NATIONAL QUALITY FORUM

Measure Submission and Evaluation Worksheet 5.0

This form contains the information submitted by measure developers/stewards, organized according to NQF's measure evaluation criteria and process. The evaluation criteria, evaluation guidance documents, and a blank online submission form are available on the submitting standards web page.

NQF #: 0179   **NQF Project:** Pulmonary Project
(for Endorsement Maintenance Review) **Original Endorsement Date:** Mar 31, 2009   **Most Recent Endorsement Date:** Mar 31, 2009

BRIEF MEASURE INFORMATION
De.1 Measure Title: Improvement in dyspnea
Co.1.1 Measure Steward: Centers for Medicare & Medicaid Services
De.2 Brief Description of Measure: Percentage of home health episodes of care during which the patient became less short of breath or dyspneic.
2a1.1 Numerator Statement: Number of home health episodes of care where the patient has less dyspnea at discharge than at start (or resumption) of care.
2a1.4 Denominator Statement: Number of home health episodes of care ending with a discharge during the reporting period, other than those covered by generic or measure-specific exclusions.
2a1.8 Denominator Exclusions: All home health episodes where at the start (or resumption) of care assessment the patient had no impairment, or the episode of care ended in transfer to inpatient facility or death at home, or was covered by the generic exclusions.
1.1 Measure Type: Outcome **2a1. 25-26 Data Source:** Electronic Clinical Data **2a1.33 Level of Analysis:** Facility **1.2-1.4 Is this measure paired with another measure?** No **De.3 If included in a composite, please identify the composite measure** (*title and NQF number if endorsed*): NA

STAFF NOTES (*issues or questions regarding any criteria*)
Comments on Conditions for Consideration:
Is the measure untested?   Yes☐   No☐   If untested, explain how it meets criteria for consideration for time-limited endorsement:
1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (*check De.5*): 5. Similar/related endorsed or submitted measures (*check 5.1*): Other Criteria:
Staff Reviewer Name(s):

1. IMPACT, OPPORTUNITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT
Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All three subcriteria must be met to pass this criterion. See guidance on evidence. *Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.*

(evaluation criteria)

1a. High Impact: H☐ M☐ L☐ I☐
(*The measure directly addresses a specific national health goal/priority identified by DHHS or NPP, or some other high impact aspect of healthcare.*)

De.4 Subject/Topic Areas (*Check all the areas that apply*):   Pulmonary/Critical Care : Dyspnea
De.5 Cross Cutting Areas (*Check all the areas that apply*):

1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers

1a.2 If "Other," please describe:

1a.3 Summary of Evidence of High Impact (*Provide epidemiologic or resource use data*):
Dyspnea interfering with activity is an important health status indicator that impacts quality of life and substantially affects a patient's ability to engage in a wide variety of activities. The etiology of dyspnea interfering with activity varies (disease-related and/or related to deconditioning from an extended time of limited activity like bedrest), but a high proportion of home health care patients are affected based on the data reported by home health care agencies where 70% of patients are reported as having some dyspnea interfering with activity.
Dyspnea interfering with activity has been identified as a risk factor for hospitalization among Medicare home care patients in one large study (n = 922) of home health care patients (Fortinsky et al, 2006). There are no more recent studies on the outcome specific to home health care.

1a.4 Citations for Evidence of High Impact cited in 1a.3: Fortinsky RH, Madigan EA, Sheehan TJ, Tullai-McGuinness S, Fenster JR. Risk factors for hospitalization among Medicare home care patients. West J Nurs Res 2006; 28(8):902-917.

1b. Opportunity for Improvement: H☐ M☐ L☐ I☐
(*There is a demonstrated performance gap - variability or overall less than optimal performance*)

1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:
Continued reporting of this measure is important for home health care agencies to use to identify the rates at which they provide care that improves dyspnea interfering with activity. There has been improvement in this measure over time, suggesting that agencies are improving care for this outcome. There was a best practice improvement package developed the Quality Improvement Organizations in the prior scope of work focused on dyspnea and interventions to improve dyspnea.

1b.2 Summary of Data Demonstrating Performance Gap (*Variation or overall less than optimal performance across providers*):
[*For Maintenance – Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.*]

HHA Ave.	58%
Std Dev	18%
Min	0%
10th-ile	33%
25th-ile	50%
Median	62%
75th-ile	70%
90th-ile	78%
Max	100%

1b.3 Citations for Data on Performance Gap: [*For Maintenance – Description of the data or sample for measure results reported in 1b.2 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included***]**
Date(s): 7/1/2010 – 6/30/2011   Data/Sample: Risk-adjusted measure as reported on Home Health Compare. Home Health Compare reports this measure for Medicare certified agencies with at least 20 quality episodes to which the measure applies that have submitted OASIS C assessments for at least 6 months of the 12 month reporting period. Of the 11,236 agencies that are listed on Home Health Compare, 8,794 agencies met the criteria for public reporting (20 episodes and 6 months of data). Information about downloading the Home Health Compare Database is available at:
http://www.medicare.gov/HomeHealthCompare/Resources/Download-Database.aspx

1b.4 Summary of Data on Disparities by Population Group: [*For Maintenance –Descriptive statistics for performance results for this measure by population group*]
Among all quality episodes 62.4% show improvement in dyspnea.

By level of Dyspnea (OASIS-C item M1400) at start of care/resumption of care:

	% Imprv Dyspnea	# of Episodes
M1400=1	50%	1,037,076
M1400=2	68%	1,113,388
M1400=3	76%	427,733
M1400=4	75%	87,197

By Age:

Age	% Imprv Dyspnea	# of Episodes
<65	61%	416,569
65-74	63%	661,527
75-84	63%	893,031
85+	61%	694,267

By Gender:

	% Imprv Dyspnea	# of Episodes
Male	62%	975,314
Female	63%	1,690,080

By Race:

	% Imprv Dyspnea	# of Episodes
White	64%	2,007,685
Black	62%	362,308
Hispanic	49%	234,745
Other	64%	64,060

Hispanic patients appear to be less likely to improve in dyspnea than non-Hispanic patients, and this disparity may merit further investigation. Additionally, patients with only mild dyspnea at start of care are less likely to show improvement than are patients with more severe dyspnea.

1b.5 Citations for Data on Disparities Cited in 1b.4: [*For Maintenance – Description of the data or sample for measure results reported in 1b.4 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*]
Data date(s): 7/1/2010 – 6/30/2011
Data/Sample: 2.67 million OASIS-C quality episodes from Medicare certified agencies ending during the 12-month observation period that meet the denominator criteria.

1c. Evidence (*Measure focus is a health outcome OR meets the criteria for quantity, quality, consistency of the body of evidence.*)
Is the measure focus a health outcome? Yes☐ No☐   If not a health outcome, rate the body of evidence.

Quantity: H☐ M☐ L☐ I☐   Quality: H☐ M☐ L☐ I☐   Consistency: H☐ M☐ L☐ I☐

Quantity	Quality	Consistency	Does the measure pass subcriterion1c?

M-H	M-H	M-H	Yes☐
L	M-H	M	Yes☐ IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise **No**☐
M-H	L	M-H	Yes☐ IF potential benefits to patients clearly outweigh potential harms: otherwise **No**☐
L-M-H	L-M-H	L	No ☐

Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service	Does the measure pass subcriterion1c? Yes☐ IF rationale supports relationship

1c.1 Structure-Process-Outcome Relationship *(Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome)*:
Dyspnea interfering with activity is an outcome measure. There are multiple etiologies for dyspnea including pulmonary and cardiac disease but also deconditioning associated with prolonged bedrest (for example). Dyspnea impairs quality of life and dyspnea interfering with activity influences the extent to which persons can care for themselves. There are interventions that can improve dyspnea in many patients (e.g. activity pacing, smoking cessation, correct use of pharmacologic agents, pursed lip breathing) and these interventions are part of what home health care agencies provide as part of their patient teaching. In some patients, however, the dyspnea interfering with activity is part of the progressive nature of the disease (e.g. end-stage COPD) and cannot be improved.

1c.2-3 Type of Evidence *(Check all that apply)*:
Selected individual studies (rather than entire body of evidence), Systematic review of body of evidence (other than within guideline development)

1c.4 Directness of Evidence to the Specified Measure *(State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population)*:
There are no studies specific to home health care practice in the US that identifies interventions to address dyspnea interfering with activity. Suter et al (2011) provide a model for care for patients with chronic obstructive pulmonary disease (COPD), but there is no evidence of testing of the model or its effectiveness.
There is evidence from a systematic review of RCTs that a home-based pulmonary rehabilitation approach for patients with COPD is an alternative to outpatient pulmonary rehabilitation (Vieira et al 2010). However, this is not a general practice within home health care for any patient groups.

1c.5 Quantity of Studies in the Body of Evidence *(Total number of studies, not articles)*: N/A — no studies specific to home health care practice in the US

1c.6 Quality of Body of Evidence (*Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events)*: There were no intervention studies specific to home health care patients identified in the literature review/environmental scan for this measure. There are interventions that agencies use to address this measure that come from clinical practice guidelines that are disease-specific (e.g. COPD).

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect)*: N/A

1c.8 Net Benefit *(Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms)*:
The benefits of providing home health care that improves dyspnea interfering with activity is likely to improve the quality of life for home health care recipients because shortness of breath is such a key indicator for how one feels and what one can do.

1c.9 Grading of Strength/Quality of the Body of Evidence. Has the body of evidence been graded? No

1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: NA

1c.11 System Used for Grading the Body of Evidence: Other

1c.12 If other, identify and describe the grading scale with definitions: NA

1c.13 Grade Assigned to the Body of Evidence: NA

1c.14 Summary of Controversy/Contradictory Evidence: NA

1c.15 Citations for Evidence other than Guidelines *(Guidelines addressed below)*:
Suter P, Hennessey B, Florez D, Newton SW. Review series: Examples of chronic care model: the home-based chronic care model: redesigning home health for high quality care delivery. Chron Respir Dis 2011; 8(1):43-52.

Vieira DS, Maltais F, Bourbeau J. Home-based pulmonary rehabilitation in chronic obstructive pulmonary disease patients. Curr Opin Pulm Med 2010; 16(2):134-143.

1c.16 Quote verbatim, <u>the specific guideline recommendation</u> *(Including guideline # and/or page #)*:
Guideline Title
(1) Nursing care of dyspnea: the 6th vital sign in individuals with chronic obstructive pulmonary disease (COPD). (2) Nursing care of dyspnea: the 6th vital sign in individuals with chronic obstructive pulmonary disease (COPD) 2010 supplement.
"Recommendation 1.0
Nurses will acknowledge and accept the patients´ self-report of dyspnea.
(Level of Evidence = IV)
Recommendation 1.1 (updated 2010)
All individuals identified as having dyspnea related to chronic obstructive pulmonary disease (COPD) will be assessed appropriately. "
(Level of Evidence = IV)

1c.17 Clinical Practice Guideline Citation: Registered Nurses´ Association of Ontario (RNAO). Nursing care of dyspnea: the 6th vital sign in individuals with chronic obstructive pulmonary disease (COPD) 2010 supplement. Toronto (ON): Registered Nurses´ Association of Ontario (RNAO); 2010 Feb. 27 p. [84 references]

1c.18 National Guideline Clearinghouse or other URL: National Guideline Clearinghouse

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? Yes

1c.20 If guideline recommendation graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: The guideline was graded by the panel that developed the guideline, the Registered Nurses´ Association of Ontario, with the names of the panel listed in the guideline and on www.guideline.gov Disclosures regarding bias were made as part of the process.

1c.21 System Used for Grading the Strength of Guideline Recommendation: Other

1c.22 If other, identify and describe the grading scale with definitions: Levels of Evidence
Ia Evidence obtained from meta-analysis or systematic review of randomized controlled trials.
Ib Evidence obtained from at least one randomized controlled trial.
IIa Evidence obtained from at least one well-designed controlled study without randomization.
IIb Evidence obtained from at least one other type of well-designed quasi-experimental study, without randomization.
III Evidence obtained from well-designed non-experimental descriptive studies, such as comparative studies, correlation studies, and case studies.
IV Evidence obtained from expert committee reports or opinions and/or clinical experiences of respected authorities.

1c.23 Grade Assigned to the Recommendation: Level IV

1c.24 Rationale for Using this Guideline Over Others: There were no guidelines that were generic to dyspnea, regardless of etiology.

The other guidelines are focused on pulmonary rehabilitation or end of life/palliative care whereas this guideline is more generic to the treatment of persons with chronic obstructive pulmonary disease, a diagnosis relevant to the home health care setting.

Based on the NQF descriptions for rating the evidence, what was the developer's assessment of the quantity, quality, and consistency of the body of evidence?
1c.25 Quantity: Low   **1c.26 Quality:** High **1c.27 Consistency:** High

Was the threshold criterion, *Importance to Measure and Report*, met?
(*1a & 1b must be rated moderate or high and 1c yes*)   Yes☐ No☐
Provide rationale based on specific subcriteria:

For a new measure if the Committee votes NO, then STOP.
For a measure undergoing endorsement maintenance, if the Committee votes NO because of 1b. (no opportunity for improvement), it may be considered for continued endorsement and all criteria need to be evaluated.

2. RELIABILITY & VALIDITY - SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

Extent to which the measure, <u>as specified</u>, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (<u>evaluation criteria</u>)
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 field. Supplemental materials may be referenced or attached in item 2.1. See <u>guidance on measure testing</u>.

S.1 Measure Web Page *(In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained)*. Do you have a web page where current detailed specifications for <u>this</u> measure can be obtained? Yes

S.2 If yes, provide web page URL: https://www.cms.gov/HomeHealthQualityInits/Downloads/HHQI-Revision1TechnicalDocumentationofMeasures.zip

2a. RELIABILITY. Precise Specifications and Reliability Testing: H☐ M☐ L☐ I☐

2a1. Precise Measure Specifications. (*The measure specifications precise and unambiguous.*)

2a1.1 Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, e.g., cases from the target population with the target process, condition, event, or outcome)*:
Number of home health episodes of care where the patient has less dyspnea at discharge than at start (or resumption) of care.

2a1.2 Numerator Time Window *(The time period in which the target process, condition, event, or outcome is eligible for inclusion)*:
CMS systems report data on episodes that end within a rolling 12 month period, updated quarterly.

2a1.3 Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, codes with descriptors, and/or specific data collection items/responses*:
Number of home health episodes from the denominator in which the value recorded for the OASIS-C item M1400 ("Dyspnea") on the discharge assessment is numerically less than the value recorded on the start (or resumption) of care assessment, indicating less impairment at discharge compared to start of care.

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured)*:
Number of home health episodes of care ending with a discharge during the reporting period, other than those covered by generic or measure-specific exclusions.

2a1.5 Target Population Category *(Check all the populations for which the measure is specified and tested if any)*: Adult/Elderly Care

2a1.6 Denominator Time Window *(The time period in which cases are eligible for inclusion)*:

CMS systems report data on episodes that end within a rolling 12 month period, updated quarterly.

2a1.7 Denominator Details (*All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses*):
All home health episodes of care (except those defined in the denominator exclusions) in which the patient was eligible to improve in dyspnea(i.e., were not at the optimal level of health status according to the "Dyspnea" OASIS-C item M1400).

2a1.8 Denominator Exclusions *(Brief narrative description of exclusions from the target population)*:
All home health episodes where at the start (or resumption) of care assessment the patient had no impairment, or the episode of care ended in transfer to inpatient facility or death at home, or was covered by the generic exclusions.

2a1.9 Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, codes with descriptors, and/or specific data collection items/responses)*:
Measure-specific exclusions:
All home health episodes where: (1) the value recorded for the OASIS-C item M1400 ("Dyspnea") on the start (or resumption) of care assessment is zero, indicating minimal or no impairment. These patients are excluded because it would be impossible for them to show measurable improvement; OR (2) the patient did not have a discharge assessment because the episode of care ended in transfer to inpatient facility or death at home[OR (3) all episodes covered by the generic exclusions.
Generic Exclusions:
a. Pediatric home health patients - less than 18 years of age.
b. Home health patients receiving maternity care only.
c. Home health clients receiving non-skilled care only.
d. Home health patients for which neither Medicare or Medicaid is a payment source.
e. The episode of care does not end during the reporting period.
f. Small and new agencies and rare conditions - the publicly-reported data on CMS' Home Health Compare web site also repress cells with fewer than 20 observations, and reports for home health agencies in operation less than six months.

2a1.10 Stratification Details/Variables *(All information required to stratify the measure results including the stratification variables, codes with descriptors, definitions, and/or specific data collection items/responses)*:
Not stratified

2a1.11 Risk Adjustment Type *(Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13)*: Statistical risk model **2a1.12 If "Other," please describe:**

2a1.13 Statistical Risk Model and Variables *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.)*:
Logistic regression models for risk adjustment were developed using three million episodes of care based on OASIS national repository data from assessments submitted between January 1, 2010 and September 30, 2010. Details of the model are available at: : https://www.cms.gov/HomeHealthQualityInits/Downloads/HHQILogisticRegressionModelsforRiskAdjustmentUpdated.pdf

2a1.14-16 Detailed Risk Model Available at Web page URL (or attachment). Include coefficients, equations, codes with descriptors, definitions, and/or specific data collection items/responses. Attach documents only if they are not available on a webpage and keep attached file to 5 MB or less. NQF strongly prefers you make documents available at a Web page URL. Please supply login/password if needed**:**
URL
https://www.cms.gov/HomeHealthQualityInits/Downloads/HHQILogisticRegressionModelsforRiskAdjustment.pdf

2a1.17-18. Type of Score: Rate/proportion

2a1.19 Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*: Better quality = Higher score

2a1.20 Calculation Algorithm/Measure Logic*(Describe the calculation of the measure score as an ordered sequence of steps*

including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.):
Calculation algorithm in technical specifications

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:
URL
https://www.cms.gov/HomeHealthQualityInits/Downloads/HHQI-Revision1TechnicalDocumentationofMeasures.zip

2a1.24 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):
NA

2a1.25 Data Source *(Check all the sources for which the measure is specified and tested)*. If other, please describe:
 Electronic Clinical Data

2a1.26 Data Source/Data Collection Instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*): OASIS-C

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment: URL
https://www.cms.gov/HomeHealthQualityInits/Downloads/HHQIOASISCAllTimePoint.pdf

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:
URL
https://www.cms.gov/OASIS/Downloads/oasisp200.zip

2a1.33 Level of Analysis (*Check the levels of analysis for which the measure is specified and tested*): Facility

2a1.34-35 Care Setting *(Check all the settings for which the measure is specified and tested)*: Home Health

2a2. Reliability Testing. (*Reliability testing was conducted with appropriate method, scope, and adequate demonstration of reliability.*)

2a2.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
The primary reliability testing for this item took place as part of the National OBQI Demonstration project when OASIS was originally designed and tested. The item has remained unchanged in its wording since the testing occurred in spring 1997 and fall 1998. In spring 1997, 41 patients from two agencies and in fall 1998, 25 patients from three different agencies were assessed by two RN level assessors who were provided training on assessment methods. The results from these studies are collectively referred to as "Study 1." Study 2 was an independent inter-rater reliability study conducted by Katherine Berg of Brown University (1999) with 144 patients ("Interim reliability report: Medicare home health case-mix project" Appendix G. Goldberg HB, Delargy D, Schmitz RJ, Moore T, Wrobel M and Berg K, Case-Mix Adjustment for a National Home Health Prospective Payment System. Second Interim Report, pp. G.3-G.25. Cambridge, MA: Abt Associates). Study 3 was a concurrent assessment of inter-rater reliability by Madigan, Tullai-McGinness, and Fortinsky (2001) with 88 patients from 21 agencies ("How to obtain meaningful and reliable results with OASIS data" Presentation at the annual meeting of the National Association for Home Care, Las Vegas, NV, October 2001).

2a2.2 Analytic Method *(Describe method of reliability testing & rationale)*:
The methodology used in Study 1 was an inter-rater reliability method whereby the patients observed were randomly selected from among new agency patients and consent was obtained from these patients. The order of assessment was alternated between the RN-level assessors, and both assessments for the patient occurred within 24-hours of each other (typically on the same day) to minimize changes in patient condition. Study 2 used more assessors with a wider range of clinical expertise and the assessments for an individual patient could occur as much as three days apart. Study 3 used the same methodology as Study 1 (i.e., 2 trained RN assessors; <24 hour revisits).

2a2.3 Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
Three different statistics were computed for Study 1: raw percent agreement, Cohen's kappa without weighting, and weighted kappa. For Study 2, only kappa values for the entire item rather than individual response options (weighted kappa) were reported. Study 3 reported both percent agreement and weighted kappas.

The inter-rater reliability (weighted kappa) values for the three studies were: 0.82 (Study 1), 0.49 (Study 2), and 0.51 (Study 3), indicating substantial agreement among the assessors for this item, based on Landis and Koch (1977) criteria (Landis JR and Koch GG "The Measurement of Observer Agreement for Categorical Data." Biometrics, 33, 159-174, March 1977).

2b. VALIDITY. Validity, Testing, <u>including all Threats to Validity</u>:   H☐ M☐ L☐ I☐

2b1.1 Describe how the measure specifications *(measure focus, target population, and exclusions)* **are consistent with the evidence cited in support of the measure focus (***criterion 1c***) and identify any differences from the evidence:**
As part of the National OBQI Demonstration project when OASIS was originally designed and tested, several tests of validity were conducted for each OASIS item, including the item for Dyspnea. The item passed each of the following validity assessments:
1)   Consensus validity by expert researcher/clinical panels for outcome measurement and risk factor measurement
2)   Consensus validity by expert clinical panels for patient assessment and care planning
3)   Criterion or convergent/predictive validity for outcome measurement/risk factor measurement
4)   Convergent/predictive validity: case mix adjustment for payment
5)   Validation by patient assessment and care planning
6)   Validation by outcome enhancement.
Descriptions for these validation assessments are contained in the accompanying descriptions taken from the "Volume 4 : OASIS Chronicle and Recommendation" OASIS and Outcome-Based Quality Improvement in Home Health Care, November 2001, Center for Health Services Research, University of Colorado Health Sciences Center, Denver, CO.

2b2. Validity Testing. (*Validity testing was conducted with appropriate method, scope, and adequate demonstration of validity.)*

2b2.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
Descriptions for these validation assessments are contained in the accompanying descriptions taken from the "Volume 4 : OASIS Chronicle and Recommendation" OASIS and Outcome-Based Quality Improvement in Home Health Care, November 2001, Center for Health Services Research, University of Colorado Health Sciences Center, Denver, CO.

2b2.2 Analytic Method *(Describe method of validity testing and rationale; if face validity, describe systematic assessment)*:
Descriptions for these validation assessments are contained in the accompanying descriptions taken from the "Volume 4 : OASIS Chronicle and Recommendation" OASIS and Outcome-Based Quality Improvement in Home Health Care, November 2001, Center for Health Services Research, University of Colorado Health Sciences Center, Denver, CO.

2b2.3 Testing Results *(Statistical results, assessment of adequacy in the context of norms for the test conducted; if face validity, describe results of systematic assessment)*:
Descriptions for these validation assessments are contained in the accompanying descriptions taken from the "Volume 4 : OASIS Chronicle and Recommendation" OASIS and Outcome-Based Quality Improvement in Home Health Care, November 2001, Center for Health Services Research, University of Colorado Health Sciences Center, Denver, CO.

POTENTIAL THREATS TO VALIDITY. (<u>*All*</u> *potential threats to validity were appropriately tested with adequate results.*)

2b3. Measure Exclusions. (*Exclusions were supported by the clinical evidence in 1c or appropriately tested with results demonstrating the need to specify them.*)

2b3.1 Data/Sample for analysis of exclusions *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
Data date(s): 7/1/2010 – 6/30/2011
Data/Sample: 5.70 million OASIS-C quality episodes from Medicare certified agencies ending during the 12-month observation period
Exclusion(s):
1.   Episodes in which the patient, at start/resumption of care, was not short of breath at any time: These patients do not have room to improve, as they had no dyspnea at the beginning of the episode.

2. Episodes that end with inpatient facility transfer: The information needed to calculate this measure is not collected if the home health episode ends in transfer or discharge to an inpatient facility. The measure cannot be calculated in excluded cases due to data limitations.

3. Episodes that end with death: The information needed to calculate this measure is not collected if the home health episode ends in death. The measure cannot be calculated in excluded cases due to data limitations.

4. Generic exclusions: As noted in the Denominator Exclusion Details (section 2a), OASIS data are only collected for particular types of patients. The exclusion of patients who are omitted from OASIS data collection (e.g., those who are non-Medicare/Medicaid, under 18, receiving maternity-related or non-skilled services only) is not based on research evidence but because the measure cannot be calculated due to data limitations.

2b3.2 Analytic Method (*Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference*):
Frequency of exclusions by type.

2b3.3 Results *(Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):*
% of quality episodes excluded: 53.2%
# total of quality episodes excluded: 3,029,960
By type:
1. Episodes in which the patient, at start/resumption of care, was not short of breath at any time:
a. # of quality episodes excluded: 1,722,757
b. % of excluded episodes: 56.9%
c. % of total quality episodes: 30.2%
2. Episodes ending with transfer/discharge to an inpatient facility:
a. # of quality episodes excluded: 1,279,755
b. % of excluded episodes: 42.2%
c. % of total quality episodes: 22.5%
3. Death exclusion:
a. # of quality episodes excluded: 27,448
b. % of excluded episodes: 0.9%
c. % of total quality episodes: 0.5%

2b4. Risk Adjustment Strategy. (*For outcome measures, adjustment for differences in case mix (severity) across measured entities was appropriately tested with adequate results.*)

2b4.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*
Model Development Process
Using the assessment data from January 1, 2010 through September 30, 2010, nearly three million episodes of care were created. This was done by linking the start of care (SOC) or resumption of care (ROC) assessment for a patient with that patient's last assessment (i.e., transfer, discharge, or death). From this analytic file, two developmental samples of 250,000 episodes of care were randomly selected, along with a validation sample of 1,000,000 episodes of care. A prediction was created using one of the two developmental samples and validated using the validation sample.

2b4.2 Analytic Method (Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables*)*:
A structured approach was used to develop the initial prediction model for each outcome. Because there were a large number of possible risk factors that needed to be considered for each outcome, the following process was used to identify unique contributing risk factors to the prediction model:
1. The risk factors were divided into six groups with approximately the same number in each group and similar content focus (e.g., ICD9-based conditions)
2. Separate logistic regressions were computed on each of these six risk factor groups and any risk factor with $p<0.200$ was identified for further review.
3. Risk factors with $p<0.200$ from the first three groups were aggregated into a new set of risk factors, as were risk factors with $p<0.200$ from the second three groups.
4. An ordinary least squares (OLS) step-wise analysis for each of these two new groups was computed. Those risk factors

that were statistically significant at p<0.01 were combined into a single group.

5.   The list of risk factors that achieved the p<0.01 level were reviewed. If one response option level of an OASIS-C item was on the list, then risk factors representing the other response option levels of that OASIS-C item were added to the list. For example, if response option levels 1 and 2 for M1800 Grooming were statistically significant at p<0.01 for a particular outcome, then response option level 3 for M1800 Grooming was added to the list.

6.   A fixed logistic regression was computed on the list of risk factors that had achieved p<0.01 and the risk factors that were added to the list because they were other response options for OASIS-C items represented on the list. Risk factors were removed if they failed to reach the p<0.01 level, unless they were a response option for an OASIS-C item where a different response option for the same OASIS-C item did meet the p<0.01 criterion. This step was repeated until only risk factors that met the criterion remained.

7.   Goodness of fit statistics (r2, C-statistic, and/or Hosmer-Lemeshow) as well as bivariate correlations between the risk factor and the outcome were computed for how well the predicted values generated by the prediction model were related to the actual outcomes.

8.   The initial models for each of the 48 outcome measures were reviewed by a team of at least three experienced home health clinicians. Each risk factor was reviewed for its "clinical plausibility" in being related to the outcome measure in the direction indicated by the coefficient in the prediction equation and its bivariate relationship. Risk factors that were not "clinically plausible" were identified for elimination.

9.   The risk factors that were deemed not "clinically plausible" were removed from the prediction model and steps 6 and 7 in this process were repeated. The resulting logistic regression equation was designated as the prediction model for the outcome.

10.   The prediction model was applied to the validation sample and goodness of fit statistics were computed. If these statistics were similar to the goodness of fit statistics computed with the development sample, the model become a "final" model. If the statistics were not similar, then alternative approaches to model building were considered or the model construction process returned to an earlier step in the process added information about risk factors that were not "clinically plausible"

2b4.3 Testing Results *(Statistical risk model: Provide quantitative assessment of relative contribution of model risk factors; risk model performance metrics including cross-validation discrimination and calibration statistics, calibration curve and risk decile plot, and assessment of adequacy in the context of norms for risk models. Risk stratification: Provide quantitative assessment of relationship of risk factors to the outcome and differences in outcomes among the strata)*:

The final prediction model for the Improvement in Dyspnea outcome measure contains 83 risk factors. See item 2a1.15. The prediction model was applied to the developmental and validation samples and predicted values
were computed. R2 and C-Statistics between the observed and predicted values were computed and are displayed in the following table.

Improvement Outcomes	Developmental		Validation	
	R2	C-Statistic	R2	C-Statistic
Improvement in Dyspnea	0.117	0.703	0.114	0.700

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment: The outcome measure is risk adjusted.

2b5. Identification of Meaningful Differences in Performance. *(The performance measure scores were appropriately analyzed and discriminated meaningful differences in quality.)*

2b5.1 Data/Sample *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
Date(s): 7/1/2010 – 6/30/2011   Data/Sample: Risk-adjusted measure as reported on Home Health Compare. Home Health Compare reports this measure for Medicare certified agencies with at least 20 quality episodes to which the measure applies that have submitted OASIS C assessments for at least 6 months of the 12 month reporting period. Of the 11,236 agencies that are listed on Home Health Compare, 8,794 agencies met the criteria for public reporting (20 episodes and 6 months of data). Information about downloading the Home Health Compare Database is available at:
http://www.medicare.gov/HomeHealthCompare/Resources/Download-Database.aspx

2b5.2 Analytic Method *(Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance)*:

Percentile data based on agency level data were analyzed to determine the inter-quartile range and the 90th vs. 10th percentile differences.

2b5.3 Results *(Provide measure performance results/scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:

HHA Ave.	58%
Std Dev	18%
Min	0%
10th-ile	33%
25th-ile	50%
Median	62%
75th-ile	70%
90th-ile	78%
Max	100%

Inter-quartile range (75th – 25th) = 70% – 50% = 20%
90th – 10th percentile = 78% – 33% = 45%
The distribution in this measure across agencies shows that there is substantial variation between the best performing and worst performing agencies. A HHA in the top decile (90th percentile or higher) has more than twice as many patients show improvement in dyspnea than an agency in the bottom decile (10th percentile or lower).

2b6. Comparability of Multiple Data Sources/Methods. (*If specified for more than one data source, the various approaches result in comparable scores.*)

2b6.1 Data/Sample *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
NA - single data source

2b6.2 Analytic Method *(Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure)*:
NA - single data source

2b6.3 Testing Results *(Provide statistical results, e.g., correlation statistics, comparison of rankings; assessment of adequacy in the context of norms for the test conducted)*:
NA - single data source

2c. Disparities in Care: H☐ M☐ L☐ I ☐ NA☐ (*If applicable, the measure specifications allow identification of disparities.*)

2c.1 If measure is stratified for disparities, provide stratified results *(Scores by stratified categories/cohorts)*: NA - not stratified

2c.2 If disparities have been reported/identified (e.g., in 1b), but measure is not specified to detect disparities, please explain:
NA

2.1-2.3 Supplemental Testing Methodology Information:
Attachment
OASIS Validity Types.pdf

Steering Committee: Overall, was the criterion, *Scientific Acceptability of Measure Properties*, met?
(*Reliability and Validity must be rated moderate or high*) Yes☐ No☐
Provide rationale based on specific subcriteria:

If the Committee votes No, STOP

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)

C.1 Intended Purpose/ Use *(Check all the purposes and/or uses for which the measure is intended)*: Public Reporting, Quality Improvement (Internal to the specific organization), Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

3.1 Current Use *(Check all that apply; for any that are checked, provide the specific program information in the following questions)*: Public Reporting, Quality Improvement with Benchmarking (external benchmarking to multiple organizations), Quality Improvement (Internal to the specific organization)

3a. Usefulness for Public Reporting: H☐ M☐ L☐ I☐
(*The measure is meaningful, understandable and useful for public reporting.*)

3a.1. Use in Public Reporting - disclosure of performance results to the public at large *(If used in a public reporting program, provide name of program(s), locations, Web page URL(s))*. If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: [*For Maintenance – If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be considered.*]
Public Reporting: Medicare Home Health Compare
http://www.medicare.gov/HomeHealthCompare/search.aspx

3a.2.Provide a rationale for why the measure performance results are meaningful, understandable, and useful for public reporting. If usefulness was demonstrated (e.g., focus group, cognitive testing), describe the data, method, and results: The CMS Center for Medicare contracted with L&M Policy Research (L&M) to help ensure that measures on the Home Health Compare (HHC) website are easy to understand and meet the needs of consumers. L&M possesses extensive knowledge of public health care issues and is experienced in qualitative and quantitative research methods and health services management and operations, including health communications. L & M also has plain language experts that are skilled in crafting straightforward language that allows CMS to provide beneficiaries, caregivers, health care professionals, and information intermediaries a better understanding of information on choice tools, such as HHC, which allows for more informed decisions on health related issues.
L&M's work during 2009-2010 with CMS includes an environmental scan of home health public reporting initiatives and a literature review of published and unpublished research relating to consumers' comprehension and use of home health quality measures. L&M independently convened its external advisory workgroup, comprised of representatives of consumer advocacy organizations, professional associations, quality improvement professionals, and experts in public reporting, to provide guidance on the organization, content, and usability of the home health measures website.

3.2 Use for other Accountability Functions (payment, certification, accreditation). If used in a public accountability program, provide name of program(s), locations, Web page URL(s):

3b. Usefulness for Quality Improvement: H☐ M☐ L☐ I☐
(*The measure is meaningful, understandable and useful for quality improvement.*)

3b.1. Use in QI. If used in quality improvement program, provide name of program(s), locations, Web page URL(s):
[*For Maintenance – If not used for QI, indicate the reasons and describe progress toward using performance results for improvement*].
Quality Improvement: Home Health Quality Initiatives
https://www.cms.gov/HomeHealthQualityInits/01_Overview.asp#TopOfPage

3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement. If usefulness was demonstrated *(e.g., QI initiative),* describe the data, method and results:
Data contained in the Home Health OBQI reports on the proportion of care episodes in which dyspnea improves provides agencies with a tool to evaluate the quality of their care and investigate how changes to processes of care related to patient education and pulmonary rehabilitation impact patient outcomes.

Overall, to what extent was the criterion, *Usability*, met? H☐ M☐ L☐ I☐
Provide rationale based on specific subcriteria:

4. FEASIBILITY

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (**evaluation criteria**)

4a. Data Generated as a Byproduct of Care Processes: H☐ M☐ L☐ I☐

4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*
Data used in the measure are:
generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition

4b. Electronic Sources: H☐ M☐ L☐ I☐

4b.1 Are the data elements needed for the measure as specified available electronically *(Elements that are needed to compute measure scores are in defined, computer-readable fields)*: ALL data elements are in a combination of electronic sources

4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources:

4c. Susceptibility to Inaccuracies, Errors, or Unintended Consequences: H☐ M☐ L☐ I☐

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:
Inaccuracies may result either due to confusion on the part of the clinician completing the OASIS or intentionally, to manipulate scores on quality measures. CMS has created and disseminated manuals and training materials to maximize accurate reporting of this data. Data accuracy could be audited through a review of medical records for evidence of relevant orders and implementation. All home health agencies serving adult, non-maternity Medicare and/or Medicaid patients must submit their OASIS assessment data to their respective state OASIS repository in a standard format. The repository software passes each incoming OASIS assessment record through an extensive set of quality edits. These include internal range and logic checks that assure that assessment items include only allowable values and that they are consistent with each other. When there are significant errors in an assessment, it is not accepted by the repository and the erroneous data are not available to be included in any published quality information. Data accuracy is also supported by the state survey process. Surveyors use OASIS to characterize each agency's caseload and to select sample patients to be interviewed. They also review and assess the accuracy of the agency's OASIS assessments. In addition, CMS payment contractors assess the accuracy of a sample of the OASIS assessments as part of their medical review processes. We are unable to provide results of these audit activities as we do not currently have access to the findings of the CMS surveyors, the data repository or CMS contractors regarding OASIS data accuracy.

4d. Data Collection Strategy/Implementation: H☐ M☐ L☐ I☐

A.2 Please check if either of the following apply (*regarding proprietary measures*):
4d.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 and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues (*e.g., fees for use of proprietary measures*):
OASIS data are collected by the home health agency during the care episode as part of the Conditions of Participation, and transmitted electronically to the state and CMS national OASIS repository. No issues regarding availability of data, missing data, timing or frequency of data collection, patient confidentiality, time or cost of data collection, feasibility or implementation have become apparent since OASIS-C was implemented 1/1/2010.

Overall, to what extent was the criterion, *Feasibility*, met? H☐ M☐ L☐ I☐
Provide rationale based on specific subcriteria:

OVERALL SUITABILITY FOR ENDORSEMENT

Does the measure meet all the NQF criteria for endorsement? Yes☐ No☐
Rationale:

If the Committee votes No, STOP.

If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.

5. COMPARISON TO 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 before a final recommendation is made.

5.1 If there are related measures *(either same measure focus or target population)* or competing measures *(both the same measure focus and same target population)*, list the NQF # and title of all related and/or competing measures:

5a. Harmonization

5a.1 If this measure has EITHER the same measure focus OR the same target population as NQF-endorsed measure(s): Are the measure specifications completely harmonized?

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

5b. Competing Measure(s)

5b.1 If this measure has both the same measure focus and the same target population as NQF-endorsed measure(s): Describe why this measure is superior to competing measures *(e.g., a more valid or efficient way to measure quality);* OR provide a rationale for the additive value of endorsing an additional measure. *(Provide analyses when possible)*:

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare & Medicaid Services, 7500 Security Boulevard , Mail Stop S3-01-02, Baltimore, Maryland, 21244-1850

Co.2 Point of Contact: Robin, Dowell, BSN, robin.dowell@cms.hhs.gov, 410-786-0060-

Co.3 Measure Developer if different from Measure Steward: Acumen LLC, 500 Airport Blvd, Suite 365, Burlingame, California, 94010

Co.4 Point of Contact: Keziah, Cook, PhD, kcook@acumenllc.com, 650-558-8882-247

Co.5 Submitter: Deborah, Deitz, BSN, Deborah_deitz@abtassoc.com, 617-520-3039-, Abt Associates Inc

Co.6 Additional organizations that sponsored/participated in measure development:
Abt Associates, Inc.
Case Western Reserve University
University of Colorado at Denver, Division of Health Care Policy and Research

Co.7 Public Contact: Robin, Dowell, BSN, robin.dowell@cms.hhs.gov, 410-786-0060-, Centers for Medicare & Medicaid Services

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 title of original measure, NQF # if endorsed, and measure steward. Briefly describe the reasons for adapting the original measure and any work with the original measure steward:

Measure Developer/Steward Updates and Ongoing Maintenance   **Ad.3 Year the measure was first released:** 2002   **Ad.4 Month and Year of most recent revision:** 01, 2010   **Ad.5 What is your frequency for review/update of this measure?** annual   **Ad.6 When is the next scheduled review/update for this measure?** 07, 2012
Ad.7 Copyright statement:
Ad.8 Disclaimers:
Ad.9 Additional Information/Comments:
Date of Submission (*MM/DD/YY*): 10/18/2011

Medical necessity determination: Items are checked that are included in an algorithm for evaluating medical necessity of home health services developed under the DHHS/ASPE study.

ELEMENT 5. Item Research, Development, Clinical, and Testing History: All but a few of the data items in the current version of OASIS have undergone considerable conceptual development, testing, refinement, and use for multiple applications in home health settings over a number of years. This section briefly highlights the research and development history of each item, indicating when and how it was used, tested, and refined over time.

ELEMENT 6. Validity: The most important types of validity undertaken in the OASIS research and development process were six in number. Each type of validity has a corresponding check box; a check mark (✓) indicates that the item under consideration underwent the indicated type of validity analysis. The six categories are:

Consensus validity by expert research/clinical panels for outcome measurement and risk factor measurement: This indicates whether an item was reviewed by panels of researchers and clinicians and was recommended for measuring patient outcomes relevant to home health care provision and quality measurement, or for risk adjustment of outcome analyses.

Consensus validity by expert clinical panels for patient assessment and care planning: This indicates whether an item was reviewed by a panel of clinical experts and was recommended for inclusion in a core set of data items for patient assessment and care planning -- for example, in addition to research project clinical panels, the Health Standards and Quality Bureau (HSQB) convened a panel consisting of HCFA staff, researchers, clinicians in a variety of disciplines, and home health industry representatives to review and possibly expand the OASIS items needed for assessment.

Criterion or convergent/predictive validity for outcome measurement/risk factor measurement: This type of validity indicates that the item has been tested empirically for use in conjunction with outcome measures or risk factors predictive of patient outcomes and, by virtue of such testing, has been found to be related to other indicators of health status and patient outcomes in a statistically significant and clinically meaningful way.

Convergent/predictive validity: Case mix adjustment for payment: This type of validity indicates that the item has been tested and is now used in the grouping algorithm that, in part, determines the per-episode payment to home health agencies for care provided under the Medicare home health benefit.

Validation by patient assessment and care planning: This type of validity indicates that the item has been used by clinicians for patient assessment and care planning in several hundred home health agencies for several years, and has been reported by practicing clinicians to be effective and useful for these purposes.