

NATIONAL QUALITY FORUM

Measure Evaluation 4.1 December 2009

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

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

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

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

Evaluation ratings of the extent to which the criteria are met

C = Completely (unquestionably demonstrated to meet the criterion)

P = Partially (demonstrated to partially meet the criterion)

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

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

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

(for NQF staff use) NQF Review #: 1357	NQF Project: Child Health Quality Measures 2010
MEASURE DESCRIPTIVE INFORMATION	
De.1 Measure Title: Outpatient hearing screening of infants who did not complete screening before hospital discharge (EHDI-1c)	
De.2 Brief description of measure: This measure assesses the proportion of all newborn infants who did not complete a hearing screen prior to discharge, who went on to receive an outpatient screen before the child was 31 days of age.	
*Numbering within the parentheses references the US national extension quality measure identifiers developed for the Use Cases published in the Integrating the Healthcare Enterprise (IHE) Quality, Research and Public Health (QRPH) EHDI Technical Framework Supplement available at www.ihe.net/Technical_Framework/index.cfm#quality	
1.1-2 Type of Measure: Process	
De.3 If included in a composite or paired with another measure, please identify composite or paired measure This measure is paired with other measures relevant to the monitoring and measurement of the early screening evaluation and intervention process.	
De.4 National Priority Partners Priority Area: Population health	
De.5 IOM Quality Domain: Effectiveness	
De.6 Consumer Care Need: Living with illness	

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. <i>Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.</i>	A YO NO

A.1 Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)? Yes	
A.2 Indicate if Proprietary Measure (as defined in measure steward agreement):	
A.3 Measure Steward Agreement: Government entity and in the public domain - no agreement necessary	
A.4 Measure Steward Agreement attached:	
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	B Y● N●
C. The intended use of the measure includes both public reporting and quality improvement. ► Actual/Planned Use: Public Reporting, Quality Improvement (Internal to the specific organization)	C Y● N●
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1 Testing: No, testing will be completed within 12 months D.2 Have NQF-endorsed measures been reviewed to identify if there are similar or related measures? Yes	D Y● N●
(for NQF staff use) Have all conditions for consideration been met? Staff Notes to Steward (if submission returned):	Met Y● N●
Staff Notes to Reviewers (issues or questions regarding any criteria):	
Staff Reviewer Name(s):	

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

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

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. 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	Eval Rating
(for NQF staff use) Specific NPP goal:	
1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, Frequently performed procedure 1a.2 1a.3 Summary of Evidence of High Impact: U.S. Preventive Services Task Force. The USPSTF recommends screening for hearing loss in all newborn infants. There is good evidence that newborn hearing screening testing is highly accurate and leads to earlier identification and treatment of infants with hearing loss. Good-quality evidence shows that early detection improves language outcomes. http://www.uspreventiveservicestaskforce.org/uspstf08/newbornhear/newbhearrs.pdf 1a.4 Citations for Evidence of High Impact: Nelson HD, Bougatsos C, Nygren P. Universal Newborn Hearing Screening: Systematic Review to Update the 2001 U.S. Preventive Services Task Force Recommendation. AHRQ Publication No. 08-05117-EF-4, July 2008. Agency for Healthcare Research and Quality (AHRQ), Rockville, MD. http://www.uspreventiveservicestaskforce.org/uspstf08/newbornhear/newbornart.pdf	1a C● P● M● N●

1b. Opportunity for Improvement	
<p>1b.1 Benefits (improvements in quality) envisioned by use of this measure: From page 194 of the 2007 Joint Committee on Infant Hearing (JCIH) Year 2007 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs(http://pediatrics.aappublications.org/cgi/content/full/120/4/898?ijkey=oj9BAleq210IA&keytype=ref&siteid=aapjournals)</p> <p>"The JCIH supports the concept of regular measurements of performance and recommends routine monitoring of these measures for interprogram comparison and continuous quality improvement. Performance benchmarks represent a consensus of expert opinion in the field of newborn hearing screening and intervention. The benchmarks are the minimal requirements that should be attained by high quality programs. Frequent measures of quality permit prompt recognition and correction of any unstable component of the EHDI process."</p>	
<p>1b.2 Summary of (data demonstrating performance gap) (variation or overall poor performance) across providers:</p> <p>http://www.cdc.gov/ncbddd/ehdi/data.htm</p>	
<p>1b.3 Citations for data on performance gap:</p> <p>"Identifying Infants with Hearing Loss --- United States, 1999--2007." CDC Morbidity and Mortality Weekly Report (MMWR). March 5, 2010 / 59(08):220-223. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5908a2.htm</p> <p>"Newborn hearing screening and follow-up: are children receiving recommended services?" Public Health Rep. 2010 Mar-Apr;125(2):199-207.</p>	
<p>1b.4 Summary of Data on disparities by population group:</p> <p>The Hispanic population is most likely not to receive the outpatient rescreen Infants born to mothers who have 12 years of education or less were less likely to obtain the rescreen. Males are less likely to receive the outpatient rescreen</p>	
<p>1b.5 Citations for data on Disparities:</p> <p>A Programmatic Analysis of a Newborn Hearing Screening Program for Evaluation and Improvement. Theses submitted to the Faculty of the Graduate School of the University of Colorado in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Vickie R Thomson. 2007.</p>	<p>1b CO PO MO NO</p>
1c. Outcome or Evidence to Support Measure Focus	
<p>1c.1 Relationship to Outcomes (For non-outcome measures, briefly describe the relationship to desired outcome). For outcomes, describe why it is relevant to the target population): Children with hearing loss who are screened for hearing loss at birth have better language outcomes at school age than those not screened. Infants identified with hearing loss through universal screening have significantly earlier referral, diagnosis, and treatment than those identified in other ways. Language outcomes at school age strengthen the case for newborn hearing screening but are also dependent on effective methods of referral, follow-up, and treatment.</p>	
<p>1c.2-3. Type of Evidence: Cohort study, Observational study, Evidence-based guideline, Expert opinion, Systematic synthesis of research</p>	
<p>1c.4 Summary of Evidence (as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome):</p> <p>U.S. Preventive Services Task Force (www.ahrq.gov/clinic/uspstf/uspstf07.htm) Year 2007 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. Joint Committee on Infant Hearing. Pediatrics 2007;120:898-921 (http://pediatrics.aappublications.org/cgi/content/full/120/4/898?ijkey=oj9BAleq210IA&keytype=ref&siteid=aapjournals)</p>	<p>1c CO PO MO NO</p>
<p>1c.5 Rating of (strength/quality of evidence) (also provide narrative description of the rating and by whom):</p>	<p>1c CO PO MO NO</p>

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

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

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

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

OR

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

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

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

Grade: B (Recommendation by the USPSTF recommends screening for hearing loss in all newborn infants.)	
1c.6 Method for rating evidence: Scientific evidence review conducted by the Oregon Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality.	
1c.7 Summary of Controversy/Contradictory Evidence: There is limited evidence about the harms of screening, with conflicting research findings regarding anxiety associated with false-positive test results. There is limited information about the harms of treatment	
1c.8 Citations for Evidence (other than guidelines):	
1c.9 Quote the Specific guideline recommendation (including guideline number and/or page number): "Outpatient screening at no later than 1 month of age should also be available to infants who were discharged before receiving the birth admission screening or who were born outside a hospital or birthing center." Page 905. Year 2007 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. Joint Committee on Infant Hearing.	
1c.10 Clinical Practice Guideline Citation: Year 2007 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. Joint Committee on Infant Hearing. Pediatrics 2007;120;898-921 (http://pediatrics.aappublications.org/cgi/content/full/120/4/898?ijkey=oj9BAleq210IA&keytype=ref&siteid=aapjournals)	
1c.11 National Guideline Clearinghouse or other URL: Newborn Screening Coding and Terminology Guide. http://newbornscreeningcodes.nlm.nih.gov/nb/sc/condition/HEAR	
1c.12 Rating of strength of recommendation (also provide narrative description of the rating and by whom):	
1c.13 Method for rating strength of recommendation (If different from USPSTF system, also describe rating and how it relates to USPSTF):	
1c.14 Rationale for using this guideline over others:	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Importance to Measure and Report?	1
Steering Committee: Was the threshold criterion, Importance to Measure and Report, met? Rationale:	1 Y O N
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	Eval Rating
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? S.2 If yes, provide web page URL:	2a- specs C P M N
2a. Precisely Specified	
2a.1 Numerator Statement (Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome): Numerator contains the number of infants born at a given facility during the time window with no documented hearing screening performed prior to patient discharge and who have been screened for	

Comment [k]: USPSTF grading system <http://www.ahrq.gov/clinic/uspstf/grade.htm>: A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. C - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. D - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. I - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

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

hearing loss as an outpatient by 30 days of age.

2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator*): The time period varies upon needs of the particular user (e.g. calendar year, quarterly, monthly) but must be the same for both the numerator and denominator.

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):

Total number with LOINC# 54109-4: Newborn hearing screen - right OR LOINC# 54108-6: Newborn hearing screen - left equals "Not performed" (LA7304-4)

AND

with "Hearing Screening Performed": evidence of hearing screening performed before the child was 31 days of age. (LOINC# 54109-4: Newborn hearing screen - right = Pass LA10392-1 OR Refer LA10393-9 AND LOINC# 54108-6: Newborn hearing screen - left= Pass LA10392-1 OR Refer LA10393-9).

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*):

Denominator contains the number of infants born at a given facility during the time window with no documented hearing screening performed prior to patient discharge.

2a.5 Target population gender: Female, Male

2a.6 Target population age range: Newborn period

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator*):

The time period varies upon needs of the particular user (e.g. calendar year, quarterly, monthly) but must be the same for both the numerator and denominator.

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):

Total number with LOINC# 54109-4: Newborn hearing screen - right OR LOINC# 54108-6: Newborn hearing screen - left equals "Not performed" (LA7304-4).

2a.9 Denominator Exclusions (*Brief text description of exclusions from the target population*): Patient deceased before the child was 31 days of age, parental refusal, or not performed due to medical exclusion.

2a.10 Denominator Exclusion Details (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*):

Joint Commission Discharge Disposition - Death Value Set (86986.v1) 1.3.6.1.4.1.33895.1.3.0.12. "Patient Deceased": Patient has expired.

LOINC# 54109-4: Newborn hearing screen - right OR LOINC# 54108-6: Newborn hearing screen - left includes "Parental refusal" (LA6644-4) OR Not performed, medical exclusion - not indicated (LA12409-1)

2a.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*):

2a.12-13 Risk Adjustment Type: No risk adjustment necessary

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*):

2a.15-17 Detailed risk model available Web page URL or attachment:

2a.18-19 Type of Score: Rate/proportion

2a.20 Interpretation of Score: Better quality = Higher score

2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps*):

(1) The time period for births included in the estimate is specified (see 2a.2, 2a.7).

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

<p>(2) All live births that occurred at a facility during the time period are selected.</p> <p>(3) Result of step 2 is filtered to remove children who died before the child was 31 days of age, cases of parental refusal, and/or cases not screened due to medical exclusion (see 2a.9, 2a.10).</p> <p>(4) Result of step 3 is filtered to be limited to the subset that (a) has been discharged from the hospital following birth, AND (b) had Newborn Hearing Screening identified as “not performed” at the time of discharge (see 2a.8). This result is saved</p> <p>The numerator is calculated using the following step:</p> <p>(5) Result of step 4 is further filtered to be limited to the subset with a hearing screening performed after discharge (see 2a.3) AND before the child was 31 days of age (see 2a.2). This result is saved as the numerator (see 2a.4).</p> <p>The denominator is calculated using the following step:</p> <p>(6) Result from Step 4 is further filtered to exclude individuals who both (a) are under the age of 31 days AND who also (b) have not received a screen following discharge. The result is saved as the denominator (see 2a.4).</p> <p>EHDI-1c is calculated using the following step:</p> <p>(7) EHDI-1c is calculated by dividing the numerator (result of step 5) by the denominator (result of step 6).</p>	
<p>2a.22 Describe the method for discriminating performance (e.g., significance testing): Method to discriminate performance is based upon jurisdictionally based statistical measurement reflecting local and national variability.</p>	
<p>2a.23 Sampling (Survey) Methodology <i>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):</i> not applicable</p>	
<p>2a.24 Data Source <i>(Check the source(s) for which the measure is specified and tested)</i> Electronic Health Records, Other, Public health data/vital statistics, Registry data</p>	
<p>2a.25 Data source/data collection instrument <i>(Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.):</i> Electronic Health/Medical Record, Public health information system</p>	
<p>2a.26-28 Data source/data collection instrument reference web page URL or attachment: URL www.hitsp.org AND www.ihe.net/Technical_Framework/index.cfm#quality AND www.cdc.gov/ncbddd/ehdi/data.htm</p>	
<p>2a.29-31 Data dictionary/code table web page URL or attachment: URL http://newbornscreeningcodes.nlm.nih.gov AND www.hitsp.org AND www.ihe.net/Technical_Framework/index.cfm#quality</p>	
<p>2a.32-35 Level of Measurement/Analysis <i>(Check the level(s) for which the measure is specified and tested)</i> Facility/Agency, Other, Population : Regional and State «measurement_level_clinician_other» «measurement_level_program_other»</p>	
<p>2a.36-37 Care Settings <i>(Check the setting(s) for which the measure is specified and tested)</i> Hospital</p>	
<p>2a.38-41 Clinical Services <i>(Healthcare services being measured, check all that apply)</i> Clinicians: Audiologist, Clinicians: Nurses, Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO), Clinicians: PT/OT/Speech</p>	
TESTING/ANALYSIS	
<p>2b. Reliability testing</p>	<p>2b CO</p>

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

<p>2b.1 Data/sample (<i>description of data/sample and size</i>): Data used in this measure are included in the EHR. As noted in the NQF draft Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties, "...the EHR will be considered the authoritative source of clinical information and legal record of care. Quality measures based on EHRs require exporting clinical information recorded by healthcare clinicians from discrete computer readable fields; therefore, measurement errors due to manual abstraction, coding by persons other than the originator, or transcription are eliminated." As these data elements are extracted from EHRs using computer programming, they "are by virtue of automation repeatable (reliable); therefore, testing at the data element level should focus on validity... reliability of data items may be bypassed if validity of data items is demonstrated." EHR data used in this measure reflect part of a national, population-based public health surveillance data collection. Data are collected at the individual-child level within each state/territory and reported nationally at an aggregated state-level to CDC. This population-based collection of EHD data has been occurring for over a decade. For the reporting period of calendar year 2007, 47 states and 2 territories reported newborn hearing screening data on a total of 3,345,629 births.</p>	<p>PO MO NO</p>
<p>2b.2 Analytic Method (<i>type of reliability</i>) & <i>rationale, method for testing</i>): As noted in 2b.1., given data are extracted from EHRs, "reliability of data items may be bypassed if validity of data items is demonstrated". (NQF draft Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties)</p>	
<p>2b.3 Testing Results (<i>reliability statistics, assessment of adequacy in the context of norms for the test conducted</i>): While the use of EHRs for data elements reflects a particular strength of this measure, "EHRs and EHR measures are new and will most likely require some adjustment of local EHR structures and recording practices to meet standards." (NQF draft Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties). This has been and will continue to be addressed in the manner recommended in the Guidance document cited above. First, nationally, CDC EHD data has and will continue to provide states and territories with a summary of results of measures reported as part of the national population-based public health data collection. This allows them to identify and address potential discrepancies. Similarly, EHD programs are and will continue to be encouraged to provide similar feedback to their reporting sources as a means of quality control and programmatic feedback. Second, state EHD programs have been and will continue to be encouraged to conduct their own reliability/validity studies, and to encourage data quality studies on the part of their reporting sources.</p>	
<p>2c. Validity testing</p>	
<p>2c.1 Data/sample (<i>description of data/sample and size</i>): Data used in this measure reflect EHR extracted information that is part of a national, population-based public health surveillance data collection. Data are collected at the individual-child level within each state/territory, and reported at state-level aggregate form nationally to CDC. This population-based collection of EHD data has been occurring for over a decade. For the reporting period of calendar year 2007, 47 states and 2 territories reported newborn hearing screening data on a total of 3,345,629 births.</p>	
<p>2c.2 Analytic Method (<i>type of validity</i>) & <i>rationale, method for testing</i>): A formal and systematic testing of face validity of the measure score as an indicator of quality has been conducted in order to serve as an acceptable indicator for validity of the measure score (NQF draft Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties). This evaluation has been conducted through the CDC EHD Data Committee.</p>	
<p>2c.3 Testing Results (<i>statistical results, assessment of adequacy in the context of norms for the test conducted</i>): Face validity has been systematically assessed by relevant stakeholders in order to assess whether the measure represents quality care for this specific topic and whether the focus of this measure is the most important aspect of quality for this specific topic (NQF draft Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties).</p>	<p>2c CO PO MO NO</p>
<p>2d. Exclusions Justified</p>	<p>2d CO</p>

Comment [k]: 8 Examples of reliability testing include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing may address the data items or final measure score.

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

Comment [k]: 9 Examples of validity testing include, but are not limited to: determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores on some other related valid measure; content validity for multi-item scales/tests. Face validity is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with BP < 140/90 is a marker of quality). If face validity is the only validity addressed, it is systematically assessed (e.g., ratings by relevant stakeholders) and the measure is judged to represent quality care for the specific topic and that the measure focus is the most important aspect of quality for the specific topic.

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

- supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;

AND

- a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus;

AND

- precisely defined and specified:
 - if there is substantial variability in exclusions across providers, the measure is specified so that

<p>2d.1 Summary of Evidence supporting exclusion(s): Not applicable -exclusions are limited to cases of infant death before the child was 31 days of age, medical exclusion or parental refusal.</p> <p>2d.2 Citations for Evidence: Not applicable - see 2d.1.</p> <p>2d.3 Data/sample (description of data/sample and size): Not applicable - see 2d.1.</p> <p>2d.4 Analytic Method (type analysis & rationale): Not applicable - see 2d.1.</p> <p>2d.5 Testing Results (e.g., frequency, variability, sensitivity analyses): Not applicable - see 2d.1.</p>	<p>PO MO NO NAO</p>
<p>2e. Risk Adjustment for Outcomes/ Resource Use Measures</p> <p>2e.1 Data/sample (description of data/sample and size): Not applicable - no risk adjustment is included</p> <p>2e.2 Analytic Method (type of risk adjustment, analysis, & rationale): Not applicable - no risk adjustment is included</p> <p>2e.3 Testing Results (risk model performance metrics): Not applicable - no risk adjustment is included</p> <p>2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: Not applicable - no risk adjustment is included</p>	<p>2e CO PO MO NO NAO</p>
<p>2f. Identification of Meaningful Differences in Performance</p> <p>2f.1 Data/sample from Testing or Current Use (description of data/sample and size): National, population-based public health surveillance data, collected at the individual-child level within each state/territory, and reported at state-level aggregate form nationally to CDC. This population-based collection of EHD data has been occurring for over a decade. For the reporting period of calendar year 2007, 47 states and 2 territories reported newborn hearing screening data on a total of 3,345,629 births.</p> <p>2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance (type of analysis & rationale): Statistical analysis comparing individual entities (provider, network of providers, state/territory) to the mean level of performance for similar entities. When appropriate, this can be limited to similar entities within a given jurisdiction (e.g., performance of a specific provider relative to other providers in a state) or nationally (e.g., mean performance across an entire state relative to other state/territories). In addition, performance can be evaluated through direct comparison to current national standards of performance (e.g., CDC National Goals, Joint Committee on Infant Hearing, Healthy People 2020.)</p> <p>2f.3 Provide Measure Scores from Testing or Current Use (description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance): For statistical analyses comparing individual entities to the mean level of performance for similar entities, performance that is 2 standard deviations below the corresponding mean can be flagged. When appropriate, this can be done both within a given jurisdiction and nationally. For example, overall performance for a low performing state may be more than 2 standard deviations below the mean for all states/territories, resulting in that state being identified. However, within that state, there may be no significant difference among providers (i.e., all are performing equally poorly). For direct comparisons to current national standards, identification will consist of (1) a determination that performance falls below the standard, and (2) a measure of the difference between observed performance and the stated standard.</p>	<p>2f CO PO MO NO</p>

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

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

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

OR

rationale/data support no risk adjustment.

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

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

Comment [k]: 14 With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74% v. 75%) is clinically meaningful; or whether a statistically significant difference of \$25 in cost for an episode of care (e.g., \$5,000 v. \$5,025) is practically meaningful. Measures with overall poor performance may not demonstrate much variability across providers.

2g. Comparability of Multiple Data Sources/Methods	
2g.1 Data/sample (description of data/sample and size): All data will be collected through Electronic Health Records - not applicable	2g CO PO MO NO NAO
2g.2 Analytic Method (type of analysis & rationale): All data will be collected through Electronic Health Records - not applicable	
2g.3 Testing Results (e.g., correlation statistics, comparison of rankings): All data will be collected through Electronic Health Records - not applicable	
2h. Disparities in Care	
2h.1 If measure is stratified, provide stratified results (scores by stratified categories/cohorts): Not applicable - measure is not stratified	2h CO PO MO NO NAO
2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans: Follow-up analysis can be performed at state and national levels based upon disparities noted in 1b.4 / 1b.5	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Scientific Acceptability of Measure Properties</i> ?	2
Steering Committee: Overall, to what extent was the criterion, <i>Scientific Acceptability of Measure Properties</i> , met? Rationale:	2 CO PO MO NO
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	Eval Rating
3a. Meaningful, Understandable, and Useful Information	
3a.1 Current Use: In use	
3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years): Healthy People 2010 objective 28-11: Increase the proportion of newborns who are screened for hearing loss by age 1 month, have audiologic evaluation by age 3 months, and are enrolled in appropriate intervention services by age 6 months. Proposed Healthy People 2020 ENT-VSL HP2020-8: Increase the proportion of newborns who are screened for hearing loss by no later than age 1 month, have audiologic evaluation by age 3 months, and are enrolled in appropriate intervention services by age 6 months.	
3a.3 If used in other programs/initiatives (If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years):	
Testing of Interpretability (Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement)	
3a.4 Data/sample (description of data/sample and size): CDC Survey (http://www.cdc.gov/ncbddd/ehdi/data.htm) Summary of 2007 National CDC EHDI Data: Number Screened = 3,345,629	3a CO PO MO NO
3a.5 Methods (e.g., focus group, survey, QI project): Hearing Screening and Follow-up Survey (HSFS): OMB No. 0920-0733	

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

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

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

3a.6 Results (qualitative and/or quantitative results and conclusions): Qualitative: "Identifying Infants with Hearing Loss --- United States, 1999–2007." CDC Morbidity and Mortality Weekly Report (MMWR). March 5, 2010 / 59(08);220-223. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5908a2.htm	
3b/3c. Relation to other NQF-endorsed measures	
3b.1 NQF # and Title of similar or related measures:	
(for NQF staff use) Notes on similar/related endorsed or submitted measures:	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 Are the measure specifications harmonized? If not, why? The Centers for Disease Control and Prevention (CDC), the HRSA Maternal and Child Health Bureau (MCHB) and the National Committee for Quality Assurance (NCQA) have submitted 2010 Child Health Quality Measures to NQF that relate to the topic of newborn screening, however the measures target different care settings and data sources. CDC, MCHB, and NCQA are collaborating to ensure the measure specifications have distinctive additive value and are harmonized.	3b CO PO MO NO NAO
3c. Distinctive or Additive Value 3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: 5.1 If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:	3c CO PO MO NO NAO
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Usability?	3
Steering Committee: Overall, to what extent was the criterion, Usability, met? Rationale:	3 CO PO MO NO
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes	
4a.1-2 How are the data elements that are needed to compute measure scores generated? Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition), Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	4a CO PO MO NO
4b. Electronic Sources	
4b.1 Are all the data elements available electronically? (elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims) Yes	4b CO PO MO NO
4b.2 If not, specify the near-term path to achieve electronic capture by most providers.	

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

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

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

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

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

4c. Exclusions	
4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications? No	4c CO PO MO NO NAO
4c.2 If yes, provide justification.	
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences	
4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results. The use of EHRs for this measure provide a number of strengths that facilitate data quality, including EHRs serving as the authoritative source of clinical information and legal record of care. Furthermore, the use of discrete, computer readable fields results in reduced measurement error that may emerge from manual abstraction, third party coding, or transcription errors. Nevertheless, potential sources of error exist and include incorrect measure, code, or logic specification, as well as incorrect programming, system structure, or data exporting code, or inconsistent field definitions across providers or users. These can be audited through quality control measures. For example, CDC EHDl provides states and territories with a summary of results of measures reported as part of the national population-based public health data collection. This allows them to identify and address potential discrepancies. Similarly, EHDl programs are encouraged to provide similar feedback to their reporting sources as a means of quality control and programmatic feedback. Furthermore, state EHDl programs are encouraged to conduct their own reliability/validity studies, and to encourage data quality studies on the part of their reporting sources.	4d CO PO MO NO
4e. Data Collection Strategy/Implementation	
4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: Requires an accurate standardized denominator and numerator to successfully determine that all infants have been accounted for and received necessary care. The limitation has been that providers have only reported on a subset of infants seen.	
4e.2 Costs to implement the measure (costs of data collection, fees associated with proprietary measures): Outpatient hearing screening of infants who did not complete screening before hospital discharge is not a proprietary measure. Many public health EHDl programs have already assumed the cost to implement and report this measure. Depending on availability, federal funds can be provided for additional public health programs to strengthen infrastructure which might be needed for this data collection.	
4e.3 Evidence for costs:	4e CO PO MO NO
4e.4 Business case documentation:	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Feasibility?	4
Steering Committee: Overall, to what extent was the criterion, Feasibility, met? Rationale:	4 CO PO MO NO
RECOMMENDATION	

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

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

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

(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.	Time-limited ●
Steering Committee: Do you recommend for endorsement? Comments:	Y● N● A●
CONTACT INFORMATION	
Co.1 Measure Steward (Intellectual Property Owner) Co.1 Organization Centers for Disease Control and Prevention, Early Hearing Detection and Intervention (EHDI), 1600 Clifton Road NE, MS E-88, Atlanta, Georgia, 30333 Co.2 Point of Contact John, Eichwald, M.A. FAAA, jeichwald@cdc.gov, 404-498-3961-	
Measure Developer If different from Measure Steward Co.3 Organization Centers for Disease Control and Prevention, Early Hearing Detection and Intervention (EHDI), 1600 Clifton Road NE, MS E-88, Atlanta, Georgia, 30333 Co.4 Point of Contact Craig, Mason, Ph.D., Craig_Mason@umit.maine.edu, 207-581-9059-	
Co.5 Submitter If different from Measure Steward POC John, Eichwald, M.A. FAAA, jeichwald@cdc.gov, 404-498-3961-, Centers for Disease Control and Prevention	
Co.6 Additional organizations that sponsored/participated in measure development On July 24, the Joint Committee on Infant Hearing (JCIH) voted unanimously to proceed with the submission these EHDI measures to NQF. Liaison representatives were present from all of the participating organizations: American Academy of Pediatrics (AAP), American Academy of Audiology (AAA), American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS), American Speech-Language-Hearing Association (ASHA), Alexander Graham Bell Association for the Deaf and Hard of Hearing, Council of Education of the Deaf (CED), and Directors of Speech and Hearing Programs in State Health and Welfare Agencies (DSHPSHWA).	
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. CDC EHDI Data Committee and the Joint Committee on Infant Hearing (JCIH) both participated in the development of EHDI quality benchmarks on which this measure is based.	
Ad.2 If adapted, provide name of original measure: Ad.3-5 If adapted, provide original specifications URL or attachment	
Measure Developer/Steward Updates and Ongoing Maintenance Ad.6 Year the measure was first released: 2000 Ad.7 Month and Year of most recent revision: 10, 2007 Ad.8 What is your frequency for review/update of this measure? Ad.9 When is the next scheduled review/update for this measure?	
Ad.10 Copyright statement:	
Ad.11 Disclaimers:	
Ad.12 -14 Additional Information web page URL or attachment: URL http://jcih.org/posstatemts.htm	
Date of Submission (MM/DD/YY): 08/30/2010	