**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.

<table>
<thead>
<tr>
<th>NQF #: 1824</th>
<th>NQF Project: Healthcare Disparities Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>(for Endorsement Maintenance Review)</td>
<td><strong>Original</strong> Endorsement Date: <strong>Most Recent</strong> Endorsement Date:</td>
</tr>
</tbody>
</table>

### BRIEF MEASURE INFORMATION

**De.1 Measure Title:** L1A: Screening for preferred spoken language for health care

**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 patient visits and admissions where preferred spoken language for health care is screened and recorded.

Hospitals cannot provide adequate and appropriate language services to their patients if they do not create mechanisms to screen for limited English-proficient patients and record patients’ preferred spoken language for health care. Standard practices of collecting preferred spoken language for health care would assist hospitals in planning for demand. Access to and availability of patient language preference is critical for providers in planning care. This measure provides information on the extent to which patients are asked about the language they prefer to receive care in and the extent to which this information is recorded.

**2a1.1 Numerator Statement:** The number of hospital admissions, visits to the emergency department, and outpatient visits where preferred spoken language for health care is screened and recorded

**2a1.4 Denominator Statement:** The total number of hospital admissions, visits to the emergency department, and outpatient visits.

**2a1.8 Denominator Exclusions:** There are no exclusions. All admissions, visits to the emergency department, and outpatient visits, including:

- Scheduled and unscheduled visits
- Elective, urgent and emergent admissions
- Short stay and observation patients
- Transfers from other facilities

**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):**
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.**

### 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):**

Immigrants comprise a large and growing segment of American society that is disproportionately low-income and uninsured (1-2). 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) (3). 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 (4) and infant mortality and are more likely to report risk factors for serious and chronic diseases such as diabetes and heart disease (5). 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 (6) and why they receive particular types of care (7-8). Patients with LEP are particularly vulnerable to miscommunication when discharged from the emergency department (9-10), and have poorer follow-up after an emergency department visit (11). LEP patients have been shown to have poorer compliance (12) and lower adherence with diabetes (13) and asthma care regimens (14-15) compared to patients who are English speakers; they also have poorer diabetes outcomes (16).

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. (17) 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 (18). LEP patients who used any interpreter were more likely than English-speaking patients to have had a mammogram over a two-year period (19). 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 (17). Patients with language barriers indicated higher levels of satisfaction with care when interpreters were used (20). Physicians who had access to the services of trained interpreters reported a significantly higher quality of patient-physician communication than physicians without these services (21).

Medical interpreters can bridge the communication gap between physician and patient (22), 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 (6). Federal civil rights legislation (23) 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) is variable and unpredictable (24).
Merely having interpreters available in a health care setting does not mean that the patients who need services will receive them. 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.

Perhaps the most substantial barrier to more widespread use of interpreters for LEP patients is the cost of the service. Studies on the marginal costs of interpreters in the health care setting are scarce, as are estimates of the overall "value" that the use of interpreters brings to the health care encounter. Studies estimate the cost of language services to be low relative to other health care costs, approximately $234-$279 per patient per year in inpatient and outpatient settings, respectively (26-27). Language services have also been shown to reduce the cost of emergency department care (28).

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 (29). 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 (22). Ad hoc interpreters can misinterpret or omit questions asked by physicians (30); family members who interpret sometimes leave the patient out of the discussion altogether, instead answering the physician's questions without consulting the patient (31).

Physicians recognize the need for trained interpreters (32) 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. (33) This process was termed “getting by” and was facilitated by the availability of ad hoc interpreters (generally family members) present with the patient.

Critical to providing effective language services is the need to identify those individuals who require such services. Substantial work has been undertaken over the past several years to identify the best ways to collect patient data on race, ethnicity and language. The proposed measure seeks to determine language preference as the principal mechanism to identify need for language services (34-35).

(9) Shapiro J, Saltzer E. Cross-Cultural Aspects of Physician-Patient Communications Patterns. Urban Health 1981;
(December):10-15.
(16) Lasater LM, Davidson AF, Steiner JF, Mehler PS. Glycemic control in English- vs. Spanish-speaking Hispanic patients with Type 2 Diabetes Mellitus. Archives of Internal Medicine 2001;161:77-82.

1b. Opportunity for Improvement: H M L I
### 1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:

Any benefits that flow from providing effective language services, through the use of a trained interpreter or a qualified bilingual provider, first require the identification of need for services. While health care organizations generally have some type of mechanism to recognize a patient who might need an interpreter, few systematically record this information, which creates opportunities for huge gaps in care.

In the Speaking Together project, hospitals generally performed extremely well on the language screening measure. This was an expected finding, because the selection process for participation in Speaking Together favored hospitals that had systems in place to screen for preferred language and capacity for data collection related to the use of language services. Nevertheless, some of the hospitals showed improvement on this measure. At one hospital, at the beginning of the collaborative, 60% of patients were screened for language preference. Through a combination of efforts such as using data to open a discussion with the leaders of registration and scheduling; training staff on the how and why of screening for language needs; programming reminders in the registration and scheduling screens to prompt staff to complete the language field; and using scripts for language screening, the measure increased to over 80%. Overall, the median value for performance on this measure was 94% at the conclusion of the collaborative. Again, we believe this is much higher than what is generally seen at hospitals since the Speaking Together hospitals were selected in part because they already collected this information in their electronic registration systems.

### 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.**

Two national surveys of hospital practices concerning the collection of language data (1-2) found that about half of hospitals (46.2%-50.2%) collect information on a patient’s language. Another national survey of hospitals found that three-quarters of hospitals indicated that they see patients with limited English proficiency (3), with most indicating that they see LEP patients every day.

### 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]


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

The evidence on disparities by race, ethnicity and language and the need for standard data collection has been described extensively (1-2). Most recently, it was summarized by the Institute of Medicine in its report: Race, Ethnicity and Language Data: Standardization for Health Care Quality Improvement (3).

The proposed measure responds to many of the issues identified in this report and other evidence on the need for better data to conduct quality improvement and to eliminate health care disparities. The proposed measure, however, is principally designed to identify need for effective language services. It can provide information for quality improvement purposes or to better understand a health care organization’s patient population. It is not designed to identify patients by a particular racial or ethnic group. Its purpose is to identify the true demand for language services provided by qualified language services individuals (e.g., trained interpreters, assessed bilingual providers). The IOM guidance includes a question about language proficiency that can be extremely helpful in understanding literacy and related issues. The proposed measure addresses the second question listed by the IOM on language needs to record the spoken language preferred for health care.

### 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]


---

**See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable**
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:

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Quality</th>
<th>Consistency</th>
<th>Does the measure pass subcriterion 1c?</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-H</td>
<td>M-H</td>
<td>M-H</td>
<td>Yes</td>
</tr>
<tr>
<td>L</td>
<td>M-H</td>
<td>M</td>
<td>Yes IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No</td>
</tr>
<tr>
<td>M-H</td>
<td>L</td>
<td>M-H</td>
<td>Yes IF potential benefits to patients clearly outweigh potential harms: otherwise No</td>
</tr>
<tr>
<td>L-M-H</td>
<td>L-M-H</td>
<td>L-M-H</td>
<td>No</td>
</tr>
</tbody>
</table>

Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service

Does the measure pass subcriterion 1c? 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 preferred language is assessed and recorded for all patients and all encounters within a health care organization. It is a fundamental measure necessary for any assessment or quality improvement work that addresses the quality or availability of language services and effective communication for patients with limited English proficiency. It is a necessary step in another proposed language services measure – L2.

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 body of evidence on the need for and current practices related to the collection of preferred language is generally included in a broader set of studies on the collection of race, ethnicity and language data. The most comprehensive assessment of these practices has been amassed in the 2009 IOM report, Chapter 4: Defining Language Need and Categories for Collection (1).

1c.5 Quantity of Studies in the Body of Evidence (Total number of studies, not articles): (1) Defining Language Need and Categories for Collection, pp 93-125, in Institute of Medicine: Race, Ethnicity and Language Data: Standardization for Health Care Quality Improvement. Washington, DC: National Academy Press, 2009. This summary and review identified more than 80 analyses and studies that address the need for better language-related data, as well as ways to collect the data in a uniform fashion.

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 uniform collection of language data is extremely high. No meaningful quality-related or quality improvement work, tracking of utilization, access-related factors or other analyses related to language services delivery can be performed without understanding in a systematic way what the language services needs of patients are.

1c.7 Consistency of Results across Studies (Summarize the consistency of the magnitude and direction of the effect): The
Principal methods for data collection and recording information are consistent across the most respected and commonly used sources of information. The original data collection methodology was developed by the Office of Management and Budget in 1977, later revised in 1997 and reaffirmed in 2000. In 2005, the Health Research and Education Trust (HRET) developed a toolkit for collection of this information that was used widely across hospitals and other health care organizations. The Joint Commission chose not to redefine the collection practices but referred to guidance from the IOM in its 2010 Requirements for Supporting Effective Communication, Cultural Competence and Patient and Family-Centered Care. In 2009, the IOM issued a report on race, ethnicity and language data as part of its subcommittee work on the IOM Committee on Future Directions for the National Healthcare Quality and Disparities Reports. The magnitude and direction of the effect is not relevant to this particular process -- namely, the uniform collection of language data.

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 endorsing a process measure for identifying patient preference for preferred language will be to make available a common measure, with clearly articulated technical specifications, to allow all interested health care organizations to collect this information in a standard fashion. No harm is associated with the collection of this information in this manner; however, it does not satisfy all that an organization may want to know or understand about its patient populations in terms of their languages spoken. For example, some organizations are interested in the language spoken in the home, which they have used in part to determine need for language services. One hospital in the Speaking Together collaborative originally collected language information this way, but found that this information alone overstated the need for an interpreter, since many individuals speak a language other than English in the home but are also fully fluent in English. The IOM recommends also collecting information about English language proficiency. We support collecting this information but do not believe it is a sufficiently precise measure when it comes to identifying the need for language services. Individuals can interpret the questions differently, and they may feel that they can speak English well in most of their interactions, yet nevertheless prefer to receive their health care in a language other than English. Health care organizations may want to consider collecting additional information about the language characteristics of their patients. We believe that the proposed measure is essential for determining need for language services.

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 has not been 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 with 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 exchange of information with the patient 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 with 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, treatment 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.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 developer’s assessment of the quantity, quality, and consistency of the body of evidence?
1c.25 Quantity: High
1c.26 Quality: High
1c.27 Consistency: High

Was the threshold criterion, Importance to Measure and Report, met?
NQF #1824 L1A: Screening for preferred spoken language for health care

(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:

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 hospital admissions, visits to the emergency department, and outpatient visits where preferred spoken language for health care is screened and recorded

2a1.2 Numerator Time Window (The time period in which the target process, condition, event, or outcome is eligible for inclusion):
Time window is a single 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):
Data Elements:
Preferred spoken language for health care
Admissions
Visits

Definitions:
Admissions: a patient health care encounter involving an inpatient stay, whether this is a direct admit to the hospital (scheduled or unscheduled) or occurs through the emergency department.

Preferred spoken language for health care: the preferred language that is stated by the patient for speaking to health care providers. This includes ASL.

Visit: patient health care encounter with a provider in the hospital emergency department, ambulatory unit or clinic.

2a1.4 Denominator Statement (Brief, narrative description of the target population being measured):
The total number of hospital admissions, visits to the emergency department, and outpatient visits.

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. All cases in the denominator are equally eligible to appear in the numerator.

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):*

**Data Elements:**
- Admissions
- Visits

**Definitions:**
Admission: a patient health care encounter involving an inpatient stay, whether this is a direct admit to the hospital (scheduled or unscheduled) or occurs through the emergency department.

Visit: patient health care encounter with a provider in the hospital emergency department, ambulatory unit or clinic.

2a1.8 **Denominator Exclusions** *(Brief narrative description of exclusions from the target population):*
There are no exclusions. All admissions, visits to the emergency department, and outpatient visits, including:
- Scheduled and unscheduled visits
- Elective, urgent and emergent admissions
- Short stay and observation patients
- Transfers from other facilities

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):*
There are no exclusions. All admissions, visits to the emergency department, and outpatient visits, including:
- Scheduled and unscheduled visits
- Elective, urgent and emergent admissions
- Short stay and observation patients
- Transfers from other facilities

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):*
Non-English Speaking Populations can be identified from screening to determine if needed language services were delivered. Clinical performance measures can be stratified by language to examine whether disparities exist among varying language groups.

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 Reported As: Aggregate numerator and denominator, monthly, stratified by language, including English, declined, or unavailable.

Numerator:
- Count the number of patient admissions and/or visits for which preferred spoken language for health care is recorded.
- Apply inclusions and exclusions.
- Stratify by language, including English, declined, or unavailable.

Denominator:
- Count the total number of patient admissions and/or visits.
- Stratify by language, including English, declined, or unavailable.

Notes for Abstraction:
- If patient refused to answer and declined is recorded, credit for screening for preferred spoken language for health care may be taken.
- If electronic systems pre-populate the language preference field, credit for screening for preferred spoken language for health care may be taken for this admission or visit.
- If a space on a document or field in an electronic system for recording language preference for health care is not populated, credit for screening for preferred spoken language for health care may not be taken.
- If the patient’s preferred written language for health care information is recorded and the preferred spoken language for health care is not recorded, credit for screening spoken language may not be taken.

Notes:
- All patients should be asked to self-identify their preferred spoken language for health care. The goal is for the patient, not the provider or registration/scheduling staff, to self-identify preferred spoken language for health care.
- Suggested screening question: “What language do you prefer to speak with your doctor or nurse?”
- American Sign Language (ASL) should be included as a preferred spoken language for health care for this measure.
- Organizational policy should specify whose preferred spoken language for health care should be documented for pediatric patients and for incapacitated adults.
- For example, Organizational policy may require that the preferred spoken language for health care for a parent, family member or caregiver is recorded in the event of a minor child or incapacitated adult.
- Some organizations pre-populate fields so that certain data are present at subsequent admissions and/or visits.
- For example, address, phone number, and insurance are often pre-populated. Some organizations also pre-populate language information fields.
- Please note: Organizational policy should specify whether preferred spoken language for health care should be asked at every admission/visit or verified periodically.
- For newborns: if it is for the birth, the newborn is excluded from the denominator. If the newborn is admitted to the hospital from day 1 forward (and the mother is not admitted to the hospital), the newborn is included in the denominator.
- For Emergency Department visits, hospitals should report all visits (i.e., all who come for care) and not just those who are admitted to the hospital.

Declined: a person who is unwilling to choose/provide a language category or cannot identify him/herself with one of the listed languages. This category is an indication that the person did NOT want to respond to the question and should not be asked again during the same visit or during a subsequent visit. [HRET]

Unavailable: a patient who is unable to physically respond, there is no available family member or caregiver to respond for the patient, or if for any reason, the demographic portion of the medical record cannot be completed. Can be called “Unknown,” “Unable to complete,” or “Other.” This category is an indication that the person could not respond to the question and can be asked again during the same visit or during a subsequent visit. [HRET]
### 2a1.21 Calculation Algorithm/Measure Logic Diagram URL or attachment:

**URL**

### 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 -- 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

### 2a1.29 Data Source/Data Collection Instrument (Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.):

Data on language preferences is generally collected at the point of admission, whether as an inpatient, emergency department patient, or ambulatory patient. Some health care organizations use registration systems that are linked with other sources of information; others use electronic health records that include registration/administrative data and clinical data.

### 2a1.27 Data Source/data Collection Instrument Reference Web Page URL or Attachment:

**URL**

### 2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:

**URL**

### 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 (2) 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 (10) 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 27,000; 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 526; Annual Admissions 27,946; Annual interpreter encounters 65,000; 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-55% Spanish; 7% Vietnamese; 4% Somali; 4% Russian; 2% 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 (10) 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

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable
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 information under "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 definitions, descriptions 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.

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); 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:


2a2.3 Testing Results (Reliability statistics, assessment of adequacy in the context of norms for the test conducted):
The proposed measure has been tested across different types of hospitals, with variation in types of language services programs and with various groups of clinical and non-clinical staff. It targets all patients receiving care at a health care organization. The measure inclusion criteria for the denominator are broad and include all patients who present in the emergency department, in ambulatory care settings, or as an inpatient in hospitals, clinics, or emergency departments. The numerator includes all patients who have been screened for language preference, whether they are English or non-English speaking. The measure has high face validity. Hospital staff who have reported on the measure view it as measuring precisely what it intends to measure -- that is, the percent of all patients who have been screened for language preference. The proposed measure has been accepted as part of the AHRQ National Measures Clearinghouse and has been used in three 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 delivery of language services (based on volume of services and total patients served).

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 measure focus and target population are consistent with IOM and Joint Commission guidance related to the collection of information on preferred language for health care. The exclusions are narrow -- the measure is designed to include all patients who seek and/or receive ambulatory, inpatient and emergency department services. The measure does not capture all information that organizations may want to know about their patients who speak languages other than English. The measure itself is consistent with the evidence about language-related data collection practices.
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 measures. All participants viewed the proposed measure as accurately reflecting the demand and need for language services, although several collected additional information at registration about other elements of patient language, including language spoken at home, English language proficiency, and preference related to written materials. Ten hospitals participated in Speaking Together and reported performance for more than 100,000 patients (this measure required that all patients be screened for language preference). Hospitals reported on a monthly basis over an 18-month period, beginning November 2006. Hospitals ranged in size from 11,500-44,000 admissions, included two children’s hospitals and teaching and non-teaching community hospitals.

2b2.2 Analytic Method (Describe method of validity testing and rationale; if face validity, describe systematic assessment):
As indicated above, the measure has strong face validity. The measure was developed after a thorough review of 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 services 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):
The only exclusions in the denominator are lab specimens and other types of registration that have a medical record number but are not attached to patients who physically came to the health care organization. All scheduled and unscheduled visits, elective, urgent and emergent admissions, short stay and observation patients, and transfers from other facilities are included in the proposed measure. The number of patients includes the number of patients who seek care from the health care organization. In the Speaking Together collaborative, this was over 100,000 patients over the course of the 18-month collaborative.

2b3.2 Analytic Method (Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference):
No analysis was performed on the exclusions.

2b3.3 Results (Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):
No statistical analysis was performed for an analysis of exclusions.

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. It is a process measure with a target of 100 percent of patients meeting the criteria associated with screening for preferred spoken language.
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.

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.

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment: No risk adjustment was performed. It is a process measure with a target of 100 percent of patients meeting the criteria associated with screening for preferred spoken language.

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 used from November 2006-February 2008 by 10 hospitals and health systems in seven states. Four public and six non-profit hospitals, including two children’s hospitals, were included in the collaborative. The 10 hospitals differed in terms of their size, scope of language services program, and geographic location in the country. All of the participating hospitals had relatively robust language services programs as a requirement for participation. At the time of the program, the hospital with the largest language services program reported a total of 63 FTE for language services staff and over 140,000 interpreter encounters annually; the smallest language services program reported eight FTE staff for language services and approximately 40,000 interpreter encounters annually. In all but one hospital, Spanish was the most common language spoken by LEP patients. Many of the hospitals had substantial numbers of patients speaking Mandarin, Portuguese, Haitian Creole, Somali, Hmong, Arabic, Russian, and Cantonese. Most also had sizeable patient populations who communicated using American Sign Language, which was included among language languages requiring effective quality improvement interventions.

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

We collected data on a monthly basis and worked closely with the hospitals to discuss challenges in data collection. We performed periodic checks of the data to determine whether the denominator matched the hospitals statistics in terms of total patients served. We made the results of the quality improvement collaborative available to the 10 participating hospitals so that they could see how hospitals with similar challenges were addressing the process of screening and implementing interventions for improvement. No additional statistical analyses were conducted to identify statistically significant and differences in performance.

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):*

Speaking Together hospitals generally performed very well on the measure. This was an expected finding, since the selection process for participation in Speaking Together favored hospitals that had systems in place to screen for preferred language and capacity for data collection related to use of language services. Five of the hospitals began the collaborative screening nearly all (>95 percent) of their patients for preferred language and continued at that level or above for the duration of the project. One hospital that initially reported high levels of reporting (>95 percent) saw consistent drops in performance; this was interpreted by the hospital and the program staff as a function of improvements in reporting and the inclusion of all sites of care in the measure. Two hospitals measured approximately 50 and 60 percent at the beginning of the collaborative and saw substantial improvement by final quarter of the project (approximately 90% and 83%, respectively). One hospital began the project at about 45% and reported inconsistent performance throughout the collaborative, ending with about 38% performance. This hospital had substantial changes and disruptions in its project team and we were not confident in its data collection on this or other measures at times during the collaborative.

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 patients were included in the measure. See other responses for information about the participating hospitals and their patient populations.

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 participating hospitals -- 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

<table>
<thead>
<tr>
<th>If applicable, the measure specifications allow identification of disparities.</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
</tr>
</tbody>
</table>

2c.1 **If measure is stratified for disparities, provide stratified results** *(Scores by stratified categories/cohorts):*
Stratification was not performed on the proposed measure for the purposes of the findings of the collaborative. The denominator includes all patients who receive care at the health care organization and the numerator includes all patients who are screened for spoken language preference. The measure provides information about patients’ needs for language services and can identify various language groups who receive care at the health care organization. This measure forms the basis for other language services measures (e.g., L2) which are stratified by language for identification of disparities and ability to target quality improvement interventions. Individual hospitals reviewed their performance on this measure by language and by site of registration as part of their quality improvement activities.

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

### 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
- Regulatory and Accreditation Programs

#### 3.1 Current Use *(Check all that apply; for any that are checked, provide the specific program information in the following questions):*
- Quality Improvement with Benchmarking
- Quality Improvement

#### 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 currently not publicly reported. Endorsement by NQF could greatly enhance efforts to include screening for preferred spoken language as a publicly reported, commonly used metric of language services quality and delivery.

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 is at the very core of an organization’s ability to provide high-quality and safe care to its patients. Organizations must understand the needs of their patients and record whether a patient prefers to receive health care in a language other than English. It is not possible to determine whether a patient receives a needed language service (See L2 measure submission) without every patient being screened for preferred language. Merely “having” interpreter services capacity does not mean that all patients who need these services are receiving them.

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.

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

Results are documented in:


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: 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 toward delivery services; and use measures to identify waste and streamline systems.
Overall, to what extent was the criterion, *Usability*, met? H □ M □ L □ I □
Provide rationale based on specific subcriteria:

<table>
<thead>
<tr>
<th>4. FEASIBILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. <em>(evaluation criteria)</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4a. Data Generated as a Byproduct of Care Processes: H □ M □ L □ I □</th>
</tr>
</thead>
</table>
| 4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*

Data used in the measure are:
- Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry). Other
- Patients asked spoken language preference for health care during registration / scheduling process

<table>
<thead>
<tr>
<th>4b. Electronic Sources: H □ M □ L □ I □</th>
</tr>
</thead>
</table>
| 4b.1 Are the data elements needed for the measure as specified available electronically *(Elements that are needed to compute measure scores are in defined, computer-readable fields):* ALL data elements are in a combination of electronic sources

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

<table>
<thead>
<tr>
<th>4c. Susceptibility to Inaccuracies, Errors, or Unintended Consequences: H □ M □ L □ I □</th>
</tr>
</thead>
</table>
| 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:
Some registration staff may assume that the patient’s spoken language and written language is the same and not ask for the written language. Education and training on collection of self reported spoken and written language is required at the facility / clinic level. Re-enforcement during training that all patients do not read (even English proficient patients) and that all languages do not have an alphabet have helped to raise an understanding of why this measure alone cannot be used to assume preference for written language. Additionally, for minors and incapacitated adults, organizations are encouraged to set policies to collect the parent or care providers preference until the patient is able. For children it may be necessary to record the language preference of each parent. |

<table>
<thead>
<tr>
<th>4d. Data Collection Strategy/Implementation: H □ M □ L □ I □</th>
</tr>
</thead>
</table>
| 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):*
The Joint Commission’s recent Patient Provider Communications Standard addresses demonstrating identification of a patient’s language preference during hospitalization. Once staffs are educated on self reported language preference for health care and data collection is established as part of the registration staff and scheduling staffs routine, this is a straightforward process. Registration systems do not always have the required fields in place and or the fields that come with the system may not be labeled ‘preferred’ but most often ‘primary’ language and these need to be changed by the organization. Organizations also benefit from having lists of the most common languages at the organization for staffs to pick from vs. relying on free text fields where typos occur. Some organizations have made the fields mandatory (cannot bypass) or have implemented pop-ups as reminders to go back and fill in data. Some have policies where once language data are collected; fields are pre-populated and verified at specific points in time so that the questions are not asked at every encounter. Once systems are established, minimal amount of time is required for asking the patients. Most organizations use language identification aids such as cards where patients point to their language. |

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 [x]

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)](https://www.qualitymeasures.hhs.gov/): 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, 202-994-8663-

**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, 202-994-8663-

**Co.5 Submitter:** Catherine, West, MS, RN, cathy.west@gwumc.edu, 202-994-8663-, Department of Health Policy, The George Washington University

**Co.6 Additional organizations that sponsored/participated in measure development:**

Interpreter Measures Field Test Hospitals

Boston Medical Center
Children's Hospital of Philadelphia

**Co.7 Public Contact:** Catherine, West, MS, RN, cathy.west@gwumc.edu, 202-994-8663-, 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.
Professor, 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
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? Annual |
| Ad.6 | When is the next scheduled review/update for this measure? 06, 2012 |
| Ad.7 | Copyright statement: © 2009 Department of Health Policy, George Washington University School of Public Health and Health Services. |
| 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 |