NATIONAL QUALITY FORUM
Driving measurable health improvements together

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

Brief Measure Information

NQF #: 2827
Corresponding Measures:
Measure Title: PointRight® Pro Long Stay(TM) Hospitalization Measure
Measure Steward: American Health Care Association
sp.02. Brief Description of Measure: The PointRight Pro Long Stay Hospitalization Measure is an MDS-based, risk-adjusted measure of the rate of hospitalization of long-stay patients (also known as "residents") of skilled nursing facilities (SNFs) averaged across the year, weighted by the number of stays in each quarter.
1b.01. Developer Rationale: In November 2013 the HHS Office of the Inspector General published a document entitled "Medicare Nursing Home Hospitalization Rates Merit Additional Monitoring" (HHS Document OEI-06-11-00040). The OIG report noted that one-quarter of Medicare nursing home residents had hospitalizations (i.e., direct discharges to acute care hospitals of Medicare residents, whether post-acute or long stay) and that these hospitalizations cost $14.3 billion – and this is for Medicare Fee for Service only.
The rates of hospitalization varied significantly between states and between SNFs with different five-star ratings, suggesting that rates could be improved substantially if facilities rendered higher-quality care. The report details reasons for hospitalization and associates hospitalization costs with these reasons. For example, hospitalizations for pneumonia cost Medicare 844 million (USD) in one year, those for urinary track infections without sepsis cost 422 million (USD), and those related to aspiration of food or vomitus cost 618 million (USD).
These three conditions alone are obvious opportunities for quality improvement: Pneumococcal pneumonia can be prevented by immunization; catheter-associated UTIs can be prevented by high quality catheter care, avoidance of unnecessary indwelling catheters, and prophylactic antibiotics where appropriate; aspiration rates can be reduced by dietary modifications, supervised eating, and therapy for addressable swallowing problems.
Even when infections develop many can be safely and effectively treated in the facility if the diagnosis is timely – reducing hospitalization rates both for the specific infection and for sepsis. Review of the OIG report suggests that reducing hospitalization costs by over $1 billion per year – for FFS Medicare beneficiaries alone – is a modest and attainable target.
A 2010 report, showed that one third of the dually eligible population in SNFs are hospitalized at least once and over a third of them can be avoidable (Walsh et al., 2010). The same study stated that in 2005, the Medicare program paid 3 billion (USD) for potentially avoidable hospitalizations, and Medicaid paid 463 million (USD). Again, these numbers demonstrate the high cost associated with hospitalizations.
CMS through its contractor RTI has developed a 30-day hospitalization rate quality measure for SNFs based on Medicare claims, and PointRight has developed one based on the MDS; both are endorsed by the NQF. However, to date no corresponding measure has been developed for long-stay residents. According to the national MDS data from CMS, there were 437,356 long nursing home stays discharged to an acute hospital in the year ending 2015 Q1. This demonstrates the importance of needing a hospitalization measure for long-stay residents,
In addition to their costs, it is known that hospitalizations are risky and potentially traumatic events for frail elderly patients, frequently associated with a decline in independent function, delirium and/or cognitive decline that may not be reversible, worsening of nutritional status and physical conditioning, and a risk of falls with injury, new pressure ulcers, and hospital-acquired infections They have also been tied to other risks associated with transitions of care such as the

increased risk of medication errors. This offers additional motivation for reducing hospitalization rates of SNF residents, further establishing the need for a comprehensive set of performance measures related to this problem, and thus for a measure focusing on long-stay residents and including all payers.

Other published studies confirm the observations and the conclusions reported by the OIG in 2013, e.g., ones from the Kaiser Foundation (Jacobson, 2010), the Commonwealth Fund (Schoen, 2013), MedPAC (MedPAC, 2012) and CMS (Walsh, 2010). Studies by Ouslander have shown that structural and process issues within SNFs have a high impact on the rate of hospitalizations (Ouslander, 2012; Ouslander, 2011), further supporting the hypothesis that hospitalization rates could be reduced by feasible changes in facilities' operations.

Jacobson, G., Neuman, T., & Damico, A. (2010). Medicare spending and use of medical services for beneficiaries in nursing homes and other long term care facilities: A potential for achieving Medicare saving and improving the quality of care. The Henry J. Kaiser Family Foundation.

MedPAC. (2012) Report to congress: Payment policy. http://medpac.gov/documents/mar12_entirereport.pdf

Ouslander, J.G., & Maslow, K. (2012). Geriatrics and the triple aim: Defining preventable hospitalizations in the long-term care population. J Am Geriatr Soc., 60(12): 2313-2318.

Ouslander, J.G., & Bersenson, R.A. (2011). Reducing unnecessary hospitalization of nursing home residents. NEJM, 356(13): 1165-1167.

Ouslander, J.G., Lamb, G., Perloe, M., Givens, J.H., Kluge, L., Rutland, T, … Saliba, D. (2010). Potentially avoidable hospitalizations of nursing home residents: Frequency, causes, and costs. J Am Geriatr Soc., 58(4): 627-635.

Ouslander, J.G., Lamb,G., Tappen, R., Herndon, L., Diaz, S., Roos, B.A., … Bonner, A. (2011). Interventions to reduce hospitalizations from nursing homes: Evaluation of the INTERACT II collaborative quality improvement project. J Am Geriatr Soc., 59(4): 745-753.

Schoen, C., Radley, D., Riley, P., Lippa, J., Berenson, J., Dermody, C., & Shih A. (2013). Health Care in the two Americas: Findings from the scorecard on the state health system performance for low-income populations. The Commonwealth Fund.
http://www.commonwealthfund.org/Publications/Fund-Reports/2013/Sep/Low-Income-Scorecard.aspx

Walsh, E.D., Freiman, M., Haber, S., Bragg, A., Ouslander, J., & Wiener, J.M. (2010) Cost drivers for dually eligible beneficiaries: Potentially avoidable hospitalization from nursing facility, skilled nursing facility, and home and community-based services waiver programs, final task 2 report. RTI International.

---

sp.12. Numerator Statement: The numerator for the measure is the sum over four quarters of the counts of hospitalizations of the quarterly denominator populations, where hospitalizations comprise discharges directly from the SNF to an acute care hospital.

sp.14. Denominator Statement: The quarterly denominator population consists of those patients present in the SNF on the first day of the quarter (the "snapshot date") who meet the criterion for long stay on that date. The denominator for a quarter is the number of patients in the quarterly denominator population. The denominator for the measure is the sum of the quarterly denominators for the four quarters in the 12 month measurement period.

The criterion for a patient's having a long stay is a cumulative length of stay in the facility of more than 100 days as of the snapshot date. The cumulative length of stay of a patient is the length of the current stay as of the snapshot date and plus the full lengths of stay of any previous stays that are linked to it. According to the criteria for linkage of stays used in the present measure, a stay in a SNF is linked to a subsequent stay in the SNF if the patient was discharged from the SNF to the community and was readmitted to the SNF within 10 days or fewer. All stays in a sequence of linked stays are included in the sum of days used to determine a patient's cumulative length of stay. In these criteria the term "community" comprises private residences and all organized settings that are primarily residential in character, including senior housing, independent living facilities, board and care homes, and assisted living facilities.

A patient can contribute multiple times to the denominator for a 12 month measure period. For example, a resident continuously present in the facility for a full year would contribute four to the denominator.

sp.16. Denominator Exclusions: There are no exclusions from the denominator; all patients in the facility on the snapshot date who meet the long stay criterion on that date are included.

Measure Type: Outcome

sp.28. Data Source: Electronic Health Records: Electronic Health Records

sp.07. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: 12/9/2016

Preliminary Analysis: Maintenance of Endorsement

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

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or 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.

The developer provides the following description for this measure:

- This is a maintenance outcome measure at the facility level that assesses the risk-adjusted rate of hospitalizations for long-stay patients of skilled nursing facilities (SNFs) averaged across the year and weighted by the number of stays in each quarter.
- The developer provides a logic model that depicts the influence of structural interventions (i.e., high staffing levels and nurse practitioner availability) and process interventions such as early detection of signs and symptoms of impending infections (e.g., pneumonia, urinary tract infections) and chronic disease exacerbation (e.g., congestive heart failure, diabetes mellitus) on hospitalizations.

Summary of prior review in 2016

- The Standing Committee acknowledged that many preventable hospitalizations (25-33 percent) and unnecessary costs ($3 billion in Medicare expenditures; $463 million in Medicaid expenditures) associated with these hospitalizations could be avoided with this long-stay measure.

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:

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

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

Guidance from the Evidence Algorithm

Health outcome or PRO (Box 1) -> Relationship between the measure health outcome and process is demonstrated by empirical data (Box 2) -> Pass

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

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

Maintenance measures – increased emphasis on gap and variation

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

- The developer provided rehospitalization statistics for SNFs nationally for the two most recent quarters (Q4 of 2019 and 2020).
    - For 2019 and 2020, the risk adjusted mean rate was 14.7 percent (2019) and 14.8 percent, respectively.
    - The standard deviation (SD) was 5.3 percent (range 0-50.9 percent, 2019) and 6 percent (range 0-53.6 percent, 2020).
- The developer provided rehospitalization rates from the American Health Care Association (AHCA) member facilities from Q1 2014 through Q4 2020.
    - The developer noted that the national average hospitalization rate has increased by 6.5 percent.

Disparities

- The developer noted that facilities with less than 5 percent of minority residents have lower risk-adjusted ProLongStay hospitalization rates (mean= 12.3 percent) compared to facilities with greater than or equal to 35 percent of minority residents (mean= 17.9 percent).
- The difference in average hospitalization rates between facilities with low (< 5 percent) and high (>= 35 percent) percentage of minorities was 5.6 percent in Q4 2020 compared to 4 percent in Q1 2014.
- The developer provided data by geographical location relative to the CDC's Social Vulnerability Index (SVI).
    - Facilities located in lower SVI counties had lower risk adjusted ProLongStay hospitalization rates.
    - The difference in average readmission rates between facilities in low and high SVI counties has decreased over time (2.9 percent [Q1 2014] and 3.8 percent [Q4 2020]).

Questions for the Committee:

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

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

Committee Pre-evaluation Comments:

1a. Evidence

- No new evidence available; the developer reports using the most recent data from 2014 for the measure

- No changes in the evidence since last evaluation. The developer cited literature/reports that outlined hospitalization rate of long-stay residents, the cost of hospitalization and three main preventable conditions that long-stay residents were hospitalized. It was noted that over one-third of the hospitalizations were preventable with better quality of care. I am not aware of any new studies or information related to this measure

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

- Disparities related to race were included and demonstrated higher rates of admission/readmission.
- Performance date demonstrated a gap in care and need for national performance measure. The developer reported national SNFs rehospitalizations statistics for 2019Q4 (n=14,823) & 2020Q4 (n=14,737). The risk adjusted mean rate was 14.7% in 2019 and 14.8% in 2020. The standard deviation in 2019 was 5.3% (0-50.9%) and in 2020 was 6.0% (0-53.6%). AHCA reported that the national average hospitalization rate increased by 6.5% from Q1 2014 to Q4 2020. Disparity data was provided and demonstrated disparity in care. The risk adjusted ProLongStay hospitalization rate was lower in facilities with less than 5% of minority residents (mean=12.3%) and higher in facilities 35% or greater minority residents (mean=17.9%). The average difference in these facilities with low and high percentage of minorities was 5.6% in Q4 2020 and 4% IN Q1 2014. The developer looked at the geographical location of facilities using CDCs Social Vulnerability Index. The average difference in readmission rates of facilities in low SVI counties and high SVI was 2.9% in Q1 2014 and 3.8% in Q4 2020.

Criteria 2: Scientific Acceptability of Measure Properties

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

Evaluators: Staff

2a. Reliability: Specifications and Testing

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

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 – less emphasis if no new testing data provided.

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.

Specifications:

- Measure Specifications are clear and precise.

Reliability Testing:

- Reliability testing at the Patient or Encounter Level
    - Agreement of Model Independent Variables
        - The developers compared the prevalence of the risk adjustment covariates between a testing sample of 2,584 SNFs and the national population using the Minimum Data Set (MDS) 3.0.

        - The developer noted that 45 percent of the risk adjustment model covariates were found to have prevalence within 5 percent of the prevalence found in the national sample.
        - The developer also noted that 65 percent (13 of 20) risk adjustment model covariates that were comparable were found to have prevalence within 10 percent of the prevalence found in the national sample.
        - The developer acknowledged that although the measure's testing sample was not a random sample of all SNF patients nationally, the model IV cohorts are sufficiently represented within the sample.
- Reliability testing at the Accountable Entity Level
    - Reliability of Rates over Time
        - The developer analyzed change from quarter to quarter in the observed and adjusted long-stay hospitalization rates.
        - The developer explained that their reasoning was that the underlying probability of an SNF's long-stay patients hospitalizing, and the characteristics of its long-stay patient population were unlikely to change greatly in a three-month period so most of the change from quarter to quarter would be due to limitations on measure reliability.
        - Correlations from one quarter to the next ranged between 0.884 to 0.894 for the parametric statistic, and 0.877 to 0.886 for the rank order statistic.
        - The developers noted that this suggests that the measure is adequately stable over short periods, but sufficiently variable to reflect clinically meaningful changes.
    - Stability of Facility Level Adjusted Rate Bootstrapping
        - The developer recalculated adjusted rates for the measure for CY 2014 using a random sample of stays. The developer then reviewed the distribution of differences between facilities' original adjusted rates and the rates calculated with the new sample.
        - The developer interpreted a distribution of differences with a small variance and a mean of zero as acceptable measure stability or reliability.
        - The developer noted that 64.8 percent of the PointRight sample had a difference in adjusted rates of less than 2 percent and only 4.6 percent of facilities had a difference greater than 5 percent. The mean difference was 0.1 percent.

Questions for the Committee regarding reliability:

- *Do you have any concerns that the measure cannot be consistently implemented (i.e., are measure specifications adequate)?*
- *The developer attests the specifications have not changed and that additional reliability testing was not conducted. Does the Committee agree that the measure is still reliable and there is no need for repeat discussion and vote on Reliability?*

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

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

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

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

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

Validity Testing

- Validity testing at the Patient or Encounter Level
    - Agreement of Model Dependent Variables
        - The developer compared the identification of hospitalizations of Medicare Fee-for-Service beneficiaries between the MDS and Medicare FFS claims.
        - The developer used 2012 MDS data, claims data, and enrollment data because it was the most recent available.
        - The developer noted that there were 241,857 discharges to an acute hospital from long stays discharges (n=15,091 SNFs)
        - The developer noted that 86 percent of hospitalizations of Medicare FFS patients identified by the MDS are confirmed by Medicare FFS claims.
        - The developer further noted that in the other direction, 98 percent of acute inpatient claims found near an MDS discharge have an MDS discharge code of acute hospitalization.
        - Overall, the developer explained that the MDS discharge assessments appear to be overstating the rate of acute hospitalizations to a moderate degree.
        - The developer noted that accuracy of the dependent variable for patients with other payers was not feasible as data for such residents is not available.
- Validity testing at the Accountable Entity Level
    - The developer performed construct validity by testing the relationship between this measure with the various components of the CMS Five-Star ratings for SNFs and its correlation with CMS long-stay quality measures.
        - The developer showed the relationship between specific long-stay quality measures and the long-stay hospitalization measure.
        - The developer also shows that higher star ratings was associated with lower adjusted long-stay hospitalization rates.
    - The developer presented the relationship between 13 CMS Long-Stay Measures and the ProLongStay Adjusted Hospitalization Rates (Table 9).
    - The developer identified five CMS Long-Stay Measures that were significantly correlated with ProLongStay Adjusted Hospitalization Rate
        - High-Risk Residents with Pressure Ulcers (0.20, $p<0.0001$)
        - Residents Who Lose Too Much Weight (0.10, $p<0.0001$)
        - Residents who Self-Report Moderate to Severe Pain (0.08, $p=0.0001$)
        - Residents Whose Need for Help with Activities of Daily Living Has Increased (0.11, $p<0.0001$)
        - Residents with a Urinary Tract Infection (0.10, $p<0.0001$)
    - The developer calculated the correlation between this measure and the Pro LongStay measure Since the original endorsement, CMS added a Medicare claims-based long-stay hospitalization measure to Care Compare (formerly Nursing Home Compare) and five stars. The developer calculated the correlation between this measure and Pro LongStay.
        - ProLongStay Adjusted Hospitalization rates had a statistically significant positive relationship with the Medicare FFS claims-based long-stay hospitalization measure

used in five-star and reported on Care Compare. The correlation coefficient was 0.770 ($p<0.001$).

Exclusions

- The developer indicated that there are no exclusions; however, the measure will not be reported for a SNF if the denominator population over the measure period's four snapshot dates is less than 30.
- The developer noted that all patients in the facility on the snapshot date who meet the long stay criterion on that date are included in the denominator.

Risk-Adjustment

- The developer noted that this measure employs four logistic regression models applied to four discrete subgroups of the denominator population to estimate the risk of any hospitalization during the quarter.
- The developer also noted that the selection of risk factors (independent variables) involved an iterative process and that the variables with the strongest univariate correlations were then used to build multivariate models.
- The multivariate models (logistic regressions for each stratum of length of stay [LOS]) were reviewed by a larger panel of clinicians and potential users of the measure.
- Variables were rejected and replaced if their coefficients were opposite to their univariate correlation with the hospitalization, or if they were viewed as potentially under the control of the SNF (i.e., creating a risk of over-adjustment).
- The developer provided the conceptual and empirical analysis to demonstrate the need for SDS adjustment.
    - The developer's overall approach was to begin with reliable and rarely missing patient-level SDS variables nominated by clinical experts.
        - Medicaid status (as a proxy for financial assets and income)
        - Black versus non-black; Hispanic/Latino versus non-Hispanic/Latino
        - Interactions of Medicaid status and race
    - The developer noted that while Black and Medicaid status was found to be significant in at least one of the fixed effects models they found minimal impact on the overall performance of the models as measured by the c-statistic.
    - Inclusion of one or both SDS variables did not impact the overall c-statistic of the models by more than .001.
    - The developer noted that in the final risk adjustment model, single-level logistic regression models coefficients of the SDS variables were forced to be the same as in the fixed-effects model. The developer explained that essentially this approach adjusts for the within-facility differences in long stay hospitalization rates associated with the SDS factor but does not adjust for the between-facility differences in long stay hospitalization rates associated with the SDS factor.
    - The developer notes that the within-facility effects are essentially those beyond those associated with facility quality differences. In all cases this made the effect of the SDS variables smaller than it would be in a single-level logistic regression that did not account for facility effects. The developer did not want to adjust away facility-level effects related to worse care at SNFs with large minority populations.
- Risk model diagnostics
    - To assess the overall performance of their risk-adjustment model, the developers compared their model coefficients to the mean coefficients from bootstrap analysis, expressed as actual values, standard deviation (S.D.) and percentage.

- The developer performed a Hosmer-Lemeshow test for the goodness of fit of the logistic regression models. The test assesses whether or not the observed event rates match expected event rates in subgroups of the model population.
- Risk-Model Discrimination Statistics:
    - Logistic Regression Model Long Stay Group 1 c-statistic = .64
    - Logistic Regression Model Long Stay Group 2, c-statistic = .63
    - Logistic Regression Model Long Stay Group 3, c-statistic = .62
    - Logistic Regression Model Long Stay Group 4, c-statistic = .63
    - Linear Regression Model Rate of all Hospitalizations, R-squared = .96

Meaningful Differences

- The distribution of change in adjusted rates was similar across all four quarters where for each quarter the average change for deciles 2 through 8 was less than +/- 3 percent. Deciles 1 and 10 had average changes greater than +/-3.5 percent.
- The distribution of differences was larger for facilities with smaller denominators, and this indicated that recommendations of clinically meaningful difference should be dependent upon facility size.

Missing Data

- The developer provided distribution data of MDS 3.0 known outcome rates across the sample, as well as the relationship between the observed rate of hospitalizations and the known outcomes rate.
- The developer noted on occasion a facility may fail to follow the deadline for submitting an MDS assessment, resulting in the inclusion of the patient in a quarterly denominator but unable to provide them with a known outcome following the snapshot date. This is vital to the measure's accuracy. In response, the developer has reviewed the known outcome rates across their sample to ensure that missing data is not a major factor.
- The developer selected a known outcome rate of 90 percent to be the minimum threshold for missing data. The median known outcome rate in their full sample of PointRight facilities was 100 percent, concluding that missing data was not an issue for the majority of facilities.
- Additionally, the developer noted a slight positive correlation between the known outcome rate and the observed hospitalization rate (Pearson correlation= 0.8, p= 0.0001; Spearman correlation= -0.006, p=0.7311).

Comparability

- The measure only uses one set of specifications for this measure.

Questions for the Committee regarding validity:

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

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

Committee Pre-evaluation Comments:

2a. Reliability-Specification

- Reliability testing done at patient or encounter level and at the accountable entity level. Patient or Encounter Level: Agreement of Model Independent Variables: A comparison of the variable prevalence between the testing sample of 2584 SNFs and the national population was performed. 45% of the risk adjustment model covariates showed prevalence within 5% of the national sample. Also 65% (13/20) risk adjustment model covariates that were comparable had 10%of prevalence in the national sample. Accountable Entity Level: Reliability of Rates over Time: The developer analyzed change from quarter to quarter in the observed and adjusted long-stay hospitalization rates. Correlations from one quarter to the next ranged between 0.884 to 0.894 for parametric statistic, and 0.877 to 0.886 for the rank order statistic. This suggests that sufficient variability to reflect clinically meaningful changes. Stability of Facility level adjusted rate bootstrapping done by recalculating adjusted rates for the CY2014 measure using random sample of stays.
- no concerns

2a2. Reliability-Testing

- No
- no concerns

2b1. Validity

- No concerns
- No

2b2-2b3. Potential threats to validity

- No concerns
- Yes. Risk adjusted strategy included in measure. Four logistic regression models applied to four discrete subgroups of the denominator population to estimate the risk of any hospitalization during the quarter. Conceptual and empirical analysis was done to demonstrate the need for SDS adjustment, however, Medicaid status and Race was significant in at least one fixed effect model but didn't impact the overall c-statistic of the models by more than .001. The developer noted that missing data could affect the measure's accuracy therefore the developer reviewed known outcome rated across their sample to ensure that missing data wouldn't be a major factor. A known outcome rate of 90% was selected to be the minimum threshold for missing data.

2b4-2b7. Potential threats to validity

- The developer reported some delays in completing assessments. This could impact validity if the number of missing assessments is higher than reported.
- Validity testing done at patient/encounter level and entity level showed meaningful differences about quality. The developer showed that higher star ratings was associated with lower adjusted long-stay hospitalization rates. Medicare claims-based long-stay hospitalization measure used in five-star and reported on Care Compare was compared to ProLongStay Adjusted Hospitalization rates. There was a statistically positive relationship. The calculated coefficient was 0.770 (p.0.001)

Criterion 3. Feasibility

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

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

- The developer noted that data elements needed to compute the measure score can be generated or collected by healthcare personnel during the provision of care.
- All data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS).
- The developer noted that computation of the measure requires a license to use software for large-scale data management and calculation of risk estimates using logistic regression models.
- The developer noted that while utilization of the measure specifications does not require a fee, there is a requirement that display, disclosure, or publication of the measure must include the measure's trademark and that the measure specifications are copyrighted by Point Right®.

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:

3. Feasibility

- Data collected can indicate appropriate measures to reduce admissions
- All data elements are already collected electronically. Other data can be collected by healthcare personnel during care.

Criterion 4: Use and Usability

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

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

4a. Use evaluates the extent to which audiences (e.g., consumers, purchasers, providers, 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
Planned use in an accountability program?	☐	Yes	☐	No	☒	NA

Accountability program details

- The developer noted that this measure is utilized in several state Medicaid programs as part of their value-based purchasing (VBP) or pay-for-performance (P4P) programs.

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 developer publishes ProLongStay rates on AHCA's Long-Term Care (LTC) Trend Tracker tool quarterly for members to track and benchmark their organization's ProLongStay performance.
- The developer publishes facility-level rates publicly on the AHCA website on a quarterly basis.
- Results are available in three Net Health PointRight solutions: Quality Measures, ScoreCard, and New Mexico VBP.
- Net Health solutions offer educational materials on-demand.
- Feedback on the measure is shared through direct conversations with Analytics, Product Management, Sales, and Client Services team members.
- Feedback is also submitted through in-application messaging, via email, and in conjunction with responses to Net Promoter Score (NPS) customer satisfaction surveys.

***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?*
- *How has those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation?*
- *How has this feedback has been considered when changes are incorporated into the measure?*

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 developer provided four figures demonstrating the improvement of hospitalization rates in New Mexico Nursing Facilities.
    - The developer noted an improvement in average performance from 15.27 percent in 2020 Q1 to 7.81 percent in Q4 2021.
    - The developer noted an improvement in median performance from 15.08 percent in 2020 Q1 to 7.01 percent in Q4 2021.

- The developer noted there is no current national value-based or pay-for-performance incentives tied to reducing long-stay hospitalizations and Medicare's national Skilled Nursing Value-Based Purchasing program only accounts for short-stay rehospitalizations.
- There has been no significant improvement in the national ProLongStay rate from 2014 to 2020.

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

- The developer did not identify any unintended consequences.

Potential harms

- The developer did not identify any potential harms.

***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:

4a. Use

- no concerns
- Measure is used in several state Medicaid programs as part of their value-based purchasing or pay-for-performance programs

4a. Usability

- no unintended consequences were noted
- This measure is used in several state Medicaid programs as part of their value-based purchasing or pay-for-performance programs. Improvement shown in hospitalization rates in New Mexico Nursing Facilities. Improvement in average performance in 2020 Q1 from 15.27% to 7.81% in Q4 2021. The median performance improved from 15.08% in 2020 Q1 to 7.01% in Q4 2021. No unintended consequences or harm noted from use of this measure.

Criterion 5: Related and Competing Measures

Related measures

- NQF # 2375 PointRight ® Pro 30™
- NQF #2510 Skilled Nursing Facility 30-Day All-Cause Readmission Measure (SNFRM)

Harmonization

- The developer noted that the measure specifications are harmonized to the extent possible.

Committee Pre-evaluation Comments:

5: Related and Competing Measures

- no additional steps needed
- Related measures NQF #2375 PointRight Pro 30 and NQF #2510 Skilled Nursing Facility 30-Day All Cause Readmission Measure. Harmonization was done to the fullest extent possible

Public and NQF Member Comments (Submitted as of Month Day, Year)

Member Expression of Support

- No members submitted an expression of support for this measure.

Comments

- No NQF member and public comments were received in advance of the Standing Committee evaluation.

Scientific Acceptability Evaluation

RELIABILITY: SPECIFICATIONS

1. **Have measure specifications changed since the last review?** ☐ **Yes** ☒ **No**
2. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**
3. **Briefly summarize any changes to the measure specifications and/or concerns about the measure specifications.**
   - There were no submitted changes to the specifications since the last updates/submission.

RELIABILITY: TESTING

4. **Did the developer conduct new reliability testing?** ☐ **Yes** ☒ **No**

   4a. If no, summarize the Standing Committee's previous feedback:
   - The Standing Committee raised concern about the dataset, but the developer confirmed that the measure is based on the Minimum Data Set (MDS) and not based on claims data.
5. **Reliability testing level:** ☒ **Accountable-Entity Level** ☒ **Patient/Encounter Level** ☐ **Neither**
6. **Reliability testing was conducted with the data source and level of analysis indicated for this measure:**

   ☒ **Yes** ☐ **No**
7. If accountable-entity level and/or patient/encounter level reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☒ **Yes** ☐ **No**
8. **Assess the method(s) used for reliability testing:**
   - Testing was completed on 2,584 facilities to verify the calculations.
   - The analyses were conducted on the MDS 3.0 database for skilled nursing facilities that purchased analytics services from PointRight.
   - Analysis was conducted based on Facility Level Descriptive Statistics and Characteristics of Patients for long stay hospitalization.
   - The developers compared the prevalence of the risk adjustment covariates between a testing sample of 2,584 SNFs and the national population.

- The distribution of MDS patient characteristic variables in the sample population and the nation generally seemed similar (Table 2) but it is unclear whether the reliability of the data elements can be inferred from Table 3 provided.

9. **Assess the results of reliability testing**
   - The results of these reliability tests showed 64.8 percent of the PointRight sample had a difference in adjusted rates of less than 2 percent and only 4.6 percent of facilities had a difference greater than 5 percent. The mean difference was 0.1percent.
   - The risk adjustment model covariates were compared to the national sample where covariates were found to have prevalence within 5 percent. In addition, 65 percent of the risk adjustment model covariates were found to have prevalence within 10 percent of the prevalence found in the national sample.
   - For larger faculties (with a denominator greater than 200) the variance distribution in differences decreased.
10. 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.

    ☒ **Yes** ☐ **No** ☐ **Not applicable**
11. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

    ☒ **Yes** ☐ **No** ☐ **Not applicable** (patient/encounter level testing was not performed)
12. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

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

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if accountable-entity 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)
13. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    Specifications precise, unambiguous, and complete (Box 1) -> Empirical reliability testing conducted using statistical tests with the measure as specified (Box 2) -> Empirical reliability testing not conducted at the accountable entity level (Box 4) ->Empirical reliability testing conducted on all critical patient/encounter data elements (Box 8) -> Method described as appropriate for assessing reliability of all elements (Box 9) -> Moderate certainty or confidence that the elements used in the measure are reliable (Box 10a) -> Moderate rating

VALIDITY: TESTING

14. **Did the developer conduct new validity testing? ☒ Yes ☐ No**

14a. If no, summarize the Standing Committee's previous feedback:

- The Standing Committee expressed concern with the inclusion of race as a variable in the risk adjustment model. Based on the discussion, the developer agreed to remove the race variable and update their measure specifications and testing results

14b. If yes, describe any differences between the new and old testing and summarize any relevant Standing Committee's feedback from the previous review:

- The developer calculated the correlation between this measure and the Pro LongStay measure Since the original endorsement, CMS added a Medicare claims-based long-stay hospitalization measure to

Care Compare (formerly Nursing Home Compare) and five stars. The developer calculated the correlation between this measure and Pro LongStay

- ProLongStay Adjusted Hospitalization rates had a statistically significant positive relationship with the Medicare FFS claims-based long-stay hospitalization measure used in five-star and reported on Care Compare. The correlation coefficient was 0.770 ($p<0.001$).

15. **Validity testing level (check all that apply):**

☒ **Accountable-Entity Level** ☐ **Patient or Encounter-Level** ☒ **Both**

NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

16. **If patient/encounter level validity testing was provided, was the method described and appropriate for assessing the accuracy of ALL critical data elements? NOTE:** Data element validation from the literature is acceptable.

☒ **Yes**

☐ **No**

- ☐ **Not applicable** (patient/encounter level testing was not performed)

17. **Method of establishing validity at the accountable-entity level:**

☐ **Face validity**

☒ **Empirical validity testing at the accountable-entity level**

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

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

☒ **Yes**

☐ **No**

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

19. **Assess the method(s) for establishing validity**

- The developer compared the identification of hospitalization events of Medicare FFS beneficiaries based on the MDS and hospitalization events based on Medicare FFS claims. The data was restricted to the MDS discharges who were enrolled in Part A Medicare when they were discharged, and patients who had a SNF claim in calendar year 2012.
- The developer tested the relationship of the measure with various components of the CSM Five-Star ratings for SNFs and its correlation with CMS's long-stay quality measure.

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

- 86 percent of hospitalizations of Medicare FFS patients were identified by the MDS and were confirmed by Medicare FFS claims.
- 98 percent of hospitalizations of acute inpatient claims were found near an MDS discharge and have an MDS discharge code of acute hospital.
- The relationship between specific long-stay quality measures and the long-stay hospitalization measure have correlation coefficients that are statistically significant at $p < .05$.
- The correlation coefficient for ProLongStay Adjusted Hospitalization rates had a statistically significant positive relationship with the Medicare FFS claims-based long-stay hospitalization measure at 0.770 ($p<0.001$).

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

- There are no exclusions from the denominator; all patients in the facility on the snapshot date who meet the long stay criterion on that date are included. However, the measure will not be reported for a SNF if the denominator population over the measure period's 4 snapshot dates is less than 30.

22. **Risk Adjustment**

22a. **Risk-adjustment method**

☐ None (only answer Question 20b and 20e) ☒ Statistical model ☐ Stratification

☐ Other method assessing risk factors (please specify)

22b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

22c. **Social risk adjustment:**

22c.1 Are social risk factors included in risk model? ☒ Yes ☐ No ☐ Not applicable

22c.2 Conceptual rationale for social risk factors included? ☒ Yes ☐ No

22c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☒ Yes ☐ No

22d.**Risk adjustment summary:**

22d.1 All of the risk-adjustment variables present at the start of care? ☒ Yes ☐ No

22d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

22d.3 Is the risk adjustment approach appropriately developed and assessed? ☒ Yes ☐ No

22d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☒ Yes ☐ No

22d.5.Appropriate risk-adjustment strategy included in the measure? ☒ Yes ☐ No

22e**. Assess the risk-adjustment approach**

- This measure employs four logistic regression models applied to four discrete subgroups of the denominator population. The four logistic regression models estimate a patient's risk for one more hospitalizations.
- Calculation of a patient's risk of any hospitalization during a quarter at risk begins by assigning the patient to one of four subgroups of the denominator population based on the duration of the patient's current stay in the SNF as of the snapshot date.
- For each group the risk of one or more discharges from the SNF directly to an acute care hospital during the quarter was estimated by a logistic regression.
- The developer notes that the selection of risk factors (independent variables) involved an iterative process.
- A panel of clinicians with extensive SNF experience recommended potential risk adjusters. The developer's overall approach was to begin with reliable and rarely-missing patient-level SDS variables nominated by the clinical experts: Medicaid status (as a proxy for financial assets and income), black versus non-black, Hispanic/Latino versus non-Hispanic/Latino, and the interactions of Medicaid status and race. The significance of these variables in predicting hospitalization rates was tested in fixed-effects logistic regression models. The developer reasoned that patient-level effects that were significant in models that included facility-specific constant terms probably reflected otherwise-unmeasured differences in baseline health status. These, and a full set of sociodemographic and contextual factors were tested for univariate relationships with hospitalizations.
- The variables with the strongest univariate correlations were then used to build multivariate models.
- The multivariate models (logistic regressions for each stratum of LOS) were reviewed by a larger panel of clinicians and potential users of the measure.

- Variables were rejected and replaced if their coefficients were opposite to their univariate correlation with the hospitalization, or if they were viewed as potentially under the control of the SNF – i.e., creating a risk of over-adjustment.

23. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**
    - The distribution of change in adjusted rates was similar across all four quarters where for each quarter the average change for deciles 2 through 8 was less than +/- 3 percent. Deciles 1 and 10 had average changes greater than +/-3.5 percent.
    - The distribution of differences was larger for facilities with smaller denominators, and this indicated that recommendations of clinically meaningful difference should be dependent upon facility size.
24. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    - The developer noted that there is only one set of specifications for this measure.
25. **Please describe any concerns you have regarding missing data.**
    - The developer noted that for individuals discharged a facility's rates could be bias due to missing data. The developers decided to exclude facilities with greater than five percent missing data from the re-hospitalization rate analyses.
    - The developer noted on occasion a facility may fail to follow the deadline for submitting an MDS assessment, resulting in the inclusion of the patient in a quarterly denominator but unable to provide them with a known outcome following the snapshot date. This is vital to the measure's accuracy. In response, the developer has reviewed the known outcome rates across their sample to ensure that missing data is not a major factor.
    - The developer selected a known outcome rate of 90 percent to be the minimum threshold for missing data. The median known outcome rate in their full sample of PointRight facilities was 100 percent, concluding that missing data was not an issue for the majority of facilities.
26. **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 accountable-entity level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if accountable-entity 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 accountable-entity level and the patient/encounter level is required; if not conducted, should rate as INSUFFICIENT.)
27. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

Threats to validity empirically assessed (Box 1) -> Empirical validity testing conducted using the measure as specified (Box 2) -> Empirical validity conducted at the accountable entity level (Box 5) -> Validity testing method described and appropriate (Box 6) -> Moderate certainty or confidence (Box 7b) -> Moderate rating

ADDITIONAL RECOMMENDATIONS

28. **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.**

Developer Submission

Criteria 1: 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

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

[Response Begins]

No

[Response Ends]

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

2021 Submission:

Updated evidence information here.

2018 Submission:

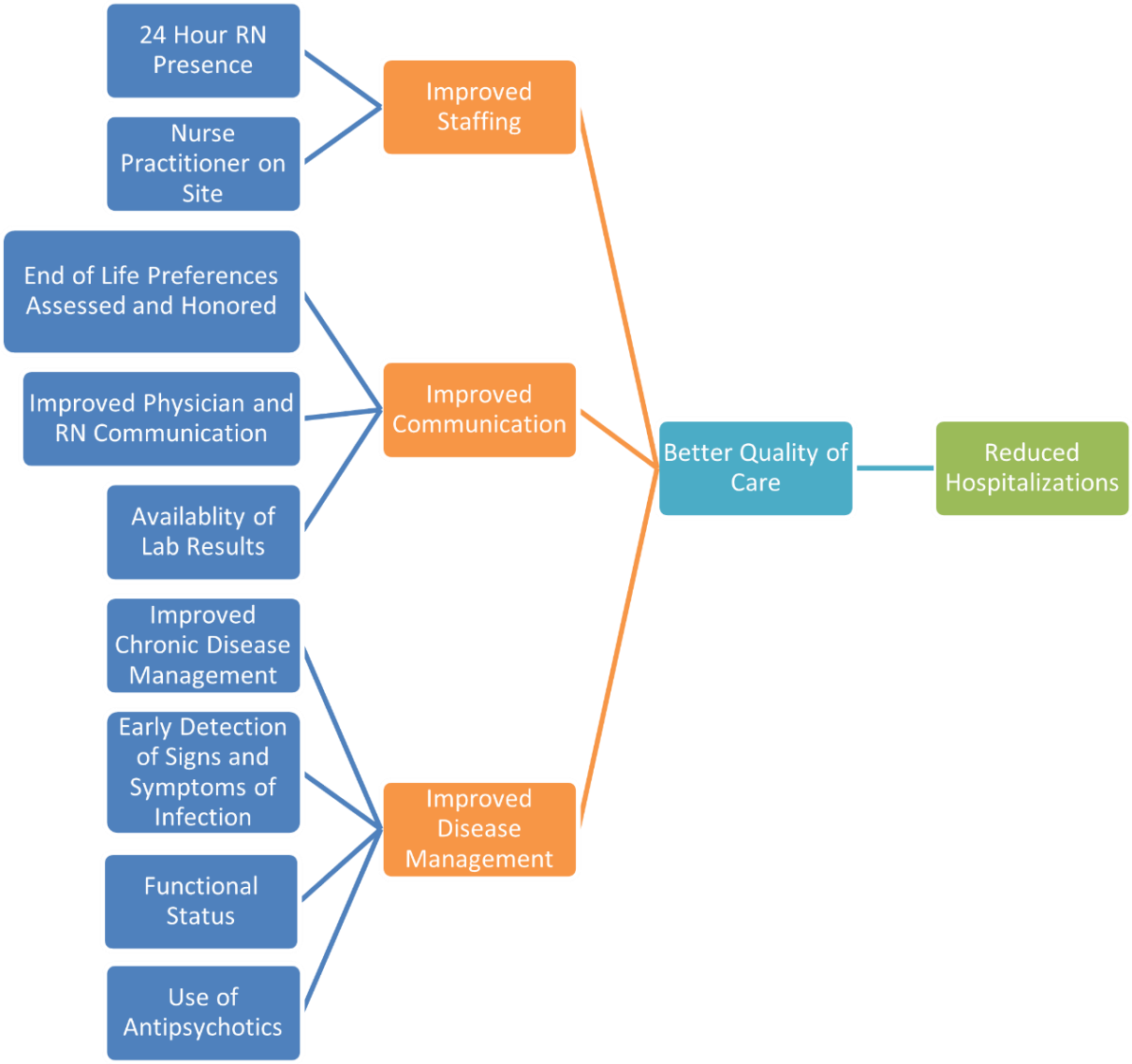
Evidence from the previous submission here.

1a.01. Provide a logic model.

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

[Response Begins]

Hospitalizations of any cause among individuals admitted to a skilled nursing facility (SNF) are the result of numerous clinical and non-clinical situations (Ouslander, 2012). However, a combination of structure, process and interventions influence the likelihood of hospitalizations more than patient acuity and condition. Structural interventions such as high staffing levels and nurse practitioner availability and process interventions such as early detection of signs and symptoms of impending infections (e.g. pneumonia, UTI) and chronic disease exacerbation (e.g. CHF, DM) can all work to decrease the incidence of hospitalizations (Ouslander, 2012; Young et al., 2011). The diagram below provides an overview of the structures and processes that can ultimately influence hospitalizations in long-term care residents.



Logic model of structure and processes that influence hospitalizations

[Response Ends]

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

Describe how and from whom input was obtained.

[Response Begins]

Studies show that hospitalizations of the elderly negatively impact mobility and function (Brown et al. 2009; Creditor, 1993). Furthermore, hospitalizations can cause complications unrelated to the primary reason for the hospitalization. These negative effects are driven by a decline in muscle strength, increased stress, and exposure to possible infections during a hospital stay, where physical movement is often restricted. Elderly individuals with dementia or memory loss can also find waking up in a hospital room disorienting. Thus, elderly individuals generally value avoiding hospitalizations when possible.

[Response Ends]

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

[Response Begins]

Quality of life and quality of care are two areas that past research initiatives have utilized to measure quality. Quality of life focuses on issues surrounding the resident's autonomy while quality of care examines the technical aspects of health care that affect the resident's quality of health outcomes such as pressure ulcer prevalence (Spilsbury et al., 2011). Previous evidence supports the theory that quality measures are beneficial to determine the rate of hospitalization among long-stay residents. Facilities that do not have a high standard on their quality measures are more likely to have higher rates of hospitalization among long-stay residents. The evidence presented below contains past findings that describe the causal relationship between clinical outcomes and quality measures among long-stay nursing home residents and the ultimate influence on long-stay resident hospitalizations.

Improving Staffing

Staffing levels within skilled nursing facilities have the ability to affect residents' quality of care and quality of life. Measures such as staffing ratios and licensed nursing staff availability have previously been determined to have a causal relationship between quality and health outcomes for long-stay residents.. Horn et al. (2005) evaluated staffing levels in relationship to residents' health outcomes. Their study provided evidence that facilities where registered nurses (RN) provided 30 to 40 minutes of care per resident per day had positive health outcomes. Higher ratios among licensed practical nurses and certified nurses' aides also provided better health outcomes. Newly admitted residents in the study were less likely to remain in the study (71.2 versus 80.8 days, $P<0.001$), and to develop pressure ulcers. In addition, in centers where registered nurses provided 30 to 40 minutes of direct care showed a decreased in adverse outcomes while improvements in care processes increased. In addition, evidence showed a reduction in catheterization, pressure ulcers, and the development of UTI (Horn et al., 2005). Quality measures for long-stay residents showed improvement when residents received 4.1 hours of direct care per day and 1.35 hours of care from licensed staff per day (Collier & Harrington, 2008).

Indicators such as quality of care deficiencies, quality of life deficiencies, in-bed time, and resident satisfaction have also been examined (Spilsbury et al., 2011). The relationship among RNs and residents typically assume a linear relationship where higher staffing numbers provides better quality of care and fewer deficiencies (Spilsbury et al., 2011). Resident outcomes included fewer resident care deficiencies within the first year of admission and reduced mortality (Collier & Harrington, 2008). Outcome measures for long-stay residents were measured by a comparison of two MDS assessments (separated by 90 days) which focused on variables such as functional improvement and weight loss (Collier & Harrington, 2008).

Facilities that have an on-site physician report lower hospitalization rates compared to facilities without a physician on-site. Young et al. (2011) reported a decrease in hospitalizations where facilities employed on-site physician assistants, nurse practitioners, and a training program for nurses' aides. However, the majority of facilities do not have an on-site physicians and rely on nursing assessments to observe residents' health and function-related problems as a strategy for management of care (Young et al., 2011). It is essential to have licensed staff members that can perform a proper assessment on residents' conditions while determining if a hospital transfer is necessary. Facilities that cannot perform the proper medical assessment and communicate their findings to a physician have a higher risk of hospitalization rates (Young et al., 2011).

Improving Communication

Effective communication between physicians and nursing staff leads to a reduction in hospitalization among long-stay residents. Both physicians and nursing staff must be trained on effective communication to provide better information about a patient's condition regarding acute conditions and end-of-life care. Effective communication reduces the number of hospitalizations and encourages physicians to treat patients in the nursing home, thus avoiding unnecessary transfers (Young et al., 2011). At the same time, physicians must be provided resources to direct care within a nursing home such as the patients' medical history, a lab, and lab results within a four hour timeframe during non-business hours. In

addition, nursing staff must be trained to provide accurate assessments of a resident's condition so that the physician may determine if a hospital transfer is necessary. Proper protocols must be put in place in order to provide the correct level of care to the resident and avoid hospitalizations (Young et al., 2011). The protocols for patient transfers must include resources for non-business hours (6 p.m.-6 a.m. and weekends) such as licensed staff on-site, in order to avoid improper hospitalization. Saliba et al. (2000) found that inappropriate transfers were more likely to occur during non-business hours because the facilities did not have the proper resources to treat the resident.

It is necessary to have conversations on advanced care planning with the resident and family members. These conversations are centered on noting the resident's preferences while they are cognitively and physically able to share their wishes. Research has demonstrated that advanced care planning improves end of life care, decreases life-sustaining treatment and prevents hospitalizations. At the same time, advanced care planning leads to an increase in the use of hospice and palliative care (Brinkman-Stoppelenburg et at., 2014). It has further been shown that having an advanced care directive can lower the rate of hospitalizations and death in a hospital (Detering & Silveira, 2015). All staff must be aware of healthcare advanced directives when discussing patient transfers to avoid inappropriate transfers and respect end-of-life wishes (Saliba et al., 2000).

Improving Disease Management

INTERACT

The Interventions to Reduce Acute Care Transfers (INTERACT) is a set of evidence-based clinical practice tools and strategies initially developed as a demonstration program to reduce hospitalization rates. The program reduced avoidable hospitalization in rates among nursing homes during the six-month implementation period (Ouslander et al., 2011). Overall, the program saw fewer complications and less morbidity from hospitalizations and reductions in Medicare expenditures (Ouslander et al., 2011). In combination with the INTERACT tools, nursing homes can employ best practices to avoid or mitigate risk factors for hospitalization among long-stay residents with chronic conditions. Potentially avoidable hospitalizations among nursing home residents are considered to be hospital admissions based on acute exacerbation of a chronic condition where preventative care could have been provided (Spector et al., 2013). Chronic conditions can be effectively managed in nursing homes if preventative measures or best practices are put into place. For example, infection control, falls prevention, and proper hygiene for residents with open sores are measures that could be utilized to reduce unnecessary hospitalization (Spector et al., 2013).

Functional Status/ADLs

Long-stay residents are more likely to demonstrate functional and behavioral impairment throughout their length of stay. Functional status is a practical outcome measure for this population, specifically in the physical and self-care domain, as long-stay residents are likely to demonstrate functional limitations (Gillen et al., 1996). Change of functional status is most likely to occur within the three month period after admission. Long-stay residents are more likely to remain stable at the same functional level and therefore, less likely to be discharged from a facility. Gillen et al. (1996) found a positive relationship between higher levels of functional impairment and higher probabilities of hospitalization and death among the long-stay population. Over half of their sample experienced a functional status change and/or two or more transitions. Activities of daily living (ADLs) dependency level is another risk factor for hospitalization and post discharge mortality (Ponzetto et al. 2003). It is essential that nursing homes provide the appropriate level of care in order for residents to maintain the same functional status. Maintaining functional status will prevent the deteriorating of health and reduce hospitalization.

Antipsychotics

Antipsychotics have been utilized to treat behavioral and psychotic symptoms in dementia patients. However, recentinitiatives have warned against the adverse effects of these drugs on the elderly. The use of antipsychotics among long-stay residents shows evidence of mixed reviews with caution for adverse effects. The typical approach to treating a health condition is a combination of pharmacological and nonpharmacological methods. There is evidence that side effects from antipsychotic use have the ability to reduce a resident's functional status and quality of life. For example, drugs with anticholinergic burden (ACB) have shown to increase cognitive and physical impairments which can lead to a rapid functional decline (Kolanowski et al., 2009).

Typically, adverse effects due to antipsychotic use may affect a resident's quality of life including depression, cognitive impairment and hospitalization. Older adults are more sensitive to adverse effects from antipsychotics, therefore, caution must be used (Frenchman, 2005). Long-stay residents who ingest high levels of ACB are more likely to be socially withdrawn from activities that require high social engagement. Sedation and confusion are common side effects associated with ACB (Kolanowski et al., 2009). Atypical antipsychotics have also been proven to have negative results on individuals with dementia. For instance, a study by Gareri et. al (2010), found that the drugs risperidone and olanzapine have been shown in increase adverse cardiovascular events in the elderly.

References

Brinkman-Stoppelenburg, A., Rietjens, J.A., Van Der Heide, A. (2014). The effects of advance care planning on end-of-life care: A systematic review. Palliat Med. 28: 1000-1025.

Brown, C.J., Roth, D.L, Allman, R.M., Sawyer, P., Ritchie, C.S., Roseman, J.M. (2009). Trajectories of life-space mobility after hospitalization. Ann Intern Med. 150(6): 372-378.

Collier, E. & Harrington, C. (2008). Staffing characteristics, turnover rates, and quality of resident care in nursing facilities. Research in Gerontological Nursing. 1(3): 157-170.

Creditor, M.C. (1993). Hazards of hospitalization of the elderly. Ann Intern Med. 118(3): 219-223.

Detering, K. & Silveira, M.J. (2015). Advanced Care Planning and Advanced Directives. UpToDate.

http://www.uptodate.com/contents/advance-care-planning-and-advance-directives .

Frenchman, I.B. (2005). Atypical antipsychotics for nursing home patients. Drugs and Aging. 22(3): 257-264.

Gareri, P., Marigliano, N. M., De Fazio, S., Lacava, R., Castagna, A., Costantino, D. S., & De Sarro, G. (2010). Antipsychotics and dementia. BMC Geriatrics, 10(Suppl 1), A93.

Gillen, P., Spore, D., Mor, V., Freiberger, W. (1996). Functional and residential status transitions among nursing home residents. Journal of Gerontology: Medical Sciences. 51A(1): M29-M36.

Hamer, S., Haddad, P.M. (2007). Adverse effects of antipsychotics as outcome measures. British Journal of Psychiatry. 191(59): 264-279.

Horn, S.D., Buerhaus, P., Bergstrom, N., Smout, R.J. (2005). RN staffing time and outcomes of long-stay nursing home residents. Journal of Nursing. 105(11): 58-70.

Intrator, O., Zinn, J., Mor, V. (2004). Nursing home characteristics and potentially preventable hospitalizations of long-stay residents. Journal of American Geriatrics Society. 52: 1730-1736.

Kahn, J.M., Werner, R.M., David, G., Have, T.R.T., Benson, N.M., Asch, D.A. (2013). Effectiveness of long-term acute care hospitalization in elderly patients with chronic critical illness. MedCare. 51(1): 4-10.

Kolanowski, A., Fick, D.M., Campbell, J., Litaker, M., Boustani, M. (2009). A preliminary study of anticholinergic burden and relationship to a quality of life indicator, engagement in activities, in nursing home residents with dementia. Journal of the American Medical Directors Association. 10(4): 252-257.

Oslander, J.G., & Maslow, K. (2012). Geriatrics and the triple aim: Defining preventable hospitalizations in the long-term care population. *J Am Geriatr Soc.*, 60(12): 2313-2318.

Ouslander, J.G., Lamb, G., Tappen, R., Herndon, L., Diaz, S., Roos, B.A., Grabowski, D.C., Bonner, A. (2011). Interventions to reduce hospitalizations from nursing homes: Evaluation of the INTERACT II collaborative quality improvement project. Journal of American Geriatrics Society 59: 745-753.

Ponzetto, M., Maero, B., Maina, P., Rosato, R., Ciccone, G., Merletti, F., Rubenstein, L.Z., Fabris, F. (2003). Risk factors for early and late morality in hospitalized older patients: The continuing importance of functional status. Journal of Gerontology: Medical Sciences. 58A(11): 1049-1054.

Saliba, D., Kington, R., Buchanan, J., Bell, R., Wang, M., Lee, M., Herbst, M., Lee, D., Sur, D., Rubenstein, L. (2000). Appropriateness of the decision to transfer nursing facility residents to the hospital. Journal of American Geriatrics Society. 48: 154-163.

Spector, W.D., Limcangco, R., Williams, C., Rhodes, W., Hurd, D. (2013). Potentially avoidable hospitalizations for elderly long-stay residents in nursing homes. Medical Care 51(8): 673-681.

Spilsbury, K., Hewitt, C., Stirk, L., Bowman, C. (2011). The relationship between nurse staffing and quality of care in nursing homes: A systematic review. International Journal of Nursing Studies. 48: 732-750.

Werner, R.M., Konetzka, R.T., Kim, M.M. (2013). Quality improvement under nursing home compare: The association between changes in process and outcome measures. MedCare. 51(7): 582-588.

Young, Y. Inandar, S., Dichter, B.S., Kilburn, H., Hannan, E.L. (2011). Clinical and nonclinical factors associated with potentially preventable hospitalizations among nursing home residents in New York state. Journal of American Medical Directors Association. 12: 364-371.

[Response Ends]

1b. Performance Gap

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

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

[Response Begins]

In November 2013 the HHS Office of the Inspector General published a document entitled "Medicare Nursing Home Hospitalization Rates Merit Additional Monitoring" (HHS Document OEI-06-11-00040). The OIG report noted that one-quarter of Medicare nursing home residents had hospitalizations (i.e., direct discharges to acute care hospitals of Medicare residents, whether post-acute or long stay), and that these hospitalizations cost $14.3 billion – and this is for Medicare Fee for Service only.

The rates of hospitalization varied significantly between states and between SNFs with different five-star ratings, suggesting that rates could be improved substantially if facilities rendered higher-quality care. The report details reasons for hospitalization and associates hospitalization costs with these reasons. For example, hospitalizations for pneumonia cost Medicare 844 million (USD) in one year, those for urinary track infections without sepsis cost 422 million (USD), and those related to aspiration of food or vomitus cost 618 million (USD).

These three conditions alone are obvious opportunities for quality improvement: Pneumococcal pneumonia can be prevented by immunization; catheter-associated UTIs can be prevented by high quality catheter care, avoidance of unnecessary indwelling catheters, and prophylactic antibiotics where appropriate; aspiration rates can be reduced by dietary modifications, supervised eating, and therapy for addressable swallowing problems.

Even when infections develop many can be safely and effectively treated in the facility if the diagnosis is timely – reducing hospitalization rates both for the specific infection and for sepsis. Review of the OIG report suggests that reducing hospitalization costs by over $1 billion per year – for FFS Medicare beneficiaries alone – is a modest and attainable target.

A 2010 report, showed that one third of the dually eligible population in SNFs are hospitalized at least once and over a third of them can be avoidable (Walsh et al., 2010). The same study stated that in 2005, the Medicare program paid 3 billion (USD) for potentially avoidable hospitalizations, and Medicaid paid 463 million (USD). Again, these numbers demonstrate the high cost associated with hospitalizations.

CMS through its contractor RTI has developed a 30-day hospitalization rate quality measure for SNFs based on Medicare claims, and PointRight has developed one based on the MDS; both are endorsed by the NQF. However, to date no corresponding measure has been developed for long-stay residents. According to the national MDS data from CMS, there were 437,356 long nursing home stays discharged to an acute hospital in the year ending 2015 Q1. This demonstrates the importance of needing a hospitalization measure for long-stay residents,

In addition to their costs, it is known that hospitalizations are risky and potentially traumatic events for frail elderly patients, frequently associated with a decline in independent function, delirium and/or cognitive decline that may not be

reversible, worsening of nutritional status and physical conditioning, and a risk of falls with injury, new pressure ulcers, and hospital-acquired infections They have also been tied to other risks associated with transitions of care such as the increased risk of medication errors. This offers additional motivation for reducing hospitalization rates of SNF residents, further establishing the need for a comprehensive set of performance measures related to this problem, and thus for a measure focusing on long-stay residents and including all payers.

Other published studies confirm the observations and the conclusions reported by the OIG in 2013, e.g., ones from the Kaiser Foundation (Jacobson, 2010), the Commonwealth Fund (Schoen, 2013), MedPAC (MedPAC, 2012) and CMS (Walsh, 2010). Studies by Ouslander have shown that structural and process issues within SNFs have a high impact on the rate of hospitalizations (Ouslander, 2012; Ouslander, 2011), further supporting the hypothesis that hospitalization rates could be reduced by feasible changes in facilities' operations.

Jacobson, G., Neuman, T., & Damico, A. (2010). Medicare spending and use of medical services for beneficiaries in nursing homes and other long term care facilities: A potential for achieving Medicare saving and improving the quality of care. The Henry J. Kaiser Family Foundation.

MedPAC. (2012) Report to congress: Payment policy. http://medpac.gov/documents/mar12_entirereport.pdf

Ouslander, J.G., & Maslow, K. (2012). Geriatrics and the triple aim: Defining preventable hospitalizations in the long-term care population. J Am Geriatr Soc., 60(12): 2313-2318.

Ouslander, J.G., & Bersenson, R.A. (2011). Reducing unnecessary hospitalization of nursing home residents. NEJM, 356(13): 1165-1167.

Ouslander, J.G., Lamb, G., Perloe, M., Givens, J.H., Kluge, L., Rutland, T, … Saliba, D. (2010). Potentially avoidable hospitalizations of nursing home residents: Frequency, causes, and costs. J Am Geriatr Soc., 58(4): 627-635.

Ouslander, J.G., Lamb,G., Tappen, R., Herndon, L., Diaz, S., Roos, B.A., … Bonner, A. (2011). Interventions to reduce hospitalizations from nursing homes: Evaluation of the INTERACT II collaborative quality improvement project. J Am Geriatr Soc., 59(4): 745-753.

Schoen, C., Radley, D., Riley, P., Lippa, J., Berenson, J., Dermody, C., & Shih A. (2013). Health Care in the two Americas: Findings from the scorecard on the state health system performance for low-income populations. The Commonwealth Fund.

http://www.commonwealthfund.org/Publications/Fund-Reports/2013/Sep/Low-Income-Scorecard.aspx

Walsh, E.D., Freiman, M., Haber, S., Bragg, A., Ouslander, J., & Wiener, J.M. (2010) Cost drivers for dually eligible beneficiaries: Potentially avoidable hospitalization from nursing facility, skilled nursing facility, and home and community-based services waiver programs, final task 2 report. RTI International.

[Response Ends]

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

Include mean, std dev, min, max, interquartile range, and scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]

AHCA has been calculating and tracking hospitalizations for all SNFs nationally for data from 2014 and updating results each quarter. Below are the basic statistics for the measure from the most recent data available; care through the 4th quarter of 2020.

Statistics	2019q4 (Pre-COVID)	2020q4
N	14,823	14,737
Risk Adjusted Mean Rate	14.7%	14.8%
Standard Deviation	5.3%	6.0%
Min-Max	0-50.9%	0-53.6%

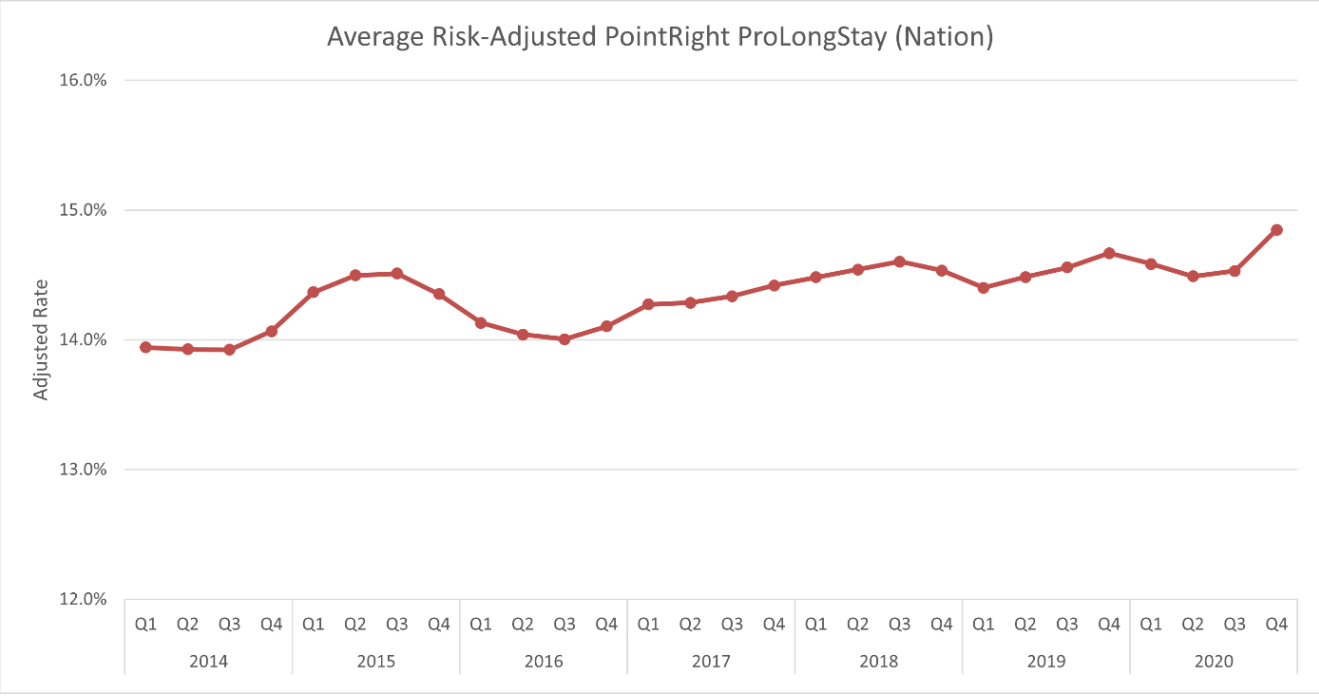
Recent Summary Statistics for ProLongStay

The distribution of the SNFs with reportable data is as follows:

Risk-Adjusted Rate Range	2019q4 (Pre-COVID) – Number of SNFs	2019q4 (Pre-COVID) – Percent of SNFs	2020q4 – Number of SNFs	2020q4 – Percent of SNFs
0-<5%	325	2.2%	513	3.5%
5-<10%	2,452	16.5%	2,660	18.1%
10-<15%	5,309	35.8%	4,687	31.8%
15-<20%	4,552	30.7%	4,078	27.7%
20-<25%	1,709	11.5%	1,987	13.5%
25-<30%	370	2.5%	621	4.2%
30-35%	86	0.6%	139	0.9%
>=35%	20	0.1%	52	0.4%

Recent Rate Distribution for ProLongStay

In addition, all AHCA member facilities have access to their own hospitalization rate updated each quarter through AHCA's Long Term Care Trend Tracker. The national average has increased from the first quarter of 2014 to the fourth quarter of 2020 by 6.5%, with the average change by state varying (see table below).



National Average ProLongStay Rate Over Time

State	2014-Q1 Rate	2020-Q4 Rate	% Change	Facilities
Nation	13.9%	14.8%	6.5%	14,735
AK	7.4%	10.0%	36.0%	19
AL	14.3%	16.2%	13.5%	223
AR	17.7%	16.9%	-4.8%	223
AZ	11.1%	13.8%	24.0%	121
CA	14.6%	16.8%	15.0%	1,133
CO	9.6%	10.1%	5.8%	208
CT	12.9%	13.8%	6.4%	210
DC	13.8%	13.0%	-5.8%	16
DE	13.9%	14.2%	2.3%	44
FL	15.4%	18.3%	18.8%	668
GA	15.2%	15.5%	2.0%	353
HI	6.8%	8.9%	30.3%	41
IA	12.9%	11.9%	-7.9%	422
ID	8.1%	8.9%	10.9%	75
IL	15.7%	16.4%	4.6%	685
IN	12.9%	13.2%	2.3%	522
KS	14.4%	15.0%	4.3%	319
KY	15.3%	16.5%	7.6%	271
LA	19.1%	20.4%	7.0%	261
MA	12.0%	14.8%	23.4%	365
MD	13.0%	13.5%	3.3%	221
ME	11.2%	9.8%	-12.5%	93
MI	13.3%	15.2%	14.3%	430
MN	11.2%	12.4%	11.0%	353
MO	14.7%	16.8%	14.2%	498
MS	19.2%	19.8%	3.3%	197
MT	10.6%	10.7%	0.9%	69
NC	13.6%	13.7%	0.4%	407
ND	11.6%	11.8%	2.1%	79
NE	12.3%	12.6%	2.2%	190

State	2014-Q1 Rate	2020-Q4 Rate	% Change	Facilities
NH	11.1%	10.0%	-9.5%	72
NJ	15.1%	16.0%	5.5%	337
NM	13.0%	13.6%	4.7%	65
NV	12.8%	14.6%	14.0%	53
NY	13.6%	12.2%	-9.8%	597
OH	13.3%	14.9%	11.9%	926
OK	17.7%	18.0%	1.8%	286
OR	11.6%	12.3%	5.8%	125
PA	13.3%	12.9%	-3.1%	654
RI	12.0%	12.4%	2.9%	79
SC	14.6%	16.9%	15.8%	171
SD	11.9%	10.7%	-9.6%	104
TN	14.5%	14.1%	-2.5%	297
TX	15.9%	17.1%	7.3%	1,166
UT	9.9%	10.4%	5.4%	85
VA	13.7%	13.1%	-4.0%	280
VT	10.2%	9.4%	-8.1%	35
WA	10.4%	10.7%	2.5%	196
WI	10.7%	12.3%	15.3%	339
WV	14.2%	12.2%	-13.6%	117
WY	9.8%	10.3%	5.2%	35

National and State Averages and Percent Change from 2014q1 to 2020q4

[Response Ends]

1b.03. If no or limited performance data on the measure as specified is reported above, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement. Include citations.

[Response Begins]

N/A

[Response Ends]

1b.04. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]

The MDS allows us to test the impact of race/ethnicity (A1000), gender (A0800), age (A0900) and Medicaid beneficiary status (A0700). All four of these variables are significant factors in explaining the risk of a long stay hospitalization, but only gender, age, and Medicaid beneficiary status are included in one or more of the four logistic regression models used to calculate the expected rate of first hospitalizations. Race/ethnicity was excluded because only black/African American as a binary variable was significant and using it as a crude proxy for other factors goes against NQF's guidance received during NQF Social Risk Trial from 2015-2017.

In our selection of sociodemographic factors to test we required that the factor be included on the MDS assessment and that it be consistently and reliably reported. These factors were age, gender, race/ethnicity and Medicaid beneficiary status as indicated by the patient's having a Medicaid number. (This does not mean that Medicaid was necessarily the payer for every day of the patient's stay – it is Medicaid eligibility that is the indicator of socioeconomic status as this means the patient has low income and few assets.) The items for occupation and education on admission MDS assessments often were missing. We rejected the option of using community-level (e.g., ZIP code based or census based) socioeconomic variables to impute socioeconomic status of individuals, both because of the high error variance implicit in that approach, and because we thought this would make the risk adjustment less acceptable to providers. In our view it would make little sense to them that the adjusted rates of hospitalization for otherwise identical facilities one city block apart would differ because one was in a different census tract or ZIP code from the other.

In testing the above-mentioned risk factors we compared their effects in both multi-level fixed effects models and in simple logistic regression models. If a multi-level fixed effects model including a given risk factor candidate explained significantly more variance in hospitalization rates than a simple logistic regression, we inferred that part of its effect was via disparities in facility performance correlated with the makeup of the facility's resident population – disparities we did not want to mitigate by risk adjustment. We attributed variance at the individual level to otherwise-unmeasured differences in baseline health status for which risk adjustment would be appropriate.

Specific risk factors were tested and utilized as follows:

A. Age. This was tested using binary variables for age ranges: <65, 65-69, 70-74, …90 or higher. Of these, only the variable indicating age of 90 or over added significantly to the explained variance to the predictive models.
B. Race/Ethnicity. Individual ethnicities (black/African American, Latino, etc.) that are listed on the MDS were tested as binary variables, as was the constructed variable White/Nonwhite. Only black/African American as a binary variable added significantly to the explained variance of logistic regression models. To determine an appropriate coefficient for the black/African American variable we tested it in a two-level fixed effects model with both facility and individual effects. In this model most of the variance due to black race was associated with the facility level – i.e., facilities with a high proportion of black residents showed worse performance after adjustment for other risk factors - but the variable remained significant at the individual level. Despite being significant, the impact was minor and we decided to exclude from our final risk adjustment model, as recommended by the NQF panel at the All-Cause Admissions and Readmissions Standing Committee Meeting on June 8th 2016.
C. Medicaid beneficiary status. As with race, Medicaid was associated with higher hospitalization rates, with most of the effect at the facility level – i.e., facilities with high proportions of Medicaid residents had worse outcomes. Nonetheless, even in the two-level fixed-effects model there was an effect of Medicaid status at the patient level. The coefficient in the simple logistic regression models were determined in fixed effects models and forced into the logistic regressions.
D. Gender. There was a strong effect of gender at the patient level. We interpreted this as totally due to health status differences associated with gender and not under the control of the facility.

Nationally, 76% of all nursing home residents are classified as White, 14% as African American, and 2% as Asian (see table below for full breakdown of race and ethnicity in 2020).

Race/Ethnicity Makeup of SNF Residents (2020q1-2020q4)

SNFs	Residents	White	African American	Asian	Hispanic	Native American/ Hawaii Pacific Islander
15,715	3,739,243	75.8%	13.7%	2.0%	5.6%	0.4%

Race/Ethnicity of SNF Residents in 2020

Stratifying the measure by race and ethnicity would result in most providers having inadequate sample size to report a hospitalization rate. Also, the measure is an all-cause readmission measure, intended to capture the overall performance of each SNF.

When we categorize facilities by their percent of minority residents, we find that facilities with fewer minorities have lower risk adjusted ProLongStay hospitalization rates. In 2020-Q4, facilities with less than 5% of minority residents had an average rate of 12.3%, while facilities with or over 35% of minority residents had an average hospitalization rate of 17.9% (See table below).

Facility-Level Race/Ethnicity x ProLongStay Hospitalization Rate

Facility-Level Race/Ethnicity	SNF Count	Average Risk-Adjusted ProLongStay Hospitalization Rate for 2014-Q1	Average Risk-Adjusted ProLongStay Hospitalization Rate for 2020-Q4
Low (<5%)	5,010 (34%)	12.3%	12.3%
Medium-Low (5-14.9%)	3,415 (23%)	13.5%	14.3%
Medium-High (15-34.9%)	3,133 (21%)	14.8%	16.3%
High (>=35%)	3,169 (22%)	16.3%	17.9%

Facility-Level Race/Ethnicity and Average ProLongStay Rates

A similar pattern is seen when we look at a facility's geographic location relative to the CDC's Social Vulnerability Index (SVI). The SVI evaluates the relative social vulnerability of counties across the United States using 15 different measures of vulnerability in 4 different themes: socioeconomic status, household composition & disability, minority status & language, and housing type & transportation. For every measure, if a county is above the 90th percentile (i.e. in the most vulnerable 10%), it is given a flag.

Facilities located in lower SVI counties had lower risk adjusted ProLongStay hospitalization rates. The difference in the average readmission rate between facilities in low (0 flags) and high (>= 4 flags) SVI counties has decreased over time. In 2020-Q4, facilities in low (0 flags) SVI counties had an average rate of 13.7%, while facilities in high (>=4 flags) SVI counties had an average rate of 17.5%. (See table below).

County-Level Social Vulnerability Index (SVI) x ProLongStay Hospitalization Rate

Facility County SVI	SNF Count	Average Risk-Adjusted ProLongStay Hospitalization Rate for 2014-Q1	Average Risk-Adjusted ProLongStay Hospitalization Rate for 2020-Q4
Low (0 Flags)	5,188 (35%)	13.2%	13.7%
Medium-Low (1 Flag)	3,727 (25%)	13.3%	14.2%
Medium-High (2-3 Flags)	3,164 (22%)	14.1%	15.4%
High (>=4 Flags)	2,597 (18%)	16.1%	17.5%

Facility County SVI and Average ProLongStay Rate

[Response Ends]

1b.05. If no or limited data on disparities from the measure as specified is reported above, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in above.

[Response Begins]

N/A

[Response Ends]

Criteria 2: Scientific Acceptability of Measure Properties

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

---

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

[Response Begins]

No

[Response Ends]

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

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

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

[Response Begins]

N/A

[Response Ends]

sp.01. Provide the measure title.

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

[Response Begins]

PointRight® Pro Long Stay(TM) Hospitalization Measure

[Response Ends]

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

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

[Response Begins]

The PointRight Pro Long Stay Hospitalization Measure is an MDS-based, risk-adjusted measure of the rate of hospitalization of long-stay patients (also known as “residents”) of skilled nursing facilities (SNFs) averaged across the year, weighted by the number of stays in each quarter.

[Response Ends]

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

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

Please do not select:

- *Surgery: General*

[Response Begins]

Other (specify)

N/A

[Response Ends]

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

[Response Begins]

Care Coordination: Readmissions

Care Coordination: Transitions of Care

[Response Ends]

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

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

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

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]

Elderly (Age >= 65)

[Response Ends]

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

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

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

Please do not select:

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

[Response Begins]

Facility

[Response Ends]

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

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

[Response Begins]

Post-Acute Care

[Response Ends]

sp.09. Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials.

Do not enter a URL linking to a home page or to general information. If no URL is available, indicate "none available".

[Response Begins]

https://www.ahcancal.org/Data-and-Research/Documents/PointRight%20Pro%20Long%20Stay%20Hospitalization%20Rate.pdf

[Response Ends]

sp.11. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

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

[Response Begins]

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

[Response Ends]

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

sp.12. State the numerator.

Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome).

DO NOT include the rationale for the measure.

[Response Begins]

The numerator for the measure is the sum over four quarters of the counts of hospitalizations of the quarterly denominator populations, where hospitalizations comprise discharges directly from the SNF to an acute care hospital.

[Response Ends]

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

sp.13. Provide details needed to calculate the numerator.

All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets.

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]

The numerator for a quarter is the number, during the quarter, of discharges from the SNF directly to an acute care hospital of patients in the denominator population for that quarter as indicated by MDS item A2100=03 'discharge status = acute hospital'. A patient in the quarterly denominator population can contribute multiple times to the quarterly numerator.

Discharges to LTACHs, IRFs, and mental hospitals are not included in the numerator, nor are acute hospital admissions directly following a discharge from the SNF to a setting other than an acute care hospital. As noted above, if a patient is discharged from a SNF directly to an acute care hospital during a quarter at risk, the hospitalization will be counted in the numerator even if the patient was discharged to a setting other than an acute care hospital earlier in that quarter.

The numerator for the measure is the sum of the quarterly numerators for the four quarters in the 12 month measurement period.

[Response Ends]

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

sp.14. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

The quarterly denominator population consists of those patients present in the SNF on the first day of the quarter (the "snapshot date") who meet the criterion for long stay on that date. The denominator for a quarter is the number of patients in the quarterly denominator population. The denominator for the measure is the sum of the quarterly denominators for the four quarters in the 12 month measurement period.

The criterion for a patient's having a long stay is a cumulative length of stay in the facility of more than 100 days as of the snapshot date. The cumulative length of stay of a patient is the length of the current stay as of the snapshot date and plus the full lengths of stay of any previous stays that are linked to it. According to the criteria for linkage of stays used in the present measure, a stay in a SNF is linked to a subsequent stay in the SNF if the patient was discharged from the SNF to the community and was readmitted to the SNF within 10 days or fewer. All stays in a sequence of linked stays are included in the sum of days used to determine a patient's cumulative length of stay. In these criteria the term "community" comprises private residences and all organized settings that are primarily residential in character, including senior housing, independent living facilities, board and care homes, and assisted living facilities.

A patient can contribute multiple times to the denominator for a 12 month measure period. For example, a resident continuously present in the facility for a full year would contribute four to the denominator.

[Response Ends]

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

sp.15. Provide details needed to calculate the denominator.

All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets.

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]

The denominator population for a quarter is a subset of the patients present in the SNF on the snapshot date (the first day of the quarter). A patient is in that subset if his or her cumulative length of stay as of the snapshot date is more than 100 days.

The cumulative length of stay is calculated by taking the length of stay of the current admission as of the snapshot date and adding the lengths of stay of any linked stays at the same SNF. The length of the current admission as of the snapshot date is the snapshot date minus the entry date for the current admission, which is MDS item A1600. A stay is linked to a subsequent stay if the patient is discharged to the community (A2100=01) and admitted to the same SNF within 10 days or less (i.e., A1600 for the second stay minus A2100 for the first stay is less than or equal to 10 days).

The denominator for a quarter is the number of residents in the denominator population for that quarter. The denominator for the measure, which reports on a full year's performance, is the sum of the denominators for the four quarters that constitute that year.

[Response Ends]

sp.16. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

There are no exclusions from the denominator; all patients in the facility on the snapshot date who meet the long stay criterion on that date are included.

[Response Ends]

sp.17. Provide details needed to calculate the denominator exclusions.

All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]

An outcome is regarded as unknown if it cannot be reasonably inferred or conservatively imputed.

The quarterly unknown outcome count is the number of patients in the quarterly denominator for whom it is not known and cannot be reasonably inferred or imputed that the patient was or was not hospitalized during the quarter. It would be known that a patient was hospitalized during the quarter if he or she had a discharge MDS with an acute care hospital as a discharge disposition. It would be known that a patient was not hospitalized during the quarter if he or she had an MDS assessment with an assessment reference date (item A2300) following the end of the quarter at risk and had an admission date (item A1600) on or prior to the snapshot date. If the patient has a discharge MDS during the quarter at risk and is subsequently readmitted to the same SNF within the same quarter it is assumed that there was a second discharge during that quarter (whether to an acute care hospital or elsewhere) if and only if there is a discharge MDS with an assessment reference date within that quarter. If there is an admission to the SNF from an acute care hospital during the quarter at risk but no preceding discharge MDS the inference is made that the preceding discharge was directly to an acute care hospital and the inferred discharge is counted in the numerator of the measure. If a patient has no MDS assessment of any kind with an assessment reference date 100 days or fewer after the latest MDS in the interval starting 10 days before the snapshot date and ending one day before the end of the quarter the patient's outcome is regarded as unknown. If the count N of patients with unknown outcomes is 10% or less of the denominator, N*0.8 is added to the numerator. If N is more than 10% of the denominator the measure is not reported.

The denominator of the annual unknown outcome rate is the sum of the four quarterly denominators. The numerator of the annual unknown outcome rate is the sum over the four quarters of the numbers of quarterly denominator patients with an unknown outcome in the quarter at risk.

[Response Ends]

sp.18. Provide all information required to stratify the measure results, if necessary.

Include the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate. Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format in the Data Dictionary field.

[Response Begins]

N/A

[Response Ends]

sp.19. Select the risk adjustment type.

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

[Response Begins]

Statistical risk model

[Response Ends]

sp.20. Select the most relevant type of score.

Attachment: If available, please provide a sample report.

[Response Begins]

Rate/proportion

[Response Ends]

sp.21. Select the appropriate interpretation of the measure score.

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

[Response Begins]

Better quality = Lower score

[Response Ends]

sp.22. Diagram or describe the calculation of the measure score as an ordered sequence of steps.

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

[Response Begins]

The formula for a facility's adjusted PointRight ProLong Hospitalization Rate is: [Observed rate of all hospitalizations]/[Expected rate of all hospitalizations]*[National average rate of all hospitalizations].

The observed and expected rates are updated quarterly and the national benchmark rate is updated annually; the national benchmark rate used in the calculation is the most recently calculated benchmark rate at the time the observed and expected rates are calculated.

Procedure for calculating the adjusted rate:

1) Calculate the observed rate.

The observed PointRight ProLong Stay Hospitalization Rate is the sum of the four quarterly numerators divided by the sum of the four quarterly denominators.

- The denominator for a quarter is the number of residents present in the facility on the first day of a calendar quarter who qualify as long stay on that day
- The numerator for a quarter is number of hospitalizations of residents in the denominator population for that quarter, where hospitalization means discharge from the SNF directly to an acute care hospital, either with no return to the SNF or with return to the SNF after at least one midnight outside the SNF.

The numerator excludes:

1. Hospitalizations occurring after a patient has been discharged somewhere other than an acute care hospital and
2. Hospitalizations at psychiatric hospitals, rehabilitation hospitals, or LTACHs.

The numerator includes:

1. "observations stays" if these involve at least one midnight away from the SNF and
2. "planned" hospitalizations.

2) Calculate the expected rate.

- Calculate the expected number of first hospitalizations of the quarterly denominator population for each of the four quarters in the measure period and sum them; multiply the sum by 1.25248 to obtain the expected number of total hospitalizations for the 12-month measure period. Divide this number by the sum of the quarterly denominators to get the expected rate for the measure period.

3) Calculate the national benchmark rate

- The national benchmark rate is the observed PointRight Pro Long Stay Hospitalization Rate for a denominator population consisting of the denominator populations for all SNFs in the largest available national sample that have complete non-discharge MDS data for all of their patients for all four quarters in the measure period and have 100% known outcomes for all patients in their denominator populations for all four quarters in the measure period. For a given member of a quarterly denominator population a known outcome means either that the patient had a discharge MDS submitted with a discharge date within the quarter and a discharge destination filled in, that the patient was readmitted from an acute care hospital during the quarter, or that the

patient had a quarterly or other MDS submitted in the 100 days following the end of the quarter that gave an admission date prior to the snapshot date for the given quarter.

Procedure for Calculating the Measure:

1. Establish a 12-month measure period comprising of four calendar quarters (each three months in length). For each quarter, the (quarterly) denominator is the number of residents who qualify as long stay for that quarter, i.e. whose cumulative length of stay as of the snapshot date (the first day of the quarter) is more than 100 days. (Cumulative length of stay is defined as the sum of the lengths of stay of the current stay and all stays linked to it.) The sum of the quarterly denominators for the four quarters constitutes the denominator for the measure period.
2. For the quarterly denominator population determine the number of (direct) acute care hospitalizations of the residents in that quarter (the quarterly numerator). The count of the hospitalizations is the quarterly numerator. The sum of the quarterly numerators for the four quarters constitutes the numerator for the measure. As noted above the count includes only admissions to acute care hospitals directly from the SNF. Planned (or presumptively planned) hospitalizations are included, as are observation stays. Hospitalizations subsequent to a discharge somewhere other than an acute care hospital, and hospitalizations at LTACHs and specialty hospitals are excluded.
3. Divide the total numerator by the total denominator to obtain the observed rate for the SNF.
4. Calculate the estimated probability of a first hospitalization for each member of each quarterly denominator population using the predictive model described above, and sum these probabilities to get the expected number of first hospitalizations per quarter for the total 12 month denominator population. Sum these expected numbers over the four quarters of the measure period to get the expected number of first hospitalizations for the measure period. Multiply this result by 1.25248 to get the expected number of total hospitalizations for the total measure period denominator population, and divide this by the total measure period denominator to get the expected PointRight Pro Long Stay Hospitalization Rate for the measure period.
5. Divide the observed rate by the expected rate and multiply by the most recent national benchmark rate to obtain the Adjusted PointRight Pro Long Stay Hospitalization Rate.

[Response Ends]

sp.25. If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

[Response Begins]

N/A

[Response Ends]

sp.28. Select only the data sources for which the measure is specified.

[Response Begins]

Assessment Data

[Response Ends]

sp.29. Identify the specific data source or data collection instrument.

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

[Response Begins]

Resident Assessment Instrument Minimum Data Set (MDS) version 3.0

[Response Ends]

sp.30. Provide the data collection instrument.

[Response Begins]

Available in attached appendix in Question 1 of the Additional Section

[Response Ends]

Attachment: Resident Assessment Instrument Minimum Data Set (MDS) version 3.0

2a. Reliability

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

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

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

No

[Response Ends]

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

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

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

Yes

[Response Ends]

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

[Response Begins]

No

[Response Ends]

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

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

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

[Response Begins]

No additional risk adjustment analysis included

[Response Ends]

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

• Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.

• All required sections must be completed.

• For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.

• If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.

• An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.

• Contact NQF staff with any questions. Check for resources at the

Submitting Standards webpage .

• For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the

2021 Measure Evaluation Criteria and Guidance .

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

2a. Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

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

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

AND

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

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

• an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration

OR

• rationale/data support no risk adjustment/ stratification.

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

OR

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

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

2b6. Analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias.

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

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

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

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

Definitions

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

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

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

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

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

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

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

2021 Submission:

Updated testing information here.

2018 Submission:

Testing from the previous submission here.

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

[Response Begins]

Assessment Data

[Response Ends]

2a.02. If an existing dataset was used, identify the specific dataset.

The dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry).

[Response Begins]

Resident Assessment Instrument Minimum Data Set (MDS) version 3.0

[Response Ends]

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

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

[Response Begins]

The PointRight® Pro Long Stay™ Hospitalization Measure risk adjustment models were fit on data from April 1, 2013 to November 30, 2014, from which covariates and the dependent variable were ascertained for a 12 month risk period between July 1, 2013 and June 30, 2014. Data covering the 12 month risk period January 1, 2014 to December 31, 2014

were utilized in various reliability and validity tests. MDS discharge status codes and Medicare Part A claims data used to demonstrate the representativeness of the PointRight® sample of MDS assessments came from CY 2013.

[Response Ends]

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

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

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

Please do not select:

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

[Response Begins]

Facility

[Response Ends]

2a.05. List the measured entities included in the testing and analysis (by level of analysis and data source).

Identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample.

[Response Begins]

The PointRight Pro Long Stay Hospitalization Measure was developed on MDS 3.0 assessments collected from skilled nursing facilities that purchased analytics services from PointRight. Modeling and estimation of risk adjustment were performed on 2,584 SNFs using four snapshots dates: Q3 2013, Q4 2013, Q1 2014 and Q2 2014. Testing and analysis were performed on 2,182 SNFs using calendar year 2014. In table 1 below, find facility level descriptive statistics on these SNFs and how they compare to the national population of skilled nursing facilities.

TABLE 1. Testing Sample Facility Level Descriptive Statistics

Metric	Nation	*	PointRight	*
*	N	%	N	%
Part of chain	8,748	56.4	2,118	83.5
For profit	10,916	69.8	2,218	87.4
Government	958	6.1	38	1.5
Hospital-based	871	5.6	12	0.5
Medicare certified facilities	15,169	97.0	2,532	99.8
Resident count less than 50	3,765	24.1	257	10.1
Resident count greater than 50, less than 110	7,994	51.1	1,420	56.0
Resident count greater than 110	3,884	24.8	860	33.9

*Cell intentionally left empty

Comparison of facility-level descriptive statistics between PointRight sample and the entire nation in 2014

The PointRight sample contained facilities of various bed counts, chain vs. independent ownership, hospital based vs. non-hospital based affiliation and for-profit vs. nonprofit designation. The PointRight sample had greater proportions of large for-profit chain facilities than the national SNF population; all provider types were sufficiently represented in the PointRight sample.

[Response Ends]

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

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

[Response Begins]

The PointRight Pro Long Stay Hospitalization Measure measures the rate of hospitalization for a SNF's Long Stay population over a 12-month measurement period. The measure sums numerators and denominators from four snapshot dates – the first day of each calendar quarter within the measure period. Using MDS from 2,584 SNFs, our dataset contained more than 150 thousand patients, present in the facility, on each of the 4 snapshot dates: Q3 2013, Q4 2013, Q1 2014 and Q2 2014. The table below presents the demographics of the denominator sample from one representative snapshot date, 4/1/2014.

TABLE 2. Characteristics of Patients

MDS Variables	MDS Variable Value Categories	Prevalence of PointRight Client Sample (N=2,584)	Prevalence of Nation (N=14,620)
Age	<65	20.1%	15.4%
	65-74	14.7%	15.5%
	75-84	27.3%	26.4%
	85-89	18.3%	19.1%
	90 or over	19.6%	23.6%
Gender	Female	69.0%	67.5%
	Male	31.0%	32.6%
Race	Asian	1.8%	1.9%
	Black or African American	16.4%	14.1%
	Hispanic Latino	5.6%	5.2%
	White	73.5%	76.2%
	Other	2.9%	3.0%
Medicaid Beneficiary	Yes	79.2%	N/A
Admission Setting	Acute Hospital	83.4%	74.7%

MDS Variables	MDS Variable Value Categories	Prevalence of PointRight Client Sample (N=2,584)	Prevalence of Nation (N=14,620)
Active Diagnosis	Anemia	29.7%	29.2%
Active Diagnosis	Asthma, COPD, or Chronic Lung Disease whether or not on oxygen	19.9%	20.4%
	Asthma, COPD, or Chronic Lung Disease on oxygen	6.0%	N/A
	Diabetes on insulin	20.1%	32.4%
	Gastroesophageal Reflux Disease (GERD) or ulcer	29.1%	33.9%
	Heart Failure	19.2%	19.3%
	Hypertension	74.3%	75.1%
	Viral Hepatitis	0.6%	0.6%
	Neurogenic bladder	2.8%	2.7%
	Renal failure or insufficiency	9.2%	10.0%
Incontinence	Total bowel incontinence	34.0%	31.4%
Medications Received	Anticoagulant within 7 days prior to ARD	16.3%	12.3%
	Antibiotics within 7 days prior to ARD	11.1%	11.0%
Symptoms	Dyspnea on exertion	7.0%	7.5%
Skin	Surgical wound(s)	1.9%	2.1%
Hospice Status	Receiving hospice care	4.9%	4.7%
Recent Treatments	IV fluid or meds within 7 days before last MDS	1.3%	1.3%
*	Oxygen in 7 days before last MDS	10.8%	11.2%

Comparison of patient characteristics between PointRight sample used in testing and the Nation

[Response Ends]

2a.07. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing.

[Response Begins]

The majority of measure reliability and validity testing was conducted on the measure development sample of 2,584 SNFs described above, which as noted above is national and provides a good representation of all major demographic categories and provider types, though it is not a random sample of all U.S. SNFs.

2022 Submission

Additional validity testing used more recent national data from 2019.

[Response Ends]

2a.08. List the social risk factors that were available and analyzed.

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

[Response Begins]

We tested black/non-black, Medicaid/non-Medicaid, and the interaction between these binary variables for their relationships with long-stay hospitalization rates. We did not utilize the patient-level variable concerning occupation, because it often is not completed on the admission MDS or the admission MDS is not available for analysis; also, it is an unreliable indicator of the patient's primary lifetime occupation. We did not use the two language-related items on the MDS, because they often are missing or unreliable.

Race/ethnicity items other than black/non-black either did not have significant patient-level effects in a fixed effects model (Hispanic/Latino) or our sample was insufficient (American Indian or Alaska Native; Native Hawaiian or Pacific Islander).

We rejected geographically-based proxy variables for two reasons:

1. We did not want to take the risk of adjusting away true disparities in care quality that might be found comparing SNFs in poorer neighborhoods with those in richer ones.
2. Particularly for the long-stay SNF population, most of whom have the SNF as their primary residence, the ZIP code or census tract of the patient is simply that of the SNF. This may not be indicative of the patients' socioeconomic status as it is for community-dwelling patients.

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.07 check patient or encounter-level data; in 2a.08 enter "see validity testing section of data elements"; and enter "N/A" for 2a.09 and 2a.10.

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

Choose one or both levels.

[Response Begins]

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

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

[Response Ends]

2a.10. For each level of reliability testing checked above, describe the method of reliability testing and what it tests.

Describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used.

[Response Begins]

Agreement of Model Independent Variables-

For the MDS data items utilized as risk adjustment covariates, a comparison of the variables prevalence between the testing sample of 2,584 SNFs and the national population was performed. Table 2, displayed above demonstrates that the PointRight client database although not a random sample of patients, is representative of the national population.

Reliability of Rates over Time -

To assess the reliability of the overall measure we analyzed change from quarter to quarter in the observed and adjusted long-stay hospitalization rates. We reasoned that a SNF's underlying probability of its long-stay patients hospitalizing, and the characteristics of its long-stay patient population, are unlikely to change greatly over a 3 month period, so that most of the change from quarter to quarter will be related to limitations on measurement reliability. Some of the reliability limitations will reflect error in the measurement of the dependent variable or risk adjustment covariates on the MDS, but most is likely to reflect changes in the characteristics of the long-stay population from one snapshot date to the next. If correlation coefficients – both parametric and non-parametric – are relatively high when consecutive quarters are compared, we infer that the combination of measurement-related variability and sampling-related variability is acceptable.

Stability of Facility Level Adjusted Rate Bootstrapping –

To further test the reliability of the measure, adjusted rates for the measure period CY 2014 were recalculated for our testing sample, where a random sample of stays was drawn with replacement for each facility. We then reviewed the distribution of differences between facilities' original adjusted rates and the rates calculated with resampling. If the distribution of differences has a small variance and a mean of 0 we can assume the measure is acceptably stable.

[Response Ends]

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

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

[Response Begins]

Agreement of Model Independent Variables -

Results found in Table 2 in section 2a.06.

Reliability of Rates over Time –

TABLE 3. Pearson Correlation Coefficients of PointRight® Pro Long Stay™ Hospitalization *Risk Adjusted Rates* by Quarter

Time Period of Adjusted Rate	Pearson Statistics	Adjusted Hospitalization Rates: Jan'14-Dec'14	Adjusted Hospitalization Rates: Oct'13-Sep'14	Adjusted Hospitalization Rates: Jul'13-Jun'14	Adjusted Hospitalization Rates: Apr'13 - Mar'14
Adjusted Hospitalization Rates: Jan'14-Dec'14	Person Correlation	1.000	0.888	0.745	0.575
	p-value		<.0001	<.0001	<.0001
	Observations	2182	2147	2124	2072

Time Period of Adjusted Rate	Pearson Statistics	Adjusted Hospitalization Rates: Jan'14-Dec'14	Adjusted Hospitalization Rates: Oct'13-Sep'14	Adjusted Hospitalization Rates: Jul'13-Jun'14	Adjusted Hospitalization Rates: Apr'13 - Mar'14
Adjusted Hospitalization Rates: Oct'13-Sep'14	Person Correlation	0.888	1.000	0.889	0.738
Adjusted Hospitalization Rates: Oct'13-Sep'14	p-value	<.0001		<.0001	<.0001
	Observations	2147	2147	2120	2067
Adjusted Hospitalization Rates: Jul'13-Jun'14	Person Correlation	0.745	0.889	1.000	0.889
	p-value	<.0001	<.0001		<.0001
	Observations	2124	2120	2124	2069
Adjusted Hospitalization Rates: Apr'13-Mar'14	Person Correlation	0.575	0.738	0.889	1.000
	p-value	<.0001	<.0001	<.0001	
	Observations	2072	2067	2069	2072

Pearson correlation statistics between quarterly risk-adjusted hospitalization rates

TABLE 4. Pearson Correlation Coefficients of PointRight® Pro Long Stay™ Hospitalization *Observed Rates* by Quarter

Time Period of Observed Rate	Pearson Statistics	Observed Hospitalization Rates: Jan'14-Dec'14	Observed Hospitalization Rates: Oct'13-Sep'14	Observed Hospitalization Rates: Jul'13-Jun'14	Observed Hospitalization Rates: Apr'13 - Mar'14
Observed Hospitalization Rates: Jan'14-Dec'14	Person Correlation	1.000	0.934	0.845	0.727
	p-value		<.0001	<.0001	<.0001
	Observations	2182	2147	2124	2072
Observed Hospitalization Rates: Oct'13-Sep'14	Person Correlation	0.934	1.000	0.942	0.851
	p-value	<.0001		<.0001	<.0001
	Observations	2147	2147	2120	2067
Observed Hospitalization Rates: Jul'13-Jun'14	Person Correlation	0.845	0.942	1.000	0.943
	p-value	<.0001	<.0001		<.0001
	Observations	2124	2120	2124	2069

Time Period of Observed Rate	Pearson Statistics	Observed Hospitalization Rates: Jan'14-Dec'14	Observed Hospitalization Rates: Oct'13-Sep'14	Observed Hospitalization Rates: Jul'13-Jun'14	Observed Hospitalization Rates: Apr'13 - Mar'14
Observed Hospitalization Rates: Apr'13-Mar'14	Person Correlation	0.727	0.851	0.943	1.000
	p-value	<.0001	<.0001	<.0001	
Observed Hospitalization Rates: Apr'13-Mar'14	Observations	2072	2067	2069	2072

Pearson correlation statistics between quarterly observed hospitalization rates

TABLE 5. Spearman Correlation Coefficients of PointRight® Pro Long Stay™ Hospitalization *Risk Adjusted Rates* by Quarter

Time Period of Adjusted Rate	Spearman Statistics	Adjusted Hospitalization Rates: Jan'14-Dec'14	Adjusted Hospitalization Rates: Oct'13-Sep'14	Adjusted Hospitalization Rates: Jul'13-Jun'14	Adjusted Hospitalization Rates: Apr'13 - Mar'14
Adjusted Hospitalization Rates: Jan'14-Dec'14	Spearman Correlation	1.000	0.879	0.742	0.582
	p-value		<.0001	<.0001	<.0001
	Observations	2182	2147	2124	2072
Adjusted Hospitalization Rates: Oct'13-Sep'14	Spearman Correlation	0.879	1.000	0.885	0.739
	p-value	<.0001		<.0001	<.0001
	Observations	2147	2147	2120	2067
Adjusted Hospitalization Rates: Jul'13-Jun'14	Spearman Correlation	0.742	0.885	1.000	0.882
	p-value	<.0001	<.0001		<.0001
	Observations	2124	2120	2124	2069
Adjusted Hospitalization Rates: Apr'13-Mar'14	Spearman Correlation	0.582	0.739	0.882	1.000
	p-value	<.0001	<.0001	<.0001	
	Observations	2072	2067	2069	2072

Spearman correlation statistics between quarterly risk-adjusted rates

TABLE 6. Spearman Correlation Coefficients of PointRight® Pro Long Stay™ Hospitalization *Observed Rates* by Quarter

Time Period of Observed Rate	Spearman Statistics	Observed Hospitalization Rates: Jan'14-Dec'14	Observed Hospitalization Rates: Oct'13-Sep'14	Observed Hospitalization Rates: Jul'13-Jun'14	Observed Hospitalization Rates: Apr'13 - Mar'14
Observed Hospitalization Rates: Jan'14-Dec'14	Spearman Correlation	1.000	0.926	0.823	0.699
	p-value		<.0001	<.0001	<.0001
Observed Hospitalization Rates: Jan'14-Dec'14	Observations	2182	2147	2124	2072
Observed Hospitalization Rates: Oct'13-Sep'14	Spearman Correlation	0.926	1.000	0.932	0.831
	p-value	<.0001		<.0001	<.0001
	Observations	2147	2147	2120	2067
Observed Hospitalization Rates: Jul'13-Jun'14	Spearman Correlation	0.823	0.932	1.000	0.935
	p-value	<.0001	<.0001		<.0001
	Observations	2124	2120	2124	2069
Observed Hospitalization Rates: Apr'13-Mar'14	Spearman Correlation	0.699	0.831	0.933	1.000
	p-value	<.0001	<.0001	<.0001	
	Observations	2072	2067	2069	2072

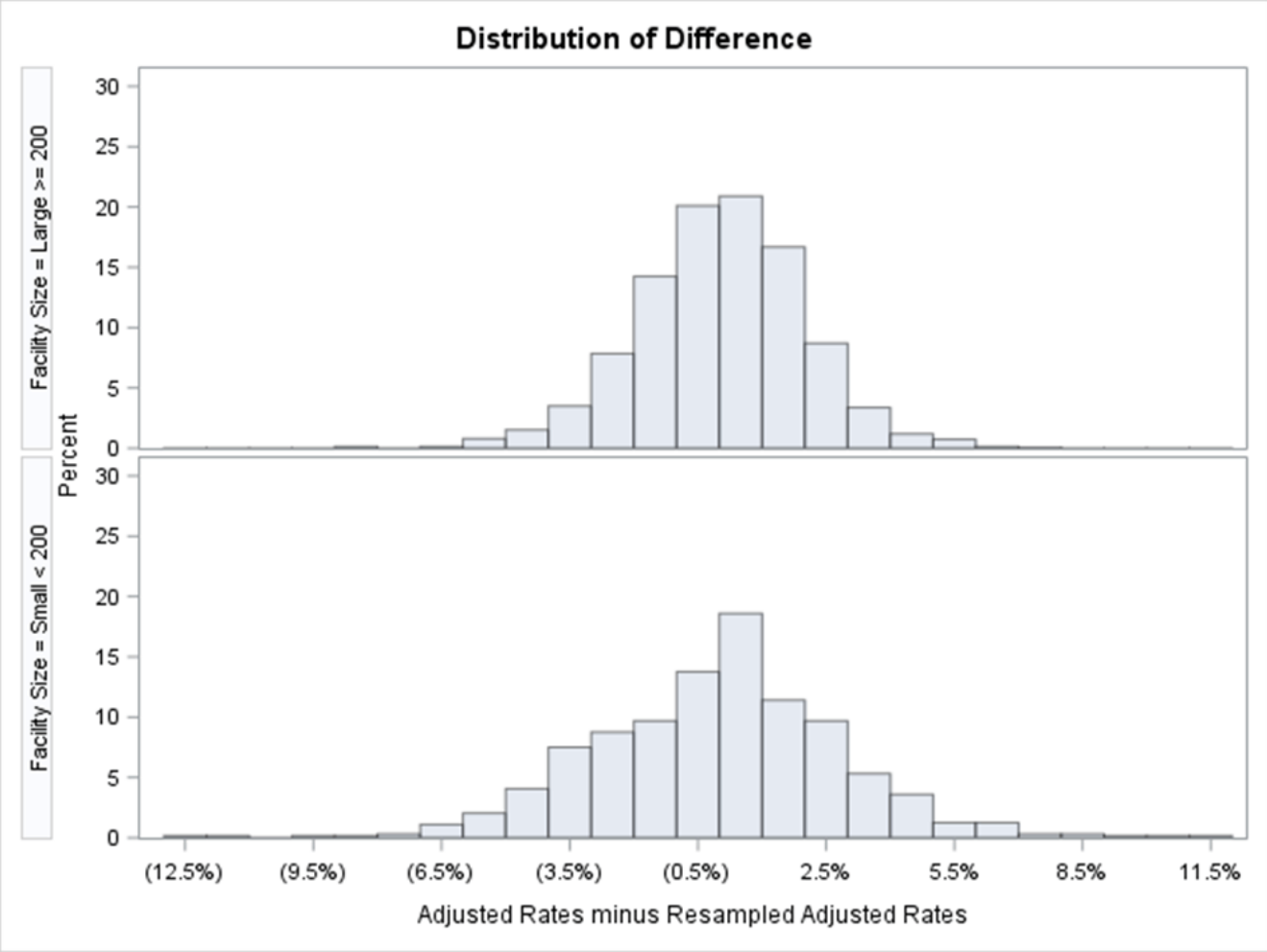
Spearman statistics between quarterly observed rates

Table 7. Distribution of Differences between Facility Adjusted Rates and Resampled Adjusted Rates

Quantile	Difference in Rates: (Adjusted Rates - Resampled Adj Rates)
100% Max	11.8%
99%	5.8%
95%	3.5%
90%	2.6%
75% Q3	1.4%
50% Median	0.0%
25% Q1	-1.3%
10%	-2.9%
5%	-3.8%
1%	-5.8%
0% Min	-12.7%

Distribution of Differences between Facility Adjusted Rates and Resampled Adjusted Rates

Figure 1. Histogram of Differences between Facility Adjusted Rates and Resampled Adjusted Rates by Facility Size



Y-Axis is Percent of Facilities, X-Axis is Difference between Adjusted and Resampled Rates. Top Histogram is of facility size greater than equal to 200. Bottom Histogram is of facility size less than 200.

[Response Ends]

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

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

[Response Begins]

Agreement of Model Independent Variables -

45% of the risk adjustment model covariates, that were comparable, were found to have prevalence within 5% of the prevalence found in the national sample. 65% (13 out of 20) risk adjustment model covariates, that were comparable, were found to have prevalence within 10% of the prevalence found in the national sample. Although the measure testing sample is not a random sample of all U.S. SNF patients, all the model IV cohorts are sufficiently represented in our sample.

Reliability of Rates Over Time –

Correlations from one quarter to the next ranged between .888 to .889 for the parametric statistic and .879 to .885 for the rank order statistic. The correlations suggest that the measure is adequately stable over short periods, but sufficiently variable to reflect clinically meaningful changes.

Stability of Facility Level Adjusted Rate Bootstrapping -

Reviewing the distribution of facility level differences between adjusted hospitalization rates and resampled adjusted rates illustrates the PointRight Pro Long Stay Hospitalization Measure has a high level of precision. 64.8% of the PointRight sample had a difference in adjusted rates of less than 2% and only 4.6% of facilities had a difference greater than 5%. The mean difference was 0.1%.

1,514 (70%) of the facilities in our sample had a denominator greater than 200 patient quarters. For these larger faculties we noticed, as expected, the variance of the distribution in differences shrinks. Smaller facilities will have less measure reliability, but we found the variance acceptable even for facilities with denominators between 30 and 200.

[Response Ends]

2b. Validity

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

[Response Begins]

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

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)

[Response Ends]

2b.02. For each level of testing checked above, describe the method of validity testing and what it tests.

Describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used.

[Response Begins]

Agreement of Model Dependent Variables -

For the dependent variable of acute care hospitalization we compared the identification of hospitalization events of Medicare FFS beneficiaries based on the MDS and hospitalization events based on Medicare FFS claims. Because MDS data include all payers, not just Part A Medicare patients, we restricted the MDS discharges to those where the patient was enrolled in Part A Medicare when they were discharged, and who had data in our extract of inpatient and other claims data for patients who had a SNF claim in calendar year 2012. We used 2012 MDS data, claims data and enrollment data (for Part A enrollment and dates of death) because it was the most recent in our database. We would not expect significant differences in the match rates for newer data, and those differences would likely be improvements, given CMS's MDS focus surveys and other MDS data quality initiatives.

Performance Measure Score- Correlation with SNF Industry Measures of Quality:

To test construct validity of the PointRight Pro Long Stay Hospitalization measure we tested the relationship of the measure with the various components of the CMS Five-Star ratings for SNFs and its correlation with CMS's long-stay Quality Measures. We hypothesized that facilities with higher star ratings would have lower adjusted long-stay hospitalization rates.

2022 Submission

Since the original endorsement, CMS added a Medicare claims-based long-stay hospitalization measure to Care Compare (formerly Nursing Home Compare) and five star. CMS did not seek NQF endorsement for this measure. Still, we calculated the correlation between this measure and ProLongStay. Additionally, we updated the analysis on the relationship between five star and ProLongStay to account for more recent data.

[Response Ends]

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

Examples may include correlations or t-test results.

[Response Begins]

Agreement of Model Dependent Variables –

TABLE 8. Agreement between MDS discharge status codes for long stays, and inpatient claims and death records

MDS Discharge Status	Total	*	According to claims and the Medicare enrollment record....	*	*	*	*	*	*	*
*	*	*	STACH/CAH	*	IRF, LTCH, Psych Hospital, or Other IP	*	Alive but no IP claim	*	Died	*
*	N	Col%	N	Row %	N	Row %	N	Row %	N	Row %
All Long Stay MDS Discharges^	332,919	100%	213,772	64%	14,762	4%	70,756	21%	33,629	10%
Acute hospital	241,857	73%	208,891	86%	6,381	3%	25,066	10%	1,519	1%
IRF, LTCH or Psych Hospital	9,957	3%	*	*	7,967	80%	*	*	*	*
Other setting	48,956	15%	3,851	8%	*	*	44,128	90%	*	*
Died	32,149	10%	*	*	*	*	*	*	31,545	98%

Agreement between MDS discharge status codes for long stays, and inpatient claims and death records

Notes: *Positive patient counts less than 11 must be blinded due to our CMS data use agreements

^A long stay discharge is defined as the patient having been in the facility for 100 days from admission to discharge

*****Cell intentionally left empty

Performance Measure Score- Correlation with SNF Industry Measures of Quality:

TABLE 9. Pearson Correlation Coefficients with CMS Long Stay Quality Measures

Quality Measure	Correlation Coefficient with ProLongStay Adjusted Hospitalization Rate	p-value
High-Risk Residents with Pressure Ulcers (Long Stay)*	0.20	<.0001
Low-Risk Residents Who Lose Control of Their Bowels or Bladder (Long Stay)	-0.02	0.3396
Residents Assessed and Appropriately Given the Pneumococcal Vaccine (Long Stay)	-0.02	0.2818
Residents Assessed and Appropriately Given the Seasonal Influenza Vaccine (Long Stay)*	-0.05	0.0265
Residents Experiencing One or More Falls with Major Injury (Long Stay)	0.04	0.078
Residents Who Have Depressive Symptoms (Long Stay)*	-0.04	0.0398
Residents Who Have/Had a Catheter Inserted and Left in Their Bladder (Long Stay)*	0.06	0.0037
Residents Who Lose Too Much Weight (Long Stay)*	0.10	<.0001
Long-Stay Residents Who Received An Antipsychotic Medication	0.02	0.2582
Residents who Self-Report Moderate to Severe Pain (Long Stay)*	0.08	0.0001
Residents Who Were Physically Restrained (Long Stay)*	0.07	0.0009
Residents Whose Need for Help with Activities of Daily Living Has Increased (Long Stay)*	0.11	<.0001
Residents with a Urinary Tract Infection (Long Stay)*	0.10	<.0001

Pearson Correlation Coefficients with CMS Long Stay Quality Measures

Notes: *Correlation coefficients that were found statistically significant at $p < 0.05$

2022 Submission

TABLE 10. Average Hospitalization Rate by Overall Five-Star Rating

Overall Five-Star Rating	Hospitalization Rate 2014-Q4	Hospitalization Rate 2019-Q3	Hospitalization Rate 2020-Q3
1	15.5%	16.7%	17.3%
2	14.1%	15.4%	15.5%
3	13.8%	14.9%	15.0%
4	13.4%	14.0%	14.0%
5	12.7%	12.2%	11.9%
Pearson Correlation	-0.142	-0.280	-0.350

Overall Five-Star Rating	Hospitalization Rate 2014-Q4	Hospitalization Rate 2019-Q3	Hospitalization Rate 2020-Q3
Coefficient (p-value)	(p < 0.001)	(p < 0.001)	(p<0.001)

Average Hospitalization Rate by Overall Five-Star Rating

TABLE 11. Average Hospitalization Rate by Survey Five-Star Rating

Survey Five-Star Rating	Hospitalization Rate 2014-Q4	Hospitalization Rate 2019-Q3	Hospitalization Rate 2020-Q3
1	14.4%	16.0%	16.1%
2	13.8%	15.0%	15.1%
3	13.4%	14.3%	14.3%
4	13.1%	13.7%	13.5%
5	13.3%	13.1%	12.6%
Pearson Correlation Coefficient (p-value)	-0.080 (p < 0.001)	-0.169 (p < 0.001)	-0.226 (p<0.001)

Average Hospitalization Rate by Survey Five-Star Rating

TABLE 12. Average Hospitalization Rate by Quality Five-Star Rating

Quality Five-Star Rating	Hospitalization Rate 2014-Q4	Hospitalization Rate 2019-Q3	Hospitalization Rate 2020-Q3
1	17.0%	18.7%	19.7%
2	15.7%	16.9%	17.4%
3	14.3%	15.4%	15.7%
4	14.1%	14.1%	14.2%
5	13.1%	12.3%	12.0%
Pearson Correlation Coefficient (p-value)	-0.127 (p < 0.001)	-0.367 (p < 0.001)	-0.429 (p<0.001)

Average Hospitalization Rate by Quality Five-Star Rating

TABLE 13. Average Hospitalization Rate by Staffing Five-Star Rating

Staffing Five-Star Rating	Hospitalization Rate 2014-Q4	Hospitalization Rate 2019-Q3	Hospitalization Rate 2020-Q3
1	14.7%	15.8%	15.8%
2	13.8%	15.5%	15.6%
3	14.1%	14.5%	14.9%

Staffing Five-Star Rating	Hospitalization Rate 2014-Q4	Hospitalization Rate 2019-Q3	Hospitalization Rate 2020-Q3
4	13.3%	13.7%	14.1%
5	11.2%	12.3%	12.3%
Pearson Correlation Coefficient (p-value)	-0.105 ($p < 0.001$)	-0.198 ($p < 0.001$)	-0.248 ($p<0.001$)

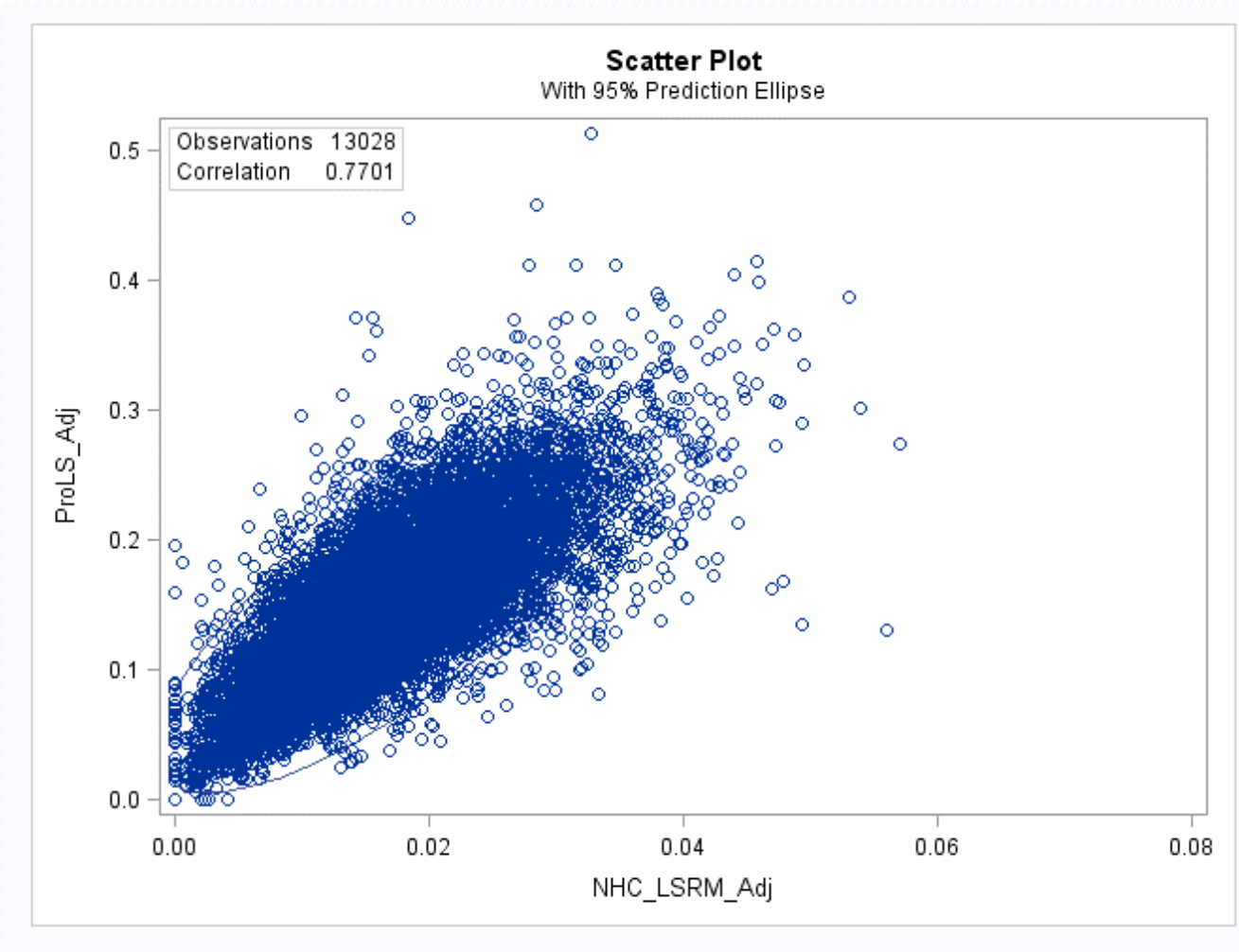
Average Hospitalization Rate by Staffing Five-Star Rating

TABLE 14. Average Hospitalization Rate by RN Staffing Five-Star Rating

RN Staffing Five-Star Rating	Hospitalization Rate 2014-Q4	Hospitalization Rate 2019-Q3	Hospitalization Rate 2020-Q3
1	14.7%	16.2%	15.0%
2	14.6%	15.4%	14.3%
3	13.9%	14.4%	13.3%
4	13.7%	13.6%	12.3%
5	12.1%	12.4%	10.2%
Pearson Correlation Coefficient (p-value)	-0.141 ($p < 0.001$)	-0.233 ($p < 0.001$)	-0.276 ($p<0.001$)

Average Hospitalization Rate by RN Staffing Five-Star Rating

Figure 2. Scatter Plot of ProLongStay Adjusted Hospitalization Rates and Nursing Home Compare Long-Stay Adjusted Hospitalization Rates (2020-Q3)



Scatter Plot of MDS-based Adjusted ProLongStay (y-axis) and Claims-based Adjusted Nursing Home Compare Long Stay Hospitalizations (x-axis). Pearson Correlation of 0.770 significant with p<0.0001

[Response Ends]

2b.04. Provide your interpretation of the results in terms of demonstrating validity. (i.e., what do the results mean and what are the norms for the test conducted?)

[Response Begins]

Critical Data Elements Agreement of Model Dependent Variables -

Table 8 presents a comparison of hospitalizations identified by claims and hospitalizations identified by the MDS, based on data from 241,857 long stay discharges for patients enrolled in Medicare Part A on discharge and existing in our extract of Medicare SNF patients in 2012, covering 15,091 SNFs, showed that 86% of hospitalizations of Medicare FFS patients identified by the MDS are confirmed by Medicare FFS claims; in the other direction, 98% (208,891 of 213,772) of acute inpatient claims found near an MDS discharge have an MDS discharge code of acute hospital. In other words, MDS discharge assessments appear to be overstating the rate of acute hospitalizations to a moderate degree. Independent confirmation of the accuracy of the dependent variable for patients with other payers was not feasible as there is no central repository of hospitalization data for such residents. Overall, the relatively high level of agreement between MDS discharge coding and claims supports the validity of the measure.

Performance Measure Score- Correlation with SNF Industry Measures of Quality:

As hypothesized the PointRight Pro Long Stay Hospitalization Measure was correlated with other measures of quality. This supports using the Pro Long Stay Adjusted rates as a measure of a SNF's quality of care.

Table 9 shows the relationship between specific long-stay quality measures and the long-stay hospitalization measure. Highlighted correlation coefficients are statistically significant at p <.05. The negative correlation with depression may, given the modest p-value, be spurious, or if genuine may suggest lower hospitalization rates in facilities that do a better job of *recognizing* the depressive symptoms known to have a higher prevalence in SNFs in rigorous epidemiologic studies than they do in the published MDS-based CMS Quality Measures.

As Tables 10 through 14 above shows, higher star ratings are associated with lower adjusted long-stay hospitalization rates, and the relationship has become stronger over time, as shown by the larger correlation coefficients in 2020-Q3 compared to the other time periods.

ProLongStay Adjusted Hospitalization rates had a statistically significant positive relationship with the Medicare FFS claims-based long-stay hospitalization measure used in five-star and reported on Care Compare. The correlation coefficient was 0.770 ($p<0.001$). In Figure 2 above the two measures were plotted against one another for 2020-Q3. We were expecting to see a strong correlation between the two measures, although it is still possible for a facility to perform well on one of the measures and poorly on the other.

Validity of MDS Assessment

The measure developers did not specifically conduct analysis of the validity of the MDS items upon which the long-stay hospitalization measure is based. The overall reliability and validity of MDS 3.0 assessment was shown to be satisfactory prior to its adoption by CMS. RAND Corporation, the developer of MDS 3.0 as a contractor to CMS published the results of its reliability and validity testing in 2008 (Saliba and Buchanan, 2008). Details, and confirmatory studies, then appeared in peer-reviewed articles, some of which are referenced below.

Saliba , D., & Buchanan, J. (2008). Development & validation of a revised nursing facility assessment tool: MDS 3.0. RAND Health Corporation.

Saliba, D. & Buchanan, J. (2012). Making the investment count: Revision of the minimum data set for nursing facility's, MDS 3.0. J Am Med Dir Assoc. 13(7), 602-610.

Saliba, D., Buchanan, J., Eldelen, M.O., Streim, J., Ouslander, J., Berlowitz, D., & Chodosh, J. (2012). MDS 3.0: Brief interview for mental status. J Am Med Dir Assoc. 13(7), 611-617.

Saliba, D., DeFilippo, S., Edelen, M.O., Kroenke, K., Buchanan, J., & Streim, J. (2012) Testing the PHQ-9 interview and observational versions (PHQ-9 OV) for MDS 3.0. J Am Dir Assoc. 13(7), 618-625.

Saliba, D., Jones, M., Streim, J., Ouslander, J., Berlowitz, D., & Buchanan, J. (2012) Overview of significant changes in the minimum data set for nursing facilities version 3.0. J Am Dir Assoc. 13(7), 595-601.

[Response Ends]

2b.05. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified.

Describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided in Importance to Measure and Report: Gap in Care/Disparities.

[Response Begins]

The PointRight ProLongStay Hospitalization Measure has been reported to SNF providers quarterly through AHCA's LTC Trend Tracker tool. To determine what amount of change in risk-adjusted rates will be considered meaningful from one quarter to the next we observed the distribution of changes amongst our testing sample. We started with risk adjusted rates covering the 12 month measurement period of October 1, 2013 to September 30, 2014 and observed the quarterly changes up until the 12 month measurement period of January 1, 2014 to December 31, 2014. We bucketed our sample

into deciles of change in adjusted rates and calculated the average change for each bucket. We preformed the same analysis on subsets of our sample, where we divided the sample into 3 groups based on denominator size.

- Large: Denominator > 400 patient quarters
- Medium: 400 >= Denominator > 200 patient quarters
- Small: 200 >= Denominator > 30 patient quarters

[Response Ends]

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

Examples may include number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined.

[Response Begins]

Quarterly Changes in Risk Adjusted Rates (All Facilities N=2,067)

Decile of Adjusted ProLongStay Rate Quarterly Changes	**Average Quarterly Differences in Adjusted ProLongStay**	*	*
*	**April 2014 to Mar 2014 minus July 2013 to June 2014**	**July 2013 to June 2014 minus Oct 2013 to Sept 2014**	**Oct 2013 to Sept 2014 minus Jan 2014 to Dec 2014**
1	-3.6%	-3.8%	-3.6%
2	-2.0%	-2.0%	-2.0%
3	-1.2%	-1.2%	-1.2%
4	-0.5%	-0.6%	-0.6%
5	0.0%	-0.1%	0.0%
6	0.5%	0.4%	0.5%
7	1.0%	1.0%	1.1%
8	1.6%	1.6%	1.7%
9	2.6%	2.4%	2.6%
10	5.4%	4.7%	4.4%

Quarterly Changes in Risk Adjusted Rates (All Facilities N=2,067)

Quarterly Changes in Risk Adjusted Rates (Large Facilities N=291)

Decile of Adjusted ProLongStay Rate Quarterly Changes	Average Quarterly Differences in Adjusted ProLongStay	*	*
*	April 2014 to Mar 2014 minus July 2013 to June 2014	July 2013 to June 2014 minus Oct 2013 to Sept 2014	Oct 2013 to Sept 2014 minus Jan 2014 to Dec 2014
1	-3.0%	-2.5%	-2.5%
2	-1.5%	-1.4%	-1.3%
3	-0.8%	-1.0%	-0.8%
4	-0.4%	-0.6%	-0.4%
5	0.0%	-0.3%	-0.1%
6	0.3%	0.2%	0.4%
7	0.7%	0.6%	0.7%
8	1.2%	1.0%	1.1%
9	1.9%	1.5%	1.9%
10	4.0%	3.1%	2.8%

* Cell intentionally left empty

Quarterly Changes in Risk Adjusted Rates (Large Facilities N=291)

Quarterly Changes in Risk Adjusted Rates (Medium Facilities N=1,107)

Decile of Adjusted ProLongStay Rate Quarterly Changes	Average Quarterly Differences in Adjusted ProLongStay	*	*
*	April 2014 to Mar 2014 minus July 2013 to June 2014	July 2013 to June 2014 minus Oct 2013 to Sept 2014	Oct 2013 to Sept 2014 minus Jan 2014 to Dec 2014
1	-3.4%	-3.6%	-3.4%
2	-2.0%	-2.0%	-1.9%
3	-1.2%	-1.2%	-1.2%
4	-0.6%	-0.7%	-0.6%
5	0.0%	-0.2%	-0.1%
6	0.4%	0.2%	0.3%
7	0.9%	0.7%	0.8%
8	1.5%	1.3%	1.4%

Decile of Adjusted ProLongStay Rate Quarterly Changes	Average Quarterly Differences in Adjusted ProLongStay	*	*
9	2.3%	2.0%	2.3%
10	4.5%	3.7%	3.9%

* Cell intentionally left empty

Quarterly Changes in Risk Adjusted Rates (Medium Facilities N=1,107)

Quarterly Changes in Risk Adjusted Rates (Small Facilities N=669)

Decile of Adjusted ProLongStay Rate Quarterly Changes	Average Quarterly Differences in Adjusted ProLongStay	*	*
*	April 2014 to Mar 2014 minus July 2013 to June 2014	July 2013 to June 2014 minus Oct 2013 to Sept 2014	Oct 2013 to Sept 2014 minus Jan 2014 to Dec 2014
1	-4.2%	-4.7%	-4.5%
2	-2.3%	-2.3%	-2.4%
3	-1.3%	-1.4%	-1.3%
4	-0.5%	-0.6%	-0.5%
5	0.1%	0.1%	0.1%
6	0.7%	0.8%	0.9%
7	1.3%	1.6%	1.7%
8	2.1%	2.4%	2.4%
9	3.3%	3.5%	3.5%
10	7.7%	6.9%	6.0%

* Cell intentionally left empty

Quarterly Changes in Risk Adjusted Rates (Small Facilities N=669)

[Response Ends]

2b.07. Provide your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities.

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

[Response Begins]

The distribution of changes in adjusted rates, for our full sample, was similar across all 4 quarters, where for each quarter the average change for deciles 2 through 8 was less than +/- 3%. Deciles 1 and 10 had average changes greater than +/-

3.5%. The distribution of differences was larger for facilities with smaller denominators and this indicated that recommendations of clinically meaningful difference should be dependent upon facility size.

We made the following recommendations as we attempted to identify changes in adjusted rates that would move a facility several deciles in our sample's distribution.

Large Facilities - 2%

Medium Facilities - 3%

Small Facilities - 4%

[Response Ends]

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

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

[Response Begins]

On occasion facilities fail to submit MDS assessments adhering to the MDS submission schedule specified by regulation. This can result in a patient being included in a quarterly denominator population but not having a known outcome during the quarter following the snapshot date. The known outcome rate for the measure period is the sum of the counts of patients over the four quarterly denominator populations that have known outcomes, divided by the measure period denominator.

Knowing the outcome of a patient in the quarter at risk is vital the measure's accuracy. We've reviewed the known outcome rates across our sample to ensure that missing data is not a major factor. For this analysis we used our full sample of 2,811 PointRight clients, before excluding facilities for having known outcome rates less than 90% (N=2,584). This distribution is provided below in the table below.

In addition to reviewing the missing data distribution across our sample, we also examined the relationship between the observed rate of hospitalizations and the known outcomes rate. For facilities with known outcomes rates between 100% and 90%, patient quarters at risk with unknown outcomes will contribute a 0.8 to the numerator. We impute 0.8 because nationally 80% of long stay patients with known outcomes are discharged to the hospital. If we are appropriately imputing the rate of hospitalization we would expect to see little to no correlation between known outcome rates and hospitalization rates. A scatter plot of the two rates is found in the figure below.

[Response Ends]

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

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

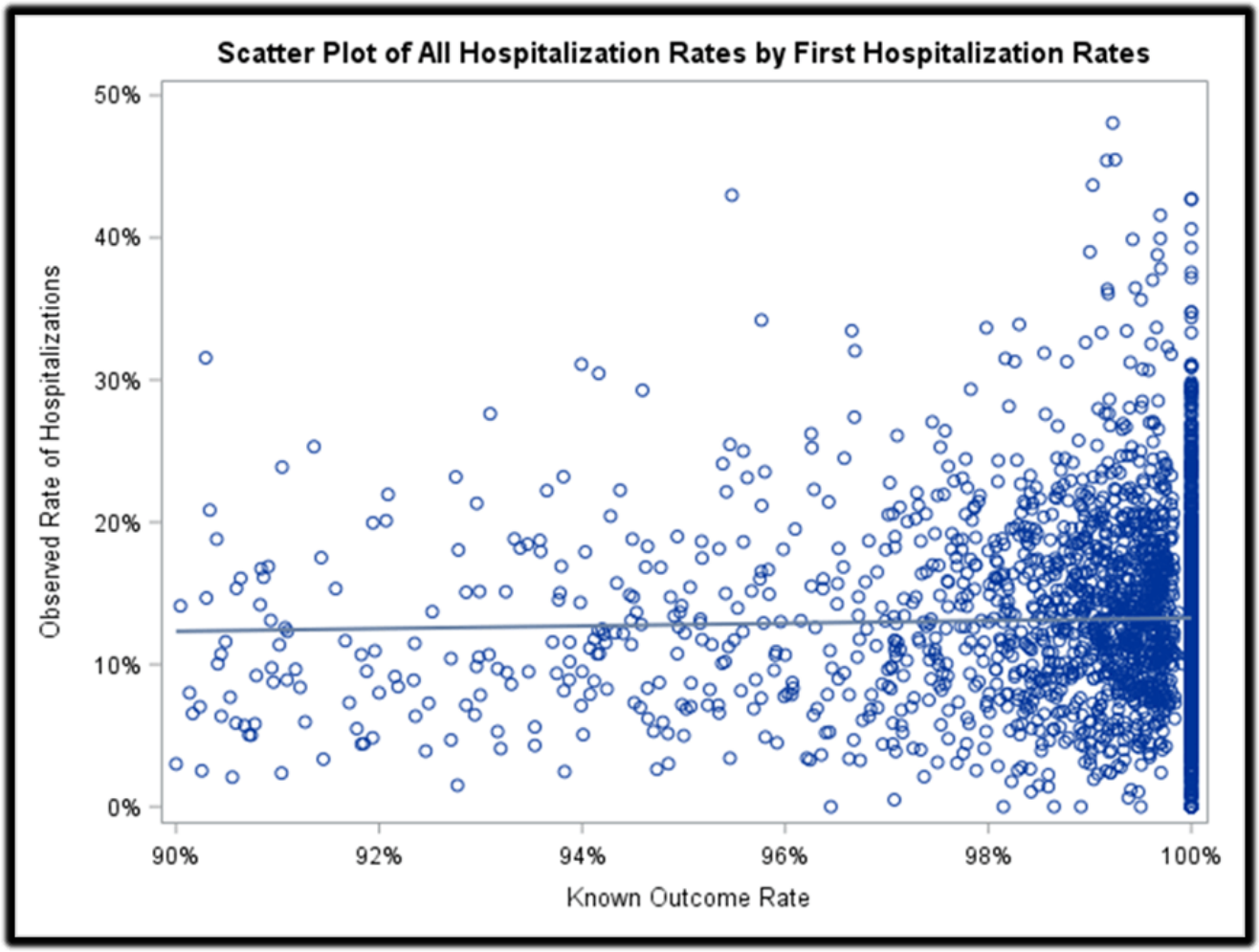
[Response Begins]

Table- Distribution of Known Outcome Rates Before Exclusion

Quantile	Estimate
100% Max	100%
99%	100%
95%	100%
90%	100%
75% Q3	100%
50% Median	100%
25% Q1	98%
10%	92%
5%	83%
1%	73%
0% Min	70%

Distribution of Known Outcome Rates Before Exclusion

Figure- Scatter Plot of Known Outcome Rates and Observed Hospitalization Rates



Scatter Plot of Known Outcome Rates (x-axis) and Observed Hospitalization Rates (y-axis)

[Response Ends]

2b.10. Provide your interpretation of the results, in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders), and how the specified handling of missing data minimizes bias.

In other words, what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis was conducted, justify the selected approach for missing data.

[Response Begins]

The median known outcome rate in our full sample of PointRight facilities was 100%, leading us to conclude missing data was not an issue for the majority of facilities. We selected a known outcome rate of 90% to the be the minimum threshold. This threshold excluded 8% of our sample and was a good balance between the availability and utility of the measure.

In examining the relationship between known outcome rates and observed hospitalization rates we do see a very slight positive correlation. The Pearson correlation coefficient was .08 with a p-value of .0001 and the Spearman correlation coefficient was -.006 with a p-value of .7311. These results lead us to believe that there is no significant bias for facilities with known outcomes rates greater than 90%.

[Response Ends]

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

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

[Response Begins]

No, there is only one set of specifications for this measure

[Response Ends]

2b.12. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications.

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

[Response Begins]

[Response Ends]

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

Examples may include correlation, and/or rank order.

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

N/A or no exclusions

[Response Ends]

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

Describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used?

[Response Begins]

There are no exclusions from the denominator; all patients in the facility on the snapshot date who meet the long stay criterion on that date are included. However, the measure will not be reported for a SNF if the denominator population over the measure period's 4 snapshot dates is less than 30.

[Response Ends]

2b.17. Provide the statistical results from testing exclusions.

Include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores.

[Response Begins]

The table below shows that the PointRight ProLongStay Hospitalization Rates are steady across time with the exception of SNFs having small denominators. The 34 facilities with denominators less than 30 experienced changes in rates of greater than 35 percentage points. Ultimately, excluding SNFs with denominators less than 30 resulted in excluding less than 1.5% of our sample.

Table- Change in Adjusted Rates from Quarter to Quarter by Denominator Size

Denominator Size	# of Facilities, CY 2014	Variable	Mean	Std Deviation	5th Pctl	95th Pctl
Denominator <30	34	Adjusted Rate CY 2014	12.1%	26.1%	0.0%	35.8%
*	*	Change in Adj Rates July'13-Jun'14 to Apr'13-Mar'14	4.0%	13.6%	-8.9%	12.9%
*	*	Change in Adj Rates Oct'13-Sep'14 to Jul'13-Jun'14	1.5%	6.7%	-7.0%	16.6%
*	*	Change in Adj Rates Jan'14-Dec'14 to Oct'13-Sep'14	2.5%	11.5%	-5.3%	25.7%
30<= Denominator <100	106	Adjusted Rate CY 2014	13.4%	6.9%	3.2%	25.0%
*	*	Change in Adj Rates July'13-Jun'14 to Apr'13-Mar'14	1.3%	4.0%	-4.9%	8.2%
*	*	Change in Adj Rates Oct'13-Sep'14 to Jul'13-Jun'14	1.3%	5.1%	-6.3%	10.0%
*	*	Change in Adj Rates Jan'14-Dec'14 to Oct'13-Sep'14	1.4%	4.4%	-4.5%	10.4%
100<= Denominator < 300	1258	Adjusted Rate CY 2014	13.9%	5.0%	6.4%	22.6%
*	*	Change in Adj Rates July'13-Jun'14 to Apr'13-Mar'14	0.3%	2.4%	-3.5%	4.2%
*	*	Change in Adj Rates Oct'13-Sep'14 to Jul'13-Jun'14	0.2%	2.5%	-3.7%	4.4%
*	*	Change in Adj Rates Jan'14-Dec'14 to Oct'13-Sep'14	0.3%	2.7%	-3.7%	5.0%
300<= Denominator < 450	506	Adjusted Rate CY 2014	14.4%	4.5%	7.7%	22.8%
*	*	Change in Adj Rates July'13-Jun'14 to Apr'13-Mar'14	0.2%	1.9%	-2.8%	3.4%
*	*	Change in Adj Rates	0.2%	1.8%	-2.7%	3.2%

Denominator Size	# of Facilities, CY 2014	Variable	Mean	Std Deviation	5th Pctl	95th Pctl
		Oct'13-Sep'14 to Jul'13-Jun'14				
*	*	Change in Adj Rates Jan'14-Dec'14 to Oct'13-Sep'14	0.4%	2.0%	-2.7%	3.8%
450 <= Denominator	197	Adjusted Rate CY 2014	13.5%	4.1%	6.2%	19.3%
*	*	Change in Adj Rates July'13-Jun'14 to Apr'13-Mar'14	0.1%	1.4%	-2.3%	2.4%
*	*	Change in Adj Rates Oct'13-Sep'14 to Jul'13-Jun'14	-0.1%	1.3%	-2.4%	2.2%
*	*	Change in Adj Rates Jan'14-Dec'14 to Oct'13-Sep'14	0.1%	1.6%	-2.6%	3.0%

* Cell intentionally left empty

Change in Adjusted Rates from Quarter to Quarter by Denominator Size

[Response Ends]

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

In other words, the value outweighs the burden of increased data collection and analysis. Note: If patient preference is an exclusion, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion.

[Response Begins]

The results show that the average change of Adjusted ProLongStay Hospitalization Rates from quarter to quarter significantly drops once a facility has a denominator greater than or equal to 30. For this reason we felt rates are unstable for SNFs with denominators less than 30 and feel validated in excluding these SNFs. This exclusion only resulted in 34 (1.5%) of facilities being excluded.

[Response Ends]

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

[Response Begins]

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

25 risk factors

[Response Ends]

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

[Response Begins]

The formula for a facility's adjusted PointRight ProLong Hospitalization Rate is: [Observed rate of all hospitalizations]/[Expected rate of all hospitalizations]*[National average rate of all hospitalizations].

Procedure for calculating the adjusted rate

1) Calculate the observed rate.

The observed PointRight Pro Long Stay Hospitalization Rate is the sum of the four quarterly numerators divided by the sum of the four quarterly denominators.

• The denominator for a quarter is the number of residents present in the facility on the first day of a calendar quarter who qualify as long stay on that day

• The numerator for a quarter is number of hospitalizations of residents in the denominator population for that quarter, where hospitalization means discharge from the SNF directly to an acute care hospital, either with no return to the SNF or with return to the SNF after at least one midnight outside the SNF.

2) Calculate the expected rate.

• Calculate the expected number of first hospitalizations of the quarterly denominator population for each of the four quarters in the measure period and sum them; multiply the sum by 1.25248 to obtain the expected number of total hospitalizations for the 12-month measurement period. Divide this number by the sum of the quarterly denominators to get the expected rate for the measure period.

The risk factors and coefficients are found in Table 17 below

3) Calculate the national benchmark rate

• The national benchmark rate is the observed PointRight Pro Long Stay Hospitalization Rate for a denominator population consisting of the denominator populations for all SNFs in the largest available national sample that have complete non-discharge MDS data for all of their patients for all four quarters in the measurement period and have 100% known outcomes for all patients in their denominator populations for all four quarters in the measure period.

For a given member of a quarterly denominator population a known outcome means either that the patient had a discharge MDS submitted with a discharge date within the quarter and a discharge destination filled in, that the patient was readmitted from an acute care hospital during the quarter, or that the patient had a quarterly or other MDS submitted in the 100 days following the end of the quarter that gave an admission date prior to the snapshot date for the given quarter.

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

Published literature

Internal data analysis

[Response Ends]

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

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

[Response Begins]

The risk adjustment model employed in the PointRight Pro Long Stay Hospitalization Rate utilizes four logistic regression models applied to four discrete subgroups of the denominator population to estimate risk of any hospitalization during a quarter at risk. This risk estimate is multiplied by a fixed factor to estimate the expected number of total hospitalization during the quarter.

Calculation of a patient's risk of any hospitalization (or equivalently, the risk of a first hospitalization) during a quarter at risk begins by assigning the patient to one of four subgroups of the denominator population based on the duration of the patient's current stay in the SNF as of the snapshot date. For each group the risk of one or more discharges from the SNF directly to an acute care hospital during the quarter is estimated by a logistic regression. The independent variables in each logistic regression model come from the patient's most recent MDS 3.0 assessment prior to the snapshot date that has the variable; most of the independent variables are common to the four models.

Our 4 logistic regression models will estimate a patient's risk of one or more hospitalizations. In order to estimate the patient's expected number of hospitalizations we multiply the risk estimate by a fixed factor. The factor was determined by modeling the relationships between the rate of first hospitalizations to the rate of all hospitalizations for a sample of 1,116 SNFs that had 100% complete reporting of their outcomes.

The decision to first segment the denominator population and then estimate logistic regression models was based on findings that the continuous length of stay at the beginning of a period at risk was the single most powerful predictor of hospitalization risk. (See Table 16). Estimating risk within strata defined by continuous length of stay gave greater predictive power than estimating a similar model without stratification, in part because risk factors operate somewhat differently within different LOS strata. The decision to estimate the risk of any hospitalization and then multiply it to get the expected rate of total hospitalizations was based on two considerations: (1) Models of binary dependent variables are much more widely known by clinical end-users of performance measures than models of dependent variables that can take multiple discrete values; this makes them more accessible and useful; (2) The relationship between SNFs' rates for first hospitalizations and their rates of all hospitalizations for their long-stay residents is extremely tight, with a linear equation accounting for 99% of the variance. Thus, modeling the risk of first hospitalizations is a rational approach to modeling of risk of all hospitalizations.

The selection of risk factors (independent variables) involved an iterative process. A panel of clinicians with extensive SNF experience recommended potential risk adjusters. These, and a full set of sociodemographic and contextual factors were tested for univariate relationships with hospitalizations. The variables with the strongest univariate correlations were then used to build multivariate models. The multivariate models (logistic regressions for each stratum of LOS) were reviewed by a larger panel of clinicians and potential users of the measure. Variables were rejected and replaced if their coefficients were opposite to their univariate correlation with the hospitalization, or if they were viewed as potentially under the control of the SNF – i.e., creating a risk of over-adjustment.

[Response Ends]

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

[Response Begins]

The initial stratification of the denominator population reflects the following analysis of the relationship of LOS on the snapshot date to the risk of a first hospitalization:

Table 16. Average Hospitalization Rates of Any (First) Hospitalization by Long Stay Groups

Long Stay Group	Continuous days form most recent admission to snapshot date	Denominator Population July 1, 2013 to June 30, 2014 (% of total sample)	Observed Rate of First Long Stay Hospitalization (patient level) July 1, 2013 to June 30, 2014	Denominator Population July 1, 2014 to June 30, 2015 (% of total sample)	Observed Rate of First Long Stay Hospitalization (patient level) July 1, 2014 to June 30, 2015
1	current LOS ≤ 100 days but cumulative days in SNF >100 days	68,222 (9.6%)	26.5%	58,805 (10.2%)	26.9%
2	100 days < LOS ≤ 181 days	100,682 (14.1%)	16.2%	78,058 (13.5%)	17.2%
3	181 days < LOS ≤ 364 days	144,749 (20.6%)	11.7%	116,052 (20.1%)	12.8%
4	LOS > 364 days	399,614 (55.7%)	7.1%	323,445 (56.1%)	7.7%

Average Hospitalization Rates of Any (First) Hospitalization by Long Stay Groups

The models applicable to each of the subgroups of the denominator population are displayed in table 17. The table also shows the prevalence of the IV in population used to estimate the models, and indicates the model C-statistic.

Table 17. PointRight ProLongStay Hospitalization Measure Logistic Regression Models

Type of Independent Variable	Independent Variable	Long Stay Group 1 (Current stay LOS <100 days but cumulative days in SNF >100) C-statistic = 0.63	Long Stay Group 1 Prevalence of Independent Variable	Long Stay Group 2 (100 days < LOS <=181 days) C-statistic = 0.63	Long Stay Group 2 Prevalence of Independent Variable	Long Stay Group 3 (181 days < LOS <=364 days) C-statistic = 0.62	Long Stay Group 3 Prevalence of Independent Variable	Long Stay Group 4 (LOS > 364 days) C-statistic = 0.63	Long Stay Group 4 Prevalence of Independent Variable
Constant	Intercept	-1.37	X	-1.99	X	-1.19	X	-1.09	X
Active Diagnoses	Anemia	0.13	37.3%	0.20	29.2%	0.21	27.8%	0.12	26.7%
*	Asthma, COPD, or Chronic Lung Disease on Oxygen	0.14	27.6%	0.20	21.5%	0.20	19.7%	0.17	16.1%
*	Diabetes on Insulin	x	x	x	x	0.27	6.3%	x	x
*	Gastroesophageal Reflux Diseas (GERD) or ulcer	x	x	x	x	x	x	0.12	22.7%
*	Heart Failure	0.14	26.7%	0.21	20.1%	0.27	18.7%	0.18	16.2%
*	Hypertension	x	x	x	x	x	x	0.23	68.6%
*	Viral Hepatitis	0.30	0.9%	x	x	x	x	x	x
*	Neurogenic bladder	0.20	4.5%	x	x	0.43	2.6%	0.27	2.3%
*	Renal failure, insufficiency, ESRD	x	x	0.18	9.3%	0.17	7.6%	0.10	6.2%
Incontinence	Total Bowel Incontinence	x	x	0.21	27.7%	0.22	29.0%	0.13	35.0%
Demographics	Age 90 or over	-0.23	16.1%	-0.13	15.9%	-0.12	16.8%	-0.08	14.7%

Type of Independent Variable	Independent Variable	Long Stay Group 1 (Current stay LOS <100 days but cumulative days in SNF >100) C-statistic = 0.63	Long Stay Group 1 Prevalence of Independent Variable	Long Stay Group 2 (100 days < LOS <=181 days) C-statistic = 0.63	Long Stay Group 2 Prevalence of Independent Variable	Long Stay Group 3 (181 days < LOS <=364 days) C-statistic = 0.62	Long Stay Group 3 Prevalence of Independent Variable	Long Stay Group 4 (LOS > 364 days) C-statistic = 0.63	Long Stay Group 4 Prevalence of Independent Variable
*	Male	0.14	37.8%	0.25	34.1%	0.21	31.5%	0.26	26.2%
Medications Received	Anticoagulan within 7 days prior to ARD	0.16	23.7%	x	x	x	x	0.06	18.3
*	Antibiotics within 7 days prior to ARD	x	x	0.27	1.2%	x	x	x	x
Stay History	Admitted from hospital (current stay)	0.40	96.4%	0.39	83.6%	0.26	82.4%	0.26	81.8%
*	In this SNF 6 months before snapshot date (any stay)	-0.33	54.4%	x	x	x	x	x	x
*	In this SNF 12 months before snapshot date (any stay)	-0.39	31.5%	x	x	x	x	x	x
*	Natural log of (length of current stay minus 100)*prevalence is LOS	x	x	-0.12	x	-0.30	x	-0.35	x
Symptoms	Dyspnea on exertion	0.18	13.8%	x	x	x	x	0.19	5.0%

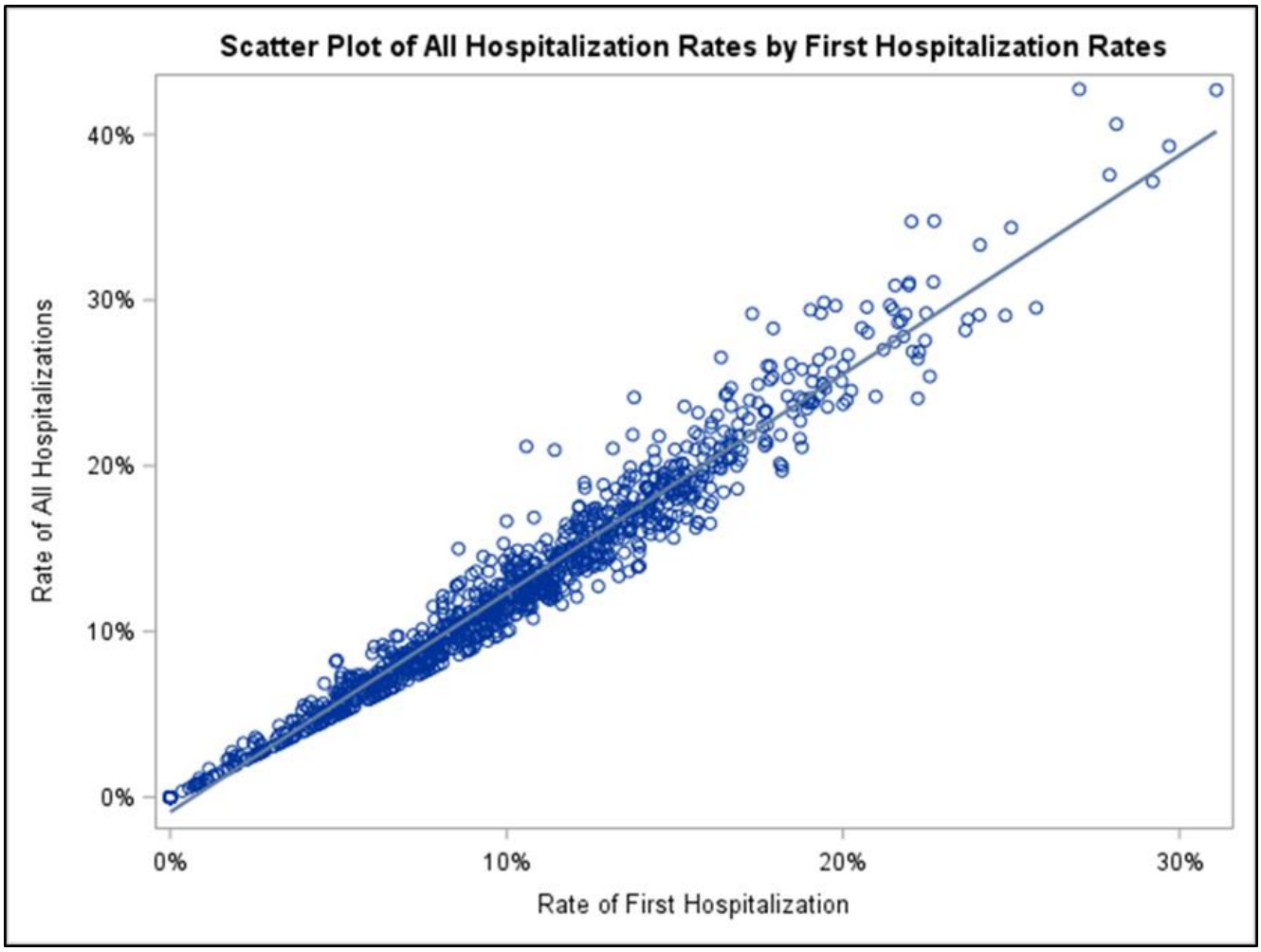
Type of Independent Variable	Independent Variable	Long Stay Group 1 (Current stay LOS <100 days but cumulative days in SNF >100) C-statistic = 0.63	Long Stay Group 1 Prevalence of Independent Variable	Long Stay Group 2 (100 days < LOS <=181 days) C-statistic = 0.63	Long Stay Group 2 Prevalence of Independent Variable	Long Stay Group 3 (181 days < LOS <=364 days) C-statistic = 0.62	Long Stay Group 3 Prevalence of Independent Variable	Long Stay Group 4 (LOS > 364 days) C-statistic = 0.63	Long Stay Group 4 Prevalence of Independent Variable
Skin	Surgical wound(s)	x	x	x	x	0.38	1.1%	0.38	0.5%
Hospice Status	Receiving hospice care	-1.31	4.7%	-1.16	5.2%	-1.20	5.2%	-1.04	4.6%
Recent Treatments	IV fluid or meds within 7 days before last MDS	0.20	5.4%	0.56	1.5%	0.38	0.9%	0.40	0.5%
*	Oxygen in 7 days before last MDS	x	x	0.38	13.8%	x	x	0.22	7.3
Socioeconomic Status	Medicaid	x	x	x	x	0.09	71.1%	0.11	79.7%

* Cell intentionally left empty

PointRight ProLongStay Hospitalization Measure Logistic Regression Models

The scatter plot below displays facility level linear regression model.

Figure 3. Scatter Plot of All Hospitalization Rates by First Hospitalization Rates



Scatter Plot of All Hospitalization Rates by First Hospitalization Rates

[Response Ends]

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

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

[Response Begins]

Our overall approach was to begin with reliable and rarely-missing patient-level SDS variables nominated by our clinical experts: Medicaid status (as a proxy for financial assets and income), black versus non-black, Hispanic/Latino versus non-Hispanic/Latino, and the interactions of Medicaid status and race. The significance of these variables in predicting hospitalization rates was tested in fixed-effects logistic regression models. We reasoned that patient-level effects that were significant in models that included facility-specific constant terms probably reflected otherwise-unmeasured differences in baseline health status. Our final risk adjustment models were single-level logistic regression models in which the coefficients on the SDS variables were forced to be the same as in the fixed-effects model. Essentially this approach adjusts for the *within-facility* differences in long stay hospitalization rates associated with the SDS factor, but does not adjust for the *between-facility* differences in long stay hospitalization rates associated with the SDS factor. The within-facility effects are essentially those beyond those associated with facility quality differences. In all cases this made

the effect of the SDS variables smaller than it would be in a single-level logistic regression that did not account for facility effects. We did not want to adjust away facility-level effects related to worse care at SNFs with large minority populations.

Both black and Medicaid status were found to be significant in at least one of the fixed effects models on our four long stay strata. Despite being significant, black and Medicaid status had minimal impact on the overall performance of the models as measured by the c-statistic. Inclusion of one or both SDS variables did not impact the overall c-statistic of the models by more than .001.

During a review with a panel of the All-Cause Admissions and Readmissions Standing Committee, the panel recommended exclusion of black from the final model. The impact of this change on the overall metrics was minimal as illustrated in table 18. Here we compare our CY 2014 sample of facilities using risk adjustment with and without black as a risk adjustor. No facilities in our sample of 2,180 SNFs moved more than 1 decile between the two models and only 5% of the sample moved 1 decile.

Table 18 Decile Ranking of Pro Long Stay Adjusted Rates with and without Black as a Covariate

Model without Black as Covariate	**Model with Black as Covariate**	*	*	*	*	*	*	*	*	*	**Total**
***	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*
0	215	3	0	0	0	0	0	0	0	0	218
1	3	209	6	0	0	0	0	0	0	0	218
2	0	6	203	9	0	0	0	0	0	0	218
3	0	0	9	202	7	0	0	0	0	0	218
4	0	0	0	7	202	9	0	0	0	0	218
5	0	0	0	0	9	201	8	0	0	0	218
6	0	0	0	0	0	8	202	8	0	0	218
7	0	0	0	0	0	0	8	206	4	0	218
8	0	0	0	0	0	0	0	4	209	5	218
9	0	0	0	0	0	0	0	0	5	213	218
Total	218	218	218	218	218	218	218	218	218	218	2180

* Cell intentionally left empty

Decile Ranking of Pro Long Stay Adjusted Rates with and without Black as a Covariate

[Response Ends]

2b.26. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (describe the steps—do not just name a method; what statistical analysis was used). Provide the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter "N/A" for questions about the statistical risk model discrimination and calibration statistics.

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

[Response Begins]

We compared our model coefficients to the mean coefficients from bootstrap analysis, expressed as actual values, standard deviation (S.D.) and percentage. Our sampling distribution consisted of 100 draws with replacement from our modeling data set.

The majority of covariates have less than 5% variation from the bootstrap mean, making the absolute value and/or the number of standard deviations clinically acceptable.

Table 19. Bootstrap Analysis Long Stay Group 1 current LOS <= 100 days but cumulative days in SNF >100 days

Variable Type	Independent Variable	Long Stay Group 1 Coefficient	Bootstrap Mean	Difference	Difference in %
Constant	Intercept	-1.367	-1.370	0.003	-0.20%
Active Diagnoses	Anemia	0.128	0.129	0.000	-0.33%
Active Diagnoses	Asthma, COPD, or Chronic Lung Disease whether or not on oxygen	0.148	0.149	-0.001	-0.51%
Active Diagnoses	Diabetes on insulin	0.202	0.202	0.000	-0.02%
Active Diagnoses	Heart Failure	0.137	0.136	0.001	0.86%
Active Diagnoses	Viral Hepatitis	0.293	0.308	-0.014	-4.85%
Active Diagnoses	Neurogenic bladder	0.204	0.206	-0.002	-1.06%
Demographics	Age 90 or over	-0.235	-0.240	0.005	-2.21%
Demographics	Male	0.139	0.137	0.002	1.32%
Medications Received	Anticoagulant within 7 days prior to ARD	0.159	0.161	-0.002	-1.11%
Stay History	Admitted from hospital (current stay)	0.397	0.401	-0.004	-1.06%
Stay History	In this SNF 6 months before snapshot date (any stay)	-0.385	-0.385	0.000	0.08%
Stay History	In this SNF 12 months before snapshot date (any stay)	-0.330	-0.330	0.000	0.00%
Symptoms	Dyspnea on exertion	0.169	0.171	-0.002	-0.95%
Hospice Status	Receiving hospice care	-1.353	-1.357	0.005	-0.35%
Recent Treatments	IV fluid or meds within 7 days before last MDS	0.200	0.201	-0.002	-0.77%

Bootstrap Analysis Long Stay Group 1 current LOS <= 100 days but cumulative days in SNF >100 days

Table 20. Bootstrap Analysis Long Stay Group 2 100 days < LOS <= 181 days

Variable Type	Independent Variable	Long Stay Group 2 Coefficient	Bootstrap Mean	Difference	Difference in %
Constant	Intercept	-2.012	-2.006	-0.006	0.29%
Active Diagnoses	Anemia	0.209	0.205	0.004	1.96%
Active Diagnoses	Asthma, COPD, or Chronic Lung Disease whether or not on oxygen	0.204	0.198	0.005	2.70%
Active Diagnoses	Diabetes on insulin	0.304	0.303	0.002	0.55%
Active Diagnoses	Heart Failure	0.211	0.211	0.000	0.14%
Active Diagnoses	Renal failure, insufficiency, or ESRD	0.187	0.187	0.000	0.01%
Incontinence	Total bowel incontinence	0.207	0.209	-0.002	-1.02%
Demographics	Age 90 or over	-0.131	-0.132	0.001	-0.76%
Demographics	Male	0.251	0.249	0.002	0.61%
Medications Received	Antibiotics within 7 days prior to ARD	0.290	0.280	0.010	3.35%
Stay History	Admitted from hospital (current stay)	0.400	0.399	0.001	0.18%
Stay History	Natural log of (Length of current stay minus 100) *prevalence is of LOS	-0.121	-0.121	0.001	-0.58%
Hospice Status	Receiving hospice care	-1.175	-1.174	-0.001	0.08%
Recent Treatments	IV fluid or meds within 7 days before last MDS	0.548	0.548	0.000	-0.06%
Recent Treatments	Oxygen in 7 days before last MDS	0.383	0.386	-0.003	-0.80%

Bootstrap Analysis Long Stay Group 2 100 days < LOS <= 181 days

Table 21. Bootstrap Analysis Long Stay Group 3 181 days < LOS <= 364 days

Variable Type	Independent Variable	Long Stay Group 3 Coefficient	Bootstrap Mean	Difference	Difference in %
Constant	Intercept	-1.193	-1.209	0.016	-1.36%
Active Diagnoses	Anemia	0.215	0.215	0.000	-0.09%
Active Diagnoses	Asthma, COPD, or Chronic Lung Disease whether or not on oxygen	0.204	0.201	0.003	1.33%
Active Diagnoses	Asthma, COPD, or Chronic Lung Disease on oxygen	0.279	0.281	-0.002	-0.76%
Active Diagnoses	Diabetes on insulin	0.346	0.347	-0.001	-0.30%

Variable Type	Independent Variable	Long Stay Group 3 Coefficient	Bootstrap Mean	Difference	Difference in %
Active Diagnoses	Heart Failure	0.271	0.270	0.001	0.38%
Active Diagnoses	Neurogenic bladder	0.434	0.436	-0.002	-0.49%
Active Diagnoses	Renal failure, insufficiency, or ESRD	0.172	0.171	0.001	0.45%
Incontinence	Total bowel incontinence	0.215	0.213	0.002	0.92%
Demographics	Age 90 or over	-0.126	-0.124	-0.002	1.91%
Demographics	Male	0.212	0.209	0.003	1.51%
Stay History	Admitted from hospital (current stay)	0.261	0.261	0.000	0.14%
Stay History	Natural log of (Length of current stay minus 100) *prevalence is of LOS	-0.296	-0.293	-0.004	1.19%
Skin	Surgical wound(s)	0.375	0.380	-0.005	-1.35%
Hospice Status	Receiving hospice care	-1.203	-1.207	0.004	-0.32%
Recent Treatments	IV fluid or meds within 7 days before last MDS	0.386	0.387	-0.001	-0.26%

Bootstrap Analysis Long Stay Group 3 181 days < LOS <= 364 days

Table 22. Bootstrap Analysis Long Stay Group 4 LOS > 364 days

Variable Type	Independent Variable	Long Stay Group 4 Coefficient	Bootstrap Mean	Difference	Difference in %
Constant	Intercept	-1.140	-1.105	-0.035	3.10%
Active Diagnoses	Anemia	0.117	0.121	-0.004	-3.48%
Active Diagnoses	Asthma, COPD, or Chronic Lung Disease whether or not on oxygen	0.163	0.172	-0.009	-5.23%
Active Diagnoses	Gastroesophageal Reflux Disease (GERD) or ulcer	0.117	0.122	-0.006	-4.89%
Active Diagnoses	Diabetes on insulin	0.289	0.298	-0.009	-2.97%
Active Diagnoses	Heart Failure	0.176	0.175	0.000	0.10%
Active Diagnoses	Hypertension	0.214	0.233	-0.019	-8.68%
Active Diagnoses	Neurogenic bladder	0.267	0.268	-0.001	-0.52%
Active Diagnoses	Renal failure, insufficiency, or ESRD	0.101	0.107	-0.005	-5.30%
Incontinence	Total bowel incontinence	0.121	0.129	-0.008	-6.47%

Variable Type	Independent Variable	Long Stay Group 4 Coefficient	Bootstrap Mean	Difference	Difference in %
Demographics	Age 90 or over	-0.066	-0.073	0.007	-10.57%
Demographics	Male	0.260	0.262	-0.002	-0.76%
Medications Received	Anticoagulant within 7 days prior to ARD	0.096	0.069	0.026	27.62%
Stay History	Admitted from hospital (current stay)	0.253	0.258	-0.005	-1.91%
Stay History	Natural log of (Length of current stay minus 100) *prevalence is of LOS	-0.350	-0.343	-0.007	2.06%
Symptoms	Dyspnea on exertion	0.289	0.298	-0.009	-2.97%
Skin	Surgical wound(s)	0.379	0.366	0.012	3.28%
Hospice Status	Receiving hospice care	-1.044	-1.050	0.006	-0.61%
Recent Treatments	IV fluid or meds within 7 days before last MDS	0.391	0.398	-0.007	-1.73%
Recent Treatments	Oxygen in 7 days before last MDS	0.218	0.221	-0.003	-1.35%

Bootstrap Analysis Long Stay Group 4 LOS > 364 days

[Response Ends]

2b.27. Provide risk model discrimination statistics.

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

[Response Begins]

Provided below are the c-statistics and r-squared values for the 5 models in the PointRight Pro Long Stay Hospitalization Measure.

Logistic Regression Model Long Stay Group 1 c-statistic = .63

Logistic Regression Model Long Stay Group 2, c-statistic = .63

Logistic Regression Model Long Stay Group 3, c-statistic = .62

Logistic Regression Model Long Stay Group 4, c-statistic = .63

Linear Regression Model Rate of all Hospitalizations, R-squared = .99

[Response Ends]

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

[Response Begins]

Table 23. Hosmer-Lemeshow Statistic Long Stay Group 1 (current LOS <= 100 days but cumulative days in SNF >100 days)

Partition for the Hosmer and Lemeshow Test	*	*	*
Group	**Total**	**First Hospitalization Occur = Yes**	*
*	*	**Observed**	**Expected**
1	6056	709	672
2	7038	1126	1174
3	7414	1455	1442
4	6824	1485	1517
5	6907	1727	1729
6	6843	1896	1912
7	6636	2046	2021
8	7087	2431	2351
9	6828	2490	2489
10	6589	2751	2809

* Cell intentionally left empty

Hosmer-Lemeshow Statistic Long Stay Group 1 (current LOS <= 100 days but cumulative days in SNF >100 days)

Table 24. Homer-Lemeshow Statistic Long Stay Group 2 (100 days < LOS <= 181 days)

Partition for the Hosmer and Lemeshow Test	*	*	*
Group	**Total**	**First Hospitalization Occur = Yes**	*
*	*	**Observed**	**Expected**
1	10068	709	706
2	10096	860	1044
3	10068	1041	1171
4	10067	1334	1311
5	10070	1549	1436
6	10077	1690	1594
7	10067	1863	1756
8	10067	2023	1972
9	10066	2328	2278
10	10036	2881	3010

* Cell intentionally left empty

Homer-Lemeshow Statistic Long Stay Group 2 (100 days < LOS <= 181 days)

Table 25. Homer-Lemeshow Statistic Long Stay Group 3 (181 days < LOS <= 364 days)

Partition for the Hosmer and Lemeshow Test	*	*	*
Group	**Total**	**First Hospitalization Occur = Yes**	*
*	*	**Observed**	**Expected**
1	14769	732	761
2	14742	1020	1106
3	14775	1175	1246
4	14770	1335	1381
5	14775	1556	1516
6	14776	1728	1679
7	14773	1951	1853
8	14775	2149	2087
9	14775	2451	2424
10	14819	3236	3279

* Cell intentionally left empty

Homer-Lemeshow Statistic Long Stay Group 3 (181 days < LOS <= 364 days)

Table 26. Hosmer-Lemeshow Statistic Long Stay Group 4 (LOS > 364 days)

Partition for the Hosmer and Lemeshow Test	*	*	*
Group	**Total**	**First Hospitalization Occur = Yes**	*
*	*	**Observed**	**Expected**
1	39961	1085	1134
2	39961	1557	1645
3	39961	1969	1945
4	39963	2135	2211
5	39977	2475	2469
6	39965	2830	2758
7	39966	3113	3094
8	39962	3656	3506
9	39961	4155	4099
10	39937	5483	5596

* Cell intentionally left empty

Hosmer-Lemeshow Statistic Long Stay Group 4 (LOS > 364 days)

Table 27. Linear Regression of All Hospitalizations by Decile of Expected Rates

Partition for the Hosmer and Lemeshow Test	*	*
Group	**Observed Rate of Total Hospitalizations**	**Expected Rate of Total Hospitalizations**

Partition for the Hosmer and Lemeshow Test	*	*
1	3.9%	8.4%
2	8.0%	11.0%
3	9.7%	12.3%
4	10.9%	13.1%
5	12.1%	13.8%
6	13.3%	14.4%
7	15.3%	15.1%
8	16.3%	15.9%
9	18.8%	16.8%
10	23.3%	18.9%

- * Cell intentionally left empty

Linear Regression of All Hospitalizations by Decile of Expected Rates

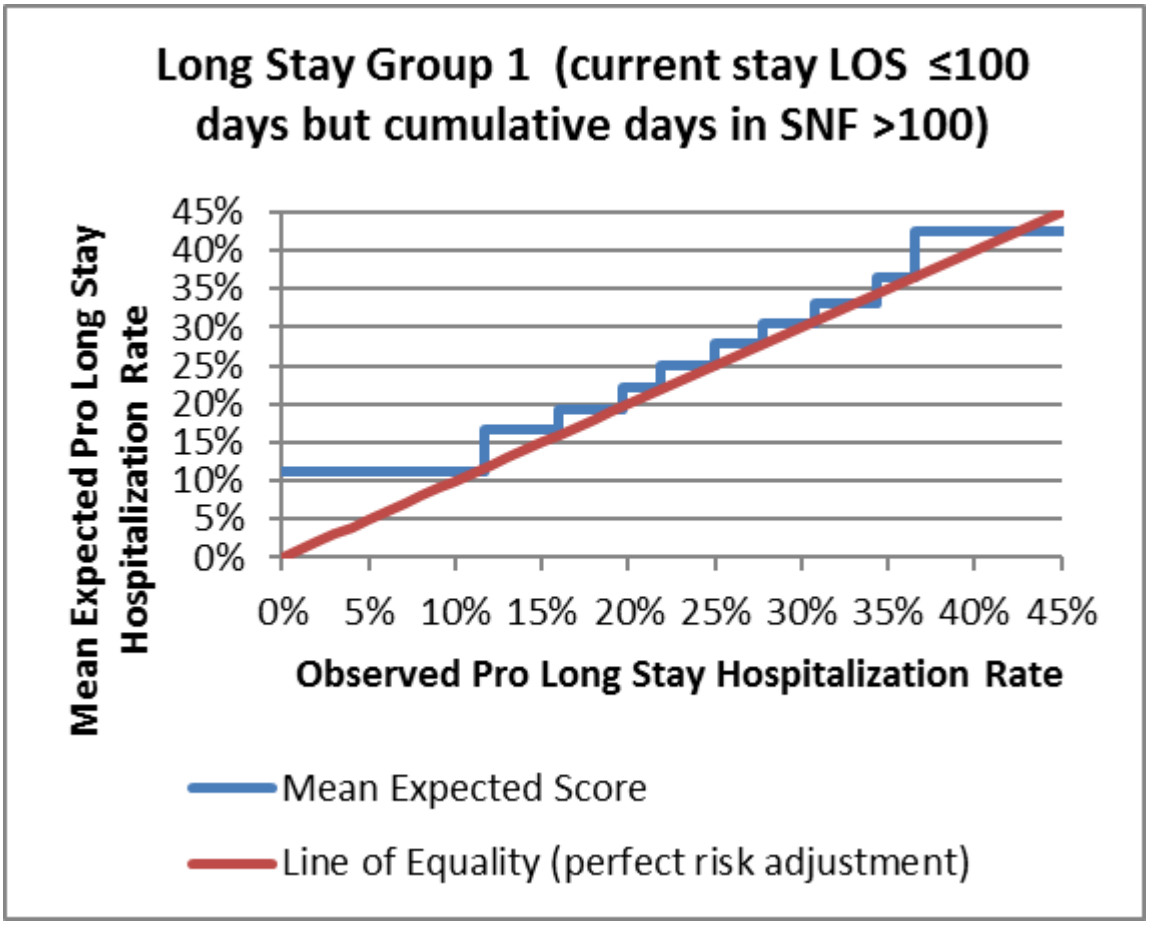
[Response Ends]

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

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

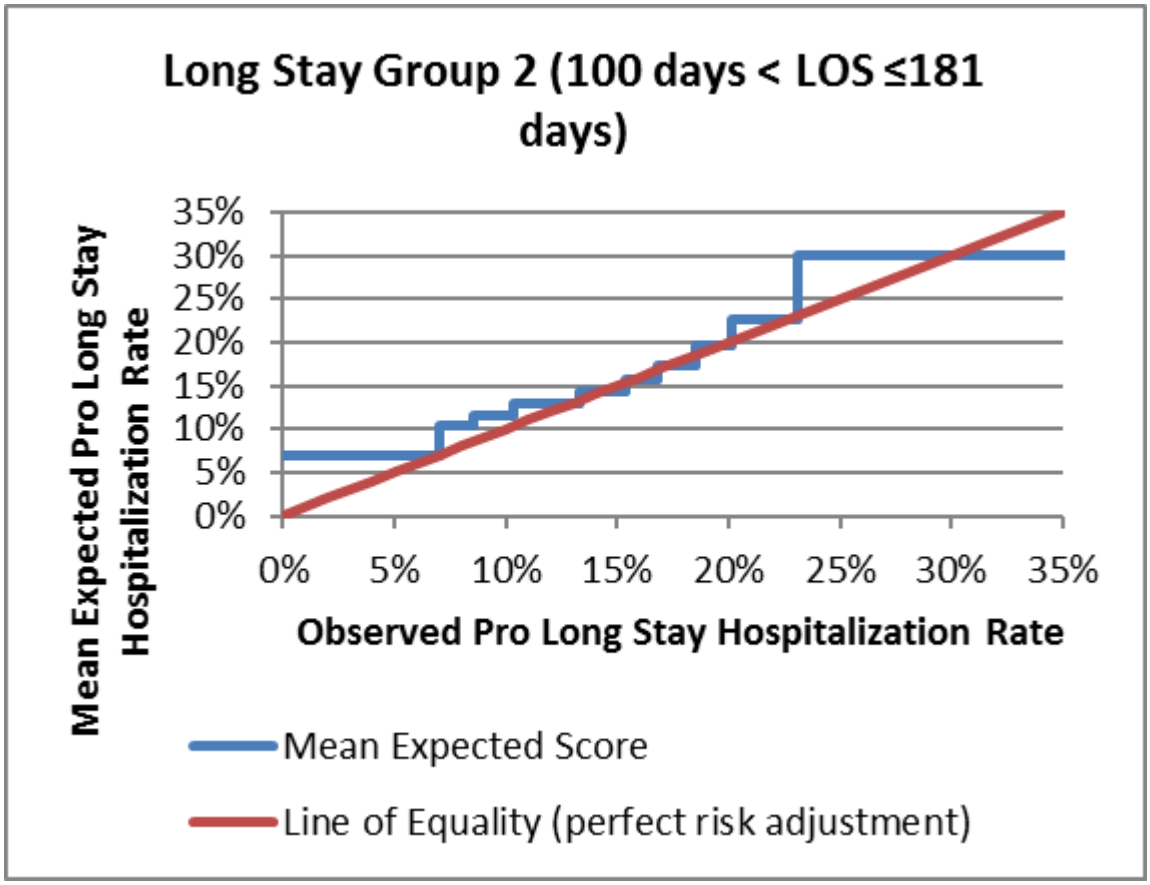
[Response Begins]

Figure 4. Calibration Plot Long Stay Group 1



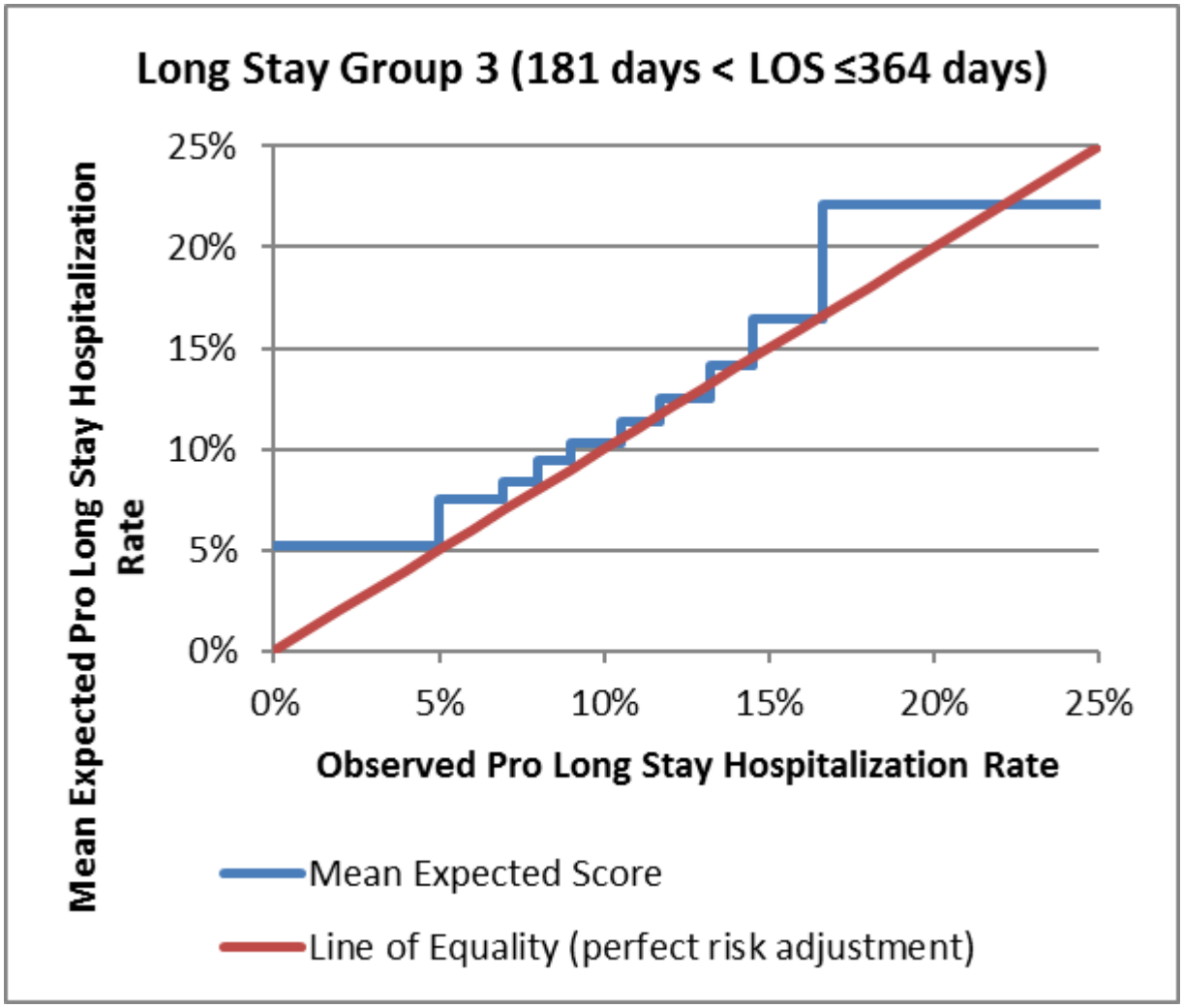
Calibration plot of long-stay group 1 (Current Stay LOS <= 100 days but cumulative days in SNF > 100)

Figure 5. Calibration Plot Long Stay Group 2



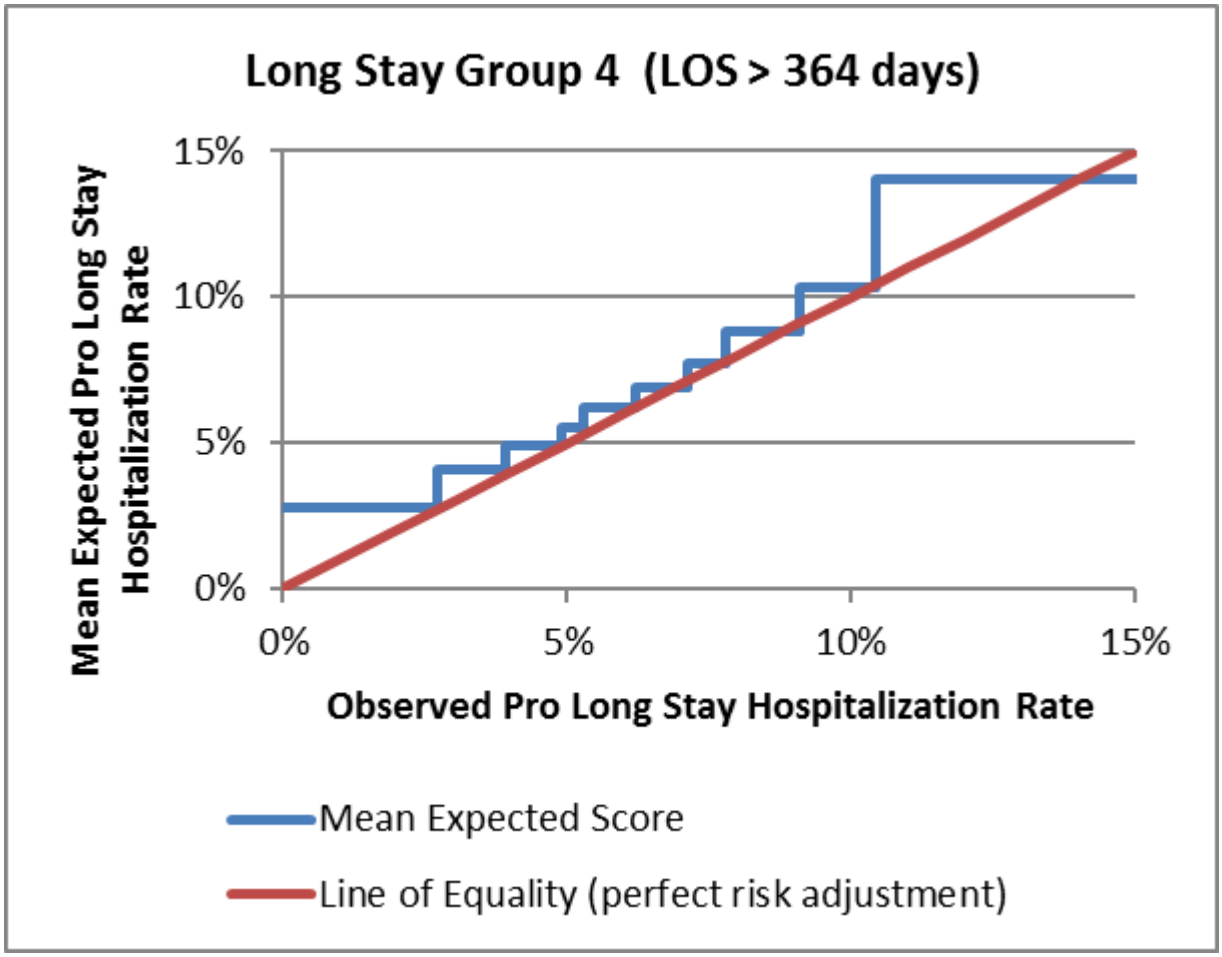
Calibration plot of long-stay group 2 (LOS between 100 and 181 days)

Figure 6. Calibration Plot Long Stay Group 3



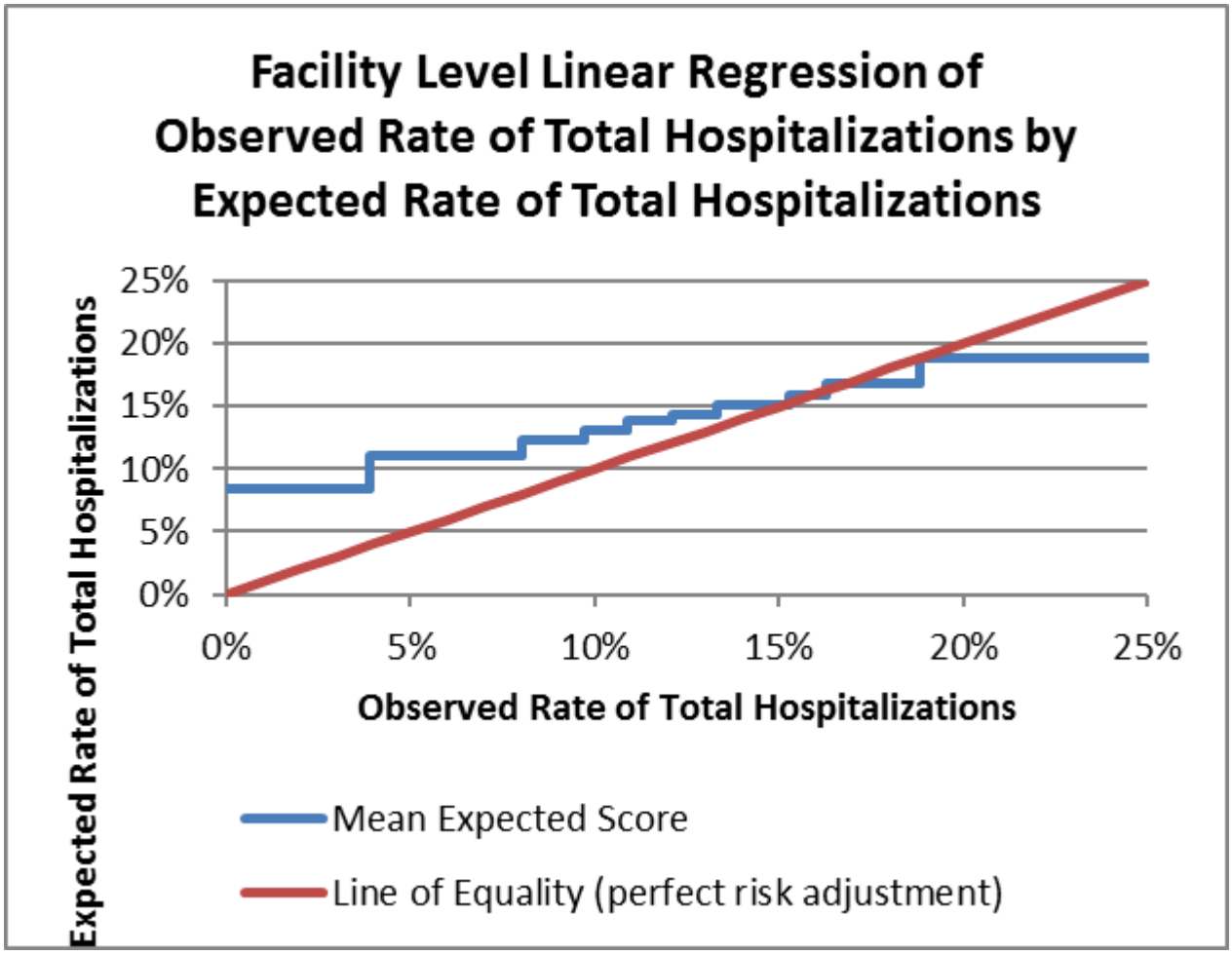
Calibration plot of long-stay group 3 (LOS between 181 and 364 days)

Figure 7. Calibration Plot Long Stay Group 4



Calibration plot of long-stay group 4 (LOS greater than 364 days)

Figure 8. Calibration Plot Linear Regression of All Hospitalizations



Calibration plot of all hospitalizations

[Response Ends]

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

[Response Begins]

N/A

[Response Ends]

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

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

[Response Begins]

N/A

[Response Ends]

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

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

[Response Begins]

N/A

[Response Ends]

Criteria 3: Feasibility

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

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

[Response Begins]

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

[Response Ends]

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

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

[Response Begins]

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

[Response Ends]

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

[Response Begins]

N/A

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]

N/A

[Response Ends]

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

[Response Begins]

Several of the decisions made in measure development were based on the objective of creating a measure that would be reliable and valid, readily computable based on data available to providers, and comprehensible and credible to providers, so that they will adopt it in their quality improvement efforts. These decisions are described in several sections above. The following specific decisions are emphasized:

1. A one year rolling measure period was selected to ensure adequate denominators (>30) for virtually all SNFs that have long-stay patients at all.

2. A “snapshot” approach was adopted rather than a complex survival model to make the measure more comprehensible to users.
3. Simple logistic regressions were selected for risk adjustment because more complex models did not perform better at the individual resident level, and we did not want to adjust for facility effects via a multilevel model.
4. Variability in hospitalization explained by race and Medicaid beneficiary status was divided into between and within facility effects, with adjustment of outcomes only for the latter. The aim was to get facilities to take responsibility for the part of sociodemographic disparities more likely to be under their control than due to otherwise unmeasured differences in baseline health status.
5. Observation of the rates of unknown outcomes in PointRight client population motivated us to exclude facilities with more than 10% unknown outcomes, and to impute hospitalization for facilities with less than 100% but 90% or more known outcomes. Doing so implies that rates will be available for 91% of all facilities. Insisting a 95% rate of known outcomes would imply that rates would not be reported for 14% of all SNFs – a problem that would limit its value for quality improvement. With 10% unknown outcomes, the maximum potential for overestimating the measure is 8% and the maximum potential for underestimating the measure is 2%. Clinically, the problem with underestimation is erroneously identifying a poor-performing facility as a good-performing facility, and then either referring more patients there or otherwise supporting the facility’s status quo. Given the distribution of the measure, a 2% improvement will not bring a facility in the worst quartile of performance to better than median performance. Thus, the 10% threshold appears to be an acceptable compromise between availability and utility of the measure.
6. Observation of the very high correlation of the rate of first hospitalizations with the rate of all hospitalizations led us to base our risk adjustment on modeling of first hospitalizations and subsequent multiplication of the results by a conversion factor. This approach is much more comprehensible to many end users than use of models with non-binary dependent variables.

[Response Ends]

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

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

Attach the fee schedule here, if applicable.

[Response Begins]

Computation of the measure requires a license to use software for large-scale data management and calculation of risk estimates using logistic regression models. These are capabilities of all typical analytics software packages used by healthcare organizations (e.g., SAS, SPSS, Stata, and R). Healthcare organizations would thus not incur additional expense to implement the measure. Utilization of the measure specifications does not require a fee. However, there is a requirement that display, disclosure or publication of the measure include the measure’s trademark (viz., PointRight® Pro Long Stay Hospitalization Measure) and that it is indicated that the measure specifications are copyrighted by PointRight®.

[Response Ends]

Criteria 4: Use and Usability

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

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

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

4a.01.

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

Name of program and sponsor

URL

Purpose

Geographic area and number and percentage of accountable entities and patients included

Level of measurement and setting

[Response Begins]

Payment Program

Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

Quality Improvement (Internal to the specific organization)

As the largest trade association for nursing homes with over 9,000 nursing home members, the American Health Care Association (AHCA) launched a multi-year quality initiative to improve nursing home quality. In the latest iteration of this quality initiative, which was launched in 2018, there was a goal for every facility to reduce long-stay hospitalizations by 10 percent, or maintain a high performance rate of 10 percent or less, by March 2021 as measured by ProLongStay. As part of the initiative, facility members and state affiliates receive routine performance reports. AHCA Quality Initiative Issue Brief - https://www.ahcancal.org/Advocacy/IssueBriefs/Issue%20Brief%20AHCA%20QI%202018-2021.pdf#search=quality%20initiative

On a quarterly basis, ProLongStay rates are updated on AHCA's LTC Trend Tracker tool for members. Independently owned members, as well as corporate members, can track and benchmark their organization's ProLongStay performance via LTC Trend Tracker.

Additionally, non-members and the public can download facility-level rates on a quarterly basis from AHCA's website -

https://www.ahcancal.org/Data-and-Research/Pages/PointRight-Downloads.aspx

Several state Medicaid value-based purchasing and pay for performance programs utilize ProLongStay. Specific examples include:

New Mexico Nursing Facility Value-Based Payment (VBP)

- Purpose – Transition from a traditional fee-for-service Medicaid reimbursement to a value-based payment model that connects payment to outcomes and quality. ProLongStay is one of the measures used in the program.
- Geographic Area – New Mexico nursing facilities
- Reference URL -

https://content.nethealth.com/New_Mexico_Nursing_Facility_Value_Based_Payment

Colorado Medicaid Nursing Facilities Pay for Performance

- Purpose – Incentivize quality improvement. ProLongStay is one of the acceptable measures to use for reporting long-stay hospitalization performance.
- Geographic Area – Colorado nursing facilities
- Reference URL -

https://hcpf.colorado.gov/nursing-facilities

Hawaii Nursing Facility Pay for Performance

- Purpose – Incentivize quality improvement. ProLongStay is one of six quality measures in the program.
- Geographic Area – Hawaii nursing facilities
- Reference URL -

https://www.ahcancal.org/Data-and-Research/Center-for-HPE/Documents/CHPE-Report-A%20Review%20of%20NH%20Medicaid%20VBP%20Programs%2002.23.2022.pdf

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Payment Program

Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

Quality Improvement (internal to the specific organization)

[Response Ends]

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

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

[Response Begins]

The American Health Care Association (AHCA) is currently implementing the PointRight® Long Stay Hospitalization measure on national MDS data, which it will then publish on its website for free public use, and also in its member data profiling and tracking tool, LTC Trend Tracker®. Once published, the measure developer and measure steward would like to see the measure adopted for regulatory and payment purposes, rather than a measure based on Fee-for-Service Medicare claims; with the increasing penetration of managed care for Medicare, Medicaid and dual eligible programs, and the significant proportion of private pay and commercial LTC insurance financing of long-term SNF care, a measure based on Medicare FFS claims alone could mischaracterize the performance of many SNFs.

[Response Ends]

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

A credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.

[Response Begins]

PointRight intends to provide the PointRight Pro Long Stay Hospitalization Measure to its customers beginning in the second half of 2016; AHCA intends to make the measure available to its members (and to other selected stakeholders) on its website in the second half of 2016. If the measure is endorsed by the NQF, AHCA and PointRight will advocate for its adoption by CMS as a publicly reported quality measure that contributes to CMS's evaluation of SNFs' clinical performance.

[Response Ends]

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

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

[Response Begins]

The PointRight® Pro Long Stay™ hospitalization measure is available in three Net Health PointRight solutions, Quality Measures, ScoreCard, and New Mexico VBP. These solutions are available to all Net Health customers who subscribe to them as part of the PointRight product that is delivered as a web-based software application. The measure is provided as part of a comprehensive quality measure feature set at the Skilled Nursing Facility (SNF) level, with comparisons to other facilities in the customer's organization, benchmarks (national average of all Net Health customer facilities), trending, and drill-down capabilities to patient-level rehospitalization information.

Over 2,000 Skilled Nursing Facilities submit MDS data to Net Health for results in the Quality Measures solution. Results for all facilities in the nation are presented in the PointRight ScoreCard solution.

The measure is used in the New Mexico Nursing Facility Value Based Purchasing (VBP) Program, a value-based payment program that incentivizes Nursing Facilities to implement quality improvement programs focused on a core set of performance metrics. All Nursing Facilities in the state that qualify for the program use the measure as part of the Pay-for-Performance (P4P) Scorecard.

[Response Ends]

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

[Response Begins]

Within the Net Health Quality Measures and NM VBP solutions, results are updated on an ongoing basis, daily in near real-time and based on the most recent MDS data submitted to Net Health by the facility. In the PointRight ScoreCard solution, where information is available for all the facilities in the nation, results are updated quarterly (long-stay resident hospitalization data for all facilities in the nation are provided by the American Health Care Association).

Educational materials are available on-demand for users of the Net Health solutions. These resources include short video tutorials explaining the measure and how it is used in the applications, and Frequently Asked Questions related to the measure calculation, its population, and data elements. In addition, clinical help desk consultation is available for clients who have specific questions.

[Response Ends]

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

[Response Begins]

Net Health customers share feedback about the measure in direct conversations with Analytics, Product Management, Sales, and Client Services team members. They also submit feedback through in-application messaging, via email, and in conjunction with their responses to Net Promoter Score (NPS) customer satisfaction surveys.

[Response Ends]

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

[Response Begins]

Net Health PointRight customers use PointRight® Pro Long Stay™ to monitor and manage their hospitalization outcomes. By evaluating their performance on long-stay resident hospitalization, they are able to conduct data-driven quality assessment and performance improvement.

Nursing Facilities have shared with Net Health that using the measure has enabled them to:

- Understand patterns and trends in long-stay resident hospitalization.
- Identify residents with repeated hospitalizations and plan specific care interventions to avoid future occurrences.
- Decrease the rate of hospitalization for their long-stay residents and thereby, improve the quality of their lives.
- Avoid loss of reimbursement from uncompensated care during hospitalizations that exceed payer bed days caps.
- Inform clinical and operational decision-making to succeed in value-based care.

In the New Mexico VBP program, relevant and actionable analytics, including the PointRight® Pro Long Stay™ hospitalization measure, have been the key to improving patient care and driving better patient outcomes through successful collaboration, communication and transparency among provider and payer stakeholders. A case study can be accessed here:

https://go.pointright.com/new-mexico-nursing-facility-value-based-payment

[Response Ends]

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

[Response Begins]

N/A

[Response Ends]

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

[Response Begins]

N/A

[Response Ends]

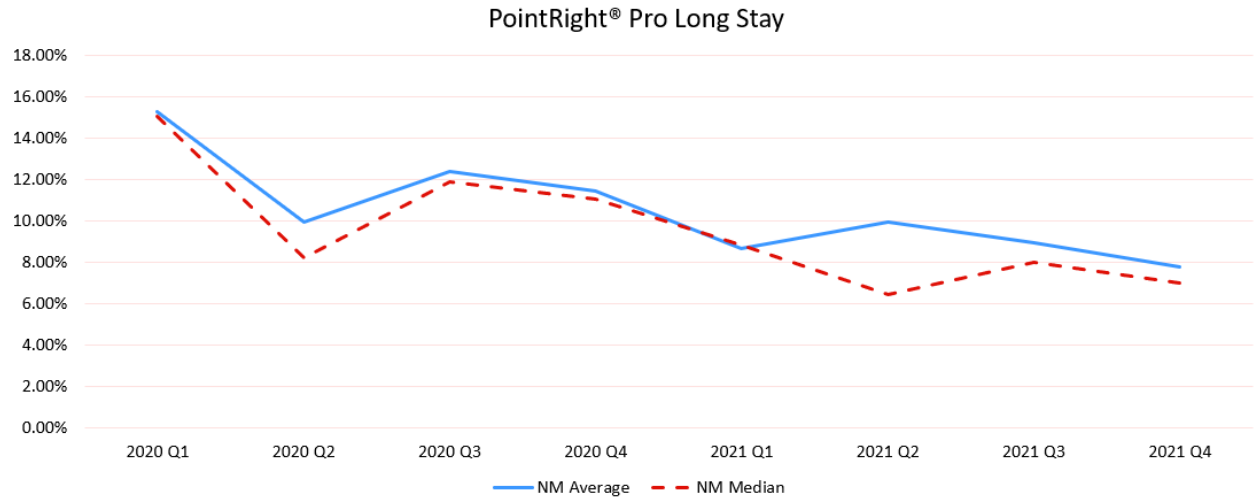
4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included). If no improvement was demonstrated, provide an explanation. If not in use for performance improvement

at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

[Response Begins]

New Mexico Nursing Facilities have demonstrated consistent improvement in their hospitalization rates, from an average of 15.27% and median rate of 15.08% in 2020 Q1 to an average rate of 7.81% and median rate of 7.01% in 2021 Q4. A chart showing the average and median rates across all Nursing Facility participants (approximately 63 facilities) for the first two program years (CY 2020 and 2021) appears below.

NMVBP - Program Years 2020 & 2021:
PointRight® Pro Long Stay™ Hospitalization



The results from the New Mexico VBP program are encouraging and highlight that when payment and quality are aligned, quality can improve. Currently, there is no national value-based or pay for performance incentives tied to reducing long-stay hospitalizations. Medicare's national Skilled Nursing Value-Based Purchasing program currently only accounts for short-stay rehospitalizations. This lack of focus on long-stay hospitalizations from a national perspective could help explain why there has been no substantial improvement in the national ProLongStay rate from 2014 to 2020.

[Response Ends]

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

[Response Begins]

No unintended consequences have been identified or are anticipated to occur as a result of this measure.

[Response Ends]

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

[Response Begins]

No unexpected benefits have been identified.

[Response Ends]

Criteria 5: Related and Competing Measures

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

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

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

(Can search and select measures.)

[Response Begins]

2510: Skilled Nursing Facility 30-Day All-Cause Readmission Measure (SNFRM)

2375: PointRight ® Pro 30™

[Response Ends]

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

(Can search and select measures.)

[Response Begins]

[Response Ends]

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

[Response Begins]

Number of Hospitalizations per 1,000 Long-Stay Resident Days by CMS and Abt Associates (Reference - https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/Downloads/Nursing-Home-Compare-Claims-based-Measures-Technical-Specifications-April-2019.pdf)

[Response Ends]

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

[Response Begins]

Yes

[Response Ends]

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

[Response Begins]

Currently there are no NQF-endorsed measures of hospitalizations for long stay nursing home residents. The only related NQF-endorsed measures (Pro30 and SNFRM) are for rehospitalizations of short stay nursing home patients.

Abt Associates developed and calculates a Medicare claims-based long stay hospitalization measure that CMS uses on Care Compare and in Five-Star ratings, but is not NQF-endorsed. Because this measure relies on Medicare claims, it can not capture all of the hospitalizations that ProLongStay can using all-payer MDS data. Because MDS and claims are required for reimbursement and federal regulatory compliance, these measures add no additional data collection burden to providers.

[Response Ends]

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

Provide analyses when possible.

[Response Begins]

There is no competing NQF-endorsed measure at this time.

With regards to the non-NQF-endorsed Abt Associates measure of long stay hospitalizations, there are two primary reasons ProLongStay is superior. The first is the use of all-payer MDS data in ProLongStay. Using MDS data allows for a wider perspective of a facility's performance, because it can capture hospitalizations that Medicare claims will miss.

The second reason ProLongStay is superior is in its ability to be more easily understood by providers and the public. ProLongStay calculates rates as percentages, while the Abt Associate measure calculates rates per 1,000 long-stay resident days. Per 1,000 resident day measures are harder for providers and the public to conceptually understand, particularly in terms of what is a bad or good rate.

[Response Ends]