NATIONAL QUALITY FORUM

Measure Submission and Evaluation Worksheet 5.0

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

NQF #: 0477 NQF Project: Perinatal and Reproductive Health Project
(for Endorsement Maintenance Review) Original Endorsement Date: Oct 24, 2008 Most Recent Endorsement Date: Oct 24, 2008

BRIEF MEASURE INFORMATION
De.1 **Measure Title:** Under 1500g infant Not Delivered at Appropriate Level of Care
Co.1.1 **Measure Steward:** California Maternal Quality Care Collaborative
De.2 **Brief Description of Measure:** The number per 1,000 livebirths of <1500g infants delivered at hospitals not appropriate for that size infant.
2a1.1 **Numerator Statement:** Liveborn infants (<1500gms but over 24 weeks gestation) born at the given birth hospital
2a1.4 **Denominator Statement:** All live births over 24 weeks gestation at the given birth hospital. NICU Level III status is defined by the State Department of Health or similar body typically using American Academy of Pediatrics Criteria.
2a1.8 **Denominator Exclusions:** Stillbirths and livebirths <24weeks gestation.
1.1 **Measure Type:** Outcome 2a1. 25-26 **Data Source:** Electronic Clinical Data : Registry, Other 2a1.33 **Level of Analysis:** Facility, Health Plan, Population : County or City, Population : National, Population : Regional, Population : State 1.2-1.4 **Is this measure paired with another measure?** No De.3 **If included in a composite, please identify the composite measure** (*title and NQF number if endorsed*): n.a.

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

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

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

De.4 Subject/Topic Areas (*Check all the areas that apply*): Perinatal
De.5 Cross Cutting Areas (*Check all the areas that apply*): Access, Population Health

1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, A leading cause of morbidity/mortality, Patient/societal consequences of poor quality, Severity of illness

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

1a.3 Summary of Evidence of High Impact (*Provide epidemiologic or resource use data*):
Although they represent less than 2% of US births, 55% of infant deaths occur among VLBW (<1500g) infants. The best outcomes for VLBW infants occur when they are delivered in appropriate level perinatal units. California and other states have shown that <1500g infants have significantly better outcomes if delvieried in a facility with immediate access to a Regional or Community Level III Neonatal Intensive Care Unit (Phibbs, 2007). A recent CDC meta-analysis of all published data indicates a RR of 1.6 (CI 1.31 to 2.46) for death for births not delivered at a Level III facility (Lasswell, 2010). This analysis was nearly identical to that by the health economist Cirian Phibbs and associates (Holmstrom, 2009). Another recent study by the CDC demonstrated significant variation at the state level. In a study of 19 states (2006 data) showed state-wide rates that varied from only 68% (California) to a high of 93% (North Dakota)(CDC-MMWR, 2010). Each state has different nomenclature to describe levels of neonatal care, so the inventory needs to be first performed at a state level. Unfortunately, in many states, regionalization of neonatal care for <1500g infants has started to unravel largely due to non-clinical factors. California has made this measure a public release quality improvement action item in an attempt to improve outcomes fo these most fragile of infants. It is not expected that the rate would ever be zero, given that some labors are very short and transfer may not be possible, but each VLBW birth born at a low level of care should engender a QI action at the least. When rates have been examined throughout California over the last 5 years, the surprise was that the majority of the <1500g infants born not at an appropriate level of care were not occurring in the distant rural areas but rather in urban areas where the alternative Level III unit was often within close proximity, i.e., 10 miles distance.

California data indicates that there are hospitals with 10 fold higher rates of VLBW infants than others at a similar perinatal level. Perinatal Profiles (2006) including this data have been distributed to every hospital in California for the last 12 years. This year, it will be publically released as part of the California Hospital Accountability and Reporting Taskforce (CHART) in association with CMQCC. Recently, March of Dimes has sponsored a series of state-wide workshops to identify barriers to transfer and found that most were related to medical practice styles and economics rather than quality of care.

1a.4 Citations for Evidence of High Impact cited in 1a.3: Centers for Disease Control (CDC). Neonatal Intensive-Care Unit Admission of Infants with Very Low Birth Weight — 19 States, 2006. Morbidty and Mortality Weekly (MMWR). 2010:59 (44);1445-7.
Holmstrom ST, Phibbs CS. Regionalization and mortality in neomatal intensive care. Pediatr Clin N Am 2009 56:617–630
Lasswell, SM, Barfield WD, Rochat RW, Blackmon L. Perinatal Regionalization for Very Low-Birth-Weight and Very Preterm Infants: A Meta-analysis. JAMA. 2010;304(9):992-1000.

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

1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:
The use of this measure at the hospital level will strongly encourage hospitals to do the safer option of maternal transport rather than neonatal transport.

1b.2 Summary of Data Demonstrating Performance Gap (*Variation or overall less than optimal performance across providers*):
[*For <u>Maintenance</u> – Descriptive statistics for performance results <u>for this measure</u> - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.*]
In a study of 19 states (2006 data) showed state-wide rates that varied from only 68% (California) to a high of 93% (North Dakota)(CDC-MMWR, 2010).

1b.3 Citations for Data on Performance Gap: [*For <u>Maintenance</u> – 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]
Centers for Disease Control (CDC). Neonatal Intensive-Care Unit Admission of Infants with Very Low Birth Weight — 19 States, 2006. Morbidty and Mortality Weekly (MMWR). 2010:59 (44);1445-7.

1b.4 Summary of Data on Disparities by Population Group: [*For <u>Maintenance</u> –Descriptive statistics for performance results for this measure by population group*]
Data on disparities is being collected in California but not yet available.

1b.5 Citations for Data on Disparities Cited in 1b.4: [*For <u>Maintenance</u> – 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*]
n.a.

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

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

Quantity	Quality	Consistency	Does the measure pass subcriterion1c?
M-H	M-H	M-H	Yes☐
L	M-H	M	Yes☐ IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise **No**☐
M-H	L	M-H	Yes☐ IF potential benefits to patients clearly outweigh potential harms: otherwise **No**☐
L-M-H	L-M-H	L	No ☐

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

1c.1 Structure-Process-Outcome Relationship *(Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome)*:
This is a process measure (selction of the site for delviery) that is directly associated with levle of neonatal mortality--i.e. those VLBW infants not born at a Level III center have significantly higher mortality and morbidity rates.

1c.2-3 Type of Evidence *(Check all that apply)*:
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)*:
A recent CDC meta-analysis of all published data indicates a RR of 1.6 (CI 1.31 to 2.46) for death for births not delivered at a Level III facility (Lasswell, 2010). This analysis was nearly identical to that by the health economist Cirian Phibbs and associates (Holmstrom, 2009). There is no difference in the focus of the literature and the focus of the healthmeasure.

1c.5 Quantity of Studies in the Body of Evidence *(Total number of studies, not articles)*: 41 studies were judged to be of good quality by the recent CDC meta-analysis and 9 were judged to be of high-quality. Of note the high quality studies were consitent with the remaining body of literature.

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)*: a) Study design--none of the studeis were RCT but multiple ones were cohort analyses with extensive analyses for confounding factors; b)Directness--These studies were direcly related to the measure with large pateint populations and spread accross the US. c) confidence intervals for the CDC

meta-anlysis were RR 1.6 with CI (1.3 to 2.4)

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect):* The large majority of studies showed consistent results but in the CDC meta-analysis there was significant heterogenity. This heterogenity disappeared when the analysis was restricted to studies that adjusted for confounding factors.

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 relative risk of neonatal death was caculated to be 1.6 (60% increase) for infants not born at a level III center. A similar RR was found for neonatal morbidity. It was judged that most of this excessmorbidity and mortality could be eliminated by delivery at a Level III center.

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: ----------

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

1c.14 Summary of Controversy/Contradictory Evidence: Some delviery centers feel that they can deliver the baby and then judge whether the baby requires a neonatal transport to a Level III center. There is no evidence to support this approach.

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

1c.16 Quote verbatim, <u>the specific guideline recommendation</u> *(Including guideline # and/or page #)*:
HP2020 Objective MICH-33 Increase the proportion of very low birth weight (VLBW) infants born at level III hospitals or subspecialty perinatal centers.

1c.17 Clinical Practice Guideline Citation:
http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=26

1c.18 National Guideline Clearinghouse or other URL:

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: ----------------

1c.23 Grade Assigned to the Recommendation:

1c.24 Rationale for Using this Guideline Over Others: no other guidelines exist for this topic.

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

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

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

2. RELIABILITY & VALIDITY - SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (**evaluation criteria**)
Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate field. Supplemental materials may be referenced or attached in item 2.1. See guidance on measure testing.

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

S.2 If yes, provide web page URL:
http://www.cmqcc.org/maternity_quality_measures/delivery_of_1500g_at_a_non_level_iii_facility

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*):
Liveborn infants (<1500gms but over 24 weeks gestation) born at the given birth hospital

2a1.2 Numerator Time Window (*The time period in which the target process, condition, event, or outcome is eligible for inclusion*):
one year

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*):
Birthweight: <1500gms; Gestational Age >=24.0 weeks; livebirth (not stillbirth)

2a1.4 Denominator Statement (*Brief, narrative description of the target population being measured*):
All live births over 24 weeks gestation at the given birth hospital. NICU Level III status is defined by the State Department of Health or similar body typically using American Academy of Pediatrics Criteria.

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

2a1.6 Denominator Time Window (*The time period in which cases are eligible for inclusion*):
typically one year to collect sufficient cases to make a meaning statistic.

2a1.7 Denominator Details (*All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses*):
All live births at the hospital>=24weeks gestation. This is easily caculated from Vital Stats data. The field used is typocally the Best Obstetric Estimate of Gestational Age.

2a1.8 Denominator Exclusions (*Brief narrative description of exclusions from the target population*):
Stillbirths and livebirths <24weeks gestation.

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*):

Vital Stats data clearly identify stillbirths and Best Obstetric Gestational Age.

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)*:
none

2a1.11 Risk Adjustment Type *(Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13)*: No risk adjustment or risk stratification   **2a1.12 If "Other," please describe:**

2a1.13 Statistical Risk Model and Variables *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.)*:
n.a.

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

http://www.cmqcc.org/maternity_quality_measures/delivery_of_1500g_at_a_non_level_iii_facility
none 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 = Lower 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.)*:
For each hospital, first the total livebirths with Gestational age >=24 weeks is calculated. Then from among these cases, those that are <1500g birth weight are counted. Missing data for birthweight and gestational age is rare in the states data that have been examined (<1%). However, if the birthweight is <1500g but >=600gm and the gesational age is missing, the case would be counted but if it is <600gm and the gestational age is missing it is excluded. THis is because the GA is more likely to be missing than usual in very early gestational ages.

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:
URL
http://www.cmqcc.org/maternity_quality_measures/delivery_of_1500g_at_a_non_level_iii_facility
not needed

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)**:**
No sampling is performed, 100% of births are examined.

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

2a1.26 Data Source/Data Collection Instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*)**:** Other-Vital Records is the best source which allows this measure to be calculated very easily.

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment: URL
http://www.cmqcc.org/maternity_quality_measures/delivery_of_1500g_at_a_non_level_iii_facility
none needed

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

2a1.33 Level of Analysis (*Check the levels of analysis for which the measure is specified and tested*): Facility, Health Plan, Population : County or City, Population : National, Population : Regional, Population : State

2a1.34-35 Care Setting (*Check all the settings for which the measure is specified and tested*): 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*)**:**
Testing was performed on all 266 maternity hospitals in California for the year 2008. Total of 556,344 births.

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

2a2.3 Testing Results (*Reliability statistics, assessment of adequacy in the context of norms for the test conducted*)**:**
After exculing hospitals with <50 births/year and all Level III facilities, the following distribution was acheived: Ther range of VLBW births in non-level III facilities was 0 to 15 per thousand with a mean of 4.8. However, the distribution of hospitals was quite skewed: the median facility was 2.4 per thousand with 25%tile at 1.6 per thousand and 75%tile at 5.3 per thousand. this indicates that many facilites do quite well--over half had only 2.4 per thousand VLBW births while 25% had 5.5 and 10% had over 10 per thousand. 61% of all the VLBW not in Level III facilites were in the top 25% of the hosptials indicating that there was clear quality improvment opportunites.

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

2b1.1 Describe how the measure specifications (*measure focus, target population, and exclusions*) **are consistent with the evidence cited in support of the measure focus (***criterion 1c***) and identify any differences from the evidence:**
The measure uses the exact specifications used in the literature.

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*)**:**
2008 California state-wide data that included 266 facilities were evaluated, with 556,344 births. There was no sampling as the measure is a small rate.

2b2.2 Analytic Method (*Describe method of validity testing and rationale; if face validity, describe systematic assessment*)**:**
THis measure is consistent with national guidelines and extensive literature.

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*)**:**
n.a.

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

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

2b3.1 Data/Sample for analysis of exclusions (*Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*)**:**
No exclusions, but it is accepted that even good centers may occasionally have a rapid delveiry of a VLBW before transport. That accounts for rates of 1 to 3 per thousand.

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

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

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

2b4.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
no risk adjustment is proposed for this measure.

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

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

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

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

2b5.1 Data/Sample *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
2008 California state-wide data that included 266 facilities were evaluated, with 556,344 births.

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

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)*:
 After exculing hospitals with <50 births/year and all Level III facilities, the following distribution was acheived: Ther range of VLBW births in non-level III facilities was 0 to 15 per thousand with a mean of 4.8. However, the distribution of hospitals was quite skewed: the median facility was 2.4 per thousand with 25%tile at 1.6 per thousand and 75%tile at 5.3 per thousand. this indicates that many facilites do quite well--over half had only 2.4 per thousand VLBW births while 25% had 5.5 and 10% had over 10 per thousand. 61% of all the VLBW not in Level III facilites were in the top 25% of the hosptials indicating that there was clear quality improvment opportunites.

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)*:
Multiple data sources were not tested.

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

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

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

2c.1 If measure is stratified for disparities, provide stratified results *(Scores by stratified categories/cohorts)*: THis analysis was not performed.

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

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 Health/Disease Surveillance, Public Reporting, Quality Improvement (Internal to the specific organization), Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

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

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

3a.1. Use in Public Reporting - disclosure of performance results to the public at large *(If used in a public reporting program, provide name of program(s), locations, Web page URL(s)).* If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: [*For Maintenance – If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be considered.*]
We will releasing this data in the next 6months thru the Californai Hospital Accountablity and Reporting Taskforce (CHART) project. THis is an interactive website for a series of quality metrics. It is also slated to be used for quality improvmetn both regionally and wiothin the hospitals.

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 main utility of public release is to drive hospital administrators to change practice at their 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): None in action at this time but Medi-Cal will be using it as a qulaity

metric for medicaid services in California.

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*].
The obstetrica ndneonatal qulaity profgrmas in California, CMQCC and CPQCC, plan to use this measure for QI activities going forward once it is released. We have also received a CDC grant to operationalize a rapid cycle maternal data center to calcualte and erelease this data in near real time.

3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement. If usefulness was demonstrated *(e.g., QI initiative),* describe the data, method and results:
The measure is simple and clear. THey are also assocaited with a clear and simple quality objective. THe results show large variation that directs QI.

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

4. FEASIBILITY

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

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

4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*
Data used in the measure are:
Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

4b. Electronic Sources: H☐ M☐ L☐ I☐

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

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

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

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:
Birth records are very accurate for birthweight. Best OB estimate of Gestational age is also good at this age range. We are in the midst of doing a repeat audit of Gestational age through out California.

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

A.2 Please check if either of the following apply (*regarding proprietary measures*):
4d.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues *(e.g., fees for use of proprietary measures)*:
This measure can be easily collected for all hospitals using state vital records. There is almost no missing data for the two data elements (birth weight and gestaional age (<1%). The other data element-facility NICU level needs data collection once per year by survey.

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

OVERALL SUITABILITY FOR ENDORSEMENT
Does the measure meet all the NQF criteria for endorsement? Yes☐ No☐   **Rationale:**
If the Committee votes No, STOP.   If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.

5. COMPARISON TO RELATED AND COMPETING MEASURES
If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure before a final recommendation is made.
5.1 If there are related measures *(either same measure focus or target population)* or competing measures *(both the same measure focus and same target population)*, list the NQF # and title of all related and/or competing measures:
5a. Harmonization
5a.1 If this measure has EITHER the same measure focus OR the same target population as NQF-endorsed measure(s): Are the measure specifications completely harmonized?   **5a.2** If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden:
5b. Competing Measure(s)
5b.1 If this measure has both the same measure focus and the same target population as NQF-endorsed measure(s): Describe why this measure is superior to competing measures *(e.g., a more valid or efficient way to measure quality);* OR provide a rationale for the additive value of endorsing an additional measure. *(Provide analyses when possible)*:

CONTACT INFORMATION
Co.1 Measure Steward (Intellectual Property Owner): California Maternal Quality Care Collaborative, Medical School Office Building, 1265 Welch Road, MS 5415, Stanford, California, 94305   **Co.2 Point of Contact:** Elliott, Main, MD, main@cmqcc.org, 650-723-6108-
Co.3 Measure Developer if different from Measure Steward: California Maternal Quality Care Collaborative, Medical School Office Building, 1265 Welch Road, MS 5415, Stanford, California, 94305   **Co.4 Point of Contact:** Elliott, Main, MD, main@cmqcc.org, 650-723-6108-
Co.5 Submitter: Elliott, Main, MD, main@cmqcc.org, 650-723-6108-, California Maternal Quality Care Collaborative
Co.6 Additional organizations that sponsored/participated in measure development: none
Co.7 Public Contact: Elliott, Main, MD, main@cmqcc.org, 650-723-6108-, California Maternal Quality Care Collaborative

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.
CPQCC (California Perinatal Qualtiy Care Colaborative) is the state neonatal quality group that worked closely with us to devlelop the project.

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: same as previously sumitted.

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.3 Year the measure was first released: 2008
Ad.4 Month and Year of most recent revision: 06, 2011
Ad.5 What is your frequency for review/update of this measure? year;y
Ad.6 When is the next scheduled review/update for this measure? 06, 2012

Ad.7 Copyright statement: there is no copyright.

Ad.8 Disclaimers: none

Ad.9 Additional Information/Comments: none

Date of Submission (*MM/DD/YY*): 10/17/2011