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 #: 1746 NQF Project: Perinatal and Reproductive Health Project
(for Endorsement Maintenance Review) Original Endorsement Date: Most Recent Endorsement Date:
BRIEF MEASURE INFORMATION
De.1 Measure Title: Intrapartum Antibiotic Prophylaxis for Group B Streptococcus (GBS)
Co.1.1 Measure Steward: Massachusetts General Hospital
De.2 Brief Description of Measure: Percentage of pregnant women who are eligible for and receive appropriate intrapartum antibiotic prophylaxis (IAP) for Group B Streptococcus (GBS)
2a1.1 Numerator Statement: All eligible patients who receive intrapartum antibiotic prophylaxis for GBS.
2a1.4 Denominator Statement: All women delivering live infants, except certain classes (described in response to 2a1.9 below) who are specifically deemed not to be at risk of vertical transmission of GBS.
2a1.8 Denominator Exclusions: Women not included in the denominator defined above, with specific exclusions as described below.
1.1 Measure Type: Process 2a1. 25-26 Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Records 2a1.33 Level of Analysis: Facility, Integrated Delivery System, 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, OPPORTUITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT
Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All three subcriteria must be met to pass this criterion. See guidance on evidence. *Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.*

(evaluation criteria)

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

De.4 Subject/Topic Areas *(Check all the areas that apply)*: Infectious Diseases, Perinatal
De.5 Cross Cutting Areas (*Check all the areas that apply)*: Disparities, Safety : Healthcare Associated Infections

1a.1 Demonstrated High Impact Aspect of Healthcare: Severity of illness

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

1a.3 Summary of Evidence of High Impact *(Provide epidemiologic or resource use data)*:
Prevention of Group B streptococcus in newborns is a nationally recognized health priority where every pregnant woman is a potential carrier of GBS, and transmission of it to newborns carries substantial risk of neonatal infection and mortality. Approximately 10%-30% of pregnant women are colonized with GBS. Classic epidemiological studies in the 1980's revealed that women with prenatal GBS colonization were >25 times more likely than women with negative cultures to deliver infants with early-onset GBS disease.

1a.4 Citations for Evidence of High Impact cited in 1a.3: ACOG Committee on Obstetric Practice. Prevention of early-onset Group B Streptococcal disease in newborns. Obstet Gynecol 2011;117:1019-27.

CDC. Prevention of perinatal Group B Streptococcal disease: revised guidelines from CDC, 2010. MMWR 2010;59:1-32.

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:
Administering appropriate antibiotic prophylaxis to this patient population significantly decreases the risk of infection to their newborn further reducing risks of complications, readmissions, morbidity, mortality and the associated costs.

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.*]
Results of the Active Bacterial Core surveillance (ABCs) system of the Centers for Disease Control (CDC), as reported in 2009 in the New England Journal of Medicine, note, "Because chemoprophylaxis guidelines differ according to gestational age, we stratified... according to term or preterm delivery. Mothers who delivered preterm were less likely to receive chemoprophylaxis when indicated than mothers who delivered at term (relative risk, 0.81; 95% CI, 0.75 to 0.87). Among women who delivered preterm and were positive for group B streptococcus, 84.5% received chemoprophylaxis. However, only 63.4% of women who delivered preterm and had unknown colonization status received intrapartum antibiotics... The rate of administration of chemoprohylaxis was high among women who delivered at term: 87.0% of women who were positive for group B streptococcus and 78.5% of women with a risk factor and unknown colonization status received intrapartum antibiotics."

In use of the current measure in Massachusetts, the Medicaid Pay for Perforance program found average compliance of 71% in FY 2008, 83% in RY 2009, and 87% in RY 2010. (Data for each rate year are based on the preceding calendar year.)

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*]
Van Dyke MK, Phares CR, Lynfield R, et al. Evaluation of universal antenatal screening for Group B Streptococcus. N Engl J Med 2009;360:2626-36.

Personal communication from MassHealth Primary Provider Network, Massachusetts EOHHS, citing data reviewed in December 16, 2010

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

for this measure by population group]
The Active Bacterial Core surveillance (ABCs) project repoted in 2009, "When stratified by race, incidence [of early-onset GBS] among black infants increased significantly (0.52 to 0.86 cases per 1,000 live births, p=0.005, whereas incidence among white infants did not change significantly (0.26 to 0.29 cases per 1,000 live births; p=0.64). When EOD incidence was stratified by gestational age, the average incidence among preterm infants during 2003-2006 was 2.8 times higher among black infants (1.79 cases per 1,000 live births) compared with white infants (0.67 cases per 1,000 live births)... Th[e] increase in EOD from 203 to 2006... was not anticipated and cannot yet be explained fully... I[ntrapartum] A[ntimicrobial] P[rophylaxis] was administered to a similar proportion of black and white mothers of term infants with EOD... evaluation of these factors will be important in determining whether the causes of increasing racial difference in EOD can be directly linked to missed opportunities for prevention."

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*]
Trends in perinatal Group B Streptococcal disease, United States 2000-2006. MMWR 2009;58:109-112

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)*:
Process of GBS risk assessment and prophylaxis mitigating risk of morbidity and mortality resulting from vertical transmission.

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

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 incidence of GBS disease and means to prevent it have been extensively studied and codified. The current measure is addressed at implementation of the part of accepted standards that is under hospital control, i.e. intrapartum antibiotic prophylaxis for appropriate candidates.

1c.5 Quantity of Studies in the Body of Evidence *(Total number of studies, not articles)*: Current CDC guidelines cite ten clinical trials and well-designed observational studies directly supporting the use of IAP in peripartum GBS prophylaxis in its currently recommended form. A large number of other studies address related questions or possible alternative therapies.

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 quality of the body of evidence is high, with well-designed randomized clinical trial results during the 1980s. Since then, observational experience following the

implementation of GBS prophylaxis guidelines has been of an 80% reduction in early-onset disease. The evidence is by now so commonly accepted that it would be difficult to obtain IRB approval for a trial in which prophylaxis was withheld.

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect):* The mainstream of expert opinion accepts the consistency of evidence as high.

A Cochrane Collaboration review of 3 trials (852 women) found that there was a statistically insignificant trend toward reduction in early-onset disease with prophylaxis, but noted the possibility of bias was high. This review was probably statistically underpowered for the study of a relatively less frequent but potentially very serious outcome.

1c.8 Net Benefit *(Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms)*:
Current CDC guidelines cite the experience of 100% efficacy in early trials, with 86%-89% in subsequent observational studies.

1c.9 Grading of Strength/Quality of the Body of Evidence. Has the body of evidence been graded? Yes

1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: The individual components of the evidence were graded by the CDC technical working group. These grades are not included in the 2010 guidelines.

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

1c.12 If other, identify and describe the grading scale with definitions: Not available.

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

1c.14 Summary of Controversy/Contradictory Evidence: As with many other antibiotic practices, the use of antimicrobial programs for GBS is so accepted in clinical practice that it would be difficult to mount significant randomized trials to assess it. In particular, one Cochrane Collaboration reviewer (Ohlsson 2009) of three trials (852 women) that there was a reduction in the risk of early onset GBS infection with the use of intrapartum prophylaxis, but that this was not statistically significant. The review itself noted that the risk of bias was high, and the studies involved 20 years old. Besides these remarks from the reviewer, it should be noted that the combined studies were underpowered to detect any but very large treatment effects; in the 1960´s early onset neonatal sepsis caused by GBS had an attack rate of 2 per 1000 live births, but a 50% fatality rate. Larsen JW, Sever JL. Group B Streptococcus and pregnancy: a review. Amer J Obstet Gynecol 2008;440-40

The question arises from time to time, has the problem of GBS transmission been resolved, so that the result of continuing antibiotic prophylaxis is to encourage the spread of resistant organisms? Phares 2008 addresses this point, "Despite increasing antibiotic use, all isolates tested susceptible to penicillin and ampicillin (the first-line agents for intrapartum prophylaxis against early onset disease) and vancomycin. However, 32% of isolates were resistant to erythromycin, clindamycin, or both. This observation underscores the importance of performing susceptibility testing..."

1c.15 Citations for Evidence other than Guidelines*(Guidelines addressed below)*:
1. AGOC Committee on Obstetric Practice. Prevention of early-onset Group B Streptococcal disease in newborns. Obstet Gynecol 2011;117:1019-27.
2. Boyer KM, Gotoff SP. Prevention of early-onset neonatal Group B Streptococcal disease with selective intrapartum chemoprophylaxis. N Engl J Med 1986;314:4665-9.
3. CDC. Trends in perinatal Group B Streptococcal disease -- United States, 2000-2006. MMWR 2009;58:109-12.
4. CDC. Prevention of perinatal Group B Streptococcal disease: revised guidelines from CDC, 2010. MMWR 2010;59:1-32.
5. Colombo DF, Lew JL, Pedersen CA, Johnson JR, Fan-Havard P. Optimal timing of ampicillin administration to pregnant women for establishing bactericidal levels in the prophylaxis of Group B Streptococcus. Am J Obstet Gynecol 2005;194:466-40.
6. Goins WP, Talbot TR, Schaffner W, et al. Adherence to perinatal Group B Streptococcal prevention guidelines. Obstet Gynecol 2010;115:1217-24.
7. Larsen JW, Sever JL. Group B Streptococcus and pregnancy: a review. Am J Obstet Gynecol 2008:440-50.
8. Matteson KA, Lievense SP, Ctanzaro B, Philipps M. Intrapartum Group B Streptococci prophylaxis in patients reporting a

penicillin allergy. Obstet Gynecol 2008;111:358-64.
9. Ohlsson A, Shah VS. Intrapartum antibiotics for known Group B streptococcal colonization. Cochrane Database of Systematic Reviews 2009;3:Art. No. CD007467. DOI 10.1002/14651858.CD007 467.pub2.
10. Phares CR, Lynfield R, Farley M, et al. Epidemiology of invasive Group B Streptococcal disease in the United States, 1999-2005. JAMA 2008;299:2056-65.
11. Van Dyke MK, Phares CR, Lynfield R, et al. Evaluation of universal antenatal screening for Group B Streptococcus. N Engl J Med 2009;360:2626-36.
12. Verani JR, Schrag SJ. Group B Streptococcal disease in infants: progress in prevention and continued challenges. Clin Perinatol 2010;37:375-92.

1c.16 Quote verbatim, the specific guideline recommendation *(Including guideline # and/or page #)*:
"Intrapartum antibiotic prophylaxis agents and dosing should be administered according to the recommendations provided (Figure 8)." (Page 17 of reference)

1c.17 Clinical Practice Guideline Citation: CDC. Prevention of perinatal Group B Streptococcal disease, revised guidelines from CDC, 2010. MMWR 59(RR-10):1-32

1c.18 National Guideline Clearinghouse or other URL: http://www.cdc.gov/groupbstrep/index.html

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? Yes

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

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

1c.22 If other, identify and describe the grading scale with definitions: Not specified

1c.23 Grade Assigned to the Recommendation: The CDC grades the various components of the recommendations separately. The key component of penicillin (alternate: ampicillin) for intrapartum prophylaxis is graded AI.

1c.24 Rationale for Using this Guideline Over Others: The CDC-sponsored working group consisted of representatives from the ACOG Committee on Obstetric Practice, the American College of Nurse-Midwives (ACNM), the AAP Committee on Infectious Diseases and Committee on the Feturs and Newborn, the American Academy of Family Practitioners (AAFP), the Society for Healthcare Epidemiology of America (SHEA), the American Society for Microbiology (ASM), and CDC´s Active Bacterial Core Surveillance system, along with other experts.

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: Moderate 1c.26 Quality: High1c.27 Consistency: Moderate

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

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

2. RELIABILITY & VALIDITY - SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

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

S.1 Measure Web Page *(In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained*). Do you have a web page where current detailed specifications for this measure can be obtained? No

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)*:
All eligible patients who receive intrapartum antibiotic prophylaxis for GBS.

2a1.2 Numerator Time Window *(The time period in which the target process, condition, event, or outcome is eligible for inclusion)*:
At the time of labor or rupture of membranes, in the absence of complicating circumstances (listed as exclusions).

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*:
Patients who receive antibiotics as recommended under current CDC guidelines. The 2010 guidelines recommend penicillin as the agent of choice, with ampicillin as an acceptable alternative. Penicillin-allergic women who do not have a history of anaphylaxis, angioedema, respiratory distress or urticaria following administration of a penicillin or a cephalosporin should antimicrobial susceptibility testing. If the culture is susceptible to clindamycin, clindamycin should be given. If the culture is resistant to clindamycin, vancomycin should be given.

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured)*:
All women delivering live infants, except certain classes (described in response to 2a1.9 below) who are specifically deemed not to be at risk of vertical transmission of GBS.

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)*:
The interval from the time of labor or membrane rupture to delivery.

2a1.7 Denominator Details (*All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses)*:
The population may be identified in two stages. The first stage identified all women delivering live infants. The second stage further restricts the eligible population on the basis of specific clinical criteria.

Identification of women giving birth to live infants is generally a straightforward task that may be accomplished in various ways. Commonly, it is done using ICD-9 principal and secondary diagnosis codes for live births as defined in the Appendices of the National Hospital Quality Measures, as they may be modified from time to time. In 2011, codes for live births are listed in Appendix A Tables 4.01, 4.02, 4.03, or 4.04 of the Specifications Manual.

This population must be further restricted on the basis of the following criteria.
• Previous infant with invasive GBS disease, or
• GBS bacteriuria during current pregnancy, or
• Positive GBS screening culture during current pregnancy* (unless a planned cesarean delivery, in the absence of labor or amniotic membrane rupture, is performed), or
• Unknown GBS status (culture not done, incomplete or results unknown) and any of the following:
o Delivery at < 37 weeks gestation**
o Amniotic membrane rupture greater than or equal to 18 hours, or
o Intrapartum temperature greater than or equal to 100.4° F (38.0° C)

*Optimal timing for prenatal GBS screening is 35-37 weeks of gestation. In the absence of culture results for this period, other available results from the 5 weeks preceding delivery should be reviewed.

**Recommendations for prophylaxis in the setting of threatened preterm delivery are presented separately by the CDC in Figures 5 and 6 of the most recent guidelines (Centers for Disease Control and Prevention. Prevention of perinatal Group B Streptococcal disease: revised guidelines from CDC, 2010. MMWR 2010;59(RR-10):1-36.) Those interested in detailed criteria and assessment of compliance for the preterm population are referred there for specifics.

2a1.8 Denominator Exclusions *(Brief narrative description of exclusions from the target population)*:
Women not included in the denominator defined above, with specific exclusions as described below.

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)*:
Excluded populations:
• Patient screened negative for GBS at 35-37 weeks of delivery.
• Patients delivering via planned cesarean sections (in the absence of labor or amniotic membrane rupture).
• Patients already on antibiotics for a pre-natal maternal infection or other prophylaxis.
• Deliveries resulting in stillbirths identified by ICD-9-CM principal and secondary diagnosis codes (in any position) of V.27.1, V27.3, V27.4, V27.6, or V27.7.

*Optimal timing for prenatal GBS screening is 35-37 weeks of gestation. In the absence of culture results for this period, other available results from the 5 weeks preceding delivery should be reviewed.

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)*:

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.)*:

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.)*:
The score is calculated by dividing the numerator by the denominator. Where sample or population sizes are limited, the Measure Steward encourages the use of reporting with confidence intervals or graphical displays using standard statistical techniques for description of measurement error. The Measure Steward discourages ranking based on statistically indistinguishable scores.

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:
URL
The locally adapted MassHealth measure has its algorithm for rate calculation at http://www.mass.gov/?pageID=eohhs2terminal&L=4&L0=Home&L1=Government&L2=Laws%2c+Regulations+and+Policies&L3=MassHealth+Regulations+and+Other+Publications&sid=Eeohhs2&b=termin

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):
Hospitals that are capable can construct systems of real-time data capture that will enable routine reporting on their complete patient populations.

In situations where resources are limited, the Measure Steward suggests that the sampling methodologies and tables in use for the National Hospital Quality Measures provide one reasonable method of balance between statistically strong sample sizes and resource requirements.

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 Clinical Data : Electronic Health Record, Paper Records

2a1.26 Data Source/Data Collection Instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*): Appropriate data sources will vary from one institution to another. Typical sources of relevant information include administrative claims, electronic records, and paper records.

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment: Attachment
MAT-1 abstraction form 2011-Q2.doc

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:
URL
The related MassHealth measure has its data dictionary at http://www.mass.gov/?pageID=eohhs2terminal&L=4&L0=Home&L1=Government&L2=Laws%2c+Regulations+and+Policies&L3=MassHealth+Regulations+and+Other+Publications&sid=Eeohhs2&b=terminalcontent&f=mas

2a1.33 Level of Analysis (*Check the levels of analysis for which the measure is specified and tested)*: Facility, Integrated Delivery System, 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)*:
The reliability of the measure has been established by the background of successful data abstraction in the multi-state ABCs project of the CDC, and in use by the hospitals of Massachusetts. The developer has conducted specific study of test-retest reliability in 20 medical records at MGH.

The current measure of the delivery phase of care reflects a subset of the overall GBS standards of the CDC and others. The data elements in question are a subset of those included in the Neonatal Infection Extended Tracking form of the CDC Active Bacterial Core surveillance (ABCs) program. The ABCs covers populations in eight states, with its own quality control program. Thus, there is extensive experience in the overall survey process from which the current hospital phase measure is drawn.

The current measure has been in use in Massachusetts Medicaid pay for performance for a few years. Each hospital´s data abstraction is validated by a third party, just as for the National Hospital Quality Measures.

2a2.2 Analytic Method *(Describe method of reliability testing & rationale)*:
In the MGH study, two abstractors independently took data from the same 20 medical records. The pass-fail results for the measure, as well as the replicability of data abstraction for key variable, were calculated.

In the MassHealth program, copies of medical records (including electronic elements) are sent to the contracted reviewer, here associated with the state Quality Improvement Organization recognized by CMS. Hospital abstraction is verified against the raw materials used. As with the NHQM validation, a score of 80% (for all sample charts in all measures) is required in order for the hospital to pass validation.

2a2.3 Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
For the study at MGH, dual abstractions resulted in the same result (pass/fail/excluded) for all 20 cases. For 1 out of the 20 cases, abstractions were not concordant for the presence of risk factors for GBS. For all of the cases, the abstractions were concordant about what antibiotics were given.

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 predictive validity of the measure may be considered as its ability to anticipate desirable outcomes on the basis of observation of unambiguous elements of clinical practice.

For this measure, the observation of the clinical practice of antibiotic administration for women with specified risk factors for GBS colonization may be
observed with little doubt and with high reliability. The literature has demonstrated in extensive trials that these elements of clinical practice tend to lead to desirable outcomes, i.e. the avoidance of subsequent infections.

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)*:

2b2.2 Analytic Method *(Describe method of validity testing and rationale; if face validity, describe systematic assessment)*:
The testing of the practice of antibiotic prophylaxis is extensive, described the the discussion of the 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)*:
Antibiotic prophylaxis is general is a universally accepted construct in health care. The large and continually updated literature on antibiotic prophylaxis, including specifically use for GBS colonization, indicates that very strong face validity in review of medical records to establish what antibiotics are given at what times to women undergoing this procedure.

The Massachusetts Executive Office of Health and Human Services notes, "Hospital stakeholders continue to be actively engaged with
MassHealth in providing input to refine maternity measure specifications consistent with evidence-based practice standards."

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)*:
By design, antibiotic prophylaxis is not administered to women without defined risk factors for GBS.

These exclusions do not bias the measure, but simply eliminate from consideration those for whom the measure was not intended.

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

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

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

2b4.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
n/a

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

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

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

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 current measure is in use across Massachusetts hospitals. As with other common rate-based measures, there is not a simply definable numeric difference representing a critical threshold. However, percentages are widely used, with audiences having some intuitive sense of what a "big" and a "little" difference is.

2b5.2 Analytic Method *(Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance)*:
As with any common rate-based grading, the raw percentages provide a basis for understanding past performance. Where rates are based on small sample sizes, they are inherently approximate. If critical interpretation is important, the Measure Steward suggests reporting of the measure with confidence intervals or graphical techniques displaying the uncertainty of measurements.

The Measure Steward does not suggest use of the simple performance scores alone to rank institutions, or to group institutions into tiers.

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)*:

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)*:
As discussed in the section on reliablity, the developer has compared paper-based and electronic record results. It is to be expected that institutions will have some differences in the format of their medical records. However, the underlying data elements -- administration of medications, times and dates, and the results of laboratory testing -- are commonly compared across different

settings.

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

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)*:

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)*:

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

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 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)*: Payment Program, 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.*]
The measure is in current use in the Massachusetts Medicaid pay for performance program, administered by the Executive Office of Health and Human Services (EOHHS). There is Internet public reporting of hospital quality in Massachusetts, which however by legislative mandate is performed by another arm of the state government. At this time, EOHHS does not have authorization to report the pay for performance results on the Internet, but there are discussions about how this might be done.

An additional wrinkle for Massachusetts Medicaid is that CMS is in the process of developing a national public reporting plan, as required under the Affordable Care Act.

The barriers to reporting the current measure are not intrinsic, but logistic, developmental, and to some extent political.

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:

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 measure is in current use in the Massachusetts Medicaid hospital pay-for-performance program.

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 measure is in use in the Massachusetts Medicaid Pay for Performance Program. AS noted in the response to 1b.2 above, scores increased from 71% in FY 2008 to 83% in FY 2009 to 87% in FY 2010. The Executive Office of Health and Human Services comments --

"MAT-1: Intrapartum Antibiotic Prophylaxis for Group B Streptococcus. Measure rates for MAT 1 achieved statistically significant increases over the three years of the program. In RY08, measure rate results (71%) were not foreseen and the expectation was that rates would be higher given CDC clinical practice guidelines are universally used. In RY09 MAT-1 measure rates improved by 12% from RY08. From RY09 to RY10, 26 hospitals had an increase in measure rates out of which 8 hospitals showed a significant increase at the 5% significance level (p<.05). Measure rate failures continue to be due to the antibiotic not being administered to an eligible mother."

(For each rate year [RY], the data is collected for the earlier calendar year´s hospital discharges.)

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:
GBS prophylaxis is a high visibility initiative of the Centers for Disease Control (CDC). Results of the various components of GBS screening and antibiotic administration are widely cited, with a clear track record of successful improvement.

Massachusetts experience with the current measure has similary found that hospitals readily understand the results and have been able to work to improve them.

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

4. FEASIBILITY

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

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

4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*
Data used in the measure are:
generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition, Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims), 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): Some data elements are in electronic sources

4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources: Data elements will usually be present in hospitals with advanced electronic systems, but others will need to review paper charts.

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

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:
The measure is generally straightforward and unambiguous. As with all guidelines, clinicians should exercise appropriate judgment in unusual cases where standard treatments may not be the most appropriate approach.

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)*:
It is of course impossible to isolate the hospital phase entirely from overall perinatal care, but in practice we have found that hospitals usually have very good information about relevant aspects of prenatal care.

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:
0004 : Initiation and Engagement of Alcohol and Other Drug Dependence Treatment: a. Initiation, b. Engagement

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): Massachusetts General Hospital, 55 Fruit Street, Boston, Massachusetts, 02114 Co.2 Point of Contact: Paul, Nordberg, M.S., pnordberg@partners.org, 617-724-8269-
Co.3 Measure Developer if different from Measure Steward: Massachusetts General Hospital, 55 Fruit Street, Boston, Massachusetts, 02114 Co.4 Point of Contact: Paul, Nordberg, M.S., pnordberg@partners.org, 617-724-8269-
Co.5 Submitter: Paul, Nordberg, M.S., pnordberg@partners.org, 617-724-8269-, Massachusetts General Hospital
Co.6 Additional organizations that sponsored/participated in measure development: Jeffrey Ecker, M.D., clinical sponsor Vincent Obstetrics Service Massachusetts General Hospital Paul Nordberg, administrative sponsor Center for Quality and Safety Massachusetts General Hospital
Co.7 Public Contact: Paul, Nordberg, M.S., pnordberg@partners.org, 617-724-8269-, Massachusetts General Hospital

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. Jeffrey Ecker, M.D., clinical sponsor Vincent Obstetrics Service Massachusetts General Hospital Dr. Ecker oversaw the development of the measure, and has collaborated in related work with colleagues across the hospitals of the Partners Healthcare system.
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: Ad.4 Month and Year of most recent revision: Ad.5 What is your frequency for review/update of this measure? As significant guideline and evidence changes emerge Ad.6 When is the next scheduled review/update for this measure? 11, 2012
Ad.7 Copyright statement:
Ad.8 Disclaimers:
Ad.9 Additional Information/Comments:
Date of Submission (*MM/DD/YY*): 10/17/2011