

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 #: 1886	NQF Project: Healthcare Disparities Project
(for Endorsement Maintenance Review)	
Original Endorsement Date: Most Recent Endorsement Date:	
BRIEF MEASURE INFORMATION	
De.1 Measure Title: Community engagement domain of Communication Climate Assessment Toolkit	
Co.1.1 Measure Steward: American Medical Association	
De.2 Brief Description of Measure: Site score on the domain of "Community Engagement" of the Communication Climate Assessment Toolkit (C-CAT), 0-100.	
2a1.1 Numerator Statement: Community Engagement component of patient-centered communication: an organization should make demonstrable, proactive efforts to understand the community it serves, including establishing relationships with community groups and developing opportunities for community members to participate in shaping organizational policies. Measure is scored based on 3 items from C-CAT patient survey and 3 items from C-CAT staff survey. Minimum n of 100 patient responses and 50 staff responses.	
2a1.4 Denominator Statement: There are two components to the target population: staff (clinical and nonclinical) and patients. Sites using this measure must obtain at least 50 staff responses and at least 100 patient responses.	
2a1.8 Denominator Exclusions: Staff respondents who do not have direct contact with patients are excluded from questions that specifically address patient contact.	
1.1 Measure Type: Patient Engagement/Experience	
2a1. 25-26 Data Source: Healthcare Provider Survey	
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): N/A	

STAFF NOTES <i>(issues or questions regarding any criteria)</i>
Comments on Conditions for Consideration:
Is the measure untested? Yes <input type="checkbox"/> No <input type="checkbox"/> 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):*

De.5 Cross Cutting Areas *(Check all the areas that apply):* Care Coordination, Disparities, Patient and Family Engagement, Safety

1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, Patient/societal consequences of poor quality

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

1a.3 Summary of Evidence of High Impact *(Provide epidemiologic or resource use data):*

Effective communication is critical to providing high quality health care and can be affected by a number of modifiable organizational factors. Validation study the measure of community engagement regarding patient-centered communication, as well as associated measures from the Communication Climate Assessment Toolkit (C-CAT) demonstrated variable performance across organizations, with better scores on the measure correlated to improved patient-reported trust in organization and patient reports of high-quality care.

- 1a.4 Citations for Evidence of High Impact cited in 1a.3:**
1. Kaplan SH, Greenfield S, Ware JE Jr. Assessing the effects of physician-patient interactions on the outcomes of chronic disease. *Med Care.* 1989;27(3 suppl):S110-S127.
 2. Flach SD, McCoy KD, Vaughn TE, Ward MM, Bootsmiller BJ, Doebbeling BN. Does patient-centered care improve provision of preventive services? *J Gen Intern Med.* 2004;19:1019-1026.
 3. Markova T, Broome B. Effective communication and delivery of culturally competent health care. *Urol Nurs.* 2007; 27:239-242.
 4. Institute of Medicine Committee on Communication for Behavior Change in the 21st Century: Improving the Health of Diverse Populations. *Speaking of Health: Assessing Health Communication Strategies for Diverse Populations.* Washington, DC: National Academies Press; 2002.
 5. Ashton CM, Haidet P, Paterniti DA, et al. Racial and ethnic disparities in the use of health services: bias, preferences, or poor communication? *J Gen Intern Med.* 2003;18:146-152.
 6. Gordon GH, Baker L, Levinson W. Physician-patient communication in managed care. *West J Med.* 1995;163:527-531.
 7. Seidel RW. How effective communication promotes better health outcomes. *JAAPA.* 2004;17:22-24.
 8. Wanzer MB, Booth-Butterfield M, Gruber K. Perceptions of health care providers' communication: relationships between patient-centered communication and satisfaction. *Health Commun.* 2004;16:363-383.
 9. Safran DG, Montgomery JE, Chang H, Murphy J, Rogers WH. Switching doctors: predictors of voluntary disenrollment from a primary physician's practice. *J Fam Pract.* 2001;50:130-136.
 10. Zolnierok KB, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. *Med Care.* 2009;47:826-834.
 11. Divi C, Koss RG, Schmaltz SP, Loeb JM. Language proficiency and adverse events in US hospitals: a pilot study. *Int J Qual Health Care.* 2007;19:60-67.
 12. The Joint Commission. *Improving America's Hospitals: The Joint Commission's Annual Report on Quality and Safety, 2007.* Oakbrook Terrace, IL: The Joint Commission; 2007.
 13. Scalise D. Clinical communication and patient safety. *Hosp Health Netw.* 2006;80:49-54.

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:

Understanding and improving communication may be a key to addressing disparities, which is an important national health policy goal.

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.]

Performance on the community engagement measure (domain) of the CCAT varies across sites. Among field-test sites, the average score was 77.8. The lowest scoring organization scored 68.3, while highest scoring organization scored 83.1.

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]

Wynia MK, Johnson M, McCoy TP, Griffin LP, Osborn CY. 2010. Validation of an Organizational Communication Climate Assessment Toolkit. Am J Med Qual. XX(X).

1b.4 Summary of Data on Disparities by Population Group: **[For Maintenance** –Descriptive statistics for performance results for this measure by population group]

Research has shown that LEP patients and patients from minority racial/ethnic groups experience communication problems more frequently than patients who speak English and those from traditionally advantaged groups. Regarding LEP patients, Flores (2005) has shown that provision of interpreters for LEP patients positively affects preventive screening rates, while those who either get no interpreter or an ad hoc interpreter have more medical tests, higher costs, and higher rates of hospitalization. Regarding patients of minority race/ethnicity, Hausmann et al (2011) have found that perceived racism is higher among African American patients than White patients, and that perceived racism negatively affects patient ratings of ease of communication (OR 0.22, 95% CI 0.07-0.67).

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]

- Hausmann L, Hannon MJ, Kresevic DM, Hanusa BH, Kwok CK, Ibrahim SA. 2011. Impact of Perceived Discrimination in Healthcare on Patient-Provider Communication. Med Care. 49;7:626-633.
- Cene CW, Roter D, Carson KA, Miller ER III, Cooper LA. The effect of patient race and blood pressure control on patient-physician communication. J Gen Intern Med. 2009;24:1057-1064.
- Weech-Maldonado R, Fongwa MN, Gutierrez P, Hays RD. Language and regional differences in evaluations of Medicare managed care by Hispanics. Health Serv Res. 2008;43:552-568.
- Flores G. 2005. The Impact of Medical Interpreter Services on the Quality of Health Care: A Systemic Review. Med Care Res Rev. 62;255. Language and regional differences in evaluations of Medicare managed care by Hispanics. Health Serv Res. 2008;43:552-568.

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 <input type="checkbox"/>
L	M-H	M	Yes <input type="checkbox"/> IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No <input type="checkbox"/>
M-H	L	M-H	Yes <input type="checkbox"/> IF potential benefits to patients clearly outweigh potential harms: otherwise No <input type="checkbox"/>
L-M-H	L-M-H	L	No <input type="checkbox"/>

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*):

The score in the domain of community engagement regarding patient-centered communication is a composite of structure, process and outcome measures from surveys addressed to patients and staff (clinical and non-clinical).

1c.2-3 Type of Evidence (Check all that apply):

Selected individual studies (rather than entire body of evidence)

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):

Evidence generated through national validation study of C-CAT instrument, including surveys of patients and staff.

1c.5 Quantity of Studies in the Body of Evidence (Total number of studies, not articles):

Body of evidence composed of 1 multi-site study involving two phases. The first phase was for psychometric testing and to refine and simplify the tools. The first-round patient surveys also included standard items about quality and trust in health care, which were used to assess the construct validity of the toolkit domains. Following the first round of field tests, 9 of the original 13 organizations agreed to perform reassessments using the refined tools to assess

variability in performance within and between organizations.

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): Study took place at 13 sites selected through a competitive process in order to achieve broad representation of sites. As such, non-random samples were used. Sites selected were in all geographic regions of the country with diverse patient and staff populations.

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

Study was conducted in 13 geographically and ethnically diverse health care organizations. Of the 13 organizations, 5 were large, urban hospitals or clinics; 4 were small, rural hospitals or clinics; and 4 were federally qualified health centers. Of the original 13 sites, 9 were involved in phase 2 in which performance variability was assessed.

Performance on the community engagement measure (domain) of the CCAT varies across sites. Among field-test sites, the average score was 77.8. The lowest scoring organization scored 68.3, while highest scoring organization scored 83.1.

1c.8 Net Benefit (Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms):

Results of study demonstrated that better scores on the measure of community engagement regarding patient-centered communication is correlated to important indicators of health care quality. Multivariate analysis showed that a 5-point increase in the measure results in more than a 50% greater odds that patients would report receiving high-quality medical care (OR 1.54, 95% CI 1.28-1.76) and a more than 25% greater odds that patients would report a belief that their medical records are kept private (OR 1.28, 95% CI 1.10-1.54). Likewise, a 5-point increase in the measure score is correlated with a more than 25% decrease in the odds a patient would believe that a mistake made in their care would be hidden from them (OR 0.73, 95% CI 0.59-0.86).

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: N/A

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

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

1c.13 Grade Assigned to the Body of Evidence: N/A

1c.14 Summary of Controversy/Contradictory Evidence: No known evidence that improved community engagement harms quality of care.

1c.15 Citations for Evidence other than Guidelines(Guidelines addressed below):

Additional evidence of the importance of leadership commitment to patient-centered communication from Improving Communication

-- Improving Care: How Health Care Organizations Can Ensure Effective, Patient-Centered Communication with People from Diverse Populations.
<http://www.ama-assn.org/resources/doc/ethics/pcc-consensus-report.pdf>

1c.16 Quote verbatim, the specific guideline recommendation (Including guideline # and/or page #):
 N/A

1c.17 Clinical Practice Guideline Citation: N/A

1c.18 National Guideline Clearinghouse or other URL: N/A

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? No

1c.20 If guideline recommendation graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias:

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

1c.22 If other, identify and describe the grading scale with definitions: N/A

1c.23 Grade Assigned to the Recommendation: N/A

1c.24 Rationale for Using this Guideline Over Others: N/A

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: Moderate 1c.27 Consistency: Moderate

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, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (**evaluation criteria**)

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 [guidance on measure testing](#).

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 this measure can be obtained? Yes

S.2 If yes, provide web page URL: <http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/the-ethical-force-program/patient-centered-communication.page>

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):

Community Engagement component of patient-centered communication: an organization should make demonstrable, proactive

efforts to understand the community it serves, including establishing relationships with community groups and developing opportunities for community members to participate in shaping organizational policies. Measure is scored based on 3 items from C-CAT patient survey and 3 items from C-CAT staff survey. Minimum n of 100 patient responses and 50 staff responses.

2a1.2 Numerator Time Window (*The time period in which the target process, condition, event, or outcome is eligible for inclusion*): Open data collection period -- ideally all C-CAT data are collected in a discreet data collection period ranging from 1-4 weeks. However, in order to achieve sufficient data sites with smaller patient populations, or those looking to collect data from a specific sub-group, may require a longer data-collection period.

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*):

Patient items:

p4 (pp4): Did hospital (clinic) staff help you find community resources?

p32 (pp32): Does the hospital (clinic) serve your community well?

p33 (pp33): Do hospital (clinic) staff come from your community?

Staff survey items:

s47: Overall, how would you rate the hospital (clinic)'s level of involvement in the local community?"

s48: Overall, how would you rate the hospital (clinic)'s efforts to help patients access community resources (e.g., assistance with medications, nutrition, insurance, legal aid, etc.)?

See field 2a1.20 for measure score calculation logic.

2a1.4 Denominator Statement (*Brief, narrative description of the target population being measured*):

There are two components to the target population: staff (clinical and nonclinical) and patients. Sites using this measure must obtain at least 50 staff responses and at least 100 patient responses.

2a1.5 Target Population Category (*Check all the populations for which the measure is specified and tested if any*): Adult/Elderly Care, Children's Health, Populations at Risk

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

A brief, discreet data-collection period is preferred. A data-collection period of between 1-4 weeks is usually sufficient to collect needed data.

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*):

Staff respondents should include all staff categories, including both clinical and non-clinical staff as well as those in roles such as building/environmental services, food services, etc. A minimum of 50 staff responses in a variety of staff categories is required to calculate the measure score. Staff surveys are made available in English and Spanish by default, with additional language available upon request. Patient respondents include all patients, with a pediatric version made available for families of minor patients. During field testing, patient surveys were available in 5 languages: English, Spanish, Chinese, Polish and Vietnamese. Currently, English and Spanish language surveys are made available by default with additional languages available upon request (language determined by organization using the C-CAT).

During field testing of the instruments, surveys were available on paper or online and during phase 1 patient surveys were also available via automated voice response systems. After very few patients replied using the voice automated system, the system was retired from use.

2a1.8 Denominator Exclusions (*Brief narrative description of exclusions from the target population*):

Staff respondents who do not have direct contact with patients are excluded from questions that specifically address patient contact.

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*):

Based on response to the first item on the staff survey ("Does your job involve direct contact with patients? yes/no"), staff respondents who do not have direct contact with patients are excluded from items that relate to direct contact with patients.

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):

N/A

2a1.11 Risk Adjustment Type (Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13): **No risk adjustment or risk stratification** 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.):

N/A

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:

2a1.17-18. Type of Score: **Non-weighted score/composite/scale**

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.):

The measure score is an average of the patient and staff components.

Calculation of patient component of measure score: each response of "never" counts as 0; each response of "sometimes" counts as 0.5; each response of "always" counts as 1.0; responses of "not sure" are excluded. A composite score for each item is calculated by summing the total response scores and dividing by the number of valid responses ("not sure" excluded); this operation is repeated for each item; an average of all patient items is calculated; this average is multiplied by 100, resulting in a 0-100 score for the patient component of "data collection."

For the staff component: each response of "very poor" counts as 0; each response of "poor" counts as 0.25; each response of "fair" counts as 0.5; each response of "good" counts as 0.75; each response of "very good" counts as 1.0; responses of "n/a" or "not sure" are excluded.

A composite score for each item is calculated by summing the total response score and dividing by the number of valid responses ("n/a" and "not sure" excluded); this operation is repeated for each item; an average of all staff items is calculated; this average is multiplied by 100, resulting in a 0-100 score for the staff component of the domain of "community engagement."

The average of the staff and patient components is obtained, resulting in the measure score for the domain of community engagement.

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:

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):

Those using the measure are encouraged to get as close to a universal sample as possible over a short period of time.

Required number of responses and response rate: To complete analysis, a minimum of 100 responses to the patient survey and 50 responses to the staff survey are required. To compare subgroups a minimum of 50 surveys from each group to be compared is required. If subgroup analysis (e.g. Spanish speaking patients, patients over age 65, etc.) is to be performed, over-sampling of the targeted group is likely necessary to achieve enough of these surveys to complete comparative analyses. Although response rates vary, we generally anticipate a 20% response rate to the patient survey, which means that to achieve 100 responses about 500

surveys will need to be distributed.

2a1.25 Data Source (Check all the sources for which the measure is specified and tested). If other, please describe:
[Healthcare Provider Survey](#)

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.): [Patient and Staff surveys of the Communication Climate Assessment Toolkit \(C-CAT\)](#)

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment: [URL](http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/the-ethical-force-program/patient-centered-communication/organizational-assessment-resources/view-surveys.page?)
<http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/the-ethical-force-program/patient-centered-communication/organizational-assessment-resources/view-surveys.page?>

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:
[Attachment](#)
[Community engagement data library.xls](#)

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): [Ambulatory Care : Clinic/Urgent Care, Ambulatory Care : Clinician Office, Hospital/Acute Care Facility](#)

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):

Measures were developed in two phases. In the first phase, a multistakeholder consensus process operating through a 13-member expert advisory panel was convened to review existing literature and propose a set of domains for organizational assessment, as well as specific organizational performance expectations in each domain. Each domain and each associated expectation was subject to a vote by the 21-member Ethical Force Oversight Body (ethicalforce.org) on whether each domain and expectation was 1) important, 2) measurable and 3) feasible to accomplish.

Adopted performance expectations were used to develop a set of communication climate assessment toolkits. The climate assessment tools were field-tested in 14 diverse organizations nationwide, with sites selected by the expert advisory panel in a competitive process intended to represent all regions of the country and a broad array of patient populations. Characteristics of survey respondents from field testing are available in Supplemental Information Section 2.1.

2a2.2 Analytic Method (Describe method of reliability testing & rationale):

In round 1, reliability was assessed by testing the internal consistency reliability of the domains, measured using Cronbach alpha. Standardized coefficients were used along with listwise deletion to optimize domain reliability. Specifically, items were systematically removed and as recalculated to determine when removing an item improved internal consistency. Finally, to assess the construct validity of the 9 domains, we examined correlations between domain scores and 3 standard measures of patient-reported quality of care and trust in health care systems. Patient reports of quality of care and trust can be correlated with a number of independent demographic factors that are unrelated to communication climate, including patient age, education, sex, race, ethnicity, and language (English vs. non-English). Because we wished to examine the effect of communication climate per se on patient-reported quality and trust, we adjusted for these demographic factors using multivariable logistic regression models. In round 2, graphical displays by sites were constructed to compare domain score variability within and between sites. All analyses were 2-sided, and statistical significance was determined at the $\alpha = .05$ level. All analyses were performed using SAS v9.1.3 (SAS Institute Inc, Cary, NC) or Stata v10 (StataCorp LP, College Station, TX).

2a2.3 Testing Results (Reliability statistics, assessment of adequacy in the context of norms for the test conducted):

The domain of community engagement showed moderate to good internal reliability. The patient survey component, which consists

of 3 items from the patient survey, displayed an internal consistency of a .64; the staff survey component, which consists of 2 items from the staff survey, displayed an internal consistency of a .78.

2b. VALIDITY. Validity, Testing, including all Threats to Validity: 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:

Measure is used at hospitals and clinics nationally. Measure was tested with patient and staff samples at geographically and ethnically diverse health care organizations nationally. The sites included in this study were not chosen at random but were selected for geographic variability and patient diversity, hence they have higher proportions of racial/ethnic minority patients and LEP patients than the national patient population. Because we analyzed our findings by these same demographic variables, in theory this should not affect the generalizeability of our findings.

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):

The measure was tested in two phases. Phase 1 consisted of psychometric testing and refinement and simplification of the tools used to collect data, as well as assessing the construct validity of the measure through inclusion of standard items about quality and trust in health care. Thirteen geographically and ethnically diverse health care organizations participated in phase 1, having been selected in a competitive selection process. In phase 1, 5,929 patients (35% LEP) and 1,229 staff (clinical and non-clinical) participated.

Phase 2 included 9 of the 13 original test sites. Phase two -- field testing -- took place between November 2007 and December 2008. In phase 2, 1,763 patients (29% LEP) and 651 staff (clinical and non-clinical) participated.

2b2.2 Analytic Method (Describe method of validity testing and rationale; if face validity, describe systematic assessment):

The measure was assessed for face, construct and content validity.

The toolkit materials were developed using a multistakeholder consensus process and refined through 2 rounds of field testing at 14 widely varying health care organizations, including 7 hospitals and 7 clinics nationwide.

For this project, a 13-member expert advisory panel on patient-centered communication (member list available at: www.EthicalForce.org) was convened to review the existing literature and propose a set of domains for organizational assessment as well as specific organizational performance expectations within each domain, of which community engagement is one. Each proposed domain and every individual performance expectation within each domain was subject to a vote by the 21-member Ethical Force Oversight Body, comprising leaders from hospital, health plan, clinician, and patient groups (membership list at www.EthicalForce.org). Using a scale of 1 to 10 (1 = not at all; 10 = completely), members voted on whether each proposed domain and performance expectation was (1) important, (2) measurable, and (3) feasible to accomplish. The mean score for adopting a domain or performance expectation was 7, and no member could vote <3 on any item; in essence, every member held a veto. This strict consensus process ensures the content validity of the domains and performance expectations.

Next, adopted performance expectations were used to develop the set of communication climate assessment tools including the coordinated surveys for patients, clinicians, nonclinical staff, and leaders. A 360° evaluation approach ensures that all views are represented in an effort to produce the most accurate and complete assessment of an organization.

The communication climate assessment tools were field tested in 14 diverse organizations nationwide. Sites were selected by the expert advisory panel in a competitive process, which aimed to represent all regions of the country and a broad array of patient populations. Following a call for nominations, more than 50 hospitals and clinics applied to serve as field test sites; 16 were selected (8 hospitals and 8 clinics, comprising 2 clinics and 2 hospitals from each of the 4 major geographic regions of the country). Of these, 14 ultimately contributed data to the field test process because 2 of the sites initially selected experienced leadership turnover or were sold and declined to participate.

The initial round of field tests was for psychometric testing and to refine and simplify the tools. The first-round patient surveys also included standard items about quality and trust in health care, which were used to assess the construct validity of the toolkit domains. Following the first round of field tests, 9 of the original 13 organizations agreed to perform reassessments using the refined tools to assess variability in performance within and between organizations.

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):

Results of study demonstrated that better scores on the measure of community engagement is correlated to important indicators of health care quality. Multivariate analysis showed that a 5-point increase in the measure results in more than a 50% greater odds

that patients would report receiving high-quality medical care (OR 1.54, 95% CI 1.28-1.76) and a more than 25% greater odds that patients would report a belief that their medical records are kept private (OR 1.28, 95% CI 1.10-1.54). Likewise, a 5-point increase in the measure score is correlated with a more than 25% decrease in the odds a patient would believe that a mistake made in their care would be hidden from them (OR 0.73, 95% CI 0.59-0.86).

POTENTIAL THREATS TO VALIDITY. (*All 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*):

N/A

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

N/A

2b3.3 Results (*Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses*):

N/A

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*):

N/A

2b4.2 Analytic Method (*Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables*):

N/A

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*):

N/A

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment: Adjusting the measure for such factors as race/ethnicity, language preference, or health literacy would have the effect of flattening important information regarding site-level performance on the measure. NQF, in a recent commission paper on health disparities and cultural competency, noted that risk adjustment by race/ethnicity "may fail to take advantage of the tremendous potential of quality measures to help eliminate disparities." For this reason, AMA does not use risk adjustment for this measure.

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*):

The measure was tested in two phases. Phase 1 consisted of psychometric testing and refinement and simplification of the tools used to collect data, as well as assessing the construct validity of the measure through inclusion of standard items about quality and trust in health care. Thirteen geographically and ethnically diverse health care organizations participated in phase 1, having been selected in a competitive selection process. In phase 1, 5,929 patients (35% LEP) and 1,229 staff (clinical and non-clinical) participated.

Phase 2 included 9 of the 13 original test sites. Phase two -- field testing -- took place between November 2007 and December 2008. In phase 2, 1,763 patients (29% LEP) and 651 staff (clinical and non-clinical) participated.

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

The measure was assessed for face, construct and content validity. The toolkit materials were developed using a multistakeholder consensus process and refined through 2 rounds of field testing at 14 widely varying health care organizations, including 7 hospitals and 7 clinics nationwide. For this project, a 13-member expert advisory panel on patient-centered communication (member list available at: www.EthicalForce.org) was convened to review the existing literature and propose a set of domains for organizational assessment as well as specific organizational performance expectations within each domain, of which community engagement is one. Each proposed domain and every individual performance expectation within each domain was subject to a vote by the 21-member Ethical Force Oversight Body, comprising leaders from hospital, health plan, clinician, and patient groups (membership list at www.EthicalForce.org). Using a scale of 1 to 10 (1 = not at all; 10 = completely), members voted on whether each proposed domain and performance expectation was (1) important, (2) measurable, and (3) feasible to accomplish. The mean score for adopting a domain or performance expectation was 7, and no member could vote <3 on any item; in essence, every member held a veto. This strict consensus process ensures the content validity of the domains and performance expectations. Next, adopted performance expectations were used to develop the set of communication climate assessment tools including the coordinated surveys for patients, clinicians, nonclinical staff, and leaders. A 360° evaluation approach ensures that all views are represented in an effort to produce the most accurate and complete assessment of an organization. The communication climate assessment tools were field tested in 14 diverse organizations nationwide. Sites were selected by the expert advisory panel in a competitive process, which aimed to represent all regions of the country and a broad array of patient populations. Following a call for nominations, more than 50 hospitals and clinics applied to serve as field test sites; 16 were selected (8 hospitals and 8 clinics, comprising 2 clinics and 2 hospitals from each of the 4 major geographic regions of the country). Of these, 14 ultimately contributed data to the field test process because 2 of the sites initially selected experienced leadership turnover or were sold and declined to participate. The initial round of field tests was for psychometric testing and to refine and simplify the tools. The first-round patient surveys also included standard items about quality and trust in health care, which were used to assess the construct validity of the toolkit domains. Following the first round of field tests, 9 of the original 13 organizations agreed to perform reassessments using the refined tools to assess variability in performance within and between organizations.

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

Results of study demonstrated that better scores on the measure of community engagement is correlated to important indicators of health care quality. Multivariate analysis showed that a 5-point increase in the measure results in more than a 50% greater odds that patients would report receiving high-quality medical care (OR 1.54, 95% CI 1.28-1.76) and a more than 25% greater odds that patients would report a belief that their medical records are kept private (OR 1.28, 95% CI 1.10-1.54). Likewise, a 5-point increase in the measure score is correlated with a more than 25% decrease in the odds a patient would believe that a mistake made in their care would be hidden from them (OR 0.73, 95% CI 0.59-0.86).

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*):

N/A

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

N/A

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*):

N/A

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*): N/A

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

The measure is not stratified for disparities, as what is measured is the role of the environment in fostering or inhibiting patient-centered communication. Further, the measure is a composite of items from the patient and staff surveys of the C-CAT. It is not possible to stratify staff responses based on the demographics of the patients particular staff respondents may provide care for. Notwithstanding the aforementioned, particular items on the patient survey of the C-CAT are regularly analyzed by demographic category to determine if an organization is providing different levels of care for different groups of patients.

2.1-2.3 Supplemental Testing Methodology Information:

URL

<http://www.ama-assn.org/resources/doc/ethics/ccat-validation.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.]

AMA allows use of measure for public reporting; however users reporting scores on one CCAT measure must report scores for all 9 measures (domains) of patient-centered communication.

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: Results are correlated to patient-reported quality of care and trust.

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): N/A

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].

<http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/the-ethical-force-program/patient-centered-communication/hospital-recognition.page?>

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:

All C-CAT measures are from 0-100, with a higher score representing better performance. In addition, sites using the C-CAT receive national average scores for each domain, enabling comparison between their performance and the average of all hospitals that have used the C-CAT tools.

Sites using the C-CAT measures for QI have found that the validated results allow them to focus limited resources to the areas most in need of attention.

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:

Other

Patient and staff surveys collect data from eligible respondents

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): Some data elements are in 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: The majority of staff data is collected through email-based surveys, with paper surveys made available for those staff members who do not have email access.

Patient surveys are distributed on paper, as many patients who are most likely to be affected by miscommunication (e.g., those with lower literacy) may not be comfortable with email-based surveys. In order to address this concern, an iPad app has been developed which includes an audio functionality that is useful to low literacy and low vision patients.

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:

As data is collected by survey, respondents' answers could potentially be inaccurate. However, psychometric testing and validation of the instruments found a high degree of reliability.

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

A.2 Please check if either of the following apply (regarding proprietary measures): Proprietary measure

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):

AMA has developed a preferred method for using the measures, in which sites wishing to use the measure work directly with a trained and licensed consultant. The consultant is able to provide guidance on preparation, data collection, and interpretation of results.

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): [American Medical Association, Institute for Ethics, 8th Floor, 515 N. State St., Chicago, Illinois, 60654](#)

Co.2 Point of Contact: [Matthew, Wynia, MD, MPH, matthew.wynia@ama-assn.org, 312-464-4980-](#)

Co.3 Measure Developer if different from Measure Steward: [American Medical Association, Institute for Ethics, 8th Floor, 515 N. State St., Chicago, Illinois, 60654](#)

Co.4 Point of Contact: [Matthew, Wynia, MD, MPH, matthew.wynia@ama-assn.org, 312-464-4980-](#)

Co.5 Submitter: [Andrew, Jager, MA, andrew.jager@ama-assn.org, 312-464-2431-, American Medical Association](#)

Co.6 Additional organizations that sponsored/participated in measure development:
[The Ethical Force Program Oversight Body led the development of the expectations that led to this measure. Funding was provided by the American Medical Association Foundation, The California Endowment, the Commonwealth Fund, and the Connecticut Health Foundation.](#)

Co.7 Public Contact: [Matthew, Wynia, MD, MPH, matthew.wynia@ama-assn.org, 312-464-4980-, American Medical Association](#)

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.

An expert advisory panel was convened to review literature on potential domains/measures, suggest measurement framework, and suggest measurable expectations within content area. Members of the expert advisory panel were:

Dennis Andrulis, PhD, MPH
Drexel University School of Public Health

David W. Baker, MD, MPH, FACP
Northwestern Memorial Hospital

David Fleming, MD
Center for Health Ethics, University of Missouri - Columbia

Elizabeth Heitman, PhD
Center for Medical Ethics, Vanderbilt University

Sharon King-Donohue, JD
National Committee for Quality Assurance

Edward L. Martinez, MS
National Assn of Public Hospitals and Health Systems

Mary A. Pittman, DrPH
Health Research and Educational Trust

Elena Rios, MD, MSPH
National Hispanic Medical Association

Stephen B. Thomas, PhD
Center for Minority Health, University of Pittsburgh

Amy Wilson, MPP
Joint Commission on Accreditation of Healthcare Organizations

Winston Wong, MD
Kaiser Permanente Community Benefit Program

Dawn E. Wood, MD, MPH
WellPoint

Mara Youdelman, JD, LLM
National Health Law Program

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: 2010

Ad.4 Month and Year of most recent revision: 04, 2011

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? 03, 2012

Ad.7 Copyright statement: ©American Medical Association, 2012.

The C-CAT's surveys, while copyrighted by the American Medical Association, are publicly accessible for viewing and for

noncommercial internal research purposes.

Ad.8 Disclaimers:

Ad.9 Additional Information/Comments: This measure is used in conjunction with additional instruments, namely a survey of executive leadership and a workbook on organizational policy; these instruments, as well as the other C-CAT instruments, are available for viewing at <http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/the-ethical-force-program/patient-centered-communication/organizational-assessment-resources/view-surveys.page?> The information collected through the executive survey and policy workbook provide important contextual information, but are not components in the calculation of the measure score. Further information is available at ethicalforce.org, or through the contact information for the measure developer and steward.

Date of Submission (MM/DD/YY): 01/18/2012

Instrument	Item Number	Item Content
Patient Survey (Pediatric Patient Survey)	p4 (pp4)	Did hospital (clinic) staff help you find community resources?
Patient Survey (Pediatric Patient Survey)	p32 (pp32)	Does the hospital (clinic) serve your community well?
Patient Survey (Pediatric Patient Survey)	p33 (pp33)	Do hospital (clinic) staff come from your community?
Staff Survey	s47	Overall, how would you rate the hospital (clinic)'s level of involvement in the local community?
Staff Survey	s48	Overall, how would you rate the hospital (clinic)'s efforts to help patients access community resources (e.g., assistance with medications, nutrition, insurance, legal aid, etc.)?

Item Scoring (aggregate scores are multiplied by 100)	Denominator	Numerator
Never=0, Sometimes=0.5, Always=1	n of respondents	n of respondents selecting each response category
Never=0, Sometimes=0.5, Always=1	n of respondents	n of respondents selecting each response category
Never=0, Sometimes=0.5, Always=1	n of respondents	n of respondents selecting each response category
Very Poor=0, Poor=0.25, Fair=0.5, Good=0.75, Very Good=1.0	n of respondents	n of respondents selecting each response category
Very Poor=0, Poor=0.25, Fair=0.5, Good=0.75, Very Good=1.0	n of respondents	n of respondents selecting each response category

Exclusions

"Not Sure"

"Not Sure"

"Not Sure"

"Not Sure", "N/A"

"Not Sure", "N/A"