of Long-Stay Residents Whose Need for Help with Daily Activities Has Increased.

b. Variation by state: Analyses investigated whether or not variation in scores on this measure was substantially attributable to state-by-state differences. If a measure is subject to variation caused by other factors beyond facility control, such as state-level payment policies or demographics, this variation can be a threat to the validity of the measure. State-by-state differences likely explain a portion of the variation in the measure due to geographic differences in quality, but a large amount of variation explained by state could be problematic.

c. Seasonality: Another potential threat to the validity of a quality measure is seasonal variation. If a quality measure score varies substantially from quarter to quarter in a consistent pattern over time corresponding to changes in seasons, it is possible that the validity of the measure is being compromised due to influences not within a nursing home's control. To address whether seasonal variation might play a role, the trend in the national mean and median for this quality measure score between FY2017 Quarter 1 and FY2019 Quarter 2 was examined.

d. Stability analysis: The extent to which relative facility rank changed on this quality measure from FY2019 Quarter 1 to FY2019 Quarter 2 was also assessed by evaluating the percentage of facilities that changed in their percentile ranking (i.e., relative quality measure score) within 1 decile, between 1 and 2 deciles, between 2 and 3 deciles, and 3 or more deciles. Dramatic changes in the quality measure score or facility rank based on the score over time may indicate measure instability, rather than true changes in quality. An important caveat is that some degree of variation in performance across time is to be expected: very poor performance in one quarter may lead to immediate changes that improve performance in subsequent quarters, and some movement in performance becomes more likely with rare event outcomes.

e. Confidence interval analysis: Proportions of facilities with scores for this measure that are significantly different from the national facility-level mean were examined and stratified by facility denominator size. For this analysis, statistical significance was determined by using 95% confidence intervals. A facility's quality measure score was significantly different from the national mean if the national mean was not included in the facility's 95% confidence interval. Because this measure is focusing on an undesirable outcome, high-performing facilities should have scores that are significantly below average, and scores of low-performing facilities should be significantly above average. The analysis was stratified by facility denominator size to examine whether this feature of the measure varies by size.

References:

1. de Jong MR, Van der Elst M, Hartholt KA. Drug-related falls in older patients: implicated drugs, consequences, and possible prevention strategies. *Ther Adv Drug Saf*. 2013;4(4):147-154. doi:10.1177/2042098613486829

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

Critical Data Elements

The kappa for gold-standard nurse assessment to facility nurse assessment of falls with major injury was 0.959.

The results of the survey on nurses who participated in the RAND study on the development and validation of MDS 3.0 indicated that 88% felt that the fall-related injury definitions were clear and 94% felt that facility falls documentation should include the information needed to complete the section.

References:

1. Saliba, D., & Buchanan, J. (2008, April). *Development and validation of a revised nursing home assessment tool: MDS 3.0*. Contract No. 500-00-0027/Task Order #2. Santa Monica, CA: Rand Corporation. Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/downloads/MDS30FinalReport.pdf.

Performance Measure Score Validity

a. Correlation with related quality measures: Among facilities that could report both measures, the analysis found a low but statistically significant positive correlation between Falls with Major Injury and Percent of Long-Stay Residents Whose Need for Help with Daily Activities has Increased, Percent of Long-Stay Residents Who Received an Antianxiety or Hypnotic Medication, Percent of Long-Stay Residents Who Received Antipsychotic Medication, Percent of Long-Stay Residents Who Have Depressive Symptoms, and Number of Outpatient Emergency Department Visits per 1,000 Long-Stay Resident Days. A low but statistically significant negative correlation between Falls With Major Injury and Percent of Long-Stay Residents Who Were Physically Restrained was observed, as well as a low but statistically significant negative correlation between Falls With Major Injury and Overall Facility Five-Star Ratings. The coefficient estimates and associated p-values are summarized in **Table 2** below.

Table 2. Correlations between NQF #0674 and other related MDS Quality Measures, Facility Five-Star Ratings and Claims-based Quality Measures (FY2019 Quarter 2)

MDS Quality Measures	Spearman Correlation	P-Value
Percent of Long-Stay Residents Whose Need for Help with Daily Activities has Increased	0.086	<.0001
Percent of Long-Stay Residents Who Were Physically Restrained	-0.033	<.0001
Percent of Long-Stay Residents Who Received Antipsychotic Medication	0.061	<.0001
Percent of Long-Stay Residents Who Received an Antianxiety or Hypnotic Medication	0.077	<.0001
Percent of Long-Stay Residents Who Have Depressive Symptoms	0.072	<.0001
Facility Five-Star Ratings	**Spearman Correlation**	**P-Value**
Overall facility ratings	-0.032	<.0001

Claims Based Quality Measures	Spearman Correlation	P-Value
Number of Outpatient Emergency Department Visits per 1,000 Long-Stay Resident Days	0.109	<.0001

Variation by State: The proportion of variation in this measure explained by the state that facilities are located in is small though significant ($p < .001$). An analysis of variance showed that only 6.43% of the overall variance in this measure can be attributed to the state in which the facility is located. The average inter-quartile range of state-level scores is 3.8%. North Dakota, Wyoming, Nebraska, and South Dakota are among some of the states that have a higher mean, median, and interquartile range for the NQF #0674 QM score compared to the national average. Washington DC, Alaska, and California are some of the states that have a lower mean, median, and interquartile range for the NQF #0674 QM score compared to the national average. The state-level average scores and percentile distributions are summarized in **Table 3** below.

Table 3. State-level NQF #0674 QM score summary (FY2019 Quarter 2)

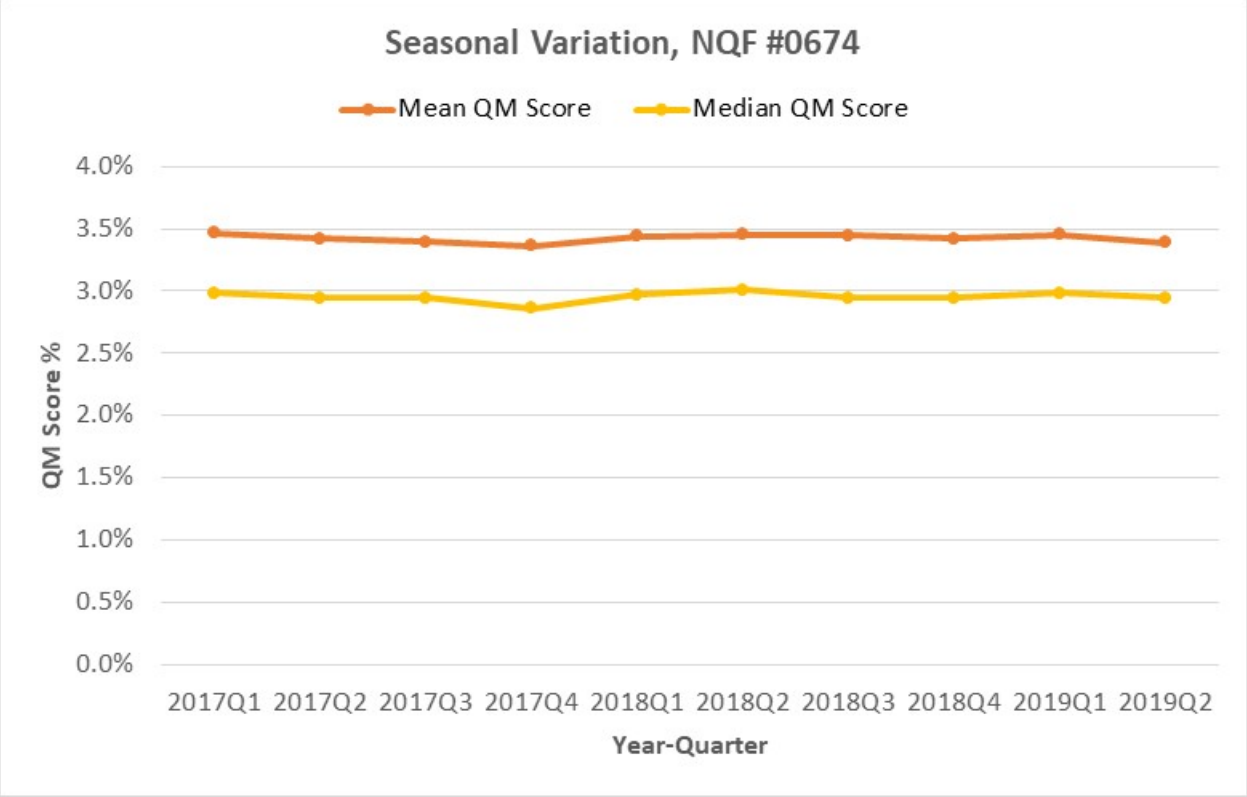
State	Number of facilities	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
WY	34	5.1%	4.2%	0.0%	2.5%	4.5%	6.8%	10.3%	4.3%
ND	78	5.0%	3.1%	0.0%	2.9%	4.7%	7.0%	9.5%	4.0%
NE	182	5.0%	3.7%	0.0%	2.6%	4.5%	7.1%	8.9%	4.6%
SD	103	5.0%	4.4%	0.0%	1.6%	4.0%	7.6%	10.0%	6.0%
MT	66	4.8%	3.8%	0.0%	2.6%	4.0%	6.4%	8.8%	3.8%
KS	304	4.8%	3.7%	0.0%	2.2%	4.3%	6.9%	9.7%	4.7%
VT	35	4.7%	4.0%	0.0%	1.8%	3.7%	8.1%	10.0%	6.3%
OK	283	4.7%	3.5%	0.0%	2.3%	4.3%	6.7%	9.3%	4.3%
NH	71	4.6%	3.1%	1.1%	2.7%	4.3%	5.9%	9.1%	3.2%
MO	491	4.4%	3.1%	0.0%	2.0%	4.1%	6.3%	8.6%	4.3%
WV	114	4.3%	2.6%	1.3%	2.5%	3.9%	5.8%	7.8%	3.3%
MN	338	4.2%	3.3%	0.0%	2.0%	3.8%	5.9%	8.5%	3.9%
ME	91	4.0%	3.3%	0.0%	1.9%	3.7%	5.9%	9.5%	4.0%
IN	505	4.0%	3.3%	0.0%	1.9%	3.3%	5.7%	7.9%	3.9%
KY	260	3.8%	2.7%	0.0%	1.9%	3.5%	5.5%	7.2%	3.5%
AR	221	3.8%	2.9%	0.0%	1.8%	3.4%	5.5%	7.3%	3.7%
RI	79	3.8%	2.9%	0.0%	1.9%	3.3%	5.7%	7.3%	3.7%
IA	415	3.7%	3.4%	0.0%	0.0%	3.2%	5.7%	8.3%	5.7%

State	Number of facilities	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
WI	325	3.7%	2.9%	0.0%	1.6%	3.4%	5.4%	8.1%	3.8%
CO	194	3.7%	2.7%	0.0%	1.9%	3.2%	5.2%	7.1%	3.3%
NC	398	3.6%	2.7%	0.0%	1.6%	3.2%	5.1%	7.2%	3.4%
NM	64	3.6%	2.6%	0.0%	1.8%	3.1%	4.9%	7.1%	3.2%
LA	261	3.5%	2.7%	0.0%	1.7%	3.1%	5.0%	7.2%	3.3%
CT	208	3.5%	2.3%	0.9%	1.9%	3.2%	4.9%	6.5%	2.9%
TN	297	3.5%	2.9%	0.0%	1.4%	3.1%	4.8%	7.3%	3.4%
OH	894	3.4%	2.9%	0.0%	1.5%	3.0%	5.0%	7.3%	3.5%
VA	265	3.4%	2.7%	0.0%	1.5%	2.9%	4.9%	6.9%	3.4%
TX	1,134	3.4%	2.8%	0.0%	1.4%	3.1%	5.0%	7.1%	3.6%
AL	223	3.4%	2.5%	0.0%	1.4%	3.3%	4.6%	6.3%	3.2%
MA	366	3.4%	2.3%	0.0%	1.6%	3.1%	4.8%	6.6%	3.1%
DE	44	3.3%	2.7%	0.0%	1.4%	2.7%	5.1%	6.9%	3.7%
PA	648	3.3%	2.7%	0.0%	1.3%	2.9%	4.6%	6.5%	3.3%
MS	197	3.2%	2.8%	0.0%	1.2%	2.6%	5.0%	7.1%	3.8%
IL	653	3.2%	2.7%	0.0%	1.2%	2.9%	4.5%	6.8%	3.3%
UT	76	3.2%	3.1%	0.0%	0.0%	2.7%	4.8%	7.7%	4.8%
SC	160	3.2%	2.8%	0.0%	1.3%	3.0%	4.5%	6.2%	3.2%
WA	186	3.2%	2.6%	0.0%	1.4%	2.7%	4.4%	6.8%	3.0%
MI	403	3.1%	2.5%	0.0%	1.4%	2.7%	4.7%	6.3%	3.3%
ID	67	3.1%	3.1%	0.0%	0.0%	3.1%	4.9%	7.1%	4.9%
OR	107	3.1%	3.2%	0.0%	0.0%	2.5%	4.8%	6.7%	4.8%
GA	349	3.0%	2.3%	0.0%	1.4%	2.6%	4.3%	6.3%	2.9%
HI	37	3.0%	3.4%	0.0%	0.8%	1.9%	3.8%	6.7%	3.1%
NY	598	2.8%	2.2%	0.0%	1.2%	2.4%	4.0%	5.7%	2.8%
FL	646	2.7%	2.2%	0.0%	1.2%	2.5%	3.9%	5.6%	2.7%
NV	49	2.7%	2.0%	0.0%	1.4%	2.7%	3.6%	4.9%	2.3%
NJ	333	2.6%	2.1%	0.0%	1.0%	2.4%	3.8%	5.3%	2.8%
AZ	112	2.6%	2.3%	0.0%	0.8%	2.3%	3.6%	6.6%	2.8%
MD	212	2.5%	2.5%	0.0%	0.0%	2.1%	3.9%	5.9%	3.9%
CA	1,087	1.8%	2.1%	0.0%	0.0%	1.4%	2.9%	4.7%	2.9%
AK	7	1.5%	1.4%	0.0%	0.0%	1.9%	2.9%	3.3%	2.9%

State	Number of facilities	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Interquartile range
DC	16	1.2%	1.2%	0.0%	0.0%	1.2%	2.2%	2.7%	2.2%

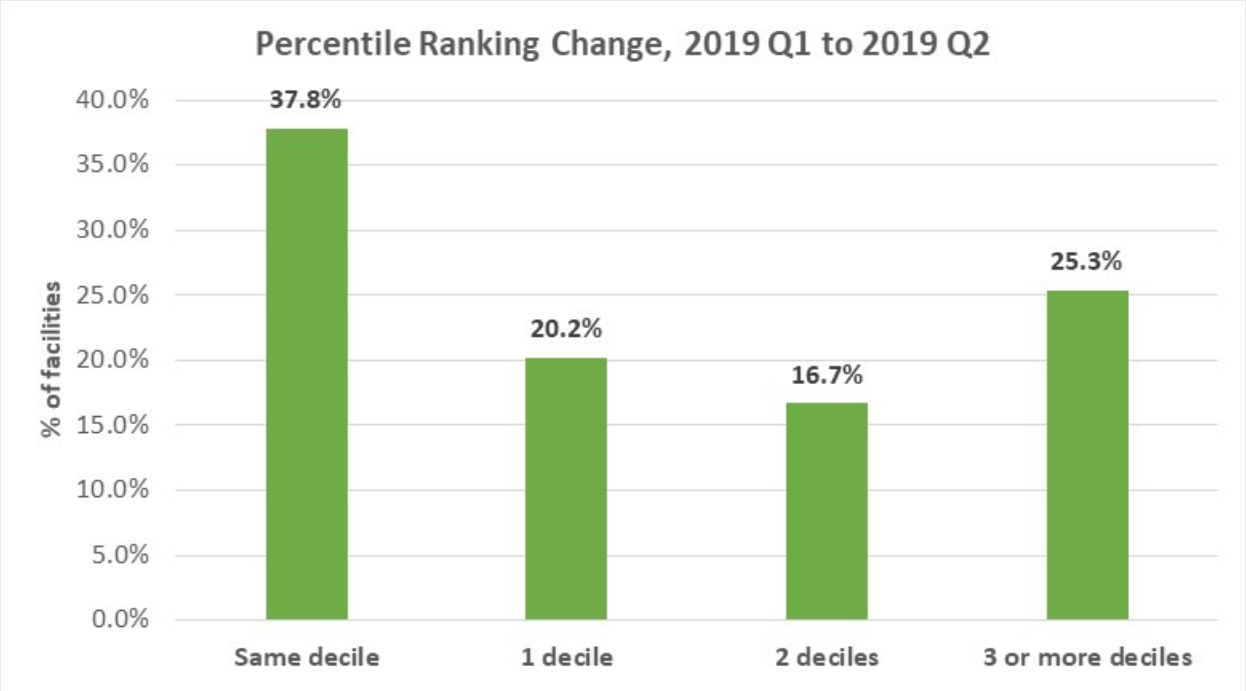
Seasonality: The seasonal variation in the measure score was examined by plotting the mean and median national level scores for each quarter from FY2017 Quarter 1 to FY2019 Quarter 2. There appears to be no perceptible seasonal variation for the mean and median of this measure. Both the mean and median scores for NQF #0674 have been fairly stable over the past several quarters. The results are presented in **Figure 1** below.

Figure 1. Trends over time for QM #0674 Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay)



b. Stability analysis: **Figure 2** illustrates the changes in facility rank by quality measure score from FY2019 Quarter 1 to FY2019 Quarter 2. Comparing ranks for these two quarters, 37.8% of facilities' percentile ranking was constant within the same decile, 20.2% of facilities changed rank within 1 decile, 16.7% changed rank within 2 deciles, and 25.3% changed rank by 3 or more deciles.

Figure 2. Decile Change in QM Ranks from FY2019 Quarter 1 to FY2019 Quarter 2, QM #0674 Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay)



c. Confidence interval analysis: Another measure of validity is performance relative to the mean, in which high-performing facilities should have scores that are significantly below average and low-performing facilities should have scores significantly above average. **Table 4** shows the proportions of facilities that scored significantly higher or lower than the national facility-level mean in FY2019 Quarter 2. For this analysis, statistical significance was determined using 95% confidence intervals. A facility's quality measure score was statistically significantly different from the national mean if the national mean was not within that facility's 95% confidence interval. This analysis was also stratified by decile of facility size based on the number of residents who qualify for the denominator count.

22.7% of facilities had a score that was statistically significantly different from the national mean with 95% confidence. Approximately 3.1% of facilities had scores that were statistically significantly higher than the national mean, and 19.5% of facilities had scores that were statistically significantly lower than the national mean. This suggests that although the measure may not distinguish among facilities across the full range of performance, it does identify an important set of facilities that are significantly poor or significantly good performers. Additionally, the proportion of facilities with scores that are significantly different from the national mean vary as a function of the number of residents included in the denominator for this measure. The percentage of facilities with scores significantly different from the mean decreases with the number of residents, except in the largest facility decile (10th decile).

Table 4. Proportion of Facilities with Scores Significantly Different from the National Facility-Level Mean, Stratified by Facility Denominator Size for NQF #0674, FY2019 Quarter 2

Decile of denominator size in residents	Number of facilities	Number of facilities with 95% confidence interval lower than national mean (%): N	Number of facilities with 95% confidence interval lower than national mean (%): %	Number of facilities with 95% confidence interval higher than national mean (%): N	Number of facilities with 95% confidence interval higher than national mean (%): %	Total number of facilities with scores significantly different from mean (%): N	Total number of facilities with scores significantly different from mean (%): %
1st Decile (20-38)	1,347	561	41.6%	40	3.0%	601	44.6%
2nd Decile (38-48)	1,456	461	31.7%	32	2.2%	493	33.9%
3rd Decile (48-57)	1,458	366	25.1%	52	3.6%	418	28.7%
4th Decile (57-66)	1,317	279	21.2%	39	3.0%	318	24.1%
5th Decile (66-76)	1,400	238	17.0%	42	3.0%	280	20.0%
6th Decile (76-86)	1,547	238	15.4%	43	2.8%	281	18.2%
7th Decile (87-97)	1,393	180	12.9%	46	3.3%	226	16.2%
8th Decile (97-112)	1,418	180	12.7%	50	3.5%	230	16.2%
9th Decile (112-139)	1,493	132	8.8%	48	3.2%	180	12.1%
10th Decile (139-767)	1,457	157	10.8%	52	3.6%	209	14.3%
Total	**14,286**	**2,792**	**19.5%**	**444**	**3.1%**	**3,236**	**22.7%**

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?*)

Performance Measure Score Validity

This measure has excellent item level validity, with no perceptible seasonal variation. Although the state-level variation is statistically significantly different from zero, it is a minimal source of variation as expected.

This measure is modestly and significantly correlated with Percent of Long-Stay Residents Whose Need for Help with Daily Activities Has Increased, Percent of Long-Stay Residents Who Received an Antianxiety or Hypnotic Medication, Percent of Long-Stay Residents Who Received Antipsychotic Medication, Percent of Long-Stay Residents Who Have Depressive Symptoms, and Number of Outpatient Emergency Department Visits per 1,000 Long-Stay Resident Days. This measure also has a modest negative correlation with Percent of Long-Stay Residents Who Were Physically Restrained and Overall Facility Five-Star Ratings. There may be several reasons for these relatively low correlations, including the strong possibility that these measures simply capture different aspects of care. Further, this measure captures only those falls that result in a significant injury. It is possible that these related quality measures have a stronger correlation with falls, but not major injury. Additionally, some of these quality measures are also low frequency measures, which could contribute to the low correlations with Falls With Major Injury. While the majority of these correlation coefficients are modest in magnitude, they are all statistically significant and in the expected direction.

The confidence interval analysis for this measure indicates that there are meaningful differences in facility-level scores for this measure. The stability analysis indicates that the facility scores and relative ranks for this measure are relatively stable from one quarter to next. Since approximately 20% of facilities have a perfect score in each quarter, the occurrence of a "never-event" in a facility that had a perfect score in a previous quarter can shift that facility's performance by more than 3 deciles in the next quarter.

2b2. EXCLUSIONS ANALYSIS

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

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

Please see **Section 2b6**. "Missing data analysis and minimizing bias for analysis of this measure's exclusions," which are only for missing data on the applicable falls items.

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*)

Please see **Section 2b6**. "Missing data analysis and minimizing bias for analysis of this measure's exclusions," which are only for missing data on the applicable falls items.

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*)

Please see **Section 2b6**. "Missing data analysis and minimizing bias for analysis of this measure's exclusions," which are only for missing data on the applicable falls items.

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

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

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

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with risk factors**

☐ **Stratification by risk categories**

☐ **Other,**

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.

Not applicable. This measure is not risk-adjusted.

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.

The measure is not risk adjusted through a statistical model nor through stratification. However, several social risk factors and clinical diagnoses for risk adjustment were explored. The discussion is presented in **Section 2b3.3a** below.

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?

Risk Adjustor Selection – Conceptual Rationale and Statistical Testing

Clinical Risk Factors

- Depression: Residents with depressive symptoms who use antidepressants may be at higher risk of experiencing a fall than residents who do not take antidepressants (Seppala et al., 2018).

- Dementia: Residents with dementia may have impairment in recognition of sensory input and coordination of movement, leading to a greater probability of experiencing a fall than residents without dementia (Gulka et al., 2019).
- Weight loss: Sudden weight loss is indicative of an underlying health issue that could result in muscle wasting and is typically associated with frailty. Frail residents are more likely to experience major injuries caused by falls (Trevisan et al, 2019).
- BMI: Research indicates that malnourished individuals aged ≥ 65 years are at higher risk for experiencing a fall than well-nourished individuals (Trevasian et al., 2019). Therefore, having a lower BMI may place a resident at higher risk for experiencing a fall with major injury than having a higher BMI.

In addition to these risk factors, Osteoporosis (I3800) was considered as a potential clinical risk factor. However, this item is only collected in the annual comprehensive MDS assessment and would have a high missing rate in the data used for these analyses.

Social Risk Factors

- Age: Older residents are more prone to falls with major injury due to their age, blurred vision, and impairment in the coordination of their movement (Ambrose, Paul, & Hausdorff, 2013).

References:

1. Seppala, L. J., Wermelink, A. M. A. T., de Vries, M., Ploegmakers, K. J., van de Glind, E. M. M., Daams, J. G., & van der Velde, N. (2018). Fall-risk-increasing drugs: A systematic review and meta-analysis: II. Psychotropics. Journal of the American Medical Directors Association, 19(4), 371e11-371e17. doi:10.1016/j.jamda.2017.12.098
2. Gulka, H. J., Patel, V., Arora, T., McArthur, C., & Iaboni, A. (2019). Efficacy and generalizability of falls prevention interventions in nursing homes: A systematic review and meta-analysis. Journal of the American Medical Directors Association,1-12. doi:10.1016/j.jamda.2019.11.012
3. Trevisan, C., Crippa, A., Ek, S., Welmer, A.-K., Sergi, G., Maggi, S., Manzato, E., Bea, J. W., Cauley, J. A., Decullier, E., Hirani, V., LaMonte, M. J., Lewis, C. E., Schott, A.-M., Orsini, N., & Rizzuto, D. (2019). nutritional status, body mass index, and the risk of falls in community-dwelling older adults: A systematic review and meta-analysis. Journal of the American Medical Directors Association,20(5), 569-82. https://doi.org/10.1016/j.jamda.2018.10.027
4. Ambrose, A. F., Paul, G., & Hausdorff, J. M. (2013). Risk factors for falls among older adults: A review of the literature. Maturitas, 75(1), 51-61. doi:10.1016/j.maturitas.2013.02.009

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?

Internal data analysis

We created binary variables for each risk factor described above as follows:

- Age: Defined from the Birth Date item (A0900) on the MDS. Age ≥ 85 is defined as 1 if the resident's age is 85 years or older at the time of the target assessment and 0 if otherwise. The birth date was not missing in any assessment in the data.

- BMI: Calculated using the Height (K0200A) and Weight (K0200B) items in the MDS. Since height is recorded as inches and weight is recorded as pounds in the MDS assessment, the BMI was calculated using the following formula:

$$\text{BMI} = \frac{\text{Weight}}{\text{Height}^2 \times 703}$$

  We then categorized the BMI into one of the following categories:

  - Underweight: If the BMI is under 18.5
  - Normal weight: If the BMI is between 20 to 25
  - Overweight: If the BMI is between 25-30
  - Obese: If the BMI is over 30

- Depression: Defined as 1 if item I5800 in the MDS has a value of 1 and 0 if otherwise.

- Dementia: Defined as 1 if item I4800 in the MDS has a value of 1 and 0 if otherwise.

- Weight-loss: Defined as being on a physician prescribed weight-loss regimen if item K0300 in the MDS has a value of 1, defined as weight loss not on a physician prescribed weight loss regimen if item K0300 in the MDS has a value of 2 and no weight loss if item K0300 in the MDS has a value of 0.

The results of the different risk-adjustment models using these risk factors are summarized in **Tables 5a-5h** below. All models were estimated as logistic models. Overall, the odds ratios of all the risk adjustment models are statistically significant at the 5% level. However, the C-statistics of all the models indicate weak model performance and suggests that the models do not have sufficient predictive ability.

Table 5a. Assessment of Alternate Risk Adjustment Specifications: Age as the covariate, NQF #0674 (FY2019 Quarter 2)

Model Covariates	Frequency of residents w/ covariate value	Frequency of residents who had falls w/ major injury	% of residents who had falls w/ major injury	Odds Ratio	95% CI	C-statistic
Candidate model: Age >= 85	407,000	16,679	4.1%	1.59	[1.556 1.627]	0.56

Table 5b. Assessment of Alternate Risk Adjustment Specifications: BMI as the covariate, NQF #0674 (FY2019 Quarter 2)

Model Covariates	Frequency of residents w/ covariate value	Frequency of residents who had falls w/ major injury	% of residents who had falls w/ major injury	Odds Ratio	95 % CI	C - statistic
Candidate model (Base case -Normal weight):	345,544	14,009	4.1%	No data	No data	0.58
BMI: Underweight	67,960	3,217	4.7%	1.18	[1.131 1.223]	No data
BMI: Overweight	271,555	8,018	3.0%	0.72	[0.700 0.740]	No data
BMI: Obese	298,160	6,360	2.1%	0.52	[0.501 0.532]	No data

**Note: Residents with an outlier BMI value less than 7.5 and BMI value greater than 82 were excluded from the risk-adjustment model*

Table 5c. Assessment of Alternate Risk Adjustment Specifications: Weight-loss as the covariate, NQF #0674 (FY2019 Quarter 2)

Model Covariates	Frequency of residents w/ covariate value	Frequency of residents who had falls w/ major injury	% of residents who had falls w/ major injury	Odds Ratio	95 % CI	C - statistic
Candidate model (Base case -No weight loss):	923,512	27,861	3.0%	No data	No data	0.53
Weight loss on physician prescribed regimen	10,872	446	4.1%	1.38	[1.251 1.515]	No data
Weight loss not on physician prescribed regimen	70,189	3,970	5.7%	1.93	[1.863 1.995]	No data

Table 5d. Assessment of Alternate Risk Adjustment Specifications: BMI and Age as the covariates, NQF #0674 (FY2019 Quarter 2)

Model Covariates	Frequency of residents w/ covariate value	Frequency of residents who had falls w/ major injury	% of residents who had falls w/ major injury	Odds Ratio	95 % CI	C - statistic
Candidate model (Base case -Normal weight and Age < 85):	174,905	5,747	3.3%	No data	No data	0.59
BMI: Underweight & Age < 85	31,465	1,242	3.9%	1.21	[1.137 1.288]	No data
BMI: Overweight & Age < 85	159,910	4,089	2.6%	0.77	[0.742 0.804]	No data
BMI: Obese & Age < 85	224,043	4,389	2.0%	0.59	[0.565 0.612]	No data
BMI: Normal weight & Age >= 85	170,639	8,262	4.8%	1.50	[1.447 1.550]	No data
BMI: Underweight & Age >= 85	36,495	1,975	5.4%	1.68	[1.598 1.775]	No data
BMI: Overweight & Age >= 85	111,645	3,929	3.5%	1.07	[1.030 1.119]	No data
BMI: Obese & Age >= 85	74,117	1,971	2.7%	0.80	[0.764 0.847]	No data

**Note: Residents with an outlier BMI value less than 7.5 and BMI value greater than 82 were excluded from the risk-adjustment model*

Table 5e. Assessment of Alternate Risk Adjustment Specifications: Weight-loss and Age as the covariates, NQF #0674 (FY2019 Quarter 2)

Model Covariates	Frequency of residents w/ covariate value	Frequency of residents who had falls w/ major injury	% of residents who had falls w/ major injury	Odds Ratio	95 % CI	C – statistic
Candidate model (Base case -No weight loss and Age < 85):	556,287	13,651	2.5%	No data	No data	0.58
Weight loss on physician prescribed regimen & Age < 85	7,316	246	3.4%	1.39	[1.219 1.575]	No data

Model Covariates	Frequency of residents w/ covariate value	Frequency of residents who had falls w/ major injury	% of residents who had falls w/ major injury	Odds Ratio	95 % CI	C – statistic
Weight loss not on physician prescribed regimen & Age < 85	38,324	1,854	4.8%	2.02	[1.923 2.124]	No data
No weight loss & Age >= 85	367,225	14,210	3.9%	1.60	[1.562 1.639]	No data
Weight loss on physician prescribed regimen & Age >= 85	3,556	200	5.6%	2.38	[2.057 2.741]	No data
Weight loss not on physician prescribed regimen & Age >= 85	31,865	2,116	6.6%	2.83	[2.700 2.967]	No data

Table 5f. Assessment of Alternate Risk Adjustment Specifications: Depression as the covariate, NQF #0674 (FY2019 Quarter 2)

Model Covariates	Frequency of residents w/ covariate value	Frequency of residents who had falls w/ major injury	% of residents who had falls w/ major injury	Odds Ratio	95 % CI	C – statistic
Candidate model (Base case -No depression):	434,805	11,323	2.6%	No data	No data	0.54
Depression	491,690	17,077	3.5%	1.35	[1.314 1.379]	No data

Table 5g. Assessment of Alternate Risk Adjustment Specifications: Dementia as the covariate, NQF #0674 (FY2019 Quarter 2)

Model Covariates	Frequency of residents w/ covariate value	Frequency of residents who had falls w/ major injury	% of residents who had falls w/ major injury	Odds Ratio	95 % CI	C – statistic
Candidate model (Base case -No dementia):	468,307	11,892	2.5%	No data	No data	0.55
Dementia	458,179	16,506	3.6%	1.43	[1.4, 1.469]	No data

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.**

The analyses in **Section 2b3.4a** indicates that the odds of fall with major injury are almost 1.6 times higher for residents over the age of 85, compared to those under the age of 85. This result seems reasonable because older residents tend to be more frail compared to younger residents. Residents who are underweight and over the age of 85 have greater odds of fall with major injury compared to normal weight residents over the age of 85. Obese and overweight residents have lower odds of fall with major injury compared to normal weight residents. We hypothesize that this is because residents with larger BMI are less likely to be mobile and also have more fatty tissue that can serve as a protectant and help prevent major injury if they were to fall. Additionally, residents with depression and dementia were also observed to have greater odds of fall with major injury. While the results of all our risk adjustment models appear to be statistically significant at the 5% level, we observed low C-statistics for these models. This suggests that our models do not have sufficient predictive ability.

It is important to note that a fall resulting in a major injury is considered to be a never event, which is a serious and costly error in the provision of healthcare services that should never happen. Therefore, the importance of this quality measure is in identifying providers with higher frequency of patients experiencing falls with major injury, which contribute to burden on the healthcare system through an increase in negative health outcomes, healthcare utilization, and costs that remain pertinent regardless of provider risk. While some providers do care for a larger proportion of patients or residents at increased risk for falls with major injury, providers with high-risk patient populations should be encouraged to adopt preventative clinical care practices instead of statistical adjustment of these providers' scores. Doing so helps to ensure that providers are held accountable for the safety of all patients and that preventative measures are implemented in a clinically appropriate and equitable manner.

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*). *Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below*.

If stratified, skip to 2b3.9

This is not applicable. This measure is not risk-adjusted.

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

This is not applicable. This measure is not risk-adjusted.

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

This is not applicable. This measure is not risk-adjusted.

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

This is not applicable. This measure is not risk-adjusted.

2b3.9. Results of Risk Stratification Analysis:

This is not applicable. This measure is not risk-adjusted.

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*)

This is 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*)

This is not applicable. This measure is not risk-adjusted.

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)*

In order to identify meaningful differences in facility performance on NQF #0674, the current variability in the facility-level quality measure scores was explored (see 2b4.2). The proportions of facilities with scores for this measure that are significantly different from the national facility-level mean were also explored and stratified by facility denominator size (see 2b1.3). For this analysis, statistical significance was determined using 95% confidence intervals: a facility's quality measure score was significantly different from the national mean if the national mean was not included in the facility's 95% confidence interval.

High-performing facilities should have scores that are significantly below average, and scores of low-performing facilities should be significantly above average. The analysis was stratified by facility denominator size to examine whether this feature of the measure varies by size.

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*)

Table 6 describes the current variability in the quality measure scores of facilities nationally. We find that the mean facility-level score for this quality measure was 3.4% in Quarter 2, 2019 with a median score of 2.9%. The interquartile range for this measure was 3.6 percentage points. Among facilities who were eligible to publicly report this measure, 19% (*k* = 2,714) had perfect scores of 0%.

Table 6. National Facility-Level Score Distribution, NQF #0674 Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay), FY2019 Quarter 2

K	Mean score	Std dev.	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	% of facilities with perfect score	Interquartile range
14,286	3.4%	2.9%	0.0%	1.3%	2.9%	4.9%	7.1%	19.0%	3.6 % points

NOTES: *k* = number of facilities that meet minimum requirements for public reporting this quality measure.

Table 4 in Section 2b1.3 above shows the proportions of facilities that scored statistically significantly higher or lower than the national facility-level mean in FY2019 Quarter 2. For this analysis, statistical significance was determined using 95% confidence intervals: a facility's quality measure score was significantly different from the national mean if the national mean was not within the facility's 95% confidence interval.

Overall, 22.7% of facilities scored significantly differently than the national mean in FY2019 Quarter 2, indicating that there are meaningful differences in facility-level scores for this measure even though the measure may not differentiate well in the middle of the distribution of facilities. The data were also stratified by the facility denominator size to examine the relationship between facility size and the reliability of facility scores. The proportions of facilities with scores that were significantly different from the national mean varied as a function of the number of residents included in the denominator for this measure. In general, the percentage of facilities with scores that were statistically significantly different from the mean decreased as the number of residents increased.

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 share of facilities with scores that differed significantly from the national mean was 22.7%. There were 19.5% of facilities with scores statistically significantly lower than the national mean and 3.1% of facilities with scores statistically significantly higher than the national mean with 95% confidence. These results suggest that the measure provides the opportunity to identify nursing homes that are both particularly good performers, as well as those that are particularly poor performers regarding injurious falls. However, there is more limited ability to distinguish among facilities in the middle of the

distribution. Additionally, the quality measure scores vary enough from the national mean that there are meaningful differences in facility-level scores for this measure.

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*)

This is 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*)

This is 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*)

This is 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 non-responders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

The rate of missing data per total number of assessments was assessed.

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*)

Based on analyses of MDS 3.0 data, missing data is not a threat to validity for this measure as very few resident episodes are excluded from the QM calculation due to missing data (An episode is excluded if the occurrence of falls was not assessed (J1800 = [-]) or if the assessment indicates that a fall occurred (J1800 = [1]) and the number of falls with major injury was not assessed (J1900C = [-])). Only 6 episodes in the FY2019 Quarter 2 long stay resident sample were excluded from the denominator for this measure because of missing responses on either J1800 or J1900.

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 non-responders) 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*)

There were too few residents excluded due to missing data to warrant concern over missing data introducing bias into the measure. Additionally, the number of excluded cases was too small to test for any kind of differences between facilities. Therefore, no further analyses were performed regarding missing data and this measure.

3. Feasibility

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

3a. Byproduct of Care Processes

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

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

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)

If other:

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 electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS)

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).

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.

The general data collection method for the MDS 3.0 is currently in operational use and mandatory for all Medicare/Medicaid certified nursing facilities.

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)*.

This is not applicable.

4. Usability and Use

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

4a. Accountability and Transparency

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

4.1. Current and Planned Use

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

Specific Plan for Use	Current Use (for current use provide URL)
*	Public Reporting Care Compare https://www.medicare.gov/care-compare/ Provider Data Catalog https://data.cms.gov/provider-data/ Care Compare https://www.medicare.gov/care-compare/ Provider Data Catalog https://data.cms.gov/provider-data/ Quality Improvement (external benchmarking to organizations) Certification And Survey Provider Enhanced Reports (CASPER) https://www.qtso.com/providernh.html Quality Improvement (Internal to the specific organization) Certification And Survey Provider Enhanced Reports (CASPER) https://www.qtso.com/providernh.html

*cell intentionally left blank

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

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

Public Reporting:

Program and sponsor: Care Compare and Provider Data Catalog/Centers for Medicare and Medicaid
Purpose: Consumer information
Geographic area and number and percentage of accountable entities and patients included: All United States Nursing Homes with Medicare-eligible long-stay residents. In quarter 2 of 2019 there were 15,215 eligible facilities and 1,033,796 residents with target assessments, and 14,286 facilities (93.9%) had sufficient sample size (20 or more long-stay residents included in the denominator) to report on this measure, and 1,012,706 residents (98.0%) were included in the calculation of this measure. Four individual quarter scores are publicly reported on Provider Data Catalog. To enhance measurement stability and reliability beyond a one-quarter measure, a four-quarter average version of the measure is publicly reported as part of the Five-Star Quality Rating System through Care Compare and Provider Data Catalog. Five-Star is a rating system CMS created to help consumers, families and care givers compare nursing homes more easily.

Quality Improvement with Benchmarking (external benchmarking to multiple organizations):
Program and sponsor: Certification and Survey Provider Enhanced Reports (CASPER)/Centers for Medicare and Medicaid
Purpose: Quality improvement
Geographic area and number and percentage of accountable entities and patients included: All United States Medicare/Medicaid certified Nursing Homes with eligible long-stay residents regardless of denominator sample size. In quarter 2 of 2019 there were 15,215 eligible facilities and 1,033,796 residents with target assessments.

Quality Improvement (internal to the specific organization):
Program and sponsor: Certification and Survey Provider Enhanced Reports (CASPER)/Centers for Medicare and Medicaid
Purpose: Quality improvement
Geographic area and number and percentage of accountable entities and patients included: All United States Medicare/Medicaid certified Nursing Homes with eligible long-stay residents regardless of denominator sample size. In quarter 2 of 2019 there were 15,215 eligible facilities and 1,033,796 residents with target assessments.

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?*)
This is not applicable; this measure is publicly reported.

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.*)

This is not applicable; this measure is publicly reported.

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.

This quality measure (NQF #0674, Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay) is part of the Nursing Home Quality Initiative (NHQI). Information on this measure is available to both nursing home providers and to the public.

All United States Medicare and/or Medicaid certified nursing home providers may view their performance results for this and other NHQI measures via the Certification and Survey Provider Enhanced Reports (CASPER) system. These CASPER MDS 3.0 QM reports are intended to provide nursing home providers with feedback on their quality measure scores, helping them to improve the quality of care delivered to their residents. CASPER MDS 3.0 reports also include Resident-Level Quality Measure Reports, which allow providers to identify the residents that trigger a particular quality measure (by scanning a column of interest and looking for the residents with an "X") and to identify residents who trigger multiple quality measures. Providers can use this information to target residents for quality improvement activities. Quality measure reports are also available to state surveyors and facility staff through the CASPER reporting system.

Consumers, including current and prospective nursing home residents and their families/caregivers, may access nursing home performance scores on this quality measure via the Care Compare website (https://www.medicare.gov/care-compare/?providerType=NursingHome) or the Provider Data Catalog (https://data.cms.gov/provider-data/). The Care Compare site reports the four-quarter average, while the Provider Data Catalog site reports the one-quarter version of the measure alongside the four-quarter average.

CMS also publishes composite quality ratings on Care Compare via the Five-Star Rating System. Five-Star features an overall quality rating of one to five stars based on nursing home performance on three domains, each of which has its own rating. The four-quarter version of this quality measure (NQF #0674, Percent of Residents Experiencing One or More Falls with Major Injury (Long Stay)) is one of the clinical measures that contribute to the rating of the Quality Measures domain of Five-Star. The Five-Star program requires the measure denominator to include at least 20 residents' assessments across four quarters of data.

Further, providers have an opportunity to review their performance prior to public reporting on the Nursing Home Compare website via Provider Preview Reports, also available through the CASPER system. These reports allow providers to view their quality measure scores for each NHQI measure, along with state and national averages for comparison, to identify potential errors in data submission or other information and request an update. These reports also allow providers to view their Five-Star rating. Detailed instructions on how to view and interpret reports, including an explanation of differences between the quality measure reports and publicly reported information, are provided in the CASPER Reporting MDS Provider Users Guide, Section 11, which can be found at the following website:
https://qtso.cms.gov/system/files/qtso/cspr_sec11_mds_prvdr_0.pdf

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.

The CASPER reports are available to providers on-demand with quality measure data updated monthly. Care Compare reports the rolling average of four quarters for the quality measure, comparing each nursing home's score to both the state and national average; providers can preview this information before it is publicly reported.

Detailed instructions on how to view and interpret reports, including an explanation of differences between the quality measure reports and publicly reported information, are provided in the CASPER Reporting MDS Provider Users Guide, Section 11, at the following website:
https://qtso.cms.gov/system/files/qtso/cspr_sec11_mds_prvdr_0.pdf

CMS provides technical users' guides (https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/Downloads/usersguide.pdf) on how the quality measures are used in the 5-star rating system, as well as a Help Line, which is accessible by telephone and email, to answer provider questions about the NHQI quality measures and reporting requirements.

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.

CMS is committed to receiving ongoing feedback on measures implemented as part of the NHQI. CMS takes into consideration feedback and input on measure performance and implementation through the appropriate sub-regulatory communication channels, including but not limited to: NQF public comment periods held as part of endorsement processes; feedback from providers submitted to the CMS quality measure support inboxes and feedback from the provider community on Open Door Forums (ODFs).

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

Upon review of all inquiries submitted to the quality measure support inbox between 10/2019 and 02/2021, those being measures raised no concerns regarding the performance and implementation of NQF 0674.

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

Upon review of all inquiries submitted to the quality measure support inbox between 10/2019 and 02/2021, one other user raised a concern regarding the overlap in implementation of NQF 0674 and the SNF QRP application of this measure. There are rare cases where a SNF stay overlaps with a LS Nursing Home stay. In such cases, the resident must be discharged from the short stay as their Medicare PPS Part A stay ends, and after discharge the resident remains in the facility. Under such circumstances, a fall with major injury could be included in both stays (SNF and LS Nursing Home) if the fall with major injury occurred before the Medicare PPS Discharge and within the target assessment period for the long-stay measure.

However, NQF 0674 and the SNF QRP application of the measure are not in the same quality reporting program. Only the LS Nursing Home measure is included in the Nursing Home Five-Star Quality Rating System.

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.

The user's feedback was not a direct comment about the implementation or performance of the LS Nursing Home falls with major injury measure, but rather a question about its overlap with the SNF QRP application of the measure. It is not concerning for the event to be flagged twice for the same facility because NHQI focuses on the quality of care for LS residents while SNF QRP focuses on the quality of care for SNF residents. Therefore, the user's comment was not taken into consideration during the Spring 2021 NQF maintenance of 0674.

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.

Progress (trends in performance results, number and percentage of people receiving high-quality healthcare)

- The national facility-level mean and median scores for the Percent of Residents Experiencing One or More Falls with Major Injury demonstrate stability from quarter to quarter (Figure 1 of NQF Testing Form). Overall, the national facility-level mean and median scores have decreased marginally and indicate a slight

improvement in performance over time. The mean score for this measure was 3.5% in quarter 1 of 2017 and the median score was 3.0%. In Q2 2019, the mean and median were 3.4% and 2.9%, respectively.

Geographic area and number and percentages of accountable entities and patients included:

- All United States Nursing Homes with Medicare-eligible long-stay residents. In quarter 2 of 2019 there were 15,215 eligible facilities and 1,033,796 residents with target assessments, and 14,286 facilities (93.9%) had sufficient sample size (20 or more long-stay residents included in the denominator) to report on this measure, and 1,012,706 residents (98.0%) were included in the calculation of this measure.

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 unexpected findings during the testing process of NQF #0674.

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

This is not applicable; there are no unexpected benefits from the implementation of NQF #0674.

5. Comparison to Related or Competing Measures

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

5. Relation to Other NQF-endorsed Measures

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

Yes

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

0101: Falls: Screening, Risk-Assessment, and Plan of Care to Prevent Future Falls

0141: Patient Fall Rate

0202: Falls with injury

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.

#0202 Falls with Injury - Acute Care Prevention of Falls (rate of inpatient falls with injury per 1,000 patient days): This measure has a similar focus as NQF #0674, but it is different because it focuses on adult acute care

inpatient and adult rehabilitation patients and is reported as a rate rather than a percentage. Additionally, this measure includes any injury from minor to major. This is an important distinction. Focusing on falls with minor injury could potentially create inappropriate incentives for nursing homes to reduce resident opportunity for mobility and independence. The selection of the outcome of falls with major injury for NQF #0674 was deliberate to reduce this potential adverse unintended consequence. #0101 Falls Screening, Risk-Assessment, and Plan of Care to Prevent Future Falls: This is a clinical process measure that assesses falls prevention in older adults. The measure has three rates: 1) screening: percentage of patients aged 65 years of age and older who were screened for future fall risk at least once within 12 months; 2) falls risk assessment: percentage of patients aged 65 years of age and older with a history of falls who had a risk assessment for falls completed within 12 months; and 3) plan of care for falls: percentage of patients aged 65 years of age and older with a history of falls who had a plan of care for falls documented within 12 months. This measure is different in that it is a process measure, rather than an outcome measure. #0141 Patient Fall Rate (Total number of patient falls [with or without injury to the patient and whether assisted by a staff member] by hospital unit during the calendar month X 1000): This measure has a similar focus as NQF #0674, but it is different because it focuses on the adult acute care inpatient and adult rehabilitation patients and does not discriminate between falls with and without injuries, which is an important distinction. Focusing on falls with minor injury could potentially create inappropriate incentives for nursing homes to reduce resident opportunity for mobility. The selection of the outcome of falls with major injury for NQF #0674 was deliberate to reduce this potential adverse unintended consequence.

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.)

This is not applicable. There are no competing measures.

Appendix

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

Attachment **Attachment:** NQF_0674_Measure_Submission_Appendix_20210402_Upload.docx

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Center for Medicare & Medicaid Services

Co.2 Point of Contact: Rebekah, Natanov, Rebekah.Natanov@cms.hhs.gov, 202-205-2913-

Co.3 Measure Developer if different from Measure Steward: Acumen LLC

Co.4 Point of Contact: Aathira, Santhosh, asanthosh@sphereinstitute.org, 650-558-8882-1256

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.

Eric Tangalos, MD, FACP, AGSF, CMD

Professor of Medicine, Mayo Clinic

American Medical Directors Association

Joseph Ouslander, MD

Professor of Clinical Biomedical Science, Associate Dean for Geriatric Programs

Charles E. Schmidt College of Biomedical Science

Paul Katz, MD, CMD

Professor of Medicine, University of Rochester School of Medicine and Dentistry

American Medical Directors Association

Sharon Sprenger, MPA, RHIA, CPHQ

Senior Advisor, Measurement Outreach, Division of Quality Measurement and Research

The Joint Commission

Stefan Gravenstein, MD

Clinical Director, Quality Improvement

Quality Partners of Rhode Island

Toby Edelman, JD

Senior Policy Attorney

Center for Medicare Advocacy

Nancy Richman, OTR/L, FAOTA

Co-Owner

Glantz/Richman Rehabilitation

Burton Silverstein, PhD

Assistant Vice President, Outcomes Management

HCR ManorCare

Christie Teigland, PhD, MA

Director Health Informatics and Research

New York Association of Homes and Services for the Aging

Terrence O´Malley, MD

Medical Director, Non-Acute Care Services

Partners HealthCare System, Inc.

Ellen Strunk, PT, MS, GCS

President and Owner

Rehab Resources and Consulting

Barbara Benson Yody

Senior Director, Outcomes Analysis

Genesis HealthCare

Keith Krein, MD

Chief Medical Officer, Health Services Division

Kindred Healthcare

David Hittle, PhD

Assistant Professor

Division of Health Care Policy and Research

University of Colorado Denver, School of Medicine

Michele F. Bellantoni, M.D.

Associate Professor of Medicine

Johns Hopkins University School of Medicine

Ann Deutsch, RN, PhD, CRRN

Clinical Research Scientist

Rehabilitation Institute of Chicago

Carol Siem, MSN, RN, BC, GNP

Clinical Consultant/Educator

University of MO, Columbia

Debra Saliba, MD, MPH

Anna and Harry Borun Chair in Geriatrics and Gerontology, UCLA

Research Physician, VA GLAHS GRECC

Director of Borun Center for Gerontological Research, UCLA/JHA

Senior Natural Scientist, RAND Health

UCLA, VA, RAND Corporation

This TEP met over 2 days in October of 2009 to review the environmental scan and deliberate on the importance and validity of potential new nursing home measures for further development.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2011

Ad.3 Month and Year of most recent revision: 04, 2015

Ad.4 What is your frequency for review/update of this measure? Every 3 years

Ad.5 When is the next scheduled review/update for this measure? 04, 2021

Ad.6 Copyright statement: n/a

Ad.7 Disclaimers: n/a

Ad.8 Additional Information/Comments: n/a