

NATIONAL QUALITY FORUM

Measure Evaluation 4.1 December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the [evaluation criteria](#) are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized): Complete all **yellow highlighted** areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: If there is no TAP or workgroup, the SC also evaluates the subcriteria (**yellow** highlighted areas).

Steering Committee: Complete all **pink** highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met

C = Completely (unquestionably demonstrated to meet the criterion)

P = Partially (demonstrated to partially meet the criterion)

M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)

N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)

NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) NQF Review #: 0190	NQF Project: Nursing Homes- 2004
MEASURE DESCRIPTIVE INFORMATION	
De.1 Measure Title: Nurse staffing hours - 4 parts	
De.2 Brief description of measure: Percentage of daily work in hours by the entire group of nurses or nursing assistants spent tending to residents	
1.1-2 Type of Measure: Structure/management	
De.3 If included in a composite or paired with another measure, please identify composite or paired measure	
De.4 National Priority Partners Priority Area:	
De.5 IOM Quality Domain:	
De.6 Consumer Care Need:	

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. <i>Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.</i> A.1 Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)? A.2 Indicate if Proprietary Measure (as defined in measure steward agreement): A.3 Measure Steward Agreement: A.4 Measure Steward Agreement attached:	A Y ☐ N ☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least	B Y ☐

every 3 years.	NO
C. The intended use of the measure includes <u>both</u> public reporting <u>and</u> quality improvement. ► Actual/Planned Use: <u>Public Reporting, Quality Improvement (Internal to the specific organization)</u>	C YO NO
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1 Testing: <u> </u> D.2 Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?	D YO NO
(for NQF staff use) Have all conditions for consideration been met? Staff Notes to Steward (if submission returned):	Met YO NO
Staff Notes to Reviewers (issues or questions regarding any criteria):	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. <i>Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)</i> <u>(1a. High Impact)</u>	Eval Rating
(for NQF staff use) <u>Specific NPP goal:</u>	
1a.1 Demonstrated High Impact Aspect of Healthcare: 1a.2	1a CO PO MO NO
1a.3 Summary of Evidence of High Impact:	
1a.4 Citations for Evidence of High Impact:	
<u>(1b. Opportunity for Improvement)</u>	
1b.1 Benefits (improvements in quality) envisioned by use of this measure:	
1b.2 Summary of <u>(data demonstrating performance gap)</u> (variation or overall poor performance) across providers:	
1b.3 Citations for data on performance gap:	
1b.4 Summary of Data on disparities by population group:	1b CO PO MO NO
1b.5 Citations for data on Disparities:	

Comment [KP]: 1a. The measure focus addresses:

- a specific national health goal/priority identified by NQF's National Priorities Partners; OR
- healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).

Comment [KP]: 1b. Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).

Comment [k]: 1 Examples of data on opportunity for improvement include, but are not limited to: prior studies, epidemiologic data, measure data from pilot testing or implementation. If data are not available, the measure focus is systematically assessed (e.g., expert panel rating) and judged to be a quality problem.

1c. Outcome or Evidence to Support Measure Focus	
1c.1 Relationship to Outcomes (For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population):	
1c.2-3. Type of Evidence:	
1c.4 Summary of Evidence (as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome):	
1c.5 Rating of strength/quality of evidence (also provide narrative description of the rating and by whom):	
1c.6 Method for rating evidence:	
1c.7 Summary of Controversy/Contradictory Evidence:	
1c.8 Citations for Evidence (other than guidelines):	
1c.9 Quote the Specific guideline recommendation (including guideline number and/or page number):	
1c.10 Clinical Practice Guideline Citation:	
1c.11 National Guideline Clearinghouse or other URL:	
1c.12 Rating of strength of recommendation (also provide narrative description of the rating and by whom):	
1c.13 Method for rating strength of recommendation (If different from USPSTF system, also describe rating and how it relates to USPSTF):	1c CO PO MO NO
1c.14 Rationale for using this guideline over others:	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Importance to Measure and Report?	1
Steering Committee: Was the threshold criterion, Importance to Measure and Report, met? Rationale:	1 YO NO
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. (evaluation criteria)	Eval Rating
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained?	2a- specs CO PO MO NO
S.2 If yes, provide web page URL:	
2a. Precisely Specified	
2a.1 Numerator Statement (Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome):	

Comment [k]: 1c. The measure focus is:

- an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed;

OR

- if an intermediate outcome, process, structure, etc., there is evidence that supports the specific measure focus as follows:

Comment [k]: 4 Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multi-step process, the step with the greatest effect on the desired outcome should be selected as the focus of measurement. For example, although assessment of immunization status and recommending

Comment [k]: 3 The strength of the body of evidence for the specific measure focus should be systematically assessed and rated (e.g., USPSTF grading system <http://www.ahrq.gov/clinic/uspstf07/methods/benefit.htm>). If the USPSTF grading system was not used, the grading system is explained including how it relates to the USPSTF grades or why it does not. However, evidence is not limited to quantitative studies and the best type of evidence depends upon the question

Comment [k]: USPSTF grading system <http://www.ahrq.gov/clinic/uspstf/grade.s.htm>: A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. C - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the

Comment [KP]: 2a. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NQF's Health Information Technology Expert Panel (HITEP).

1. RN hours per resident per day 2. LPN/LVN hours per resident per day 3. CNA hours per resident per day 4. Total number of nursing staff hours per resident per day Numerator: Average daily work in hours by the entire group of nurses or nursing assistants.
2a.2 Numerator Time Window (The time period in which cases are eligible for inclusion in the numerator):
2a.3 Numerator Details (All information required to collect/calculate the numerator, including all codes, logic, and definitions):
2a.4 Denominator Statement (Brief, text description of the denominator - target population being measured): Total number of residents.
2a.5 Target population gender: 2a.6 Target population age range:
2a.7 Denominator Time Window (The time period in which cases are eligible for inclusion in the denominator):
2a.8 Denominator Details (All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions):
2a.9 Denominator Exclusions (Brief text description of exclusions from the target population):
2a.10 Denominator Exclusion Details (All information required to collect exclusions to the denominator, including all codes, logic, and definitions):
2a.11 Stratification Details/Variables (All information required to stratify the measure including the stratification variables, all codes, logic, and definitions):
2a.12-13 Risk Adjustment Type: No risk adjustment necessary
2a.14 Risk Adjustment Methodology/Variables (List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method):
2a.15-17 Detailed risk model available Web page URL or attachment:
2a.18-19 Type of Score:
2a.20 Interpretation of Score:
2a.21 Calculation Algorithm (Describe the calculation of the measure as a flowchart or series of steps):
2a.22 Describe the method for discriminating performance (e.g., significance testing):
2a.23 Sampling (Survey) Methodology If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate):
2a.24 Data Source (Check the source(s) for which the measure is specified and tested)

Comment [k]: 11 Risk factors that influence outcomes should not be specified as exclusions.
12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

2a.25 Data source/data collection instrument (Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.):	
2a.26-28 Data source/data collection instrument reference web page URL or attachment:	
2a.29-31 Data dictionary/code table web page URL or attachment:	
2a.32-35 Level of Measurement/Analysis (Check the level(s) for which the measure is specified and tested) «measurement_level_clinician_other» «measurement_level_program_other»	
2a.36-37 Care Settings (Check the setting(s) for which the measure is specified and tested) Nursing home (NH) /Skilled Nursing Facility (SNF)	
2a.38-41 Clinical Services (Healthcare services being measured, check all that apply)	
TESTING/ANALYSIS	
2b. Reliability testing	
2b.1 Data/sample (description of data/sample and size):	
2b.2 Analytic Method (type of reliability) & rationale, method for testing):	
2b.3 Testing Results (reliability statistics, assessment of adequacy in the context of norms for the test conducted):	2b CO PO MO NO
2c. Validity testing	
2c.1 Data/sample (description of data/sample and size):	
2c.2 Analytic Method (type of validity) & rationale, method for testing):	
2c.3 Testing Results (statistical results, assessment of adequacy in the context of norms for the test conducted):	2c CO PO MO NO
2d. Exclusions Justified	
2d.1 Summary of Evidence supporting exclusion(s):	
2d.2 Citations for Evidence:	
2d.3 Data/sample (description of data/sample and size):	
2d.4 Analytic Method (type analysis & rationale):	
2d.5 Testing Results (e.g., frequency, variability, sensitivity analyses):	2d CO PO MO NO NAO
2e. Risk Adjustment for Outcomes/ Resource Use Measures	2e

Comment [KP]: 2b. Reliability testing demonstrates the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.

Comment [k]: 8 Examples of reliability testing 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 may address the data items or final measure score.

Comment [KP]: 2c. Validity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.

Comment [k]: 9 Examples of validity testing include, but are not limited to: determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores.

Comment [KP]: 2d. Clinically necessary measure exclusions are identified and must be:

- supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;
- AND
- a clinically appropriate exception (e.g., patient indication) to eligibility for the measure.

Comment [k]: 10 Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.

Comment [KP]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present

2e.1 Data/sample (description of data/sample and size):	CO PO MO NO NAO
2e.2 Analytic Method (type of risk adjustment, analysis, & rationale):	
2e.3 Testing Results (risk model performance metrics):	
2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:	
2f. Identification of Meaningful Differences in Performance	
2f.1 Data/sample from Testing or Current Use (description of data/sample and size):	
2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance (type of analysis & rationale):	2f CO PO MO NO
2f.3 Provide Measure Scores from Testing or Current Use (description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance):	
2g. Comparability of Multiple Data Sources/Methods	
2g.1 Data/sample (description of data/sample and size):	
2g.2 Analytic Method (type of analysis & rationale):	2g CO PO MO NO NAO
2g.3 Testing Results (e.g., correlation statistics, comparison of rankings):	
2h. Disparities in Care	
2h.1 If measure is stratified, provide stratified results (scores by stratified categories/cohorts):	2h CO PO MO NO NAO
2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Scientific Acceptability of Measure Properties?	2
Steering Committee: Overall, to what extent was the criterion, Scientific Acceptability of Measure Properties, met? Rationale:	2 CO PO MO NO
3. USABILITY	
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. (evaluation criteria)	Eval Rating
3a. Meaningful, Understandable, and Useful Information	
3a.1 Current Use:	3a CO PO MO NO
3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly	

Comment [k]: 13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than adjusting out differences.

Comment [KP]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

Comment [k]: 14 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% v. 75%) 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 poor performance may not demonstrate much variability across providers.

Comment [KP]: 2g. If multiple data sources/methods are allowed, there is demonstration they produce comparable results.

Comment [KP]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender);OR rationale/data justifies why stratification is not necessary or not feasible.

Comment [KP]: 3a. Demonstration that information produced by the measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives). An important outcome that may not have an identified improvement strategy still can be useful for informing quality improvement by identifying the need for and stimulating new approaches to improvement.

<i>reported, state the plans to achieve public reporting within 3 years):</i>	
3a.3 If used in other programs/initiatives (If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years): Testing of Interpretability (Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement) 3a.4 Data/sample (description of data/sample and size): 3a.5 Methods (e.g., focus group, survey, QI project): 3a.6 Results (qualitative and/or quantitative results and conclusions):	
3b/3c. Relation to other NQF-endorsed measures	
3b.1 NQF # and Title of similar or related measures:	
(for NQF staff use) Notes on similar/related endorsed or submitted measures:	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 Are the measure specifications harmonized? If not, why?	3b CO PO MO NO NAO
3c. Distinctive or Additive Value 3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: 5.1 If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:	3c CO PO MO NO NAO
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Usability?	3
Steering Committee: Overall, to what extent was the criterion, Usability, met? Rationale:	3 CO PO MO NO
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes	4a
4a.1-2 How are the data elements that are needed to compute measure scores generated?	CO PO MO NO
4b. Electronic Sources	4b
	CO

Comment [KP]: 3b. The measure specifications are harmonized with other measures, and are applicable to multiple levels and settings.

Comment [k]: 16 Measure harmonization refers to the standardization of specifications for similar measures on the same topic (e.g., *influenza immunization* of patients in hospitals or nursing homes), or related measures for the same target population (e.g., eye exam and HbA1c for *patients with diabetes*), or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are dictated by the evidence. The dimensions of harmonization can include numerator, denominator, exclusions, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.

Comment [KP]: 3c. Review of existing endorsed measures and measure sets demonstrates that the measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare, is a more valid or efficient way to measure).

Comment [KP]: 4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery. (e.g., BP recorded in the electronic record, not abstracted from the record later by other personnel; patient self-assessment tools, e.g., depression scale; lab values, meds, etc.)

Comment [KP]: 4b. The required data elements are available in electronic sources. If the required data are not in existing electronic sources, a credible, near-term path to electronic collection by most providers is specified and clinical data elements are specified for transition to the electronic health record.

4b.1 Are all the data elements available electronically? (<i>elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims</i>)	PO MO NO
4b.2 If not, specify the near-term path to achieve electronic capture by most providers.	
4c. Exclusions	
4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?	4c CO PO MO NO NAO
4c.2 If yes, provide justification.	
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences	
4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.	4d CO PO MO NO
4e. Data Collection Strategy/Implementation	
4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:	
4e.2 Costs to implement the measure (<i>costs of data collection, fees associated with proprietary measures</i>):	
4e.3 Evidence for costs:	4e CO PO MO NO
4e.4 Business case documentation:	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Feasibility?	4
Steering Committee: Overall, to what extent was the criterion, Feasibility, met? Rationale:	4 CO PO MO NO
RECOMMENDATION	
(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.	Time-limited O
Steering Committee: Do you recommend for endorsement? Comments:	YO NO AO
CONTACT INFORMATION	
Co.1 Measure Steward (Intellectual Property Owner) Co.1 Organization Centers for Medicare & Medicaid Services, 7500 Security Boulevard , Mail Stop S3-01-02, Baltimore, Maryland,	

Comment [KP]: 4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.

Comment [KP]: 4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.

Comment [KP]: 4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

21244-1850
Co.2 <u>Point of Contact</u> Edward Q., Garcia III, MHS, Health Policy Analyst, MMSNQF@hsag.com, 410-786-6738-
Measure Developer If different from Measure Steward Co.3 <u>Organization</u>
Co.4 <u>Point of Contact</u>
Co.5 Submitter If different from Measure Steward POC Barry, Straube, barry.straube@cms.hhs.gov, 410-786-6841-, Centers for Medicare & Medicaid Services
Co.6 Additional organizations that sponsored/participated in measure development
ADDITIONAL INFORMATION
Workgroup/Expert Panel involved in measure development Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
Ad.2 If adapted, provide name of original measure: Ad.3-5 If adapted, provide original specifications URL or attachment
Measure Developer/Steward Updates and Ongoing Maintenance Ad.6 Year the measure was first released: Ad.7 Month and Year of most recent revision: Ad.8 What is your frequency for review/update of this measure? Ad.9 When is the next scheduled review/update for this measure?
Ad.10 Copyright statement:
Ad.11 Disclaimers:
Ad.12 -14 Additional Information web page URL or attachment:
Date of Submission (MM/DD/YY): 01/01/0001