

# Establishment of Gestational Age

## Section 1. Basic Measure Information

### 1.A. Measure Name

Establishment of Gestational Age

### 1.B. Measure Number

0169

### 1.C. Measure Description

**Please provide a non-technical description of the measure that conveys what it measures to a broad audience.**

This measure is designed to assess the percentage of patients who had gestational age of the fetus estimated by ultrasound at or prior to 20 weeks (20 weeks initially estimated by date of last menstrual period [LMP]). The measure was developed by the American Medical Association (AMA)-convened Physician Consortium for Performance Improvement (PCPI<sup>®</sup>), which is a key member of the Pediatric Measurement Center of Excellence (PMCoE). The PMCoE is funded by AHRQ and includes the following consortium members: the American Academy of Pediatrics, American Board of Pediatrics, American Board of Medical Specialties, Northwestern University, Truven Health Analytics (formerly Thomson Reuters), Children's Hospital and Health System-Milwaukee, Medical College of Wisconsin, and the AMA.

### 1.D. Measure Owner

The AMA-convened PCPI is the measure owner.

### 1.E. National Quality Forum (NQF) ID (if applicable)

Not applicable.

### 1.F. Measure Hierarchy

**Please note here if the measure is part of a measure hierarchy or is part of a measure group or composite measure. The following definitions are used by AHRQ:**

- 1. Please identify the name of the collection of measures to which the measure belongs (if applicable). A collection is the highest possible level of the measure hierarchy. A collection may contain one or more sets, subsets, composites, and/or individual measures.**

Prenatal/Perinatal Performance Measurement Set

- 2. Please identify the name of the measure set to which the measure belongs (if applicable). A set is the second level of the hierarchy. A set may include one or more subsets, composites, and/or individual measures.**

Not applicable.

- 3. Please identify the name of the subset to which the measure belongs (if applicable). A subset is the third level of the hierarchy. A subset may include one or more composites, and/or individual measures.**

Not applicable.

- 4. Please identify the name of the composite measure to which the measure belongs (if applicable). A composite is a measure with a score that is an aggregate of scores from other measures. A composite may include one or more other composites and/or individual measures. Composites may comprise component measures that can or cannot be used on their own.**

Not applicable.

## **1.G. Numerator Statement**

Patients who had gestational age of the fetus estimated by ultrasound at or prior to 20 weeks (20 weeks initially estimated by date of LMP).

## **1.H. Numerator Exclusions**

None.

## **1.I. Denominator Statement**

All patients, regardless of age, who gave birth during a 12-month period and were seen at least once for prenatal care.

## **1.J. Denominator Exclusions**

None.

## **1.K. Data Sources**

**Check all the data sources for which the measure is specified and tested.**

Electronic health record (EHR).

**If other, please list all other data sources in the field below.**

## **Section 2: Detailed Measure Specifications**

**Provide sufficient detail to describe how a measure would be calculated from the recommended data sources, uploading a separate document (+ Upload attachment) or a link to a URL. Examples of detailed measure specifications can be found in the CHIPRA Initial Core Set Technical Specifications Manual 2011 published by the Centers for Medicare & Medicaid Services. Although submission of formal programming code or algorithms that demonstrate how a measure would be calculated from a query of an appropriate electronic data source are not requested at this time, the availability of these resources may be a factor in determining whether a measure can be recommended for use.**

Please see supporting documents for full eSpecifications and coding spreadsheets. The measure specifications include the following components: (1) a text description of the measure; (2) the Data Requirements Table, which outlines the data elements that are required for the measure, including the identification of the clinical vocabularies applicable to a given data element, the NQF Quality Data Model category and State, as well as the timing parameters for each data element; (3) a visual flow diagram that uses Boolean logic to identify the initial patient population, exclusions, denominator, numerator and exceptions included in the measure; (4) measure calculation; and (5) value sets for each of the data elements. The measure specifications provide the required information to collect the data needed to calculate the quality measure.

## **Section 3. Importance of the Measure**

**In the following sections, provide brief descriptions of how the measure meets one or more of the following criteria for measure importance (general importance, importance to Medicaid and/or CHIP, complements or enhances an existing measure). Include references related to specific points made in your narrative (not a free-form listing of citations).**

### **3.A. Evidence for General Importance of the Measure**

**Provide evidence for all applicable aspects of general importance:**

- Addresses a known or suspected quality gap and/or disparity in quality (e.g., addresses a socioeconomic disparity, a racial/ethnic disparity, a disparity for Children with Special Health Care Needs (CSHCN), a disparity for limited English proficient (LEP) populations).**
- Potential for quality improvement (i.e., there are effective approaches to reducing the quality gap or disparity in quality).**
- Prevalence of condition among children under age 21 and/or among pregnant women**
- Severity of condition and burden of condition on children, family, and society (unrelated to cost)**
- Fiscal burden of measure focus (e.g., clinical condition) on patients, families, public and private payers, or society more generally, currently and over the life span of the child.**

- **Association of measure topic with children’s future health – for example, a measure addressing childhood obesity may have implications for the subsequent development of cardiovascular diseases.**
- **The extent to which the measure is applicable to changes across developmental stages (e.g., infancy, early childhood, middle childhood, adolescence, young adulthood).**

The use of ultrasonography to assess for potential fetal abnormalities, confirm the site of pregnancy within the uterus, and determine gestational age is considered the standard of care. Also, the use of ultrasound scanning during the first trimester is correlated with reduced post-term labor induction rates as compared to second trimester ultrasound scanning.

A critical factor in assessing infant mortality risk is gestational age, since it has been shown that for any constant birth weight the mortality rate decreases as gestational age increases. Accurate gestational age is also critical to the timing of birth and decisions made for many procedures related to birth. Patients with an accurate gestational age established are at lower risk of having procedures at inappropriate times during the pregnancy.

### **3.B. Evidence for Importance of the Measure to Medicaid and/or CHIP**

**Comment on any specific features of this measure important to Medicaid and/or CHIP that are in addition to the evidence of importance described above, including the following:**

- **The extent to which the measure is understood to be sensitive to changes in Medicaid or CHIP (e.g., policy changes, quality improvement strategies).**
- **Relevance to the Early and Periodic Screening, Diagnostic and Treatment benefit in Medicaid (EPSDT).**
- **Any other specific relevance to Medicaid/CHIP (please specify).**

This measure would fill a gap in the Medicaid and CHIP programs’ core set of children’s health care quality measures aimed at providing services and treatment to promote healthy birth and prevent premature birth. This measure is important to Medicaid and CHIPRA because it expands the core set of measures beyond their current use. The measure will provide a mechanism to help assess how many women in the Medicaid population have an accurate measure of gestational age, which can help prevent adverse maternal and neonatal outcomes. This measure is of particular importance for CHIPRA in that it is high impact with Medicaid patients, and it addresses concerns related to both mother and baby. We encourage the use of this measure by physicians, other health care professionals, and health care systems and health plans, where appropriate. This clinical performance measure is designed for practitioner and/or system-level quality improvement to achieve better outcomes for maternity care patients and their babies.

### **3.C. Relationship to Other Measures (if any)**

**Describe, if known, how this measure complements or improves on an existing measure in this topic area for the child or adult population, or if it is intended to fill a specific gap in an existing measure category or topic. For example, the proposed measure may enhance an existing measure in the initial core set, it may lower the age range for an existing adult-**

**focused measure, or it may fill a gap in measurement (e.g., for asthma care quality, inpatient care measures).**

There are no other measures that address this topic.

## **Section 4. Measure Categories**

**CHIPRA legislation requires that measures in the initial and improved core set, taken together, cover all settings, services, and topics of health care relevant to children. Moreover, the legislation requires the core set to address the needs of children across all ages, including services to promote healthy birth. Regardless of the eventual use of the measure, we are interested in knowing all settings, services, measure topics, and populations that this measure addresses. These categories are not exclusive of one another, so please indicate "Yes" to all that apply.**

**Does the measure address this category?**

- a. Care Setting – ambulatory: Yes.**
- b. Care Setting – inpatient: No.**
- c. Care Setting – other – please specify: No.**
- d. Service – preventive health, including services to promote healthy birth: Yes.**
- e. Service – care for acute conditions: No.**
- f. Service – care for children with acute conditions: No.**
- g. Service – other (please specify): No.**
- h. Measure Topic – duration of enrollment: No.**
- i. Measure Topic – clinical quality: Yes.**
- j. Measure Topic – patient safety: Yes.**
- k. Measure Topic – family experience with care: No.**
- l. Measure Topic – care in the most integrated setting: No.**
- m. Measure Topic other (please specify): No.**
- n. Population – pregnant women: Yes.**
- o. Population – neonates (28 days after birth) (specify age range): No.**
- p. Population – infants (29 days to 1 year) (specify age range): No.**
- q. Population – pre-school age children (1 year through 5 years) (specify age range): No.**
- r. Population – school-aged children (6 years through 10 years) (specify age range): No.**
- s. Population – adolescents (11 years through 20 years) (specify age range): No.**
- t. Population – other (specify age range): No.**
- u. Other category (please specify): Not applicable.**

## **Section 5. Evidence or Other Justification for the Focus of the Measure**

**The evidence base for the focus of the measures will be made explicit and transparent as part of the public release of CHIPRA deliberations; thus, it is critical for submitters to**

specify the scientific evidence or other basis for the focus of the measure in the following sections.

## **5.A. Research Evidence**

Research evidence should include a brief description of the evidence base for valid relationship(s) among the structure, process, and/or outcome of health care that is the focus of the measure. For example, evidence exists for the relationship between immunizing a child or adolescent (process of care) and improved outcomes for the child and the public. If sufficient evidence existed for the use of immunization registries in practice or at the State level and the provision of immunizations to children and adolescents, such evidence would support the focus of a measure on immunization registries (a structural measure).

Describe the nature of the evidence, including study design, and provide relevant citations for statements made. Evidence may include rigorous systematic reviews of research literature and high-quality research studies.

The following evidence statements are quoted verbatim from the referenced clinical guidelines:

### **Ultrasonography in Pregnancy (ACOG, 2009)**

(Level A):

- Ultrasound examination is an accurate method of determining gestational age, fetal number, viability, and placental location.
- Gestational age is most accurately determined in the first half of pregnancy.
- Ultrasonography can be used in the diagnosis of many major fetal anomalies.
- Ultrasonography is safe for the fetus when used appropriately.

(Level B):

- Ultrasonography is helpful in detecting fetal growth disturbances.
- Ultrasonography can detect abnormalities in amniotic fluid volume.

(Level C):

- The optimal timing for a single ultrasound examination in the absence of specific indications for a first trimester examination is at 18–20 weeks of gestation.

The benefits and limitations of ultrasonography should be discussed with all patients.

### **Reference**

American College of Obstetricians and Gynecologists (ACOG). Ultrasonography in pregnancy. ACOG Practice Bulletin No. 101. *Obstet Gynecol* 2009; 113:451-61. Available at <http://www.acog.org/About-ACOG/ACOG-Departments/Deliveries-Before-39-Weeks/ACOG-Clinical-Guidelines>. Accessed October 25, 2016.

## **5.B. Clinical or Other Rationale Supporting the Focus of the Measure (optional)**

**Provide documentation of the clinical or other rationale for the focus of this measure, including citations as appropriate and available.**

See Section 3.

## **Section 6. Scientific Soundness of the Measure**

**Explain the methods used to determine the scientific soundness of the measure itself. Include results of all tests of validity and reliability, including description(s) of the study sample(s) and methods used to arrive at the results. Note how characteristics of other data systems, data sources, or eligible populations may affect reliability and validity.**

### **6.A. Reliability**

**Reliability of the measure is the extent to which the measure results are reproducible when conditions remain the same. The method for establishing the reliability of a measure will depend on the type of measure, data source, and other factors.**

**Explain your rationale for selecting the methods you have chosen, show how you used the methods chosen, and provide information on the results (e.g., the Kappa statistic). Provide appropriate citations to justify methods.**

#### **Reliability Testing Results – Inter-rater Reliability Approach**

Inter-rater reliability is used to assess the reliability of the measure based on results from two independent reviewers trained in the same way reviewing the same patient record. To perform inter-rater reliability testing, we created an electronic data collection tool and trained the reviewers (raters) on its use. We selected a random sample of 75 patients for review. The reviewers separately reviewed every sampled patient and collected all data elements necessary for computation of the performance measure (contained on the electronic data collection tool). Data analyses were conducted using SAS/STAT software, version 9.3 (SAS Institute, Cary, NC).

For analysis of kappa, we used the technique described by Landis and Koch (1977):

- 0.0–.20 = slight agreement
- .21–.40 = fair agreement
- .41–.60 = moderate agreement
- .61–.80 = substantial agreement
- .81–1.0 = almost perfect agreement

A kappa less than zero would indicate agreement worse than that expected by chance (Bartko, Carpenter, 1976).

The test results are as follows. We calculated a kappa statistic to assess measure reliability. There were 75 observations from one site used for the reliability analysis.

The kappa statistic results find that there is almost perfect agreement between the two reviewers (0.90). The observed agreement percentage is 0.96 (Table 1).

**Table 1. Inter-rater Reliability Results**

Reviewer 1	Reviewer 2		Total
	NO	YES	
NO	18	0	18
YES	3	54	57
Total	21	54	75

  

Kappa	0.90
95% Lower Confidence Limit	0.78
95% Upper Confidence Limit	1.00
Observed Agreement Percentage	0.96

### Performance Rate

The performance rate is determined by having both reviewers agree that the measure was met for the patient. For this sample, the performance rate was 72 percent (54/75).

### References

Bartko JJ, Carpenter WT. (1976). On the methods and theory of reliability. *J Nerv Ment Dis.*163:307–317.

Landis J.R, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33(1):159–74.

## 6.B. Validity

**Validity of the measure is the extent to which the measure meaningfully represents the concept being evaluated. The method for establishing the validity of a measure will depend on the type of measure, data source, and other factors.**

*Explain your rationale for selecting the methods you have chosen, show how you used the methods chosen, and provide information on the results (e.g., R2 for concurrent validity).*

The measure was assessed for content validity and face validity. Evidence of content validity is provided by looking for agreement among subject matter experts. The performance measure was assessed for content validity by a panel of 24 expert workgroup members during the development process. This subject matter expert panel included representation from measure methodologists, patient advocacy groups, and the following clinical specialties: anesthesiology, family practice, geriatric medicine, maternal fetal medicine, neonatology, nurse midwife, obstetrics and gynecology, and perinatal nursing. Additional input on the content validity of draft measures was obtained through a 30-day public comment period and by also soliciting comments from a panel of consumer, purchaser, and patient representatives convened by the PCPI specifically for this purpose. All comments received were reviewed by the expert workgroup, and the measure was adjusted as needed.



The expert panel members also assessed the measure face validity through an online survey. The survey introduction provided the following definition of face validity: Face validity is the extent to which an empirical measurement appears to reflect that which it is supposed to “at face value.” Face validity of an individual measure poses the question of how well the definition and specifications of an individual measure appear to capture the single aspect of care or health care quality as intended. The expert panel was asked to rate their agreement with the following statement: The scores obtained from the measure as specified will accurately differentiate quality across providers. A 5-point Likert scale was used in the survey (1=Strongly Disagree; 2=Disagree; 3=Neither Disagree nor Agree; 4 = Agree 5=Strongly Agree).

The survey results show that for the Establishment of Gestational Age measure, the mean score was 4.31; 84.6 percent (11/13) of respondents agree or strongly agree that the scores obtained from the measure as specified will accurately differentiate quality across providers; and no respondents disagree or strongly disagree that the scores obtained from the measure as specified will accurately differentiate quality across providers.

## **Section 7. Identification of Disparities**

**CHIPRA requires that quality measures be able to identify disparities by race, ethnicity, socioeconomic status, and special health care needs. Thus, we strongly encourage nominators to have tested measures in diverse populations. Such testing provides evidence for assessing measure’s performance for disparities identification. In the sections below, describe the results of efforts to demonstrate the capacity of this measure to produce results that can be stratified by the characteristics noted and retain the scientific soundness (reliability and validity) within and across the relevant subgroups.**

### **7.A. Race/Ethnicity**

**Recognizing that individuals with differing races and ethnicities make up a diverse population of individuals with needs of varying complexity, please describe the results of any efforts to demonstrate the capacity of this measure to produce results that stratify by race and ethnicity.**

We included race and ethnicity as Supplemental Data Elements to collect for each measure to allow for the stratification of measure results by these variables to assess disparities and initiate subsequent quality improvement activities. The Centers for Disease Control and Prevention (CDC) value sets for race and ethnicity are referenced in the measure specifications to collect race and ethnicity information, which is the requirement for race and ethnicity outlined in the Centers for Medicare & Medicaid Services (CMS) Blueprint (CMS, 2015).

### **7.B. Special Health Care Needs**

Not applicable for this measure.

### **7.C. Socioeconomic Status**

We include payer as a Supplemental Data Element to collect for each measure to allow for the stratification of measure results by these variables to assess disparities and initiate subsequent

quality improvement activities addressing identified disparities, consistent with recent national efforts to standardize the collection of socioeconomic status data. The Payment Typology value set (Bernstein, 2007) is referenced in the measure specifications to collect payer information, which is the requirement for payer outlined in the CMS (2015) Blueprint.

## **7.D. Rurality/Urbanicity**

Future measure testing and implementation should collect data on the location of the patient and provider populations in order to stratify performance and test for variation by location.

## **7.E. Limited English Proficiency (LEP) Populations**

We include preferred language as a Supplemental Data Element to collect for each measure to allow for the stratification of measure results by these variables to assess disparities and initiate subsequent quality improvement activities, as noted above.

## **References**

Bernstein A. Source payment typology code set. Hyattsville, MD: National Center for Health Statistics; 2007. Available at [https://www.healthit.gov/archive/archive\\_files/FACA%20Hearings/2010/2010-03-23%20Vocabulary%20Task%20Force/4Bernstein\\_testimony032310.pdf](https://www.healthit.gov/archive/archive_files/FACA%20Hearings/2010/2010-03-23%20Vocabulary%20Task%20Force/4Bernstein_testimony032310.pdf). Accessed October 25, 2016.

Centers for Medicare & Medicaid Services (CMS). A blueprint for the CMS Measures Management System. Baltimore, MD: CMS; 2015. Available at <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Blueprint111.pdf>. Accessed October 25, 1026.

# **Section 8. Feasibility**

**Feasibility is the extent to which the data required for the measure are readily available, retrievable without undue burden, and can be implemented for performance measurement. Using the following sections, explain the methods used to determine the feasibility of implementing the measure.**

## **8.A. Data Availability**

**1. What is the availability of data in existing data systems? How readily are the data available?**

### **Data Element Table (DET) Tool**

The PMCoE Center of Excellence adopted the AMA-PCPI testing methodology which uses the Data Element Table (DET) tool to assess the availability of the data and the technical feasibility and implementation feasibility of the measures. The DET is an Excel workbook designed to capture information that will determine whether or not it is feasible for each site to collect the data for the measures. It is structured to collect meta-data about each data element necessary to

construct each measure stored in the EHR. It will also collect information related to the integrity and validity of data collection. Specifically, the DET is designed to capture the following information:

**Data element information:** Whether or not the data element is captured in the HER and the data source application, primary user interface data location, data type, coding system, unit of measure, frequency of collection, and calculability within the measure context.

**Measure integrity information:** An assessment by the testing site as to what degree the measure, as specified, retains the originally stated intent of the measure.

**Measure validity information:** An assessment by the testing site as to what degree the scores obtained from the measure, as specified, will accurately differentiate quality performance across providers.

The DETs collected responses used to assess technical and implementation feasibility for each measure. Measure technical feasibility was defined as “Can my EHR do this?” Measure implementation feasibility was defined as “Will workflow be used consistently?” The responses were captured in the form of a rating using the following responses:

- “Feasible. Can do today.”
- “Feasible with workflow mod/changes to EHR.”
- “Non-feasible. Unable to do today.”

This information was entered from drop-down options pertaining to the specific criteria and in free text fields for questions related to specific workflow and EHR configurations. The free text fields and specific narrative questions provide qualitative feedback from the sites that can be factored into the overall feasibility grade for the measure. The DET is completed by staff at each testing site. After the completion of the DET by the testing sites, a determination can be made as to which of the measures are relevant for each specific site. For some sites, all of the measures in the Perinatal/Prenatal Measurement Set may be collected, for others it may be only a few. Once the completed DETs were submitted by the test sites, the PMCoE project team conducted quality assurance review of the DETs to ensure the data were complete and ready for analysis. A series of analyses were subsequently performed to characterize the feasibility, integrity, and face validity of the measures being tested.

Feasibility testing was conducted at an urban, public hospital. The test site reported that their EHR could capture all seven data elements in code, text, or Boolean format. The elements are entered in free text.

**2. If data are not available in existing data systems or would be better collected from future data systems, what is the potential for modifying current data systems or creating new data systems to enhance the feasibility of the measure and facilitate implementation?**

**Measure Technical Feasibility and Implementation Feasibility**

The measure technical feasibility assessment determined how many of the total measure data elements are feasible data elements to collect. A “feasible data element” is one that can be captured by the test site EHR system. The sites assessed technical feasibility for the measure based on the following rating scale:

- “Feasible. Can do today.”
- “Feasible with workflow mod/changes to EHR.”
- “Non-feasible. Unable to do today.”

The sites also used this scale to assess measure implementation feasibility. Implementation feasibility represents the site’s ability to implement the measure using current workflows and EHRs and addresses issues of projected data reliability related to the consistency with which providers document and capture the data elements needed to implement the measure.

The technical feasibility and implementation feasibility were rated the same for each of the measures. For example, if the technical feasibility of a measure was rated as “Feasible. Can do today,” its implementation feasibility was also rated as “Feasible. Can do today.”

The site that evaluated the technical and implementation feasibility for this measure selected the rating of “Feasible with workflow mod/changes to EHR.” Although the data elements were captured in free text, the information provided by the site indicates that changes to their EHR would be required in order for them to calculate the measure.

## **8.B. Lessons from Use of the Measure**

### **1. Describe the extent to which the measure has been used or is in use, including the types of settings in which it has been used, and purposes for which it has been used.**

We are not aware of any project or setting where this measure has been used or is currently in use.

### **2. If the measure has been used or is in use, what methods, if any, have already been used to collect data for this measure?**

We are not aware of any project or setting where this measure has been used or is currently in use.

### **3. What lessons are available from the current or prior use of the measure?**

We are not aware of any project or setting where this measure has been used or is currently in use.

## **Section 9. Levels of Aggregation**

**CHIPRA states that data used in quality measures must be collected and reported in a standard format that permits comparison (at minimum) at State, health plan, and provider**

levels. Use the following table to provide information about this measure's use for reporting at the levels of aggregation in the table.

For the purpose of this section, please refer to the definitions for provider, practice site, medical group, and network in the Glossary of Terms.

If there is no information about whether the measure could be meaningfully reported at a specific level of aggregation, please write "Not available" in the text field before progressing to the next section.

*Level of aggregation (Unit) for reporting on the quality of care for children covered by Medicaid/ CHIP†:*

*State level\* Can compare States*

**Intended use:** Is measure intended to support meaningful comparisons at this level?  
(Yes/No)

Yes.

**Data Sources:** Are data sources available to support reporting at this level?

Yes.

**Sample Size:** What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Dependent on specific area.

**In Use:** Have measure results been reported at this level previously?

No.

**Reliability & Validity:** Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

**Unintended consequences:** What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

*Other geographic level: Can compare other geographic regions (e.g., MSA, HRR)*

**Intended use:** Is measure intended to support meaningful comparisons at this level?  
(Yes/No)

Yes.

**Data Sources:** Are data sources available to support reporting at this level?

Yes.

**Sample Size:** What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Dependent on specific area.

**In Use:** Have measure results been reported at this level previously?

No.

**Reliability & Validity:** Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

**Unintended consequences:** What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

**Medicaid or CHIP Payment model:** Can compare payment models (e.g., managed care, primary care case management, FFS, and other models)

**Intended use:** Is measure intended to support meaningful comparisons at this level? (Yes/No)

No.

**Data Sources:** Are data sources available to support reporting at this level?

No.

**Sample Size:** What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Not applicable.

**In Use:** Have measure results been reported at this level previously?

No

**Reliability & Validity:** Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

**Unintended consequences:** What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

**Health plan\*:** Can compare quality of care among health plans.

**Intended use:** Is measure intended to support meaningful comparisons at this level? (Yes/No)

Yes.

**Data Sources:** Are data sources available to support reporting at this level?

Yes.

**Sample Size:** What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

All health plans serving pregnant women will have a sufficient sample size to support reliable performance measurement and to distinguish between good and poor performance.

**In Use:** Have measure results been reported at this level previously?

No.

**Reliability & Validity:** Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

**Unintended consequences:** What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

**Provider Level**

**Individual practitioner:** Can compare individual health care professionals

**Intended use:** Is measure intended to support meaningful comparisons at this level?  
(Yes/No)

Yes.

**Data Sources:** Are data sources available to support reporting at this level?

Yes.

**Sample Size:** What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Dependent on the size of the practice. All large OB provider practices serving pregnant women will have a sufficient sample size to support reliable performance measurement and to distinguish between good and poor performance.

**In Use:** Have measure results been reported at this level previously?

No.

**Reliability & Validity:** Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

**Unintended consequences:** What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

*Provider Level*

*Hospital: Can compare hospitals*

**Intended use: Is measure intended to support meaningful comparisons at this level?  
(Yes/No)**

No.

**Data Sources: Are data sources available to support reporting at this level?**

No.

**Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?**

Dependent on the size of the OB practice in the hospital. All large hospitals serving pregnant women will have a sufficient sample size to support reliable performance measurement and to distinguish between good and poor performance.

**In Use: Have measure results been reported at this level previously?**

No.

**Reliability & Validity: Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?**

No.

**Unintended consequences: What are the potential unintended consequences of reporting at this level of aggregation?**

Unknown.

*Provider Level*

*Practice, group, or facility:\*\* Can compare: (i) practice sites; (ii) medical or other professional groups; or (iii) integrated or other delivery networks*

**Intended use: Is measure intended to support meaningful comparisons at this level?  
(Yes/No)**

Yes.

**Data Sources: Are data sources available to support reporting at this level?**

Yes.

**Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?**

All large provider practices serving pregnant women will have a sufficient sample size to support reliable performance measurement and to distinguish between good and poor performance.

**In Use: Have measure results been reported at this level previously?**



No.

**Reliability & Validity:** Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

**Unintended consequences:** What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

## **Section 10. Understandability**

**CHIPRA states that the core set should allow purchasers, families, and health care providers to understand the quality of care for children. Please describe the usefulness of this measure toward achieving this goal. Describe efforts to assess the understandability of this measure (e.g., focus group testing with stakeholders).**

The AMA-PCPI has worked collaboratively on this measure set with the AMA-PCPI-Consumer Purchaser Panel (CPP), which comprises representatives from the patient, consumer, and purchaser communities. The panel strongly supports this measure and applauds its inclusion at the level of the individual clinician. The CPP notes that this important measure can help to reduce adverse maternal and neonatal outcomes, as well as reduce medical costs. In addition, the workgroup comprised member representatives from consumer groups, patient advocacy groups, and a health plan.

## **Section 11. Health Information Technology**

**Please respond to the following questions in terms of any health information technology (health IT) that has been or could be incorporated into the measure calculation.**

### **11.A. Health IT Enhancement**

**Please describe how health IT may enhance the use of this measure.**

The use of health IT in the collection and calculation of this measure allows for the clinical data to be used to assess measure results. The use of clinical data is more desirable compared to administrative data due to the increased granularity of information that can be collected.

### **11.B. Health IT Testing**

**Has the measure been tested as part of an electronic health record (EHR) or other health IT system?**

Yes.

**If so, in what health IT system was it tested and what were the results of testing?**

A second phase of reliability testing of the measure also occurred at the same sites where feasibility testing was inducted. This approach utilized parallel forms reliability testing, where measure data elements and performance from an automated report from the EHR were compared to those data from a manual review of the EHR—that is, comparison to the gold standard.

### **11.C. Health IT Workflow**

**Please describe how the information needed to calculate the measure may be captured as part of routine clinical or administrative workflow.**

See Section 8.A.

### **11.D. Health IT Standards**

**Are the data elements in this measure supported explicitly by the Office of the National Coordinator for Health IT Standards and Certification criteria (see [healthit.hhs.gov/portal/server.pt/community/healthit\\_hhs\\_gov\\_\\_standards\\_ifr/1195](http://healthit.hhs.gov/portal/server.pt/community/healthit_hhs_gov__standards_ifr/1195))?**

Yes.

**If yes, please describe.**

We used the following standards in the development of our EHR specifications: The Quality Data Model (QDM), developed by the National Quality Forum, the vocabulary recommendations named by the Health IT Standards Committee (of the Office of the National Coordinator for Health IT), (e.g., SNOMED, RxNorm, LOINC), and also referenced in the CMS Blueprint (2015). The vocabulary standards used in the specifications are consistent with those recommendations proposed for Stage II of the CMS EHR incentive program (Meaningful Use). Another available standard is the HL7 Health Quality Measure Format (HQMF), an XML-based structured document to express a quality measure specification. The HQMF is used for specifications included in the Meaningful Use program and also references the QDM. The specifications provided for this measure have not been incorporated into the HQMF eMeasure format; however, the information included in the specifications serves as a foundation for the HQMF—that is, the PCPI electronic specification outlines the requirements to develop the HQMF.

### **11.E. Health IT Calculation**

**Please assess the likelihood that missing or ambiguous information will lead to calculation errors.**

It is highly likely that missing data or ambiguous information stored in the EHR will lead to calculation errors. The specifications provided for this measure are designed to query the EHR in order to obtain the data required for the measure calculation.

### **11.F. Health IT Other Functions**

**If the measure is implemented in an EHR or other health IT system, how might implementation of other health IT functions (e.g., computerized decision support systems in an EHR) enhance performance characteristics on the measure?**

A Clinical Decision Support (CDS) system is an example of how a health IT function may enhance performance on this measure. A CDS template or prompt will facilitate documentation of the data required for the measure at the time the physician is treating the patient. Capturing these data at the point of care ensures the required data are present at the time the retrospective query is performed for measure calculation.

## **Section 12. Limitations of the Measure**

**Describe any limitations of the measure related to the attributes included in this CPCF (i.e., availability of measure specifications, importance of the measure, evidence for the focus of the measure, scientific soundness of the measure, identification of disparities, feasibility, levels of aggregation, understandability, health information technology).**

The measure may have limited utilization due to the limited adoption of EHRs, particularly among practices treating the Medicaid population. However, the vocabulary standards used in the specifications are as proposed for Stage II of the CMS EHR incentive program (Meaningful Use), so its usability is expected to be enhanced by increased participation in this program. As adoption of EHRs increases, utilization of this measure should also increase.

## **Section 13. Summary Statement**

**Provide a summary rationale for why the measure should be selected for use, taking into account a balance among desirable attributes and limitations of the measure. Highlight specific advantages that this measure has over alternative measures on the same topic that were considered by the measure developer or specific advantages that this measure has over existing measures. If there is any information about this measure that is important for the review process but has not been addressed above, include it here.**

This measure should be selected because it expands the core set of measures beyond their current use. The measure will provide a mechanism to help assess the appropriateness of deliveries and prevent adverse neonatal outcomes. This measure is of particular importance for CHIPRA in that it is high impact with Medicaid patients and addresses concerns related to both mother and baby. Additionally, since this measure has full eSpecifications, it can be a candidate for future inclusion in the EHR Incentive Program for Meaningful Use. Our EHR specifications follow the standards in the Quality Data Model (QDM), developed by the National Quality Forum, the vocabulary recommendations named by the Health IT Standards Committee (of the Office of the National Coordinator for Health IT), (e.g., SNOMED, RxNorm, LOINC), and also referenced in the CMS Blueprint (2015). The vocabulary standards used in the specifications are a part of Stage II of the CMS EHR incentive program (Meaningful Use).

## **Section 14: Identifying Information for the Measure Submitter**

**First Name:** Ramesh

**Last Name:** Sachdeva, MD, PhD, MBA, FAAP  
**Title:** Professor of Pediatrics (Critical Care)  
**Organization:** Medical College of Wisconsin  
**Mailing Address:** 9000 West Wisconsin Avenue, MS-681  
**City:** Milwaukee  
**State:** WI  
**Postal Code:** 53226  
**Telephone:** 414-266-3022  
**Email:** rsachdeva@chw.org

**The CHIPRA Pediatric Quality Measures Program (PQMP) Candidate Measure Submission Form (CPCF) was approved by the Office of Management and Budget (OMB) in accordance with the Paperwork Reduction Act.**

**The OMB Control Number is 0935-0205 and the Expiration Date is December 31, 2015.**

## **Public Disclosure Requirements**

**Each submission must include a written statement agreeing that, should U.S. Department of Health and Human Services accept the measure for the 2014 and/or 2015 Improved Core Measure Sets, full measure specifications for the accepted measure will be subject to public disclosure (e.g., on the Agency for Healthcare Research and Quality [AHRQ] and/or Centers for Medicare & Medicaid Services [CMS] websites), except that potential measure users will not be permitted to use the measure for commercial use. In addition, AHRQ expects that measures and full measure specifications will be made reasonably available to all interested parties. "Full measure specifications" is defined as all information that any potential measure implementer will need to use and analyze the measure, including use and analysis within an electronic health record or other health information technology. As used herein, "commercial use" refers to any sale, license or distribution of a measure for commercial gain, or incorporation of a measure into any product or service that is sold, licensed or distributed for commercial gain, even if there is no actual charge for inclusion of the measure. This statement must be signed by an individual authorized to act for any holder of copyright on each submitted measure or instrument. The authority of the signatory to provide such authorization should be described in the letter.**

AHRQ Publication No. 14(17)-P009-15-EF  
February 2017