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 #: 1828 NQF Project: Healthcare Disparities Project
(for Endorsement Maintenance Review) Original Endorsement Date: Most Recent Endorsement Date:
BRIEF MEASURE INFORMATION
De.1 Measure Title: L3: Patient wait time to receive interpreter servcies
Co.1.1 Measure Steward: Department of Health Policy, The George Washington University
De.2 Brief Description of Measure: This measure is used to assess the percent of encounters where the wait time for an interpreter is 15 minutes or less. Patients and providers report resistance or reluctance to using interpreter services due to long wait times or delays in obtaining an interpreter upon request. As interpreter services continue to evolve, many hospitals across the country have adopted standards for wait times for interpreter encounters. This measure provides information on the extent to which interpreter services are able to respond to requests for service within a reasonable amount of time, defined here as within 15 minutes.
2a1.1 Numerator Statement: The number of interpreter encounters in which the wait time is 15 minutes or less for the interpreter to arrive
2a1.4 Denominator Statement: The total number of interpreter encounters, stratified by language.
2a1.8 Denominator Exclusions: Exclusions • Encounters with bilingual providers and/or other bilingual workers/employees • Encounters with outside vendor telephone interpreters • Encounters with outside vendor video interpreters
1.1 Measure Type: Process 2a1. 25-26 Data Source: Administrative claims, Electronic Clinical Data : Electronic Health Record, Paper Records 2a1.33 Level of Analysis: Clinician : Group/Practice, 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*):

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, OPPORTUITY, 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)*: Disparities, 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)*:
The Institute of Medicine identified timeliness as one of six domains of quality of health care in its 2001 landmark report, Crossing the Quality Chasm (1). In terms of language services delivery, timely care can reduce wait times and delays for those those who receive and those who give care. Because of a number of barriers to the delivery of language services (including availability, knowledge about the benefits, and cost), long wait times can result in patients and providers moving forward with encounters without adequate communication from a trained interpreter.

Immigrants comprise a large and growing segment of American society that is disproportionately low-income and uninsured (2-3). Among the immigrant population, barriers to accessible and high-quality health care are only exacerbated for those who do not speak English fluently. In the U.S., 24 million individuals speak English "less than very well" and are said to be limited English proficient (LEP) (4). For this growing segment of the population, poor health status and diminished access to health care are frequent challenges. As members of racial, ethnic or linguistic minorities, persons with LEP experience disproportionately high rates of infectious disease (5) and infant mortality and are more likely to report risk factors for serious and chronic diseases such as diabetes and heart disease (6). Many of the challenges associated with delivering care to LEP populations result from communication barriers inherent in the LEP patient-provider interaction. Patients who speak languages other than English can have difficulties understanding their diagnosis (7) and why they receive particular types of care (8-9). Patients with LEP are particularly vulnerable to miscommunication when discharged from the emergency department (10-11), and have poorer follow-up after an emergency department visit (12). LEP patients have been shown to have poorer compliance (13) and lower adherence with diabetes (14)and asthma care regimens (15-16)compared to patients who are English speakers; they also have poorer diabetes outcomes (17).

Having an interpreter to facilitate communication between patients and health professionals can mitigate many of the disparities in care that LEP patients regularly face. LEP patients who are provided with an interpreter have more preventive and primary care visits and fill more prescriptions, compared to LEP patients who do not use an interpreter. (18) Having an interpreter can level the playing field for LEP patients with diabetes, whose care was found to be better than or equal to care received by non-LEP patients with diabetes (19). LEP patients who used any interpreter were more likely than English-speaking patients to have had a mammogram over a two-year period (20). In a study of the impact of interpreter services on low-income LEP patients, the availability of trained interpreters was associated with LEP patients having more office visits and filling more prescriptions, as well as reducing disparities related to flu vaccinations and fecal occult blood testing (18). Patients with language barriers indicated higher levels of satisfaction with care when interpreters were used (21). Physicians who had access to the services of trained interpreters reported a significantly higher quality of patient-physician communication than physicians without these services (22).

Medical interpreters can bridge the communication gap between physician and patient (23), yet interactions between patients with LEP and health professionals frequently occur without the services of an interpreter. No published studies estimate the frequency of interpreter use among LEP patients in the health care setting, although there is evidence of substantial underutilization in the emergency department and across ambulatory and inpatient services (7). Federal civil rights legislation (24) requires health care providers that receive any federal funds (including Medicare and Medicaid reimbursement) to provide language access for LEP patients, although federal oversight of the health care industry's compliance with these rules is extremely limited. In practice, LEP

patients' access to interpreter services is variable and unpredictable (26).

Merely having interpreters available in a health care setting does not mean that the patients who need services will receive them (26). Structural, logistical, and financial barriers are just a few of the impediments to effective use of services. Interpreter services require coordination across components of health systems when, for example, physicians or other health professionals schedule interpreters for LEP patients or coordinate physicians' schedules to match LEP patients with bilingual providers. Interpreter services can require equipment such as dual handsets for telephone interpretation that allow the patient and physician to participate in a conversation without passing the telephone back and forth, thereby disrupting the flow of the interaction. In-person and telephone interpreting also entail additional costs.

Health care organizations routinely place patients at risk for poor quality care by turning to untrained individuals to facilitate communication for LEP patients. Health care organizations use an alternative to the trained medical interpreter. The "ad hoc" interpreter is probably the vehicle used most frequently to allow LEP patients to communicate with health professionals. Ad hoc interpreters are not trained interpreters, but rather friends, family members, staff members and other individuals who present with the patient or are called upon to serve in the interpreter role (27). They are untrained in terms of clinical knowledge or appropriate methods to interpret in a health care setting. Their utility comes from being able to converse with the patient in the patient's language as well as having some level of English proficiency. Though ad hoc interpreters appear to be free to the health system, their use is not without its own set of costs, especially in terms of high rates of clinically significant medical errors. Ad hoc interpreters can misinterpret or omit questions asked by physicians (28); family members who interpret sometimes leave the patient out of the discussion altogether, instead answering the physician's questions without consulting the patient (29).

Physicians recognize the need for trained interpreters (30) but may opt to move forward with ad hoc interpreters nevertheless. In a study of resident physicians in urban teaching hospitals with excellent interpreter services, residents described a process of risk assessment in which the perceived value of communication was evaluated against their own constraints in terms of the additional time and processes associated with involving a trained interpreter.(31) This process was termed "getting by" and was facilitated by the availability of ad hoc interpreters (generally family members) present with the patient.

1a.4 Citations for Evidence of High Impact cited in 1a.3:

(1) Institute of Medicine. Crossing the Quality Chasm. (Washington, DC: National Academy Press), 2001.
(2) U.S. Bureau of the Census. American Community Survey: language spoken at home (table S1601). 2005.
(3)Derose JP, Bahney BW, Lurie N, Escarce JJ. Review: Immigrants and health care access, quality and cost. Medical Care Research and Review 2009;66:355-408.
(4) U.S. Census Bureau, Selected Social Characteristics in the United States: 2006-2008. American Community Survey, 2009. Available at http://factfinder.census.gov.
(5) National Center for Health Statistics. Health, United States, 2010: With Special Feature on Death and Dying. Hyattsville, MD. 2011.
(6) Institute of Medicine. Unequal treatment: Confronting racial and ethnic disparities in health care. Washington, DC: National Academies Press; 2003.
(7) Baker DW, Parker RM, Williams MV, Coates WC, Pitkin K. Use and Effectiveness of Interpreters in an Emergency Department. Journal of the American Medical Association 1996;275(10):783-788.
(8) Cass A, A Lowell, M Christie, PL Snelling, M Flack, B Marrnganyin, I Brown. Sharing the True Stories: Improving Communication between Aboriginal Patients and Healthcare Workers. Medical Journal of Australia 2002; 176(10):466-70.
(9) Crane JA. Patient Comprehension of Doctor-Patient Communication on Discharge from the Emergency Department. Journal of Emergency Medicine 1997; 15(1):1-7.
(10) Shapiro J, Saltzer E. Cross-Cultural Aspects of Physician-Patient Communications Patterns. Urban Health 1981; (December):10-15.
(11) Kazzi Bonacruz G, Cooper C. Barriers to the Use of Interpreters in Emergency Room Paediatric Consultations. Journal of Paediatric Child Health 2003; 39(4):259-63.
(12) Sarver J, Baker DW. Effect of Language Barriers on Follow-up Appointments after an Emergency Department Visit. Journal of General Internal Medicine 2000; 15(4):256-64.
(13) Ku L, Waidman T. How race/ethnicity, immigration status and language affect health insurance coverage, access to care and quality of care among the low-income population. Final Report. Washington, DC: Kaiser Family Foundation, Publication #4132, 2003.
(14) Karter AJ, Ferrara J, Darbinian LM, Ackerson, JV Selby. Self-monitoring of blood glucose: Language and financial barriers in a

managed care population with diabetes. Diabetes Care 2000;23(4):477-83.
(15) Manson A. Language concordance as a determinant of patient compliance and emergency room use in patients with asthma. Medical Care 1988;26(12):1119-28.
(16) Apter AJ, Reisine ST, Afflect G, Barrows E, ZuWallack RI. Adherence with twice daily dosing of inhaled steroids. American Journal of Respiratory and Critical Care Medicine 1988:157:1810-17.
(17) Lasater LM, Davidson AF, Steiner JF, Mehler PS. Glycemic control in English- vs. Spanish-speaking Hispanic patinets with Type 2 Diabetes Mellitus. Archives of Internal Medicine 2001;161:77-82.
(18) Jacobs EA, Lauderdale DS, Meltzer D, Shorey JM, Levinson W, Thisted RA. Impact of interpreter services on delivery of health care to limited-English proficient patients. Journal of General Internal Medicine 2001;16:468-74.
(19) Tocher TM, Larson E. Quality of diabetes care for non-English-speaking patients. A comparative study. Western Journal of Medicine 1998;168:504-11.
(20) David RA, Rhee B. The impact of language as a barrier to effective health care in an underserved urban Hispanic community. Mt Sinai J Med 1998;Oct/Nov 65(5,6):393-397.
(21) Kuo D, Fagan MJ. Satisfaction with methods of Spanish interpretation in an ambulatory care clinic. Journal of General Internal Medicine 1999;14(9):457-50.
(22) Hornberger JC, Gibson CD, Wood W, Dequeldre C, Corso I, Palla B, Bloch DA. Eliminating language barriers for non-English-speaking patients. Medical Care 1996;34:845-56.
(23) Flores G, Laws MB, Mayo SJ, Zuckerman B, Abreu M, Medina L, Hardt EJ. Errors in Medical Interpretation and their Potential Clinical Consequences in Pediatric Encounters. Pediatrics 2003;111(1):6-14.
(24) Title VI of the Civil Rights Act of 1964 (42 U.S.C. § 2000d); 65 FR 50121, August 16, 2000.
(24) Flores G, Torres S, Holmes LJ, Salas-Lopez D, Youdelman MK, Tomany-Korman SC. Access to hospital interpreter services for limited English proficient patients in New Jersey: a statewide evaluation. J Health Care Poor Underserved 2008;19(2):391-415.
(26) Regenstein M. Measuring and improving the quality of hospital language services: insights from the Speaking Together collaborative. JGIM 2007;22 Suppl 2:356-9.
(27) Leanza Y, Boivin I, Rosenberg E. Interruptions and resistance: a comparison of medical consultations with family and trained interpreters. Soc Sci Med. 2010;70(12):1888-95.
(28) Ebdin P, OJ Carey, Bhatt A, Harrison B. The bilingual consultation. Lancet 1988;1:347.
(29) Marcos LR. Effects of interpreters on the evaluation of psychopathology in non-English-speaking patients. American Journal of Psychiatry 1979;136:171-74.
(30) Leman P. Interpreter use in an inner city accident and emergency department. Journal of Accident and Emergency Medicine 1997;14:98-100.
(31) Diamond LC, Schenker Y, Curry L, Bradley EH, Fernandez A. Getting By: Underuse of Interpreters by Resident Physicians. Journal of General Internal Medicine 2008;24(2):256-62.

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:
The measure can lead to improvements in quality in terms of timeliness of interpreter services. Measuring the percent of encounters when the patient waits 15 minutes or less can set a benchmark for improvement, identify particular locations or services where wait times are particularly long, and can reveal disparities in terms of the timeliness of services for patients speaking certain languages.

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.*]
One of the few data sources demonstrating a performance gap related to timeliness of language services is the results from the Speaking Together collaborative (1). Overall, hospitals in the collaborative performed very well on the measure, with 90% of the patients waiting 15 minutes or less for language services provided by on-site interpreters, telephone services, remote simultaneous interpretation or video. Performance varied by language, however. For example, one hospital learned that Chinese and Vietnamese-speaking patients were much less likely to receive timely language services compared to Spanish and Portuguese speaking patients.

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*]
(1) Regenstein M, Huang J, West C, Trott J, Mead H, Andres E. Improving the Quality of Language Services Delivery: Findings

from a Hospital Quality Improvement Initiative. Journal for Health Care Quality, forthcoming March 2012.

1b.4 Summary of Data on Disparities by Population Group: [*For Maintenance –Descriptive statistics for performance results for this measure by population group*]
Patients with LEP have been shown to experience longer wait times for certain services (1). Patients with limited-English proficiency have longer waiting times to see a physician in the emergency department (2) and delays in time to appendectomy (3).

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*]
(1) Betancourt J. Improving Quality and Achieving Equity: A Guide for Hospital Leaders. Boston, MA: The Disparities Solution Center, 2009.
(2) Wilper AP, Woolhandler S, Lasser KE, et al. Waits to see an emergency department physician: U.S. trends and predictors, 1997-2004. Health Aff (Millwood). Mar-Apr 2008;27(2):w84-95.
(3) Bickell NA, Hwang U, Anderson RM, Rojas M, Barsky CL. What affects time to care in emergency room appendicitis patients? Med Care. Apr 2008;46(4):417-422.

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)*:
The proposed measure is a process measure that tracks the extent to which interpreter services are provided in a timely fashion. Timeliness is one of the most commonly tracked measure to determine whether language services are operating in an effective and efficient manner and reaching patients and providers without potentially harmful delays. The proposed measure would provide common definitions and inclusion/exclusion criteria for health care organizations that want to track this measure and possibly compare performance against national or local benchmarks.

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)*:
The proposed measure provides a common way to track timeliness of services. Substantial evidence supports the delivery of language services although the body of evidence does not generally describe the timeliness of the delivery of language services. This tends to be an internal metric used by organizations to determine their own performance in service delivery. A large body of evidence also identifies access barriers faced by patients with patients with limited English proficiency. If services are not timely and patients and providers face excessive delays in the delivery of language services, ad hoc or otherwise untrained interpreters could be used (or no interpreter at all) to faciliate communication.

1c.5 Quantity of Studies in the Body of Evidence *(Total number of studies, not articles)*: The Flores review included 36 studies and the systematic review by Jacobs included over 150 studies, including many identified in the Flores review. Another systematic review (Bauer A, Alegria M. The Impact of Patient Language Proficiency and Interpreter Service Use on the Quality of Psychiatric Care: A Systematic Review. Psychiatr Serv. 2010 August;61(8):765-773) identified 26 studies meeting specific inclusion criteria.

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*): The evidence supporting the use of interpreters to facilitate communication between patients with limited English proficiency and their providers is very strong. Few studies specifically address the timeliness of language services, although this is clearly an important factor in making language services accessible in a health care setting.

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect):* Because of the variability in the design of the studies, and the fact that the field has not used common definitions for the delivery of language services, it is difficult to summarize the consistency of the magnitude and direction of the effect. In general, interpreter services improve safety, quality of the encounter, result in fewer errors, improve use of prevention and primary care services, conserve resources by avoiding repeated tests, and improve patient satisfaction. Timely delivery of language services is critical to assuring access to these important services.

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 net benefit of effective language services is positive -- the evidence is extremely persuasive on this point. However, no common definitions are used and therefore it is difficult to determine what patients are getting when the studies report language services versus no language services. The proposed measure will provide a common metric for the field to measure the extent to which interpreter services are provided in a timely fashion.

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:

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

1c.12 If other, identify and describe the grading scale with definitions: Systematic literature reviews. Evidence was not graded.

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

1c.14 Summary of Controversy/Contradictory Evidence: N/A

1c.15 Citations for Evidence other than Guidelines*(Guidelines addressed below)*:
N/A

1c.16 Quote verbatim, the specific guideline recommendation *(Including guideline # and/or page #)*:
JOINT COMMISSION STANDARDS FOR PATIENT-CENTERED COMMUNICATION

PC.02.01.21 The hospital effectively communicates wtih patients when providing care, treatment and services.

Rationale for PC.02.01.21
Identifying the patient´s oral and written communication needs is an essential step in determining how to facilitate the exchance of information with the patinet during the care process. Patients may have hearing or visual needs, speak or read a language other than English, experience difficulty understanding health information, or be unable to speak due to their medical condition or

treatment. Additionally, some communication needs may change during the course of care. Once the patient´s communication needs are identified, the hospital can determine the best way to promote two-way communication between the patient and his or her providers in a manner than meets the patient´s needs. This standard complements R1.01.01.01 EP 5 (patient´s right to and need for effective communication); R1.01.01.03 EP2 (provision of language interpreting and translation services); and R1.01.01.03 EP 3 (meeting needs of patients wtih vision, speech, hearing, or cognitive impairments).
EP 1
The hospital identifies the patient´s oral and written communication needs, including the patient´s preferred language for discussing health care. Note 1: Examples of communication needs include the need for personal devices such as hearing aids or glasses, language interpreters, communication boards, and translated or plain language materials.
EP 2
The hospital communicates with the patient during the provision of care, treatment and services in a manner that meets the patient´s oral and written communication needs.

RC.02.01.01 The medical record contains information that reflects the patient´s care, tratment and services.
EP 1
The medication record contains the following demographic information: The patient´s communication needs including preferred language for discussing health care.

NATIONAL QUALITY FORUM COMPREHENSIVE FRAMEWORK AND PREFERRED PRACTICES FOR MEASURING AND REPORTING CULTURAL COMPETENCY: A CONSENSUS REPORT (NQF, 2009)
Domain 2: Integration into Management Systems and Operations
Preferred Practice 9: Implement language access planning in any area where care is delivered.
Domain 3: Patient-Provider Communication
Preferred Practice 12: Offer and provide language access resources in the patient's primary written and spoken language at no cost, at all points of contact, and in a timely manner during all hours of operation, and provide both verbal offers and written notices informing patients of their right to receive language assistance services free of charge.
Preferred Practice 13: Determine and document the linguistic needs of a patient or legal guardian at first points of contact, and periodically assess them throughout the healthcare experience.
Preferred Practice 14: Maintain sufficient resources for communicating with patients in their primary written and spoken languages through qualified/competent interpreter resources, such as competent bilingual or multilingual staff, staff interpreters, contracted interpreters from outside agencies, remote interpreting services, credentialed volunteers, and others, to ensure timely and high-quality communication.

1c.17 Clinical Practice Guideline Citation: The Joint Commission. Advancing Effective Communication, Cultural Competences, and Patient- and Family-Centered Care: A Roadmap for Hospitals. Oakbrook Terrace, IL: The Joint Commission, 2010.

1c.18 National Guideline Clearinghouse or other URL:
http://www.jointcommission.org/assets/1/6/aroadmapforhospitalsfinalversion727.pdf

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 <u>developer's assessment</u> of the quantity, quality, and consistency of the body of evidence?

1c.25 Quantity: High 1c.26 Quality: Moderate1c.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.gwumc.edu/sphhs/departments/healthpolicy/dhp_publications/pub_uploads/dhpPublication_3870218A-5056-9D20-3D6DA9069C41BB77.pdf

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)*:
The number of interpreter encounters in which the wait time is 15 minutes or less for the interpreter to arrive

2a1.2 Numerator Time Window *(The time period in which the target process, condition, event, or outcome is eligible for inclusion)*:
The time window is an encounter or point in time. All cases in the denominator are equally eligible to appear in the numerator

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*:
The number of interpreter encounters in which the wait time is 15 minutes or less for the interpreter to arrive.
Inclusions
The number of interpreter encounters in which the wait time is 15 minutes or less for the interpreter to arrive, for encounters provided by:
- On-site interpreter encounters with hospital operated interpreters, and on-site contract and/or agency interpreters
- Hospital operated telephone interpreters
- Hospital operated video interpreters
- Scheduled and unscheduled interpreter encounters

Note: Stratified by language.
Exclusions
- Encounters where the wait time is greater than 15 minutes for interpreter to arrive
- Encounters with bilingual providers and/or other bilingual hospital workers/employees
- Encounters with outside vendor telephone interpreters and/or outside vendor video interpreters.

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured)*:
The total number of interpreter encounters, stratified by language.

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, Maternal Care, Populations at Risk

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

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)*:
The total number of interpreter encounters, stratified by language.
Inclusion
The number of interpreter encounters provided by:
- On-site interpreter encounters with hospital operated interpreters, and on-site contract and/or agency interpreters
- Hospital and outside vendor telephone interpreters
- Hospital and outside vendor video interpreters
- Scheduled and unscheduled interpreter encounters

Note: Stratified by language.

2a1.8 Denominator Exclusions *(Brief narrative description of exclusions from the target population)*:
Exclusions
- Encounters with bilingual providers and/or other bilingual workers/employees
- Encounters with outside vendor telephone interpreters
- Encounters with outside vendor video interpreters

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)*:
Data elements:
Preferred spoken language for health care
Interpreter encounters

Data Collection Approach: Retrospective data sources for required elements include interpreter services department logs and medical records.

Data Accuracy/Data Completeness: Variation may exist in data recording practices; therefore, data recording practices may require evaluation, monitoring and training to ensure consistency.

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)*:
Aggregate measure provides information about timeliness of interpreter services and is a disparities-related measure since long waits for interpreters cause delays only for patients who speak languages other than English. Hospitals can stratify measure by language, location or type of service, mode of interpreting, scheduled versus unscheduled appointments

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.)*:
None

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: 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.)*:
Data calculated as aggregate numerator and denominator, monthly, stratified by language. Includes multiple modes of interpretation; excludes interpreters provided by outside vendors because of difficulties implementing quality improvement change cycles with these service providers.

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:
URL
http://www.gwumc.edu/sphhs/departments/healthpolicy/dhp_publications/pub_uploads/dhpPublication_3870218A-5056-9D20-3D6DA9069C41BB77.pdf

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):
Measure includes all admissions and visits for those requesting interpreter services -- it is not based on a sample or survey.

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

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.*): Data sources for required elements include interpreter services logs and medical records.

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment: URL
http://www.gwumc.edu/sphhs/departments/healthpolicy/dhp_publications/pub_uploads/dhpPublication_3870218A-5056-9D20-3D6DA9069C41BB77.pdf

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:
URL
http://www.gwumc.edu/sphhs/departments/healthpolicy/dhp_publications/pub_uploads/dhpPublication_3870218A-5056-9D20-3D6DA9069C41BB77.pdf

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

2a1.34-35 Care Setting *(Check all the settings for which the measure is specified and tested)*: Ambulatory Care : Clinic/Urgent Care, 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)*:
The measure was pilot tested in one inpatient and in one outpatient care setting in two large metropolitan hospitals October 2006. The measure was used by the 10 grantee hospitals in the Speaking Together National Language Services Collaborative from November 2006 - May 2008. Ten hospitals reported data monthly on 40,000 - 60,000 patients seen in inpatient and ambulatory

care settings. Hospitals ranged in size from 11,500 - 44,000 admissions, included 2 children´s hospitals and were comprised of both academic teaching and non-teaching community hospitals.
The measures specifications were revised based on the learning from the Speaking Together Collaborative and input from the participating hospitals.

- Hospital A - New York NY; Public hospital; Beds 771; Annual Admissions 26.068; Annual interpreter encounters 58,962; Percent of interpreter encounters in top 5 languages-60% Spanish;26% Mandarin; 6% Cantonese; 3%Polish; 2% French

- Hospital B - Cambridge MA; Public hospital; Beds 350; Annual Admissions 15,263; Annual interpreter encounters 140, 556; Percent of interpreter encounters in top 5 languages-55% Brazilian Portuguese; 24% Spanish; 7% Haitian Creole; 2% Euro Portuguese; 2% Hindi

- Hospital C - Minneapolis MN; Public hospital; Beds 434; Annual Admissions 22,117; Annual interpreter encounters 120,000; Percent of interpreter encounters in top 5 languages-60% Spanish; 12% Somali; 4% Russian; 3% Hmong; 3% Hmong; 1% Laotian

- Hospital D - Phoenix AZ; Non-profit hospital; Beds 285; Annual Admissions 11,712; Annual interpreter encounters 48,043; Percent of interpreter encounters in top 5 languages->99% Spanish

- Hospital E - St. Paul, MN; Non-profit hospital; Beds 399; Annual Admissions 22,827; Annual interpreter encounters 28,887; Percent of interpreter encounters in top 5 languages- 50% Spanish; 12% Hmong; 10% Somali; 9% Vietnamese; 4% ASL

- Hospital F – Rochester, NY; Non-profit hospital; Beds 973; Annual Admissions 36,321; Annual interpreter encounters 14,885; Percent of interpreter encounters in top 5 languages-46$ Spanish; 35% ASL; 3% Vietnamese; 2% Russian ; 2% Arabic

- Hospital G – Seattle, WA; Non-profit hospital; Beds 250; Annual Admissions 11,608; Annual interpreter encounters 40,690; Percent of interpreter encounters in top 5 languages-55% Spanish; 7% Vietnamese; 4% Somali; 4% Russian; 2% Cantonese

- Hospital H – Sacramento, CA; Public hospital; Beds 526; Annual Admissions 27,946; Annual interpreter encounters 65,000; Percent of interpreter encounters in top 5 languages-58% Spanish;20%Russian; 8% Mien; 5% Hmong; 5% Cantonese
- Hospital I – Worcester, MA; Non-profit hospital; Beds 731; Annual Admissions 44,231; Annual interpreter encounters 59,134; Percent of interpreter encounters in top 5 languages-62% Spanish; 13% Portuguese; 7% Vietnamese; 5% Albanian; 3% ASL

- Hospital J – Ann Arbor, MI; Non-profit hospital; Beds 802; Annual Admissions 42,811; Annual interpreter encounters 21,503; Percent of interpreter encounters in top 5 languages-22% Spanish; 18% Chinese; 14% Japanese; 12% Arabic; 10% Russian

2a2.2 Analytic Method *(Describe method of reliability testing & rationale)*:
Development of Interpreter Services Performance Measures
In 2006, the Robert Wood Johnson Foundation funded Speaking Together: National Language Services Network, an 18-month national program aimed at improving the
delivery of language services through the use of quality improvement techniques. Ten hospitals were selected through an open, competitive solicitation to participate in the program. The 10 hospitals were: Bellevue Hospital Center (New York, NY); Cambridge Health Alliance (Cambridge, MA); Hennepin County Medical Center (Minneapolis, MN); Phoenix Children's Hospital (Phoenix, AZ); Regions Hospital (St. Paul, MN); The University of Rochester—Strong Memorial Hospital (Rochester, NY); Seattle Children's Hospital and Medical Center, (WA); the University of California Davis Medical Center (Sacramento, CA); the University of Massachusetts Memorial Medical Center (Worcester, MA); and, University of Michigan Health System (Ann Arbor, MI).

Because the field of language services did not have commonly used language performance measures, the Speaking Together National Program Office (NPO) at the George Washington University developed a set of performance measures for language services for use throughout the learning collaborative. As a starting point for measures development for the field, the Speaking Together NPO made an explicit decision to initially focus on signed and spoken interpreter services measures with a plan to develop measures for written (translation) services at a later date. The Speaking Together NPO employed a multi-stage process to

identify and develop a set of measures for signed and spoken interpreter services:

Stage 1: Identifying a framework for quality: The Speaking Together NPO used the Institute of Medicine's (IOM's) six dimensions of quality, as articulated in Crossing the quality chasm: A new health system for the 21st century, as a framework for developing language service performance measures. These dimensions (safety, timeliness, effectiveness, efficiency, equity, and patient-centeredness) are outlined in Figure 1.

Figure 1: IOM Domains of Quality, Adapted for Language Services
Domain Principle
Safe: Avoiding injuries to patients from the language assistance that is intended to help them.
Timely: Reducing waits and sometimes harmful delays for both those who receive and those who give care.
Effective: Providing language services based on scientific knowledge that contribute to all who could benefit, and refraining from providing services to those not likely to benefit.
Efficient: Avoiding waste, including waste of equipment, supplies, ideas, and energy.
Equitable: Providing language assistance that does not vary in quality because of personal characteristics such as language preference, gender, ethnicity, geographic location, and socioeconomic status.
Patient-Centered: Providing language assistance that is respectful of and responsive to individual patient preferences, needs, culture and values, and ensuring that patient values guide all clinical decisions.

Stage 2: Reviewing the relevant literature: The Speaking Together NPO conducted extensive literature searches to support the development of evidence-based measures and identify key quality concerns related to the delivery of language services in hospitals and other health care settings.

Stage 3: Interviewing experts: The Speaking Together NPO interviewed experts in the field of language services and directors of established hospital-based interpreter services programs to help identify issues related to quality of language services and potentially valuable performance measures. For a full listing of the contributors, please see Additional.

Stage 4: Identifying a framework for organizational change: The Speaking Together NPO used Nerenz and Neil's Performance Measures for Health Care Systems (2001) as a guidepost to look across an organization and identify how care is organized and delivered. Using this framework, we identified components of language and interpreter services that address significant and important quality issues pertinent to the delivery of language services and identified measurable events as potentially valuable performance measures.

Stage 5: Developing the measures: Using the frameworks mentioned above, as well as information from the literature and interviews, the Speaking Together NPO developed a set of 10 draft process measures for review and field testing.

Stage 6: Getting feedback on the draft measures: The Speaking Together NPO assembled a panel of experts in language services, who have contributed greatly to the literature in the field, to review the 10 draft performance measures and evaluate them according to uniform evaluation criteria.

Stage 7: Meeting with clinicians and interpreters services directors: The draft measures were reviewed by an expert panel consisting of medical directors, physician leaders and interpreter services directors who convened in Washington, DC, in September 2006 to review the 10 draft measures and evaluate each according to its importance to quality, feasibility in terms of data collection, clarity and accuracy of description. (For a full listing of the contributors, please see Additional.) The expert panel recommended the following 5 of the 10 measures for implementation in acute care hospitals and outpatient settings:

- The percent of patients who have been screened for their preferred spoken language.
- The percent of LEP patients receiving initial assessment and discharge instructions from assessed and trained interpreters or from bilingual providers assessed for language proficiency.
- The percent of encounters where the patient wait time for interpreter is 15 minutes or less.
- The percent of time interpreters spend providing medical interpretation in clinical encounters with patients.
- The percent of encounters interpreters wait less than 10 minutes to provide interpreter services to provider and patient.

Stage 8: Field testing the measures: Two hospitals with established language services programs participated in a week-long pilot test of the recommended performance measures, gathering information on the feasibility of data collection, usefulness of data

reporting formats, and barriers and challenges associated with successful data collection and submission. (Please note: The two pilot sites were not part of the 10 Speaking Together grantee hospitals.)

Stage 9: Implementing the measures: The 10 Speaking Together grantee hospitals used the measures throughout the 18-month learning collaborative, applying quality improvement methodologies to improve the delivery of interpreter services. The Speaking Together hospitals reported data (stratified by language) on the measures to the NPO monthly for the duration of the 18-month program. Hospitals also provided information about data collection challenges, feedback on the data abstraction instructions, data variables and definitions in monthly reports, at on-site visits with the NPO, during monthly conference calls, and at the 4 collaborative meetings.

Stage 10: Revising and refining data collection specifications: The NPO revised the measures based on the learnings from the Speaking Together collaborative then convened a panel of language services experts to review the measures revisions for clarity and accuracy of descriptions, definitions and abstraction instructions. The panel was comprised of medical directors and quality improvement specialists from 5 Speaking Together hospitals. (For a full listing of the contributors, please see Additional.) Revisions to the 5 measures were largely centered on clarifying numerator and denominator descriptions, clarifying inclusions and exclusions descriptions and defining data elements. The work in this stage has allowed us to standardize the measures and to create standardized technical specifications.

The measures were also used in another quality improvement project, the Aligning Forces for Quality: Language Quality Improvement Collaborative
From July 2009-October 2010, the measures were used in the Aligning Forces for Quality Language Quality Improvement Collaborative (LQIC). As in Speaking Together, the LQIC hospitals reported monthly data, stratified by language, on the measures to the NPO. Hospitals also provided information about data collection challenges, feedback on the data abstraction instructions, data variables and definitions in monthly reports, at on-site visits with the NPO, during monthly conference calls, and at 2 collaborative meetings. The 9 LQIC hospitals were: Beaumont Hospitals (Royal Oak, MI)l; Central Maine Medical Center (Lewiston, ME); Cincinnati Children's Hospital (Cincinnati, OH); Harborview Medical Center (Seattle, WA); Mercy Hospital—State Street Campus (Portland, ME); Oakwood Hospital & Medical Center (Dearborn, MI); St. Joseph Hospital (Eureka, CA); St. Joseph Mercy Oakland—Trinity Health (Pontiac, MI); and, Valley Medical Center (Renton, WA).

References:
Graham, C., Ivey, S.L., Neuhauser, L. From Hospital to home: Assessing the transitional care needs of vulnerable seniors. The Gerontologist. Feb 2009: 49(1): 23-33

Institute of Medicine. Crossing the quality chasm: A new health system for the 21st century. National Academies Press. 2001: 51-53.

Nerenz, D. and N. Neil. Performance Measures for Health Care Systems. Commissioned paper for the Center for Health Management Research, 2001.

Regenstein, M., Huang, J., West, C., Mead, H., Trott, J., Stegun, M. Hospital language services: Quality improvement and performance measures. Advances in Patient Safety: New Directions and Alternative Approaches. Agency for Healthcare Research and Quality. Rockville, MD. July 2008; Vol. 1-4: AHRQ Publication Nos. 08-0034 (1-4).

2a2.3 Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
In the Speaking Together collaborative, performance on the proposed measure was consistently high for the majority of participants from baseline, with two hospitals demonstrating improvement by 5 or more percentage points during the project. More than 90 percent of the time, patients waited 15 minutes or less for language services, provided via on-site interpreter, telephonic services, remote simultaneous interpretation, or video. Much of the improvement on L3 stemmed from an examination of timeliness across languages and attempts to bring timely services to patients, regardless of the language spoken. For example, by tracking L3 performance, one hospital learned that Chinese and Vietnamese speaking patients were much less likely to receive timely language services compared to Speanish and Portuguese speaking patients. As a result of initiatives such as revising interpreter schedules based on peak times and increasing access to remote interpreting methods for infrequent languages, nights and weekends, performance for all language groups was consistently high. By the end of the collaborative, nearly 100 percent of encounters with Vietnamese speaking patients occurred with the patient waiting 15 minutes or less.

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:
The proposed measure has been tested across different types of hospitals, with different language services staffing arrangements and different practices in terms of the mode of interpretation used. The measure focuses on patients with limited English proficiency who use interpreter services and the extent to which they receive timely language services.
Prior to the development of the measure, no standards existed in the language services field to define timely language services. The measure development process included lengthy discussions with experts in the field, language services professionals and clinicians to determine the waiting period that would be acceptable in terms of a quality measurement.
The proposed measure has high face validity. Hospital staff who reported on the measure view it as measuring precisely what it intends to measure -- that is, the percent of patients who have wait times for an interpreter of 15 minutes or less. Hospitals in the Speaking Together collaborative and the subsequent round of quality improvement work in the Aligning Forces for Quality LQIC considered the measure -- set at a 15-minute wait time as indicating timely services -- to be an appropriate and reasonable target.
The proposed measure has been accepted as part of the AHRQ National Measures Clearinghouse and has been used in two rounds of quality improvement work in inpatient and outpatient settings. The proposed measure has strong construct validity and is consistent with data collection processes related to the delivery of language services.

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)*:
Participants in the Speaking Together collaborative were interviewed throughout the learning network about their progress with the performance measures, including the proposed measure. All participants viewed the proposed measure as accurately reflecting the effectiveness of language services delivery. Ten hospitals participated in Speaking Together and reported performance for more than 50,000 patients with limited English proficiency. At the beginning of the collaborative, the ten hospitals reported that they had approximately 250 FTE staff devoted to interpreter services. This provided a large group of interpreters and a large group of patients to test the measure´s validity, usefulness and reliability.

2b2.2 Analytic Method *(Describe method of validity testing and rationale; if face validity, describe systematic assessment)*:
As indicated above, the proposed measure has strong face validity. The measure was developed after a thorough revuew if the literature, structured interviews with providers of language services in hospitals and health systems (generally directors of interpreter services) as well as the clinical staff who worked with interpreters (generally directors of ambulatory services or other service lines). We convened an expert panel of interpreter services and ambulatory service directors using a Delphi panel to systematically review the proposed measure on specific review criteria. We pilot tested the measure, as approved by the expert panel, in one large acute care hospital with substantial numbers of staff interpreters and high demand for language servics and a children´s hospital with similar characteristics. No substantial changes to the proposed measure were required following the pilot test. We used the proposed measure throughout the Speaking Together learning network. Following the Speaking Together learning network, we convened representatives from our initial expert panel as well as some Speaking Together participants to review the validity, usefulness and adequacy of the proposed measure. The group strongly supported the use of the proposed measure with no substantive modifications.

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)*:
We did not conduct statistical tests of the adequacy of the measure.

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)*:
For the numerator, we excluded encounters where the wait time for the interpreter to arrive was greater than 15 minutes. We also excluded encounters with bilingual providers/staff and encounters with outside vendor telephone interpreters and outside video interpreters. For the denominator, we excluded encounters with bilingual providers/staff and encounters with outside vendors of telephone and video interpreting. We did not track excluded encounters.

2b3.2 Analytic Method (*Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference)*:
We did not track or analyze excluded encounters. Contract telephone and video interpretations were excluded because they involve interpreters who are not employed by the hospitals. The measure focuses on the timeliness of interpreter services and excludes language services provided by bilingual clinicians or staff.

2b3.3 Results *(Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):*
We did not perform analyses on excluded encounters.

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)*:
No risk adjustment was performed. This is a process measure that includes all interpreter encounters that meet the inclusion criteria.

2b4.2 Analytic Method (*Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables)*:
No risk adjustment was used. Performance was stratified by language, which enabled health care organizations to determine whether services were timely for patients speaking languages other than English, as a group of LEP patients, and by the individual language spoken.

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)*:
No risk adjustment was used.

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment: No need for risk adjustment. This is a process measure that applies to all interpreters and LEP patient encounters who meet the measure inclusion/exclusion criteria. The measure does not differentiate between modes of interpretation (as long as they are provided by hospital-based interpreters). Hospital-based telephone interpreters may have higher rates of timely language services delivery compared to on-site interpreters.

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)*:
In the Speaking Together collaborative, performance on the proposed measure was consistently high for the majority of participants from baseline, with two hospitals demonstrating improvement by 5 or more percentage points during the project. More than 90 percent of the time, patients waited 15 minutes or less for language services, provided via on-site interpreter, telephonic services, remote simultaneous interpretation, or video. Performance improvement on the measure was dramatic for the proposed measure. At the beginning of the collaborative, performance ranged from a low of about 25 percent at one hospital to a high of over 90 percent at two hospitals. By the end of the collaborative, even the lowest performing hospitals had wait times of 15 minutes or less for nearly 90 percent of patients.

Much of the improvement on L3 stemmed from an examinatino of timeliness across languages and attempts to bring timely services to patients, regardless of the language spoken. For example, by tracking L3 performance, one hospital learned that Chinese and Vietnamese speaking patients were much less likely to receive timely language services compared to Speanish and Portuguese speaking patients. As a result of initiatives such as revising interpreter schedules based on peak times and increasing access to remote interpreting methods for infrequent languages, nights and weekends, performance for all language groups was consistently high. By the end of the collaborative, nearly 100 percent of encounters with Vietnamese speaking patients occurred with the patient waiting 15 minutes or less.

2b5.2 Analytic Method *(Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance)*:
We tracked performance on a monthly basis across the 10 hospitals. We considered a hospital to have improved performance if it showed improvement of at least five percentage points.

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)*:
Performance on L3 was consistently high for the majority of participants from baseline, with two hospitals demonstrating improvement by 5 or more percentage points during the collaborative. More the 90 percent of the time, patients waited 15 minutes or less for language services (provided via on-site interpreter, telephonic services, remote simultaneous interpretation, or video). Much of the impovement on L4 stemmed from an examination of timeliness acorss languages and attempts to bring timely services to patients, regardless of the language spoken. For example, by tracking L3 performance, one hospital laerned that Chinese and Vietnamese speaking patients were much less likely to receive timely language services compared to Spanish and Portuguese speaking patients. As a result of initiatives such as revising interpreting schedules based on peak times and increasing access to remote interpreting methods for infrequent languages, nights and weekends, performance for all language groups was consistently high. By the end of the collaborative, nearly 100 percent of encounters with Vietnamese speakng patients ooccured with the patient waiting 15 minutes or less.

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)*:
All interpreters and encounters that met the inclusion criteria were in the proposed measure.

2b6.2 Analytic Method *(Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure)*:
Rates were calculated for each participating hospital by quarter/year. All hospital scores were available to all hospitals in the collaborative -- no additional testing for comparability of scores was conducted.

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)*:
Performance within the collaborative was highly transparent, with hospitals able to view other hospitals´ performance to allow them to learn strategies and techniques from more successful performers. No additional statistical analyses were conducted.

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)*: Hospitals stratified the results of the measure by language. We did not track performance by language at the program office level.

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

2.1-2.3 Supplemental Testing Methodology Information:

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)

3.1 Current Use *(Check all that apply; for any that are checked, provide the specific program information in the following questions)*: Public Reporting, Regulatory and Accreditation Programs, 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.*]
The measure is not publicly reported. Endorsement from NQF could establish a common platform for public reporting as a quality measure.

The results of the collaborative and the proposed measure can be viewed in the follow publicly available publications:

Regenstein M, Huang J, West C, Trott J, Mead H, Andres E. Improving the quality of language services delivery: Findings from a hospital quality improvement initiative. Journal for Healthcare Quality 2012: Forthcoming, March 2012.

Regenstein, M. (2007). Measuring and improving the quality of hospital language services: insights from the Speaking Together collaborative. Journal of General Internal Medicine, 22 Suppl 2, 356-59.

Regenstein, M., Huang, J., West, C., Mead, H., Trott, J. & Stegun, M. (2008). Hospital language services: quality improvement and performance measures. Advances in patient safety: new directions and alternative approaches. Rockville, MD: Agency for Healthcare Research and Quality.

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 measure provides information about the wait times patients experience when accessing language services resources.The timeliness of these resources is absolutely critical to the provision of high-quality care for persons with limited English proficiency. However, most hospitals and other health care organizations do not know (and do not track in any consistent way) whether their language services resources are timely and this can reveal disparities in terms of the timeliness of services for patients speaking certain languages.

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): The Joint Commissions Standards for Patient Provider Communication

http://www.jointcommission.org/assets/1/6/ARoadmapforHospitalsfinalversion727.pdf

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*].
Organizations were able to use the measures to improve the delivery of interpreter services to LEP populations. Across 3 collaboratives/learning networks (Speaking Together; Aligning Forces for Quality Languguage Quality Improvement Collaborative; and the Aligning Forces for Quality Language Hospital Quality Network Improving Language Servcies program)organizations have used the measures and results to better identify patients needing services and identifythe organizations true demand for servcies;

measures progress towards a delivery services; and use measures to identify waste and streamline systems.

Results are documented in:
Regenstein M, Huang J, West C, Trott J, Mead H, Andres E. Improving the quality of language services delivery: Findings from a hospital quality improvement initiative. Journal for Healthcare Quality: Forthcoming, March 2012.

Regenstein, M. (2007). Measuring and improving the quality of hospital language services: insights from the Speaking Together collaborative. Journal of General Internal Medicine, 22 Suppl 2, 356-59.

Regenstein, M., Huang, J., West, C., Mead, H., Trott, J. & Stegun, M. (2008). Hospital language services: quality improvement and performance measures. Advances in patient safety: new directions and alternative approaches. Rockville, MD: Agency for Healthcare Research and Quality.

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:
The measures have been tested extensively and feedback has been sought on several occasions from individuals who are most familiar with on-the-ground interpreter services in hospital settings. Their feedback indicates that they are meaningful, easy to understand, and extremely useful for quality improvement.

As part of the Speaking Together program, we commissioned a series of focus groups with patients at the participating hospitals. On average, 3 focus groups were held with different language groups who used or would potentially use language services to facilitate their communication with the participating health care organization. Patients clearly indicated that they recognize the need for high-quality language services and that they understand that their care and safety are compromised without these services. They identified barriers to receiving these services at all of the hospitals.

Organizations were able to use the measures to improve the delivery of interpreter services to LEP populations. Across 3 collaboratives/learning networks (Speaking Together; Aligning Forces for Quality Language Quality Improvement Collaborative; and the Aligning Forces for Quality Language Hospital Quality Network Improving Language Services program) organizations have used the measures and results to better identify patients needing services and identify the organizations true demand for services; measure progress towards delivery services; and use measures to identify waste and streamline systems.

Results are documented in:
Regenstein M, Huang J, West C, Trott J, Mead H, Andres E. Improving the quality of language services delivery: Findings from a hospital quality improvement initiative. Journal for Health Care Quality: Forthcoming, March 2012.

Regenstein, M. (2007). Measuring and improving the quality of hospital language services: insights from the Speaking Together collaborative. Journal of General Internal Medicine, 22 Suppl 2, 356-59.

Regenstein, M., Huang, J., West, C., Mead, H., Trott, J. & Stegun, M. (2008). Hospital language services: quality improvement and performance measures. Advances in patient safety: new directions and alternative approaches. Rockville, MD: Agency for Healthcare Research and Quality.

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,

Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry), Other
Retrospective data sources for required elements include interpreter services department logs and medical records

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: Organizations keep these data in a variety of places. In some organizations, clinicians place an order for the interpreter (similar to ordering any other medical treatment)in the computer order entry system. This effectively record the time of order. Interpreters or clinicians record of interpreting in the record. Recording of time of request and when interpreter available varies when telephonic or video interpretation is used. Some organizations do not have electronic order entry and dispatchers or other administrative personnel record the time of interpreter request on electronic spreadsheets at a central organization.

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:
The measure is straight forward and instructions were provided in group settings and 1:1 on the rationale for the measures, specific variables, definitions and calculation of the measure. A detailed specifications manual was created.

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)*:
These are new data collection and not required for any other programs.
Organizations may not have all the data elements for this measure and need to collaborate with clinicians and with information technology departments to create fields for collecting and for generating reports on measure performance. Paper tracking was most often used for this measure with the use of various records and tracking forms to generate the measure. The data for on-site interpreters, over the phone/vide and contract / agency interpreters was often kept indifferent reports. Once systems are in place the data collection and reporting take minimal time.

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)*: N/A

CONTACT INFORMATION
Co.1 Measure Steward (Intellectual Property Owner): Department of Health Policy, The George Washington University, 2121 K Street NW, Suite 200, Washington, District Of Columbia, 20037 Co.2 Point of Contact: Catherine, West, MS, RN, cathy.west@gwumc.edu, 201-994-8663-
Co.3 Measure Developer if different from Measure Steward: Department of Health Policy, The George Washington University, 2121 K Street NW, Suite 200, Washington, District Of Columbia, 20037 Co.4 Point of Contact: Catherine, West, MS, RN, cathy.west@gwumc.edu, 201-994-8663-
Co.5 Submitter: Catherine, West, MS, RN, cathy.west@gwumc.edu, 201-994-8663-, Department of Health Policy, The George Washington University
Co.6 Additional organizations that sponsored/participated in measure development: Stage 8: Interpreter Measures Measure Field Test Hospitals Boston Medical Center Children's Hospital of Phildelphia
Co.7 Public Contact: Catherine, West, MS, RN, cathy.west@gwumc.edu, 201-994-8663-, Department of Health Policy, The George Washington University

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. Measure Contributor List Lead Developer of Language Services Performance Measures Marsha Regenstein, Ph.D., M.C.P. Department of Health Policy The George Washington University Stages 1, 2, 3, 4, and 5: Interpreter Measures Speaking Together Program Staff Jennifer Huang, M.S.

Research Scientist

Holly Mead, Ph.D.
Assistant Professor

Marsha Regenstein, Ph.D., M.C.P.
Director, Speaking Together

Jennifer Trott, M.P.H.
Research Associate

Catherine West, M.S., R.N.
Senior Research Scientist

Stage 6: Interpreter Measures
Speaking Together Performance Measures Reviewers and Contributors
Wilma Alvarado-Little
University at Albany, SUNY
Albany, NY

Oscar Arocha, M.M.
Boston Medical Center
Boston, MA

Rochelle Ayala, M.D.
Memorial Healthcare System
Hollywood, FL

Sang-ick Chang, M.D.
San Mateo Medical Center
San Mateo, CA

Lou Hampers, M.D., M.P.H.
The Children's Hospital Denver
Denver, CO

Anita Hunt
Regional Medical Center at Memphis
Memphis, TN

Wendy Jameson
California Health Care Safety Net Institute
Oakland, CA

Bret A. McFarlin, D.O.
Broadlawns Medical Center
Des Moines, IA

Gloria Garcia Orme, RN, MS
San Francisco General Hospital
San Francisco, C.A.

Melinda Paras
CEO, Paras and Associates

Albany, CA

Martine Pierre-Louis, M.P.H.
Harborview Medical Center
Seattle, WA

Angelique Ramirez, M.D.
Parkland Health & Hospital System
Dallas, TX

Cynthia Roat
National Council on Interpreting in Health Care

Matt Wynia, M.D., M.P.H.
American Medical Association

Speaking Together Advisors

Bruce Siegel, M.D., M.P.H.
George Washington University
Washington, DC

Richard A. Wright, M.D., M.P.H., F.A.C.P.E.
Wright Consulting

Stage 7: Interpreter Measures

Speaking Together: Expert Clinicians and Interpreter Service Directors

Oscar Arocha, M.M
Boston Medical Center
Boston, MA

Rochelle Ayala, M.D.
Memorial Healthcare System
Hollywood, FL

Sang-ick Chang, M.D.
San Mateo Medical Center
San Mateo, CA

Anita Hunt
Regional Medical Center at Memphis
Memphis, TN

Bret A. McFarlin, D.O.
Broadlawns Medical Center
Des Moines, IA

Gloria Garcia Orme, RN, MS
San Francisco General Hospital
San Francisco, C.A.

Martine Pierre-Louis, M.P.H.
Harborview Medical Center
Seattle, WA

Angelique Ramirez, M.D.
Parkland Health & Hospital System
Dallas, TX

Stage 10: Interpreter Measures

Language Services Performance Measures: Reviewers and Contributors

Dena Brownstein, MD
Seattle Children's Hospital
Seattle, WA

Maribet McCarty, PhD, RN
Regions Hospital
St. Paul, MN

Kathy Miraglia, MS
University of Rochester Medical Center
Rochester, NY

Sally Moffat, RN
Phoenix Children´s Hospital
Phoenix, AZ

Sarah Rafton, MSW
Children´s Hospital & Regional Medical Center
Seattle, WA

Loretta Saint-Louis, PhD
Cambridge Health Alliance
Somerville, MA

Sidney Van Dyke, MA
Regions Hospital
St. Paul, MN

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: 2006
Ad.4 Month and Year of most recent revision: 08, 2009
Ad.5 What is your frequency for review/update of this measure? Annula
Ad.6 When is the next scheduled review/update for this measure? 06, 2012

Ad.8 Disclaimers:

Ad.9 Additional Information/Comments: The measures were accepted for the NQMC Web site and are at http://www.qualitymeasures.ahrq.gov/about/inclusion-criteria.aspx. This NQMC summary was completed by ECRI Institute on May

17, 2010. The information was verified by the measure developer on July 2, 2010.
Date of Submission (*MM/DD/YY*): 01/17/2012