NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 0469

Corresponding Measures: 0469e

De.2. Measure Title: PC-01 Elective Delivery

Co.1.1. Measure Steward: The Joint Commission

De.3. Brief Description of Measure: This measure assesses patients with elective vaginal deliveries or elective cesarean births at >= 37 and < 39 weeks of gestation completed. This measure is part of a set of four nationally implemented measures that address perinatal care (PC-01: Elective Delivery, ePC-01: Elective Delivery; PC-02: Cesarean Birth, ePC-02: Cesarean Birth will be added as an eCQM 1/1/2020; PC-05: Exclusive Breast Milk Feeding, ePC-05: Exclusive Breast Milk Feeding; PC-06 Unexpected Complications in Term Newborns was added 1/1/2019).

PC-01: Elective Delivery is one of three measures in this set that have been re-engineered as eCQMs (ePC-01 Elective Delivery, ePC-02 Cesarean Birth and ePC-05 Exclusive Breast Milk Feeding).

A reduction in the number of non-medically indicated elective deliveries at >=37 to <39 weeks gestation results in a substantial decrease in neonatal morbidity and mortality, as well as a significant savings in health care costs. In addition, the rate of cesarean sections should decrease with fewer elective inductions resulting in decreased length of stay and health care costs (AAFP, 2000).

The measure will assist health care organizations (HCOs) to track non-medically indicated early term elective deliveries and reduce the occurrence.

American Academy of Family Physicians. (2000). Tips from Other Journals: Elective induction doubles cesarean delivery rate, 61, 4.Retrieved December 29, 2008 at: http://www.aafp.org/afp/20000215/tips/39.html.

1b.1. Developer Rationale: For almost 3 decades, the American College of Obstetricians and Gynecologists (ACOG) and the American Academy of Pediatrics (AAP) established and followed the standard requiring 39 completed weeks gestation prior to ELECTIVE delivery, either vaginal or operative (ACOG, 1996). A survey conducted in 2007 of almost 20,000 births in HCA hospitals throughout the U.S. carried out in conjunction with the March of Dimes at the request of ACOG revealed that almost one-third of all babies delivered in the United States are electively delivered with 5% of all deliveries in the U.S. delivered in a manner violating ACOG/AAP guidelines. Most of these are for convenience and result in significant short-term neonatal morbidity, such as neonatal intensive care unit admission rates of 13- 21% (Clark et al., 2009).

According to Glantz (2005), compared to spontaneous labor, elective inductions result in more cesarean births and longer maternal length of stay. The American Academy of Family Physicians (2000) also notes that elective

induction doubles the cesarean delivery rate. Repeat elective cesarean births before 39 weeks gestation also result in higher rates of adverse respiratory outcomes, mechanical ventilation, sepsis and hypoglycemia for the newborns (Tita et al., 2009).

A reduction in the number of non-medically indicated elective deliveries at >=37 to <39 weeks gestation will result in a substantial decrease in neonatal morbidity and mortality, as well as a significant savings in health care costs. In addition, the rate of cesarean sections should decrease with fewer elective inductions resulting in decreased length of stay and health care costs.

The measure will assist health care organizations (HCOs) to track non-medically indicated early term elective deliveries and reduce the occurrence.

Sources

- American College of Obstetricians and Gynecologists. (November 1996). ACOG Educational Bulletin.
- American Academy of Family Physicians. (2000). Tips from Other Journals: Elective induction doubles cesarean delivery rate, 61, 4.Retrieved December 29, 2008 at: http://www.aafp.org/afp/20000215/tips/39.htm
- Clark, S., Miller, D., Belfort, M., Dildy, G., Frye, D., & Meyers, J. (2009). Neonatal and maternal outcomes associated with elective delivery. [Electronic Version]. Am J Obstet Gynecol. 200:156.e1-156.e4.
- Glantz, J. (Apr.2005). Elective induction vs. spontaneous labor associations and outcomes. [Electronic Version]. J Reprod Med. 50(4):235-40.
- Tita, A., Landon, M., Spong, C., Lai, Y., Leveno, K., Varner, M, et al. (2009). Timing of elective repeat cesarean delivery at term and neonatal outcomes. [Electronic Version]. NEJM. 360:2, 111-120.

S.4. Numerator Statement: Patients with elective deliveries with ICD-10-PCS Principal Procedure Code or ICD-10-PCS Other Procedure Codes for one or more of the following:

- Medical induction of labor as defined in Appendix A, Table 11.05 while not in Labor prior to the procedure
- Cesarean birth as defined in Appendix A, Table 11.06 and all of the following:
- not in Labor
- no history of a Prior Uterine Surgery

S.6. Denominator Statement: Patients delivering newborns with >= 37 and < 39 weeks of gestation completed with ICD-10-PCS Principal or Other Procedure Codes for delivery as defined in Appendix A, Table 11.01.1 and with ICD-10-CM Principal Diagnosis Code or ICD-10-CM Other Diagnosis Codes for planned cesarean birth in labor as defined in Appendix A, Table 11.06.1.

S.8. Denominator Exclusions: ICD-10-CM Principal Diagnosis Code or ICD-10-CM Other Diagnosis Codes for conditions possibly justifying elective delivery prior to 39 weeks gestation as defined in Appendix A, Table 11.07 include the following:

- History of prior stillbirth
- Less than 8 years of age
- Greater than or equal to 65 years of age
- Length of Stay >120 days
- Gestational Age < 37 or >= 39 weeks or UTD

De.1. Measure Type: Process

S.17. Data Source: Electronic Health Records, Other, Paper Medical Records

S.20. Level of Analysis: Facility, Other

IF Endorsement Maintenance – Original Endorsement Date: Oct 24, 2008 **Most Recent Endorsement Date:** Oct 25, 2016

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not Applicable

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since **the prior evaluation.**

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☐ **Yes** ☒ **No**

Summary of prior review in 2016

- American College of Obstetricians and Gynecologists (ACOG) Practice Bulletin No. 107—systematic review of literature (evidence not graded, recommendation Level II).
- Non-medically indicated elective deliveries prior to 39 weeks gestation can result in negative neonatal outcomes and a reduction can lead to improved maternal and fetal outcomes, decreased length of stay and fetal morbidity and mortality. Developer submitted research results that shown that a hard-stop policy preventing nonmedically indicated deliveries at a hospital contributed to a reduction from 8.2% to 1.7% (P=0.007). This reduction is reported to be greater than in other soft policies, but specific data are not reported for comparison.
- The developer reported Quantity = High; Quality = Moderate; Consistency = High

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

- The developer attests that there is no evidence showing adverse outcomes associated with a decrease in non-medically indicated early term elective deliveries. However, the developer did offer new evidence with this submission, a 2019 ACOG Committee Opinion—Avoidance of nonmedically indicated early-term deliveries and associated neonatal morbidities. The additional evidence supports the use of the measure.
  - Five recommendations by ACOG and Society of Fetal Medicine made, including: Nonmedically indicated delivery, including cesarean delivery, inductions of labor, and cervical ripening should not occur before 39 0/7 weeks of gestation (evidence not graded, no recommendation level/grade).
  - Systematic review of literature (24 studies) found late-preterm and early-term children have lower performance scores across (a range of cognitive and educational measures compared with their full-term peers. Non-respiratory morbidities also are increased in early-term deliveries; documentation of fetal pulmonary maturity does not justify an early nonmedically indicated deliver

Question for the Committee:

- *The developer provided updated evidence for this measure that is directionally consistent and strengthens the evidence for the previous NQF review. Does the Committee agree the evidence basis for the measure has not changed and there is no need for repeat discussion and vote on Evidence?*

Guidance from the Evidence Algorithm

Process measure based on systematic review (Box 3) → Evidence not graded (Box 7) → Developer assesses Quantity: high; Quality: moderate; Consistency: high (Box 9) → Moderate.

Preliminary rating for evidence: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer provided CY 2018 data (1,616 hospitals, 139,213 patients), the most recent full year of data. The results are as follows:
  - Mean: 1.7%, SD 2.8%
  - IQR: 2.5%
  - Deciles (0,10,20,30,40,50,60,70,80,90,100): 0%, 0%, 0%, 0%, 0%, 1.4%, 2.2%, 3.0%, 4.8%, 29.0%
- The developer stated that although performance has been improving, 8.7% of hospitals have not yet met the desired mark of 5%, which is based on recommendations from a technical advisory panel convened by the developer for this measure.
- The developer noted that the national aggregate rate has declined over time:
  - 2010: 18.8%
  - 2014: 3.3%
  - 2015: 2.3%

- 2016: 1.9%
- 2017: 1.7%
- 2018: 1.6%

Disparities
Based on 2018 discharges:

- Performance by age category

Age	Rate (%)
<20	1.57
20-24	1.25
25-29	1.50
30-34	1.77
35-39	2.28
40	3.90

- Performance by Hispanic ethnicity

Hispanic Ethnicity	Rate (%)
No	1.64
Yes	1.79

- Performance by race

Race	Rate (%)
White	1.77
African American	1.49
American Indian	1.06
Asian	1.00
Pacific Islander	1.34
Unable to Determine	1.77

Questions for the Committee:

- *Is there a gap in care and/or disparities that warrant a national performance measure? Of note, the performance rates on the eCQM differ significantly than this measure (mean 17.6%).*
- *Is performance on this measure "topping out"?*

Preliminary rating for opportunity for improvement: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Importance to Measure and Report
Comments:

** Not aware of new evidence/studies

** There is moderate evidence to support this measure and balancing measures do not show negative outcomes. I'm not aware of new evidence.

** Appears to be direct evidence from systematic reviews and society guidelines. I am not aware of un-cited evidence.

** New study ARRIVE trial shows that Cesarean rates are NOT higher in elective inductions. However, the timing of the elective induction is still important for neo natal outcomes. Background and justification for the measure should be updated.

** The evidence has not changed substantially and I agree that there is no need for repeat discussion and vote on Evidence.

1b. Performance Gap

Comments:

** yes, performance data was provided and has been declining over time.

** There is current performance provided and performance continues to show a gap in care in some facilities.

** Performance gap seems to have narrowed significantly since the initial measure. Not sure if this is still needed?

** There is variability in compliance with the measure and variation within patient characteristics.

** There are still opportunities for improvement on this measure.

1b. Disparities

Comments:

** yes, data on subgroups provided and very minor variation among groupings but advocate for continuing to measure.

** There are disparities in performance, mostly noted around age.

** Disparities by age, other sub-group differences seem less pronounced

** Disparities are noted. However, the direction of the disparities is unusual and warrants further investigation.

** I would like to see an analysis of significance of variation based on age, ethnicity, and race.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- Are you satisfied with the reliability testing for the measure?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- Are you satisfied with the validity analyses for the measure?

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 0496

Measure Title: PC-01 Elective Delivery

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☐ **Electronic Health Data** ☒ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☒ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

Submission document: "MIF_0469" document, items S.1-S.22

***NOTE**: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**
    - No concerns

RELIABILITY: TESTING

Submission document: "MIF_0469" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☒ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

    ☐ **Yes** ☐ **No**
6. **Assess the method(s) used for reliability testing**

    Submission document: Testing attachment, section 2a2.2

    - Developer evaluated score-level reliability results using a beta-binomial model (signal to noise) on a data set of 1,616 hospitals with a median number of deliveries of 1,227; the median number of denominator cases was 59. Per Adams (2009), reliability scores range from 0.0 to 1.0. A score of zero implies that all variation is attributed to measurement error (i.e., noise), whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance (across hospitals).
    - Because of the switch from ICD-9 to ICD-10 codes, the developer also examined reliability by comparing results from the 2017 data submission that utilized ICD-9 codes to the results from the 2018 data submission that utilized ICD-10 codes. The developer used a pairwise comparison of previous ICD-9 codes and current ICD-10 codes to illustrate the continued reliability of the measure within this new coding system.
    - These tests are appropriate for purposes of reliability testing and conform to NQF guidance.
7. **Assess the results of reliability testing**

    Submission document: Testing attachment, section 2a2.3

    For the 2020 submission, the developer reports the following signal to noise reliability statistics:
    - Average: 0.78
    - Median: 1.0
    - 10^{th}-90^{th} percentile across hospitals: 0.32-1.0
    - The developer noted that, in general, a score of 0.7 or higher suggests the measure has adequate reliability and that these results suggest the measure has high reliability for most of the hospitals.

    For the 2020 submission, the developer reported the following with respect to ICD-9 vs. ICD-10 codes:

Comparison between the ICD-9 and ICD-10 number of numerator cases, denominator cases, and observed rates

		N	Mean	Std. Dev.	Min	Q1	Median	Q3	Max	Pairwise Difference	P-Value
Numerator	ICD-9	2071	0.3071	0.8446	0	0	0	0	13	.0132	0.5432
	ICD-10	1998	0.3233	0.957	0	0	0	0	23		
Denominator	ICD-9	2071	18.799	20.945	1	8	12	21	291	-0.6876	0.0032
	ICD-10	1998	18.408	19.991	1	8	12	21	269		

Rate	ICD-9	2071	0.01782	0.05411	0	0	0	0	1	0.0008	0.5675
	ICD-10	1998	0.0179	0.04947	0	0	0	0	0.5		

- The developer reported that pairwise comparisons were not statistically significant difference for the numerator. Statistically significant difference was found for the denominator cases, but the developer reported that a mean difference of 0.68 is not clinically meaningful. The developer concluded that the results demonstrate that the measure continues to distinguish between hospitals' performance levels as a result of quality of care, rather than chance, and that the change in coding did not affect reliability.

8. **Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities?** NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)

9. **Was the method described and appropriate for assessing the reliability of ALL critical data elements?**

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

   ☒ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

    ☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    - The developer provided a score-level reliability statistic indicating moderate reliability.
    - The developer was thorough and demonstrated no difference in reliability from the switch from ICD-9 to ICD-10 codes.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2

    - In the current submission, three exclusions (at least one condition possibly justifying elective delivery; gestational age <37 or >=39; history of stillbirth) were empirically tested for impact. The developer provided a rationale for each exclusion and the percentage lost to the exclusions, which are not mutually exclusive.
    - No concerns.

- In its previous submission, the developer noted exclusions that were not derived directly from the evidence and the justification for them.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4

    - To demonstrate meaningful differences in performance, the developer calculated a funnel plot for the annual hospital rates of the measure, where the observed measure is plotted against a measure of its precision, so that the control limits form a 'funnel' around the target outcome. It superimposes the 95 per cent (≈two standard deviation) and 99.8 per cent (≈three standard deviation) prediction limits over this plot around the overall measure rate; those rates lying outside the confidence limits are identified as outliers. (Spiegelhalter, DJ. Funnel plots for comparing institutional performance. *Statistics in Medicine*. 2005; 24:1185–1202.)
    - The developer reported that 52 hospitals were identified as outliers with rates beyond the two standard deviation upper limit, and 9 hospitals were identified as outliers with rates beyond the three standard deviation limits—the upper confidence rate limit for a hospital with the median denominator size of 59 is 15.1% for a 95% confidence interval (2 standard deviations) and 27.8% for a 99.8% confidence interval (3 standard deviations).
    - The developer stated that the results indicate significant differences in performance among hospitals and an appreciable number of hospitals that are not within the expected level of variability and differ significantly from the mean overall rate.
    - No concerns.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**

    Submission document: Testing attachment, section 2b5

    - Not applicable

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6

    - The developer reports that hospitals submitting data with any missing data are not accepted.

16. **Risk Adjustment**

    16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

    16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

    ☐ Yes ☐ No ☒ Not applicable

    16c. **Social risk adjustment:**

    16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

    16c.2 Conceptual rationale for social risk factors included? ☒ Yes ☐ No

    16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

    16d.**Risk adjustment summary:**

    16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

    16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

    16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

    16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration)

☐ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

- The measure is not risk adjusted; the developer did not provide a rationale for this approach.
- The measure is not risk adjusted for social risk factors; the developer provided a rationale for this approach.
- The Committee may wish to discuss these approaches with the developer.

VALIDITY: TESTING

17. **Validity testing level: ☒ Measure score ☐ Data element ☐ Both**
18. **Method of establishing validity of the measure score:**

☐ Face validity

☒ Empirical validity testing of the measure score

☐ N/A (score-level testing not conducted)

19. **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

- The developer used construct validity to calculate correlations of this measure with other measures of perinatal quality.
  - For 0469, a lower rate is higher quality. The developer hypothesized that it would correlate negatively to other perinatal care measures where a high rate is desirable (e.g., [e]PC-05 Exclusive Breast Milk Feeding, e0469) and correlate positively to perinatal care measures where a low rate is desirable (e.g., PC-02 Cesarean Birth).
- The developer stated a correlation of 0.1 to 0.3 was considered weak, 0.3 to 0.5 was considered moderate, and >0.5 was considered strong.

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

- The developer reported the following:

Table of Correlations

Measure	PC-01	PC-02	PC-05	ePC-01	ePC-05
PC-01-Elective Delivery	1				
PC-02-Cesarean Birth	0.133192	1			
PC-05-Exclusive Breast Milk Feeding	-0.02553	-0.28103	1		
ePC-01-Elective Delivery	0.008936	0.108322	0.022812		
ePC-05-Exclusive Breast Milk Feeding	0.040365	-0.17522	0.748033	-0.45737	1

- The developer noted that except for the correlation between PC-01 and ePC-05, the correlations were in the expected direction.
- It further stated that the correlation of PC-01 and ePC-05, although expected to be in the negative direction, was not significantly greater than zero, nor was the correlation between PC-01 and PC-05. It pointed out that these measures (PC-01 vs PC-05 and ePC-05) evaluate two different

populations, mothers and babies, and therefore two different aspects of perinatal care, which are apparently not correlated.

- Further, the developer stated that the perinatal care measures used in this analysis are assessing different components of perinatal care and would not be expected to be more than weakly correlated, since perinatal care quality is a multidimensional quantity.

21. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

    Submission document: Testing attachment, section 2b1.

    ☒ **Yes**

    ☐ **No**

    ☐ **Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**

    NOTE that data element validation from the literature is acceptable.

    Submission document: *Testing attachment, section 2b1.*

    ☐ **Yes**

    ☐ **No**

    ☒ **Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

    ☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developer's approach to demonstrating validity.**

- The developer provided data related to exclusions and data on meaningful differences. Missing data are not an issue.
- The developer's construct validity was appropriate, but a stronger approach would have been to correlate the measure with a more global measure (e.g., CMS's Five-Star system or HCAHPS).

ADDITIONAL RECOMMENDATIONS

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**

- No concerns.

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability – Specifications

Comments:

** No concerns

** No concerns related to reliability.

** no concerns

** none

** Because some of the indicators for delivery are subjective and not well-defined, there is a risk that some deliveries will be labeled "medically indicated" when other clinicians would not code them with a medical indication.

2a2. Reliability – Testing

Comments:

** No concerns

** No

** no concerns

** none

** See 2a1

2b1. Validity –Testing

Comments:

** No concerns

** No concerns.

** no concerns.

** none

** I would like to know if the rate of "medically indicated" deliveries has changed over time. Has there been a shift from "not medically indicated" to "medically indicated" that belies some of the stated improvements? I am concerned that some of the Conditions Possibly Justifying Elective Delivery (i.e., logistic or psychosocial indicators) are subjective and vague and may be selected by the clinician to justify an otherwise non-medically indicated early delivery.

2b2-3. Exclusions/Risk Adjustment

Comments:

** Agree with rationale provided for the risk-adjustment variables. No concerns

** Exclusions are consistent with the evidence. From a social risk perspective, especially regarding distance from the facility and time of year, variations in performance can be explained considering these factors. This measure is not risk-adjusted.

** no concerns

** Impossible to include all exclusions but the major ones are there

** The developer states: "There is no compelling evidence available supporting association between social risk factors and this measure." With our current knowledge of institutional racism and implicit bias, I would like to see a deeper analysis of the data around social risk factors.

2b4-7. Threats to Validity/Meaningful Differences/Comparability of Performance Scores/Missing Data

Comments:

** no concerns, agree to not accept when data is missing

** There are meaningful differences identified between delivering facilities. Missing data is not accepted in this measure.

** no concerns re: threats to validity

** no

** "The developer reports that hospitals submitting data with any missing data are not accepted." Please describe any follow-up to reports that are not accepted due to missing data.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- The developer noted that the data are generated by and used by healthcare personnel during the provision of care, coded by someone other than person obtaining original information, and/or abstracted from a record by someone other than person obtaining original information.
- The developer noted that some data elements are in defined fields in electronic sources.
- The developer indicated that although most data elements are regularly captured in electronic medical records, this measure is still warranted for those that do not use electronic medical records or do not have data fields that are structured for this measure.
- An eCQM version of this measure is available.

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *What are the burdens of data collection that may not be mentioned in the submission? Are there any issues with the data collection strategy?*

Preliminary rating for feasibility: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

Comments:

** data elements should be routinely captured in EHR and on claims submitted. No concerns.

** Data elements collected are used during care delivery. No concerns regarding data collection strategy at this point in the duration of the measure.

** may depend on the type of EHR, but seems feasible to collect this data

** none

** None

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No** ☐ **UNCLEAR**

The developer reported the measure is publicly reported, as follows:

- Quality Check® (a public reporting site of The Joint Commission)
- Hospital Compare (CMS)

The developer reported the measure is part of the following accountability programs:

- Hospital Inpatient Quality Reporting (payment program; CMS)
- Hospital Value-based Purchasing Program (CMS)
- Hospital Accreditation Program (The Joint Commission)
- Perinatal Care Certification (The Joint Commission)

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- The developer stated that those being measured are provided with performance results in a quarterly or annual report. The developer also noted that its Continuous Customer Engagement (CCE) program with embedded quality improvement tools allows measure users to engage in quality improvement initiatives and obtain guidance on improvement.
- The developer mentioned that measure users are able to submit questions or comments about measure implementation and engage with the developer through a dashboard.
- The developer also noted that all feedback is tracked and considered. Additionally, measure specifications are reviewed twice a year and updates are made as needed based on feedback from the measure users, input from the TAP, changes in the guidelines, or changes in clinical practice.
- Feedback can be submitted to a web tool for response typically within 8 business hours. Frequent topics are developed into educational webinars for measure users to join. Those with particularly high measure gaps are encouraged to join.
- The developer stated that minor modifications have been made based upon this feedback, but does not indicate what those modifications have been. It also noted that hospitals have provided feedback that required data elements are generally available in the medical record and that the specifications are robust and easy to understand.

- The developer indicated that statistical questions on this measure were related to how it is to be publicly reported in 2020. Additionally, the developer stated that "queries submitted via the automated feedback system have decreased significantly for this measure in the past three years."

Additional Feedback

- This measure has not been reviewed by Measure Applications Partnership

***Questions for the Committee**:*

- *Can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Has the measure been vetted in real-world settings by those being measured or others to your satisfaction?*

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

The developer provided the following trend data (national aggregate rate):

- 2015: 2.3%
- 2016: 1.9%
- 2017: 1.7%
- 2018: 1.6%

The developer also noted that 8.7% of facilities do not meet the goal of 5%, which was identified by the developer's Technical Expert Panel.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

- The developer identified three unexpected findings and the actions it took to mitigate them.
- First, cases with prior uterine surgery were inappropriately failing the measure. The developer adapted the algorithm by adding a new data element to check for prior uterine surgery before checking for cesarean birth to retain these patients.
- Second, patients who did not receive prenatal care were inappropriately included in the measure denominator, as the gestational age data element was abstracted as unable to be determined (UTD). The developer adapted the algorithm to remove these UTD cases from the measure population in order to avoid penalizing hospitals for this.
- Finally, some hospitals reported higher rates due to small denominator populations as a result of sampling. The developer added additional data sources to the list of acceptable sources for identifying all cases to increase the denominator.

Potential harms

- No potential harms were identified.

Additional Feedback:

- No additional feedback was submitted via QPS, MAP, or other sources.

***Questions for the Committee**:*

- *Can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Has the developer adequately addressed the known, unexpected findings?*
- *Are you aware of other unintended benefits or consequences of the measure that have not discussed?*

Preliminary rating for Usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1-2. Use - Accountability and Transparency/Feedback
Comments:

** Publicly reported via Joint Commission and CMS. Quarterly and annual feedback. No concerns

** Measure is being publicly reported through CMS and TJC. This measure is part of HIQR, VBP, Hospital Accreditation Program, and the Perinatal Care Certification. Those being measured are provided results and are able to provide feedback on the measure.

** while not transparent in the process and changes, appears to have incorporated feedback

** should not be combined with other measures but should stand on its own

** Data have been made available to those being measured and to the public.

4b1. Usability – Improvement/ Benefits vs. harms/ Transparency
Comments:

** Using additional sources of data will appropriately identify the appropriate denominators.

** Visibility and transparency into performance can help to drive improvement. No harm was identified.

** no unintentional harms identified

** hospitals with complex high risk patients that are not included in the exceptions may appear to have a higher rate of elective induction <39 weeks when in fact they are providing good care. It is unlikely that this would cause a hospital to be an outlier as these are also hospitals with large denominators

** The measure provides potential for benefit and little risk of harm.

Criterion 5: Related and Competing Measures

Related or competing measures

- The developer indicated there are no related or competing measure, but in fact this measure is related to Measure 0469e, the eCQM version.

Harmonization

- The developer stated that this measure is harmonized to the extent feasible given the differences in data source, and justifies this due to the electronic nature of the related measure.

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

Comments:

** Is related to the eCQM measure and harmonized to the extent feasible

** No related or competing measure identified.

** e measure, appears to be harmonized

** none

** This measure is intended to be harmonized with #0469e.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of June 15, 2020

- No comments received
- **Of the 0 NQF members who have submitted a support/non-support choice:**
  - 0 support the measure
  - 0 do not support the measure

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

2020_nqf_evidence_attachment_PC01_0469_.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0469
Measure Title: PC-01 Elective Delivery
IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title
Date of Submission: April 8, 2020

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☐ Outcome: Click here to name the health outcome
- ☐ Patient-reported outcome (PRO): Click here to name the PRO
  PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health- related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

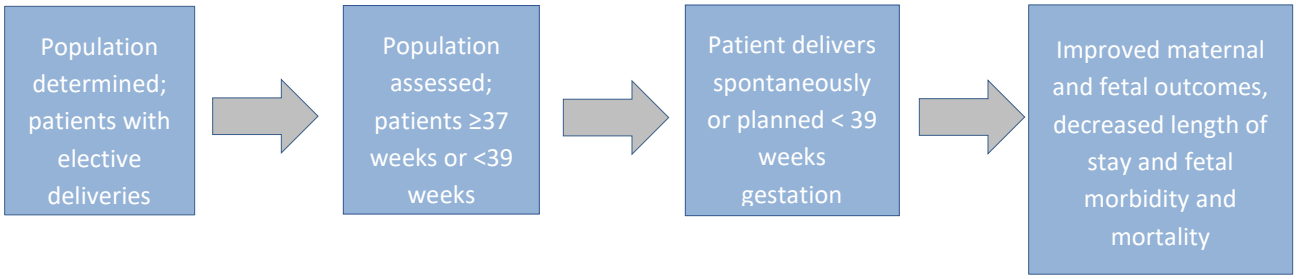
☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☐ Process: Patients with elective vaginal deliveries or elective cesarean births at >= 37 and < 39 weeks of gestation completed
- ☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.



The intent of the measure is to decrease the number of elective deliveries before 39 weeks>> population determined; patients with elective deliveries >> population assessed; patients ≥37 weeks and < 39 weeks>> patient delivers spontaneously or planned delivery <39 weeks gestation >> improved maternal and fetal outcomes; decreased length of stay and fetal morbidity and mortality.

1a.3 Value and Meaningfulness: If this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)
Not applicable
****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.
Not applicable

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.
What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (publication in the table)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other See 1a.4

2020 Submission

Source: • Title • Author • Date • Citation, including page number • URL	**Title:** Avoidance of nonmedically indicated early-term deliveries and associated neonatal morbidities. **Author:** Borders, E.B., Birsner, M.L., Gyanmfi-Bannerbaum, C. **Date:** 2019 **Citation:** Borders, E.B., Birsner, M.L., Gyanmfi-Bannerbaum, C. (2019). Avoidance of nonmedically indicated early-term deliveries and associated neonatal morbidities. *American College of Obstetricians and Gynecologists Committee Opinion, 133:2,* e156-163. **URL:** https://www.acog.org/-/media/Committee-Opinions/Committee-on-Obstetric-Practice/co765.pdf?dmc=1&ts=20191119T1727071568
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	**From guideline abstract:** A recent systematic review found that late-preterm and early-term children have lower performance scores across a range of cognitive and educational measures compared with their full-term peers. Because non-respiratory morbidities also are increased in early-term deliveries, documentation of fetal pulmonary maturity does not justify an early nonmedically indicated delivery. Amniocentesis for the determination of fetal lung maturity should not be used to guide the timing of delivery, even in sub optimally dated pregnancies. The American College of Obstetricians and Gynecologists (ACOG) and the Society for Maternal–Fetal Medicine make the following recommendations: 1. Nonmedically indicated delivery, including cesarean delivery, inductions of labor, and cervical ripening should not occur before 39 0/7 weeks of gestation. 2. Implementation of a policy to decrease the rate of nonmedically indicated deliveries before 39 0/7 weeks of gestation has been found to decrease the number of these deliveries and, as a result, improve overall neonatal outcomes. 3. Avoidance of a nonmedically indicated delivery before 39 0/7 weeks of gestation is distinct from, and should not result in, an increase in expectant management of patients with medical indications for delivery before 39 0/7 weeks of gestation.

	4. Indications for delivery before 39 0/7 weeks of gestation should be documented clearly and discussed with the patient. 5. Because nonrespiratory morbidities also are increased in early-term deliveries, documentation of fetal pulmonary maturity does not justify an early nonmedically indicated delivery. Amniocentesis for the determination of fetal lung maturity should not be used to guide the timing of delivery, even in sub optimally dated pregnancies.
Grade assigned to the evidence associated with the recommendation with the definition of the grade	No grades of evidence were assigned to the recommendations.
Provide all other grades and definitions from the evidence grading system	Not applicable
Grade assigned to the recommendation with definition of the grade	Not applicable
Provide all other grades and definitions from the recommendation grading system	Not applicable
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	**Quantity:** This committee opinion by the American College of Obstetricians and Gynecologists has identified 24 studies that helped to support the evidence in reducing early elective deliveries as presented by the committee. Of those studies, there were retrospective cohort studies (11), cohort studies (8), observational studies (2), sequential ecological studies (2), population-based prospective cohort analysis (1). **Quality:** Information on the overall quality of evidence across the studies is not provided; although, this committee opinion discusses the evidence supporting the reduction of elective deliveries >= 37 and < 39 weeks of gestation. As states have effectively reduced early elective delivery, multiple studies using national population level data have shown that even as the gestational age at term has increased in response to efforts to reduce early elective delivery, these efforts have not adversely affected stillbirth rates nationally or even in states with the greatest reductions in early elective delivery. There is no documented evidence regarding controversy related to the reduction of non-medically indicated early term elective deliveries. A review of studies also supports the use of quality improvement interventions to further reduce the number of such deliveries. **Quantity:** High **Quality:** Moderate **Consistency:** Moderate

Estimates of benefit and consistency across studies	Estimates of benefit and consistency across the studies are not provided; although, this committee opinion discusses the evidence supporting the reduction of elective deliveries >= 37 and < 39 weeks of gestation. **Benefits:** The central topic for the measure is to reduce elective deliveries >= 37 and < 39 weeks of gestation completed. The target population for the performance measure is consistent with the body of evidence supporting the reduction of elective deliveries. Implementation of a policy to decrease the rate of nonmedically indicated deliveries before 39 0/7 weeks of gestation has been found to decrease the number of these deliveries and, as a result, improve overall neonatal outcomes. A recent study examined the implementation of three approaches to this issue: 1) a hard-stop policy, which prohibited nonmedically indicated deliveries at the hospital level; 2) a soft-stop policy, in which health care providers agreed not to perform nonmedically indicated deliveries before 39 weeks of gestation; and 3) an education program that informed health care providers about the risks associated with delivery before 39 weeks of gestation. Overall, these approaches contributed to a greater than 50% reduction in the rate of nonmedically indicated early-term deliveries, regardless of the policy used (28). However, the reduction was the greatest in the hard-stop policy group, with a reduction from 8.2% to 1.7% (P=.007). The reduction was slightly less in the soft-stop policy group, with a reduction from 8.4% to 3.3% (P=.025), and the least in the educational approach group, with a reduction from 10.9% to 6.0% (P=.135), which was not statistically significant. **Consistency:** These studies demonstrate that a reduction in nonmedically indicated early-term and late-preterm deliveries can be achieved. Studies clearly have shown short-term and long-term outcomes are improved for infants born at full term (39 0/7–40 6/7 weeks of gestation) versus late preterm (34 0/7–36 6/7 weeks of gestation) or early term (37 0/7–38 6/7 weeks of gestation). The studies referenced in the committee opinion show a strong support of evidence that early elective deliveries pose a greater risk of unfavorable neonatal outcomes.
What harms were identified?	There have been no harms identified as a result of the implementation of the studies. However, the following harms listed below may occur as a result of early elective delivery.

	Neonatal Morbidities Associated with Early-Term Delivery: • Respiratory distress syndrome • Transient tachypnea of the newborn • Ventilator use • Pneumonia • Respiratory failure • Neonatal intensive care unit admission • Hypoglycemia • 5-minute Apgar score less than 7 • Neonatal mortality In a large cohort of planned term deliveries (defined as deliveries not initiated by labor or ruptured membranes) during a 3-month period in 27 hospitals across the United States, neonatal intensive care unit (NICU) admission rates were higher among neonates delivered in the early-term period. A comparison of NICU admission rates for neonates delivered at 37 weeks of gestation or 38 weeks of gestation with those for neonates delivered at 39 weeks of gestation revealed that 31% of 17,794 deliveries had no medical indication. Admission to the NICU, which can be dependent on a variety of factors, was required for 17.8% of infants delivered without medical indication at 37 weeks of gestation and for 8% delivered without medical indication at 38 weeks of gestation, compared with 4.6% of infants delivered at 39 weeks of gestation or beyond (P<.001 for deliveries at 38 weeks of gestation and 39 weeks of gestation). Harms identified with early term delivery are listed above.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	As a result of the literature search there have been no new studies conducted since this publication that would change the conclusions from the referenced recommendations.
2016 submission	
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	American College of Obstetricians and Gynecologists Practice Bulletin No 107. (2009, August). Induction of labor. Obstetrics & Gynecology, 386-397. Retrieved from https://www.mnhospitals.org/Portals/0/Documents/patientsafety/Perinatal/acog--practice_bulletin_107_2009.pdf The American College of Obstetricians and Gynecologists the nation´s leading group of professionals providing health care for women. Practice Bulletins provide obstetricians and gynecologists with current information on established techniques and clinical management guidelines. The American College of Obstetricians and Gynecologists (the College) continuously surveys the

	field for advances to be incorporated in these series and monitors existing bulletins to ensure they are current. Individual bulletins are withdrawn from and added to the series on a continuing basis and reaffirmed periodically.
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	Clinical Considerations and Recommendations Page 389 What are the indications and contraindications to induction of labor? Indications for induction of labor are not absolute but should take into account maternal and fetal conditions, gestational age, cervical status, and other factors. Following are examples of maternal or fetal conditions that may be indications for induction of labor: • Abruptio placentae • Chorioamnionitis • Fetal demise • Gestational hypertension • Preeclampsia, eclampsia • Premature rupture of membranes • Post term pregnancy • Maternal medical conditions (e.g., diabetes mellitus, renal disease, chronic pulmonary disease, chronic hypertension, antiphospholipid syndrome) • Fetal compromise (e.g., severe fetal growth restriction, isoimmunization, oligohydramnios) Labor also may be induced for logistic reasons, for example, risk of rapid labor, distance from hospital, or psychosocial indications. In such circumstances, at least one of the gestational age criteria in the box should be met, or fetal lung maturity should be established. A mature fetal lung test result before 39 weeks of gestation, in the absence of appropriate clinical circumstances, is not an indication for delivery. The individual patient and clinical situation should be considered in determining when induction of labor is contraindicated. Generally, the contraindications to labor induction are the same as those for spontaneous labor and vaginal delivery. They include, but are not limited to, the following situations: • Vasa previa or complete placenta previa • Transverse fetal lie • Umbilical cord prolapse • Previous classical cesarean delivery • Active genital herpes infection • Previous myomectomy entering the endometrial cavity What criteria should be met before the cervix is ripened or labor is induced?

	Assessment of gestational age and consideration of any potential risks to the mother or fetus are of paramount importance for appropriate evaluation and counseling before initiating cervical ripening or labor induction. The patient should be counseled regarding the indications for induction, the agents and methods of labor stimulation, and the possible need for repeat induction or cesarean delivery. Although prospective studies are limited in evaluating the benefits of elective induction of labor, nulliparous women undergoing induction of labor with unfavorable cervices should be counseled about a twofold increased risk of cesarean delivery (Level II-2). In addition, labor progression differs significantly for women with an elective induction of labor compared with women who have spontaneous onset of labor (Level II-2). Allowing at least 12–18 hours of latent labor before diagnosing a failed induction may reduce the risk of cesarean delivery (Level II-2, 3). Additional requirements for cervical ripening and induction of labor include assessment of the cervix, pelvis, fetal size, and presentation. Monitoring FHR and uterine contractions is recommended as for any high-risk patient in active labor. Although trained nursing personnel can monitor labor induction, a physician capable of performing a cesarean delivery should be readily available. The American College of Obstetricians and Gynecologists the nation´s leading group of professionals providing health care for women. Practice Bulletins provide obstetricians and gynecologists with current information on established techniques and clinical management guidelines. The American College of Obstetricians and Gynecologists (the College) continuously surveys the field for advances to be incorporated in these series and monitors existing bulletins to ensure they are current. Individual bulletins are withdrawn from and added to the series on a continuing basis and reaffirmed periodically.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Although grading of the evidence was not determined during our systematic review, it was determined that the guideline developers accounted for a balanced representation of information, looked beyond one specialty group or discipline, and provided information that was accessible and met the requirements set out in this measure maintenance form.
Provide all other grades and definitions from the evidence grading system	Not applicable
Grade assigned to the **recommendation** with definition of the grade	Yes The American College of Obstetricians and Gynecologists Level II
Provide all other grades and definitions from the recommendation grading system	USPSTF

Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The central topic for the measure is the reduction of elective deliveries at >= 37 and < 39 weeks of gestation completed. The target population for the performance measure is consistent with the body of evidence supporting the reduction of elective deliveries. Quantity- No randomized-control trials (RCTs) were identified for early-term elective deliveries. RCTs were only identified for post-term elective deliveries versus expectant management. Given the current amount of population data available on the harms of early term and late pre-term delivery, it would be unethical to conduct such a study. Several studies were identified which were retrospective cohort or prospective observational in design examining thousands of births and the potential for adverse outcomes for both mother and newborn. In addition, several recent studies were identified addressing quality improvement interventions that were successful in reducing non-medically indicated early term elective deliveries. Quality- The quality of evidence supporting the reduction in the number of non-medically indicated elective deliveries is moderate. It is noteworthy to examine the fact that randomized control trials cannot be conducted, as one cannot randomly select women to agree to an elective delivery at < 39 weeks gestation. As previously noted, both ACOG and AAP have had guidelines in place for a number of years which do not support non-medically indicated elective deliveries at > 39 weeks gestation. Several studies consistently document increased morbidity associated with elective delivery before 39 weeks. The studies note that elective deliveries performed at < 39 weeks carry significant risk for the newborn (odds ratios 2.0-3-0 compared to newborns born between 39-41 weeks). In spite of the fact that all studies reviewed were either retrospective or prospective cohort studies, no study design flaws were noted. There is no documented evidence regarding controversy related to the reduction of non-medically indicated early term elective deliveries. A review of recent studies also supports the use of quality improvement interventions to further reduce the number of such deliveries. **Quantity: High** **Quality: Moderate** **Consistency: High**
Estimates of benefit and consistency across studies	The body of evidence consistently supports the benefit of reduction of non-medically indicated early term elective deliveries. All studies show an increase in the

	number of neonatal morbidities associated with early term deliveries, subsequent reduction of elective non-medically indicated deliveries reduces harm to the neonate. All studies demonstrated similar findings related to the direction of effect, though the magnitude varied from study to study, i.e., 8-17.8% increase in NICU admissions, rates of adverse respiratory outcomes, mechanical ventilation, newborn sepsis, hypoglycemia, admission to the NICU and hospitalization of 5 days or more increased by a factor of 1.8 to 4.2. and the incidence of transient tachypnea of the newborn, respiratory distress syndrome (RDS) and persistent pulmonary hypertension of the newborn were 3.1%, 0.25% and .17% respectively. As described before, elective deliveries performed at =>39 weeks gestation results in improved maternal and neonatal outcomes and will result in a substantial decrease in cesarean sections and neonatal morbidity, as well as substantial savings in health care costs. A recent study showed that by waiting until 39 weeks gestation, the NICU admissions fell from 12.8% to 5.9%, RDS fell from 3.7% to 0.9%, newborn sepsis fell from 7.0% to 2.5% and hospitalization > 5 days fell from 9.1% to 3.6%. This same study estimated that one-half million newborn intensive care unit days could be avoided in the U.S. population were a national rate of 1.7% to be achieved, with cost savings approaching $1 billion annually. There is no documented evidence regarding controversy related to the reduction of non-medically indicated early term elective deliveries. A review of recent studies also supports the use of quality improvement interventions to further reduce the number of such deliveries.
What harms were identified?	Not applicable
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Not applicable

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

Not applicable

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Not applicable

1a.4.2 What process was used to identify the evidence?

Not applicable

1a.4.3. Provide the citation(s) for the evidence.

Not applicable

From previous submission: Citations from Evidence Other Than Guidelines

- American Academy of Family Physicians. (2000). Tips from Other Journals: Elective induction doubles cesarean delivery rate, 61, 4. Retrieved September 16, 2011 at: http://www.aafp.org/afp/20000215/tips/39.html.
- American College of Obstetricians and Gynecologists. (November 1996). ACOG Educational Bulletin.
- American College of Obstetricians and Gynecologists (ACOG) Practice Bulletin No 107. (2009). Induction of labor. Obstetrics & Gynecology. 114(2). 386-97.
- Clark, S., Miller, D., Belfort, M., Dildy, G., Frye, D., & Meyers, J. (2009). Neonatal and maternal outcomes associated with elective delivery. [Electronic Version]. Am J Obstet Gynecol. 200:156. e1-156.e4.
- Clark, S., Frye, D., Meyers, J., Belfort, M., Dildy, G., Kofford, S et al. (2010). Reduction in elective delivery at <39 weeks of gestation: comparative effectiveness of 3 approaches to change and the impact on neonatal intensive care admission and stillbirth. Am J Obstet Gynecol. 203:449. e1-6.
- Davidoff, M., Dias, T., Damus, K., Russell, R., Bettegowda, V.R., Dolan, S., et al. (2006). Changes in the gestational age distribution among U.S. singleton births; impacts on rates of late preterm birth, 1992-2002. Semin Perinatol. Feb;30(1):8-15.
- Engle, W.A. & Kominiarek, M.A. (2008). Late preterm infants, early term infants, and timing of elective deliveries. Clin Perinatol. 35:325-41.
- Glantz, J. (Apr.2005). Elective induction vs. spontaneous labor associations and outcomes. [Electronic Version]. J Reprod Med. 50(4):235-40.
- Tita, A., Landon, M., Spong, C., Lai, Y., Leveno, K., Varner, M, et al. (2009). Timing of elective repeat cesarean delivery at term and neonatal outcomes. [Electronic Version]. NEJM. 360:2, 111-120.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

<u>If a COMPOSITE</u> (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

For almost 3 decades, the American College of Obstetricians and Gynecologists (ACOG) and the American Academy of Pediatrics (AAP) established and followed the standard requiring 39 completed weeks gestation prior to ELECTIVE delivery, either vaginal or operative (ACOG, 1996). A survey conducted in 2007 of almost 20,000 births in HCA hospitals throughout the U.S. carried out in conjunction with the March of Dimes at the request of ACOG revealed that almost one-third of all babies delivered in the United States are electively delivered with 5% of all deliveries in the U.S. delivered in a manner violating ACOG/AAP guidelines. Most of these are for convenience and result in significant short-term neonatal morbidity, such as neonatal intensive care unit admission rates of 13- 21% (Clark et al., 2009).

According to Glantz (2005), compared to spontaneous labor, elective inductions result in more cesarean births and longer maternal length of stay. The American Academy of Family Physicians (2000) also notes that elective induction doubles the cesarean delivery rate. Repeat elective cesarean births before 39 weeks gestation also

result in higher rates of adverse respiratory outcomes, mechanical ventilation, sepsis and hypoglycemia for the newborns (Tita et al., 2009).

A reduction in the number of non-medically indicated elective deliveries at >=37 to <39 weeks gestation will result in a substantial decrease in neonatal morbidity and mortality, as well as a significant savings in health care costs. In addition, the rate of cesarean sections should decrease with fewer elective inductions resulting in decreased length of stay and health care costs.

The measure will assist health care organizations (HCOs) to track non-medically indicated early term elective deliveries and reduce the occurrence.

Sources

- American College of Obstetricians and Gynecologists. (November 1996). ACOG Educational Bulletin.
- American Academy of Family Physicians. (2000). Tips from Other Journals: Elective induction doubles cesarean delivery rate, 61, 4.Retrieved December 29, 2008 at: http://www.aafp.org/afp/20000215/tips/39.htm
- Clark, S., Miller, D., Belfort, M., Dildy, G., Frye, D., & Meyers, J. (2009). Neonatal and maternal outcomes associated with elective delivery. [Electronic Version]. Am J Obstet Gynecol. 200:156.e1-156.e4.
- Glantz, J. (Apr.2005). Elective induction vs. spontaneous labor associations and outcomes. [Electronic Version]. J Reprod Med. 50(4):235-40.
- Tita, A., Landon, M., Spong, C., Lai, Y., Leveno, K., Varner, M, et al. (2009). Timing of elective repeat cesarean delivery at term and neonatal outcomes. [Electronic Version]. NEJM. 360:2, 111-120.

1b.2. Provide performance scores on the measure as specified (<u>current and over time</u>) at the specified level of analysis. *(<u>This is required for maintenance of endorsement</u>. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

For this 2020 Submission, CY 2018 is the most recent full year of data. The results are as follows:

CY 2018 Statistics:

Number of hospitals: 1616

Total Number of Patients: 139,213

Mean (SD): 1.7% (2.8%)

IQR: 2.5%

Deciles (0,10,20,30,40,50,60,70,80,90,100): 0%, 0%, 0%, 0%, 0%, 1.4%, 2.2%, 3.0%, 4.8%, 29.0%

Although performance has been improving, there are still 9% of hospitals reporting rates higher than the goal of 5%.

The Joint Commission requires submission of this measure for accreditation purposes as part of the ORYX Performance Measurement requirements. In January 2016, the Joint Commission required hospitals with greater than 300 live births to submit the PC-01 measure.

Correction

The trend for this measure shows a decrease in the rates from 2015-2018. Although performance has been improving, there are still 9% of hospitals reporting rates higher than the goal of 5%.

2015-2.3%

2016-1.9%

2017-1.7%

2018-1.6%

2016 Submission

Early term elective deliveries are still being performed; however, the performance gap has narrowed over time. A goal of 5% or less based on recommendations from the Perinatal Care (PC) Technical Advisory Panel (TAP) should be achievable. The PC core measures were added as a new core measure set in 2010 for hospitals to select to meet their ORYX performance measurement requirement for Joint Commission accreditation. At that time, approximately 164 hospitals reported the data with an average measure rate of 18.8% (n=11,843 patients). In January 2014, The Joint Commission required mandatory reporting of the PC measure set for all accredited hospitals with 1100 births or more annually. 1388 hospitals reported the data with an average rate of 3.4% (n=130,882 patients). It is important to note that a performance gap of 3.7% existed for the 90th percentile of hospitals performing at 8.7% (if 5% is considered goal performance). The threshold for mandatory reporting was recently lowered to 300 births annually effective January 2016. The new reporting requirement will now capture approximately 80% of all accredited birthing hospitals. As a result, the rates may increase with the addition of approximately 821 more hospitals reporting data. Below is the specified level of analysis for PC-01 beginning with discharges April 1, 2010 through December 31, 2014.

- 2Q 2010: 11,843 denominator cases; 2,231 numerator cases; 164 hospitals; 18.8% national aggregate rate; 0.17827 mean of hospital rates; 0.12745 standard deviation; 33.3% 90th percentile rate; 23.7% 75th percentile rate/upper quartile; 15.5% 50th percentile rate/median rate; 9.0% 25th percentile rate/lower quartile; and 4.7% 10th percentile rate.
- CY 2011: 1,3907 denominator cases; 1,892 numerator cases; 166 hospitals; 13.6% national aggregate rate; 0.13998 mean of hospital rates; 0.13183 standard deviation; 31.5% 90th percentile rate; 18.3% 75th percentile rate/upper quartile; 9.8% 50th percentile rate/median rate; 5% 25th percentile rate/lower quartile; and 1.5% 10th percentile rate.
- CY 2012: 1,3404 denominator cases; 1,081 numerator cases; 170 hospitals; 8.0% national aggregate rate; 0.08296 mean of hospital rates; 0.09555 standard deviation; 21.2% 90th percentile rate; 10.8% 75th percentile rate/upper quartile; 4.9% 50th percentile rate/median rate; 2.6% 25th percentile rate/lower quartile; and 0% 10th percentile rate.
- CY 2013: 1,4880 denominator cases; 658 numerator cases; 200 hospitals; 4.4% national aggregate rate; 0.05737 mean of hospital rates; 0.10193 standard deviation; 13.9% 90th percentile rate; 7.6%% 75th percentile rate/upper quartile; 2.6% 50th percentile rate/median rate; 0% 25th percentile rate/lower quartile; and 0% 10th percentile rate.
- CY 2014: 130,882 denominator cases; 4,331 numerator cases; 1388 hospitals; 3.3% national aggregate rate; 0.03406 mean of hospital rates; 0.04647 standard deviation; 8.7% 90th percentile rate; 4.5% 75th percentile rate/upper quartile; 2.1% 50th percentile rate/median rate; 0% 25th percentile rate/lower quartile; and 0% 10th percentile rate.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

Not Applicable

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in*

care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.

For 2018 discharges:

Rates by Age category

Age	Rate (%)
<20	1.57
20-24	1.25
25-29	1.50
30-34	1.77
35-39	2.28
40+	3.90

Rates by Hispanic Ethnicity

Hispanic Ethnicity	Rate (%)
No	1.64
Yes	1.79

Rates by Race

Race	Rate (%)
White	1.77
African American	1.49
American Indian	1.06
Asian	1.00
Pacific Islander	1.34
Unable to Determine	1.77

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

2. Reliability and 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. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

Perinatal Health

De.6. Non-Condition Specific *(check all the areas that apply):*

Safety : Complications, Safety : Overuse

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

Women

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

https://manual.jointcommission.org/releases/TJC2020A2/PerinatalCare.html

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

Attachment **Attachment:** PC01AppendixATJCTablesv2020A2.xlsx

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

Yes

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

On an annual basis, the chart abstracted measures maintained by The Joint Commission undergo an annual update to revise specifications based on updated research and clinical information or standards changes. The following changes have been made to the measure specifications:

- Updated data element Prior Uterine Surgery: Added notes in order to clarify abstraction of prior uterine surgeries.
- Algorithm changed to check History of Stillbirth at the end to ease abstraction burden.
- Appendix A - ICD-10 Code Tables: Revised to reflect the ICD-10 code updates for Fiscal Year (FY) 2020, effective for discharges October 1, 2019
- Data element Gestational Age:

o Notes for abstraction and Suggested Data Sources have been updated and reordered to clarify, reduce burden of abstraction and align with the eCQM 0469e measure specifications.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

<u>IF an OUTCOME MEASURE</u>, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Patients with elective deliveries with ICD-10-PCS Principal Procedure Code or ICD-10-PCS Other Procedure Codes for one or more of the following:

• Medical induction of labor as defined in Appendix A, Table 11.05 while not in Labor prior to the procedure

• Cesarean birth as defined in Appendix A, Table 11.06 and all of the following:

•not in Labor

•no history of a Prior Uterine Surgery

S.5. 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, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

<u>IF an OUTCOME MEASURE,</u> describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Four data elements are used to calculate the numerator:

1. ICD-10-PCS Other Procedure Codes - The International Classification of Diseases, Tenth Revision, Procedure Coding System code that identifies significant procedures performed other than the principal procedure during this hospitalization.

2. ICD-10-PCS Principal Procedure Code - The International Classification of Diseases, Tenth Revision, Procedure Coding System code that identifies the principal procedure performed during this hospitalization. The principal procedure is the procedure performed for definitive treatment rather than diagnostic or exploratory purposes, or which is necessary to take care of a complication.

3. Labor- Documentation that the patient was in labor prior to induction and/or cesarean birth. Allowable values: Yes or No/UTD.

4. Prior Uterine Surgery- Documentation that the patient had undergone prior uterine surgery which includes: a prior classical cesarean birth defined as a vertical incision into the upper uterine segment, a prior myomectomy, a prior uterine surgery resulting in a perforation of the uterus due to an accidental injury, a history of a uterine window or thinning or defect of the uterine wall noted during prior uterine surgery or during a past or current ultrasound, a history of uterine rupture requiring surgical repair, a history of a cornual ectopic pregnancy, a history of a transabdominal cerclage, or a history of metroplasty and/or prior removal of vestigial horn with entry into the uterine cavity.

Allowable Values: Yes or No/UTD

Patients are eligible for the numerator population with ICD-10-PCS Other Procedure Codes or ICD-10-PCS Principal Procedure Code for medical induction or with ICD-10-PCS Other Procedure Codes or ICD-10-PCS Principal Procedure Code for cesarean birth when the allowable value equals “no” for the data elements Labor and Prior Uterine Surgery.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

Patients delivering newborns with >= 37 and < 39 weeks of gestation completed with ICD-10-PCS Principal or Other Procedure Codes for delivery as defined in Appendix A, Table 11.01.1 and with ICD-10-CM Principal Diagnosis Code or ICD-10-CM Other Diagnosis Codes for planned cesarean birth in labor as defined in Appendix A, Table 11.06.1.

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets –*

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Seven data elements are used to calculate the denominator:

1. Admission Date – The month, day, and year of admission to acute inpatient care.
2. Birthdate - The month, day, and year the patient was born.
3. Discharge Date – The month, day, and year the patient was discharged from acute care, left against medical advice, or expired during the stay.
4. Gestational Age – Documentation of the weeks of gestation completed at the time of delivery. Allowable Values: 1-50 or UTD.
5. History of Stillbirth – Documentation that the patient had prior history of stillbirth. Allowable Values: Yes or No/UTD
6. ICD-10-CM Other Diagnosis Codes - The International Classification of Diseases, Tenth Revision, Clinical Modification codes associated with the other or secondary diagnoses for this hospitalization.
7. ICD-10-CM Principal Diagnosis Code - The International Classification of Diseases, Tenth Revision, Clinical Modification diagnosis code that is primarily responsible for the admission of the patient to the hospital for care during this hospitalization.

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

ICD-10-CM Principal Diagnosis Code or ICD-10-CM Other Diagnosis Codes for conditions possibly justifying elective delivery prior to 39 weeks gestation as defined in Appendix A, Table 11.07 include the following:

- History of prior stillbirth
- Less than 8 years of age
- Greater than or equal to 65 years of age
- Length of Stay >120 days
- Gestational Age < 37 or >= 39 weeks or UTD

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

- Patients with ICD-10-CM Principal Diagnosis Code or Other Diagnosis Codes for conditions for possibly justifying elective delivery are excluded.
- The patient age in years is equal to the Admission Date minus the Birthdate. Patients less than 8 years of age or greater or equal to 65 years of age are excluded.
- Length of stay (LOS) in days is equal to the Discharge Date minus the Admission Date. If the LOS is greater than 120 days, the patient is excluded.
- Patients with a Gestational Age less than 37 weeks or equal to or greater than 39 weeks or UTD are excluded from the measure.
- Patients with a prior history of stillbirth are excluded from the measure.

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and*

the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)

Not Applicable

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Lower score

S.14. Calculation Algorithm/Measure Logic (*Diagram or 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; time period for data, aggregating data; risk adjustment; etc.)*

1. Start processing. Run cases that are included in the PC-Mother Initial Patient Population and pass the edits defined in the Transmission Data Processing Flow: Clinical through this measure.

2. Check ICD-10-CM Principal or Other Diagnosis Codes

a) If at least one of the ICD-10-CM Principal or Other Diagnosis Codes is on Table 11.07, the case will proceed to a Measure Category Assignment of B and will not be in the Measure Population. Stop processing.

b) If none of the ICD-10-CM Principal or Other Diagnosis Code is on Table 11.07, continue processing and proceed to Gestational Age.

3. Check Gestational Age

a) If Gestational Age is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

b) If Gestational Age is less than 37 or greater than or equal to 39 or equal to a Not Unable to Determine Value, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Stop Processing.

c) If Gestational Age is greater than or equal to 37 and less than 39, continue processing and proceed to Check History of Stillbirth.

4. Recheck ICD-10-CM Principal or Other Diagnosis Codes

a) If at least one of the ICD-10-CM Principal or Other Diagnosis Code is on Table 11.06.1, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop processing.

b) If none of the ICD-10-CM Principal or Other Diagnosis Code is on Table 11.06.1, continue processing and proceed to ICD-10-CM Principal or Other Procedure Codes.

5. Check ICD-10-PCS Principal or Other Procedure Codes

a) If at least one of the ICD-10-PCS Principal or Other Procedure Codes is on Table 11.05, continue processing and proceed to Labor

b) If Labor is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

c) If Labor equals No, the case will proceed to a Measure Category Assignment of E and will be in the Numerator Population. Stop Processing.

d) If none of the ICD-10-CM Principal Procedure Codes is on Table 11.05, continue processing and proceed to recheck ICD- 10-PCS Principal or Other Procedure Codes.

6. Recheck ICD-10-PCS Principal or Other Procedure Codes

a) If none of the ICD-10-PCS Principal or Other Procedure Code is on Table 11.06, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop Processing.

b) If at least one of the ICD-10-PCS Principal or Other Procedure Code is on Table 11.06, continue processing and proceed to Labor.

7. Check Labor

a) If Labor is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

b) If Labor equals Yes, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop processing.

c) If Labor equals No, continue processing and proceed to Prior Uterine Surgery.

8. Check Prior Uterine Surgery

a) If Prior Uterine Surgery is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

b) If Prior Uterine Surgery equals Yes, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop processing.

c) If Prior Uterine Surgery equals No, the case will proceed to a Measure Category Assignment of E and will be in the Numerator Population. Stop Processing. Gestational Age.

9. Check History of Stillbirth (as of 1/1/2019 this check moves to last position)

a) If History of Stillbirth is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

b) If History of Stillbirth is Yes, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Stop Processing.

c) If History of Stillbirth is No, continue processing and proceed to recheck ICD-10- CM Principal Procedure or Other Diagnosis Codes.Gestational Age.

4. Check Gestational Age

a. If Gestational Age is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

b. If Gestational Age is less than 37 or greater than or equal to 39 or equal to a Not Unable to Determine Value, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Stop Processing.

c. If Gestational Age is greater than or equal to 37 and less than 39, continue processing and proceed to recheck ICD-10-CM Principal Procedure or Other Diagnosis Codes.

5. Recheck ICD-10-CM Principal or Other Diagnosis Codes

a. If at least one of the ICD-10-CM Principal or Other Diagnosis Code is on Table 11.06.1, the case will proceed to a Measure Category Assignment of D and will

be in the Measure Population. Stop processing.

b. If none of the ICD-10-CM Principal or Other Diagnosis Code is on Table 11.06.1, continue processing and proceed to ICD-10-CM Principal or Other Procedure Codes.

6. Check ICD-10-PCS Principal or Other Procedure Codes

a. If all of the ICD-10-PCS Principal or Other Procedure Codes are missing, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop Processing.

b. If at least one of the ICD-10-PCS Principal or Other Procedure Codes is on Table 11.05, continue processing and proceed to Labor

i. If Labor is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

ii. If Clinical Trial equals Yes, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop Processing.

iii. If Labor equals No, the case will proceed to a Measure Category Assignment of E and will be in the Numerator Population. Stop Processing.

c. If none of the ICD-9-CM Principal Procedure Codes is on Table 11.05, continue processing and proceed to recheck ICD-10-PCS Principal or Other Procedure Codes.

7. Recheck ICD-10-PCS Principal or Other Procedure Codes

a. If none of the ICD-10-PCS Principal or Other Procedure Code is on Table 11.06, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop Processing.

b. If at least one of the ICD-10-PCS Principal or Other Procedure Code is on Table 11.06, continue processing and proceed to Labor.

8. Check Labor

a. If Labor is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

b. If Labor equals Yes, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop processing.

c. If Labor equals No, continue processing and proceed to Spontaneous Rupture of Membranes.

9. Check Prior Uterine Surgery

a. If Prior Uterine Surgery is missing, the case will proceed to a Measure Category Assignment of X and will be rejected. Stop Processing.

b. If Prior Uterine Surgery equals Yes, the case will proceed to a Measure Category Assignment of D and will be in the Measure Population. Stop processing.

c. If Prior Uterine Surgery equals No, the case will proceed to a Measure Category Assignment of E and will be in the Numerator Population. Stop Processing.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

The initial patient population includes patients admitted to the hospital for inpatient acute care if they have: ICD-10-PCS Principal or Other Procedure Code as defined in Appendix A, Table 11.01.1, a Patient Age (Admission Date – Birthdate) >= 8 years and < 65 and a Length of Stay (Discharge Date - Admission Date) = 120 days. The sample is taken randomly as follows for a monthly sample:

• Average monthly Initial Patient Population >= 501 results in a minimum random sample size of 101.

• Average monthly Initial Patient Population 126 – 500 results in a minimum random sample size of 20% of the population size.

• Average monthly Initial Patient Population 25 – 125 results in a minimum random sample size of 25.

• Average monthly Initial Patient Population < 25 results in no sampling; 100% Initial Patient Population required

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

Not Applicable

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Electronic Health Records, Other, Paper Medical Records

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

Starting in 2020, hospitals will use the Direct Data Submission Platform for submission of chart abstracted measures. Thus, in 2020, organizations have one place to submit both eCQM and chart abstracted data. The goal of the Direct Data Submission Platform is to ease the burden and expense of submission and empower organizations with data for quality improvement.

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Facility, Other

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Inpatient/Hospital

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

Not Applicable

2. Validity – See attached Measure Testing Submission Form

2020_nqf_testing_attachment_PC01_0469_final-637227327320109820.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 0469

Measure Title: PC-01 Elective Delivery

Date of Submission: January 3, 2020

Type of Measure:

□Outcome (*including PRO-PM*)	□Composite – ***STOP – use composite testing form***
☐Intermediate Clinical Outcome	□Cost/resource
□Process *(including Appropriate Use)*	□Efficiency
□Structure	

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing, (e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	**Measure Tested with Data From:**
□abstracted from paper record	□abstracted from paper record
□claims	□claims

□registry	□registry
□abstracted from electronic health record	□abstracted from electronic health record
□eMeasure (HQMF) implemented in EHRs	□eMeasure (HQMF) implemented in EHRs
□ other: Click here to describe	□ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*). Not Applicable

1.3. What are the dates of the data used in testing?

2020 Submission
Testing of measure score reliability and validity was performed using data from hospital discharges occurring in 2018.

2016 Submission
This submission included initial testing of measure score reliability and validity was performed using data from hospital discharges occurring in 1Q2011. Also, updates to validity and exclusions used 2015 data.

1.4 What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, healthplan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	**Measure Tested at Level of:**
□individual clinician	□individual clinician
□group/practice	□group/practice
□hospital/facility/agency	□hospital/facility/agency
□health plan	□health plan
□ other: Click here to describe	□ other: Click here to describe

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

2020 Submission
This measure assesses the proportion of patients delivering newborns with >= 37 and < 39 weeks of gestation completed with an elective delivery. The intended use of the measure is to assess the quality of perinatal care in hospitals across the population.

Entities in reliability and validity testing: Results were calculated from Joint Commission data that included 1616 hospitals submitting the measure using 2018 discharges and had greater than or equal to 30 denominator cases, the minimum sample size required for public reporting. The hospitals were geographically diverse and varied in size.
1616 health care organizations representing various types, locations and sizes:
330 For Profit, 1103 Not for Profit, 183 Government
565 >=300 beds; 777 100-300 beds; 274 <100 beds
299 Rural; 1317 Urban

153 Major Teaching; 800 Minor Teaching; 663 Non-Teaching

2016 Submission
The PC measure set has been in national use since the 2nd quarter of 2010. It is a requirement of participation in the ORYX initiative that data on all measures in the set are collected. (ORYX is the term used by The Joint Commission to describe the component of the hospital accreditation program which requires data collection and reporting on standardized national performance measures.) Demographics of organizations collecting and reporting data on these measures are as follows:
163 health care organizations representing various types, locations and sizes:
10 For Profit, 91 Not for Profit, 46 Military Facilities, 9 County, 2 State, 5 Other
15 >=500 beds; 29 250-499 beds; 50 100-249 beds; 69 <100 beds
Located in: AE, AK, AL, AP, AR, AZ, CA, DO, DC, FL, GA, HI, ID, IL, IN, KS, KY, LA, MA, MD, MI, MN, MO, MS, MT, NC, NE, NV, NY, OH,OK, PA, PR, RI, SC, TN, TX, VA, WA, WI, WV
26 performance measurement systems

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

2020 Submission
Patients in reliability and validity testing: Data are summarized at the hospital level. Below is a description of the sample. It includes number of hospitals included in Joint Commission data, the median initial population size, and the median denominator size for the measure across hospitals.

Median denominator size for Elective Delivery, 2018 (Patients included in sample=14,184)

Number of Hospitals	Median number of deliveries	Median number of denominator cases
1616	1227	59

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

2020 Submission
No differences in the data used for reliability and validity testing.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

2020 Submission
No patient-level sociodemographic variables are used in the measure and none were available for analysis. There is no compelling evidence available supporting association between social risk factors and this measure.

2a2. RELIABILITY TESTING
***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data element; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☐ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

2020 Submission

Reliability testing of performance measure score

We utilized the Beta-binomial model (Adams 2009) to assess how well one can distinguish the performance of one hospital from another. Conceptually, the Beta-binomial model measures the ratio of signal to noise. The signal is the proportion of the variability in the measured performance that can be explained by real differences in performance. The Beta-binomial model is an appropriate model when estimating the reliability of simple pass/fail rate measures as is the case with most Joint Commission measures. Reliability scores range from 0.0 to 1.0. A score of zero implies that all variation is attributed to measurement error (i.e., noise), whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance (across hospitals).

Adams, J.L. The Reliability of Provider Profiling: A Tutorial. Santa Monica, California: RAND Corporation. TR-653-NCQA, 2009

Comparison of ICD-9 to ICD-10 codes

Reliability was measured by comparing results from the 2017 data submission that utilized ICD-9 codes to the results from the 2018 data submission that utilized ICD-10 codes. Summary statistics for the number of numerator cases, number of denominator cases, and observed rates are presented. Hospital data were also matched by each year in each of the three attributes and a paired t-test was used to determine statistical significance for each attribute.

This measure was adapted from NQF-endorsed measure 0469 Elective Delivery Prior to 39 Completed Weeks Gestation. As such, initial data reliability would have been addressed during the original endorsement. The Joint Commission will be conducting additional reliability studies on this measure as well as the entire PC measure set beginning in October 2011.

Currently, hospitals are supported in their data collection and reporting efforts by 26 contracted performance measurement system (PMS) vendors. It is a contractual requirement of Joint Commission listed vendors that the quality and reliability of data submitted to them by contracted health care organizations must be monitored on a quarterly basis. In addition, The Joint Commission analyzes these data by running 17 quality tests on the data submitted into ORYX. (ORYX is the term used by The Joint Commission to describe the component of the hospital accreditation program which requires data collection and reporting on standardized national performance measures). The following is a list of the major tests done on the submitted ORYX data, taken from the 2011 ORYX Performance Measurement System Requirements manual.

- Transmission of complete data
- Usage of individual core measure data received: To understand if the HCO provides the relevant service to treat the relevant population
- Investigation of aberrant data points
- Verification of patient population and sample size
- Identification of missing data elements
- Validation of the accuracy of target outliers
- Data integrity

- Data corrections

Data Element Agreement Rate:
Inter-rater reliability testing methodology utilized by contracted performance measure system vendors as outlined in the contract is as follows:

- All clinical data elements and all editable demographic elements are scored.
- All measure data are re-abstracted with originally abstracted data having been blinded so that the re-abstraction is not biased.
- Re-abstracted data are compared with originally abstracted data on a data element by data element basis. A data element agreement rate is calculated. Clinical and demographic data are scored separately, and an overall agreement rate is computed.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

2020 Submission

Reliability testing of performance measure score
Reliability statistic for the measure: patients with elective vaginal deliveries or elective cesarean births at >= 37 and < 39 weeks of gestation completed:

Average: 0.78
Median: 1.0
10th-90th percentile across hospitals: 0.32 – 1.0

Comparison of ICD-9 to ICD-10 codes

Comparison between the ICD-9 and ICD-10 number of numerator cases, denominator cases, and observed rates

		N	Mean	Std. Dev.	Min	Q1	Median	Q3	Max	Pairwise Difference	P-Value
Numerator	ICD-9	2071	0.3071	0.8446	0	0	0	0	13	.0132	0.5432
	ICD-10	1998	0.3233	0.957	0	0	0	0	23		
Denominator	ICD-9	2071	18.799	20.945	1	8	12	21	291	-0.6876	0.0032
	ICD-10	1998	18.408	19.991	1	8	12	21	269		
Rate	ICD-9	2071	0.01782	0.05411	0	0	0	0	1	0.0008	0.5675
	ICD-10	1998	0.0179	0.04947	0	0	0	0	0.5		

Data element agreement rates were reported to The Joint Commission for 1Q11 This reflects the findings of 108 hospitals, comprising 13,279 records (100% sample). The following table delineates calculated agreement rates for individual data elements that are used to compute measure rates for PC-01.

Data Elements with a Mismatch	Total Numerator	Total Denominator	Rate
Active Labor	33	35	94.29%
Gestational Age	639	712	89.75%

These agreement rates are considered to be well within acceptable levels.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

2020 Submission

In general, a score of 0.7 or higher suggests the measure has adequate reliability. The results suggest the measure has high reliability for most of the hospitals.

Pairwise comparisons were not statistically significant for the number of numerator cases and the observed rates between matched hospitals between 2017 data (ICD-9) and 2018 data (ICD-10) with p-values greater than 0.05. For the number of denominator cases, there was statistical significance (p=0.0032) but the mean difference was only 0.68 case which is not clinically meaningful. This suggests that there are no differences in reliability of the measure using the previous ICD-9 coding and the current ICD-10 coding.

2b1. VALIDITY TESTING-new

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☐ **Performance measure score**

☐ **Empirical validity testing**

☐ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

2020 Submission

Correlations of the measure with other measures of perinatal care quality. Since a low measure rate for PC-01 is desirable, this measure is hypothesized to correlate negatively to other perinatal care measures where a high rate is desirable (PC-05, ePC-05, ePC-01) and positively to perinatal care measures where a low rate is desirable (PC-02).

A correlation of 0.1 - 0.3 was considered weak, 0.3 - 0.5 was considered moderate, and over 0.5 was considered strong.

Since the measure has been in national use, continued face validity of the measure has been determined through analysis of feedback from measure users. The Joint Commission provides a web-based application with which measure users can provide feedback regarding appropriateness of measure specifications, request clarification of specifications, and/or provide other comments pertinent to the measure. This feedback is systematically, continually, reviewed in order to identify trends and to identify areas of the measure specifications that require clarification or revision. Additionally, Joint Commission staff continually monitors the national literature and environment in order to assess continued validity of this measure.

As noted previously, The Joint Commission is currently performing reliability site visits. A component of these visits will include focus group interviews with hospital staff working with the PC measures to obtain feedback regarding the validity of the measures and suggestions for further refinement of the specifications.

ICD-9 to ICD-10 Conversion Process:

The goal was to convert ICD-9 to ICD-10 equivalent codes, consistent with the clinical intent of the original measure specifications. The Joint Commission worked with a certified coding expert throughout the conversion process. The 3M Coding Conversion Tool was utilized, including forward mapping of ICD-9 codes to

ICD-10 codes as well as reverse mapping from ICD-10 to ICD-9 to ensure appropriateness. MSDRGs and instructions in the tabular index were also examined to ensure appropriate code mapping. Crosswalks comprising ICD-9 codes mapped to ICD-10 codes were created and reviewed by members of the Technical Advisory Panel, CMS subcontractors, and performance measurement system vendors prior to being posted for a 12 month public comment period. Feedback from the field indicated that the crosswalks generally were mapped correctly. Minor modifications to the code tables were made as needed. Final code tables were published in early 2015, well in advance of the mandated date of October 1, 2015.

Perinatal Care (PC) Initial Patient Population
The PC measure set is unique in that there are two distinct Initial Patient Populations within the measure set, mothers (PC-01, PC-02, PC-03) and newborns. (PC-04, PC-05).

Subpopulation Mothers
Patients admitted to the hospital for inpatient acute care are included in the PC Mother Initial sampling group if they have: ICD-9-CM Principal or Other Diagnosis Code as defined in Appendix A, Tables 11.01, 11.02, 11.03, or 11.04, a Patient Age (Admission Date Birthdate) >= 8 years and < 65 and a Length of Stay (Discharge Date - Admission Date) ≤ 120 days. PC-01- Elective Delivery belongs to the above population.

The data used to measure the validity of the PC measure are comprised of data from the third and fourth quarters of 2014, and the first and second quarters of 2015. 1,345 hospitals submitted 2,695,467 inpatient records for all the elected PC measures. The hospitals included in the analysis reported one year of data and had 30 or more denominator cases in the analysis period.
Measure convergent validity for PC-01 was assessed using hospitals patient level data from The Joint Commission warehouse. Measure specifications, including population identification, numerator and denominator statements, exclusions, and data elements and their definitions were found to be understandable, retrievable, and relevant in previous validity testing.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)-

2020 Submission

Correlations with other measures of perinatal care quality.

Table of Correlations

Measure	PC-01	PC-02	PC-05	ePC-01	ePC-05
PC-01-Elective Delivery	1				
PC-02-Cesarean Birth	0.133192	1			
PC-05-Exclusive Breast Milk Feeding	-0.02553	-0.28103	1		
ePC-01-Elective Delivery	0.008936	0.108322	0.022812		
ePC-05-Exclusive Breast Milk Feeding	0.040365	-0.17522	0.748033	-0.45737	1

Analysis of feedback obtained via our automated feedback system reveals slightly more than 180 submissions regarding specifications for this measure since its implementation in 2010. Predominant themes of these submissions involved questions regarding clarification of the data elements Active Labor and Gestational Age with respect to both definitions and the calculation of gestational age and the order of priority sources to

retrieve the data. Additional notes for abstractors were added to the data elements for clarification. In addition, the data elements Active Labor and Spontaneous Rupture of Membranes were moved from the numerator population to the denominator population and the algorithm was revised in order to capture all deliveries in the denominator population. Additional ICD-9-CM diagnosis codes were added to Table 11.07 to update exclusions based on consultation with the perinatal care experts. The gestational age range for the denominator statement and included population was also revised to exclude patients with a gestational age of 39 weeks of gestation completed, since the upper range for gestational age for 38 weeks ends at 38 6/7 weeks gestation.

Overall descriptive statistics for sub population MOTHER:

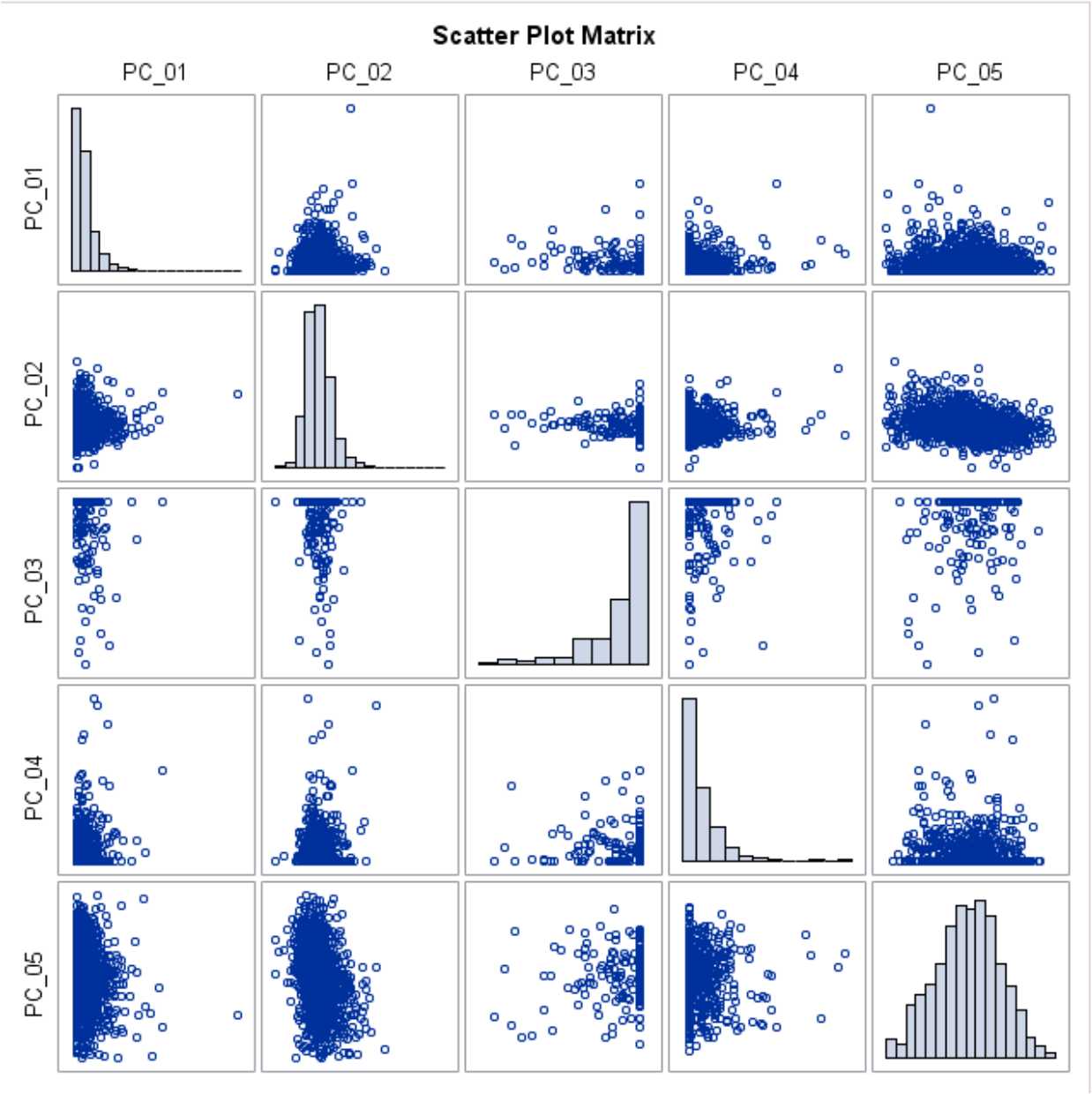
N=1,345 hospitals
n = 2,695,467 records submitted

Descriptive statistics for PC- 01 measure:
N = 1,237 hospitals
n = 1,130,083
Mean: 2.74%
Min = 0%
Percentile 10%: 0%
Percentile 25%: 0%
Median: 1.7%
Percentile 75%: 3.7%
Percentile 90%: 6.7%
Max = 51.2%

Simple Statistics

Variable	N	Mean	Std Dev	Median	Minimum	Maximum
PC_01	1237	0.02753	0.03803	0.01734	0	0.51240
PC_02	1345	0.26287	0.07974	0.25410	0	1.00000
PC_03	162	0.97762	0.03311	0.99425	0.84615	1.00000
PC_04	523	0.05267	0.08432	0.02203	0	0.66129
PC_05	1352	0.49198	0.19284	0.50190	0.00317	1.00000

Spearman Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations					
	PC_01	PC_02	PC_03	PC_04	PC_05
PC_01	1.00000 1237	0.06843 0.0163 1231	-0.26960 0.0006 159	0.10724 0.0169 496	-0.03538 0.2137 1237
PC_02	0.06843 0.0163 1231	1.00000 1345	-0.18318 0.0196 162	0.02807 0.5218 523	-0.32009 <.0001 1343
PC_03	-0.26960 0.0006 159	-0.18318 0.0196 162	1.00000 162	-0.03117 0.7030 152	0.07729 0.3283 162
PC_04	0.10724 0.0169 496	0.02807 0.5218 523	-0.03117 0.7030 152	1.00000 523	-0.03560 0.4165 523
PC_05	-0.03538 0.2137 1237	-0.32009 <.0001 1343	0.07729 0.3283 162	-0.03560 0.4165 523	1.00000 1352



The Spearman rank-order correlation is a nonparametric measure of association based on the ranks of the data values by measure PC-01 and hospitals. We used this methodology because of the skewness of the distribution of the measure rates.

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

2020 Submission

Except for the correlation between PC-01 and ePC-05, the directions of the correlations were in the expected

direction. The correlation of PC-01 and ePC-05, although expected to be in the negative direction, was not significantly greater than zero, nor was the correlation between PC-01 and PC-05. These measures (PC-01 vs PC-05 and ePC-05) evaluate two different populations, mothers and babies, and therefore two different aspects of perinatal care, which are apparently not correlated. The perinatal care measures used in this analysis are assessing different components of perinatal care and would not be expected to be more than weakly correlated since perinatal care quality is a multidimensional quantity.

The correlation of PC-01 with the other PC measures in the PC measure set indicates that the correlations (with the exception of PC-05), although in the expected direction and statistically significant, are relatively weak. Although 90% of the hospital measure rates fall between 0 and 6.7%, there are still a number of hospitals with measure rates significantly greater than 6.7%, indicating that the performance of hospitals on this measure are not uniformly acceptable.

2b2. EXCLUSIONS ANALYSIS

NA ☐no exclusions — *skip to section 2b3*

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

2020 Submission

Our testing addresses exclusions, as shown in Table 5.

Measure Exclusions

Exclusion	Rationale	Measure Denominator lost due to exclusion
At least one condition possibly justifying elective delivery	Table 11.07 contains diagnosis codes for medical conditions that are reasons to perform an early term medical induction and/or cesarean delivery.	57.9%
Gestational age < 37 or >=39	The denominator population is limited to patients > 37 to < 39 weeks of completed gestation. Patients with UTD for gestational age typically have had no prenatal care. Babies delivered at 37 to 38 6/7 weeks have a higher risk of complications than those delivered after 39 weeks.	72.6%
History of stillbirth	To exclude patients with a history of stillbirth, this condition is a reason for early delivery, however it is not captured in an ICD-10 code.	0.4%

Note: The exclusions presented in this table are not mutually exclusive. For example, a discharge that falls under exclusions 1 and 2 would appear in both places in this table.

We tested whether the exclusions affected overall performance score denominators.

Measure exclusions that were not derived directly from the evidence are presented below. Please note that these are population exclusions that are necessary to ensure consistency in all measures in this 4-measure set. These exclusions were analyzed for frequency of occurrence. An issue that is of great concern to users of this measure is that due to the presence of exceptions to the measure, attainment of a 0% measure rate is not possible. Because of the role of this measure in the current Joint Commission accreditation process this is

especially troubling to measure users. This concern is the basis for a number of the non-evidence-based exclusions to these measures. Additional reasons for these population exclusions are enumerated in our response to section 2b1.1 above. The following measure exclusions that were not derived directly from the evidence are as follows:

1. Patients with LOS <120 days
2. Patients less than 8 years of age or greater than or equal to 65 years of age
3. Patients enrolled in clinical trials

There were 1,134,640 admissions selected from the initial cohort. From among the 1,134,640 admissions in 1,237 hospitals, the descriptive statistics are given below.

The following exclusions were analyzed by subpopulation and measure for frequency and variability across providers:

- ICD-9-CM Principal Diagnosis Code or ICD-9-CM Other Diagnosis Codes for conditions possibly justifying elective delivery prior to 39 weeks gestation as defined in Appendix A, Table 11.07
- Less than 8 years of age
- Greater than or equal to 65 years of age
- Length of stay > 120 days
- Enrolled in clinical trials
- Gestational Age < 37 or >= 39 weeks or UTD

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

2020 Submission

Exclusion	Rationale	Measure Denominator lost due to exclusion
At least one condition possibly justifying elective delivery	Table 11.07 contains diagnosis codes for medical conditions that are reasons to perform an early term medical induction and/or cesarean delivery.	57.9%
Gestational age < 37 or >=39	The denominator population is limited to patients > 37 to < 39 weeks of completed gestation. Patients with UTD for gestational age typically have had no prenatal care. Babies delivered at 37 to 38 6/7 weeks have a higher risk of complications than those delivered after 39 weeks.	72.6%
History of stillbirth	To exclude patients with a history of stillbirth, this condition is a reason for early delivery, however it is not captured in an ICD-10 code.	0.4%

Number and percent of denominator remaining after exclusions

PC-01 denominator before	PC-01 denominator after	Percent after exclusions

exclusions	exclusions	
1,569,198	149,998	9.6%

The percentiles for the hospital percent after exclusions had the following values for the 10th, 25th, 50th, 75th and 90th percentiles respectively: 6.2%, 7.8%, 10.1%, 12.7%, and 15.7%.

N=353,671
1. Patients who have a length of stay (LOS) greater than 120 days =0%
2. Patients less than 8 years of age or greater than or equal to 65 years of age=0%
3. Patients enrolled in clinical trials =0.04%

Exclusion Subpopulation 1 - PC-01
ICD-9-CM Principal Diagnosis Code or ICD-9-CM Other Diagnosis Codes for conditions possibly justifying elective delivery prior to 39 weeks gestation as defined in Appendix A, Table 11.07
Exclusion: No observations noted
Less than 8 years of age
Exclusion: Included in the initial population exclusion
Greater than or equal to 65 years of age
Exclusion: Included in the initial population exclusion
Length of Stay >120 days
Exclusion: Included in the initial population exclusion

Exclusion: Enrolled in Clinical Trials
Overall Number of Occurrences n = 748
Overall Occurrence Percentage: 0.07%
Minimum: 0%
10th Percentile: 0%
Median: 0%
90th Percentile: 0.062%
Maximum: 28%

Exclusion: Gestational Age < 37 or gestational Age = >39 weeks or UTD
Overall Number of Occurrences n = 851,258
Overall Occurrence Percentage: 84.9%
Minimum 0.29%
10th Percentile: 69.17%
Median: 75.2%
90th Percentile: 79.2%
Maximum: 84.8%

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

2020 Submission

We tested several exclusions in order to understand the impact on the denominator. All exclusions are necessary to ensure the construct validity of the measure and all have a clinical rationale. The exclusions have an appreciable impact on those cases included in the denominator of the measure. It should be noted that this high number is expected since it is not the population of interest as defined by this measure. Although the initial patient population is used as the basis for sampling for the measure, it is not the population of interest. Therefore, in the specifications these exclusions from the initial patient population have been incorporated into the denominator definition.

The frequency of exclusions is high for "Gestational Age < 37 or gestational Age = >39 weeks or UTD" Occurrence with an overall percentage equal 75%. The high percentage is justified by the scope of measure PC-01. The difference between the 10th and 90th percentiles of the distribution of exclusion rates is narrow indicating that the occurrence is random and likely would not bias performance results.

It is believed that all of the exclusions should be retained for the following reasons:

Exclusion: ICD-9-CM Principal Diagnosis Code or ICD-9-CM Other Diagnosis Codes for conditions possibly justifying elective delivery prior to 39 weeks gestation as defined in Appendix A, Table 11.07
Rationale: Rationale: Table 11.07 contains diagnosis codes for medical conditions that are reasons to perform an early term medical induction and/or cesarean delivery.

Exclusion: Patients who have a Length of stay greater than 120 days
Rationale: Included for this measure in order to harmonize with other CMS/Joint Commission aligned measures.

Exclusion: Patients enrolled in a Clinical Trial
Rationale: Only capture patients not enrolled in clinical trials studying pregnant patients or newborns.

Exclusion: Patients with Gestational Age < 37 or >=39 weeks or UTD
Rationale: The denominator population is limited to patients > 37 to < 39 weeks of completed gestation. Patients with UTD for gestational age typically have had no prenatal care.

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b4.

2b3.1. What method of controlling for differences in case mix is used?
☐ **No risk adjustment or stratification**
☐ **Statistical risk model with** Click here to enter number of factors **risk factors**
☐ **Stratification by** Click here to enter number of categories **risk categories**
☐ **Other,** Click here to enter description

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.
Not applicable

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not applicable

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)

Not applicable

Also discuss any "ordering" of risk factor inclusion; for example, are social risk factors added after all clinical factors?

Not applicable

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

- ☐ **Published literature**
- ☐ **Internal data analysis**
- ☐ **Other (please describe)**

Not applicable

2b3.4a. What were the statistical results of the analyses used to select risk factors?

Not applicable

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

Not applicable

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Not applicable

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Not applicable

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Not applicable

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Not applicable

2b3.9. Results of Risk Stratification Analysis:

Not applicable

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)
Not applicable

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)
Not applicable

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE
2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified

(*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

2020 Submission
To demonstrate meaningful differences in performance, The Joint Commission calculated a funnel plot (Spiegelhalter 2004) for the annual hospital rates of the measure. In a funnel plot, the observed measure is plotted against a measure of its precision, so that the control limits form a 'funnel' around the target outcome. The 95 per cent (≈2 standard deviation) and 99.8 per cent (≈3 standard deviation) prediction limits are then superimposed over this plot around the overall measure rate. Those rates lying outside the confidence limits are identified as outliers.

Spiegelhalter, DJ. Funnel plots for comparing institutional performance. Statistics in Medicine 2005; 24:1185–1202.

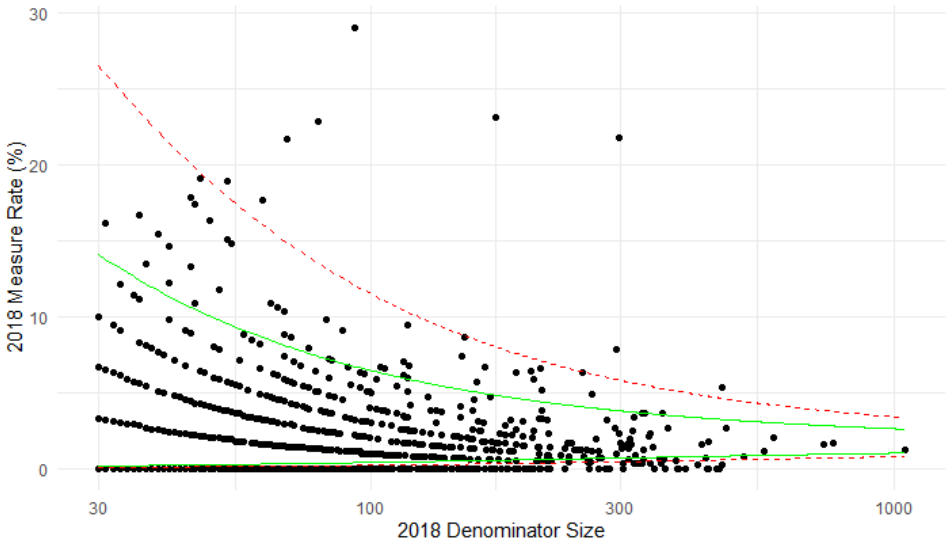
The method used to analyze meaningful differences in performance at The Joint Commission is Target Analysis. The object of target analysis is to compare a health care organization's (HCO) data against a comparative norm for the purpose of evaluating performance improvement opportunities. When an organization's performance level is statistically significantly different from a comparative norm, it is considered a statistical deviation. A statistical deviation may be desirable or undesirable depending on the "direction of improvement" of the measure.
There are two components to the target analysis methodology used at The Joint Commission. Given the national average for a performance measure, a target range is constructed. Using generalized linear mixed models' methodology (also known as hierarchical models), a predicted estimate of an HCO's performance, with a corresponding 95% confidence interval, is generated. This confidence interval is compared to the target range, to determine the HCOs' rating. The estimate of the organization's true performance is based on both the data from that organization and on data from the entire set of reporting organizations.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

2020 Submission
Using the funnel plot, 52 hospitals were identified as outliers with rates beyond the 2 standard deviation

upper limit and 9 hospitals were identified as outliers with rates beyond the 3 standard deviation limits. The upper confidence rate limit for a hospital with the median denominator size of 59 is 15.1% for a 95% confidence interval (2 SD) and 27.8% for a 99.8% confidence interval (3 SD).

Funnel Plot for PC-01:



PC-01 Distribution of Rates
2018 Data:
Scores on this measure: N=1616, Mean 1.7%, SD 2.8%
10th Percentile= 0%
25th Percentile= 0%
50th Percentile= 0%
75th Percentile= 2.5%
90th Percentile= 4.8%

PC-01 Distribution of Outliers
2011 1st Quarter Data:
Scores on this measure: N=160, Mean 13.6%, SD 0.1594
10th Percentile= 0%
25th Percentile= 0%
50th Percentile= 9%
75th Percentile= 19%
90th Percentile= 34%

156 (97.5%) Neutral – results not significantly different from target range
4 (2.5%) Unfavorable - results statistically significantly lower than the national rate

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

2020 Submission
The results indicate that there is significant hospital variability in rates and an appreciable number of

hospitals that are not within the expected level of variability.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2020 Submission
This submission is for the chart-based measure version of the eCQM measure 0469e, which has been submitted as a separate measure.

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)
Not applicable

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)
Not applicable

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)
Not applicable

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2020 Submission
The measure has been collected since 2011 and hospitals transmitting data with missing data on any of the critical data elements are not accepted.

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)
Not applicable

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of*

various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each)
Not applicable

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i*.e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)
Not applicable

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

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)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*)
Update this field for **maintenance of endorsement**.

Some data elements are in defined fields in electronic sources

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

The Joint Commission recognizes that not all hospitals currently have the capacity to abstract the electronic version of this measure, so continues to offer this chart abstracted version which allows for data capture from unstructured data fields. All data elements needed to compute the PC-01 performance measure score have been retooled for capture from electronic sources. Annual updates are performed to match the eCQM specifications to the current version of the chart-abstracted specifications.

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (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.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

At the present time, hospitals using this performance measure generally collect measure data via manual review of the EMR, data derived from vital records reports received from state or local departments of public health, delivery logs or clinical information systems or a combination. Collected data are submitted to The Joint Commission on a quarterly basis, as described previously. Specifications for this measure are freely available to anyone who wishes to use the measure. Feedback from hospitals using this measure indicates that required data elements are generally available in the medical record, and measure specifications are robust and easy to understand. As feedback from measure users indicate the need for clarification or revision of measure specifications, the Joint Commission reviews and implements as needed.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

There are no fees or licensing requirements to use the Joint Commission performance measures, all of which are in the public domain.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)

	Public Reporting Quality Check® http://www.qualitycheck.org/consumer/searchQCR.aspx Hospital Compare https://www.medicare.gov/hospitalcompare/search.html Quality Check® http://www.qualitycheck.org/consumer/searchQCR.aspx Hospital Compare https://www.medicare.gov/hospitalcompare/search.html Payment Program Hospital Inpatient Quality Reporting Program https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html Hospital Value Based Purchasing Program https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/index.html?redirect=/hospital-value-based-purchasing/ Regulatory and Accreditation Programs Hospital Accreditation Program http://jointcommission.org Hospital Accreditation Program http://jointcommission.org Quality Improvement (Internal to the specific organization) Perinatal Care Certification http://www.jointcommission.org/certification/perinatal_care_certification.aspx

4a1.1 For each CURRENT use, checked above (update for <u>maintenance of endorsement</u>), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

Name of program and sponsor: The Joint Commission Hospital Accreditation Program

- Purpose: An accreditation program that recognizes hospitals that meet standard requirements to provide safe and effective patient care.
- Geographic area and number and percentage of accountable entities and patients included: Nationwide; 3895 Joint Commission-accredited hospitals (2019)

Name of program and sponsor: Quality Check®- The Joint Commission

- Purpose: A public website that allows consumers to: search for accredited and certified organizations by city and state, by name or by zip code (up to 250 miles); find organizations by type of service provided within a geographic area; download free hospital performance measure results; and, print a list of Joint Commission certified disease-specific care programs and health care staffing firms.
- Geographic area and number and percentage of accountable entities and patients included: Nationwide; 3895 Joint Commission-accredited hospitals (2019)

Name of program and sponsor: The Joint Commission Perspective's: The Official Newsletter of the Joint Commission. (2019). The joint commission recognizes 20 years of ORYX performance measure reporting; look

back at the 20-year evolution of performance measure reporting and review the ORYX chart-abstracted measure results for 2017 and 2018, 39, 10.

- Purpose: The Perspective's article provides authoritative, accurate, and timely information about revisions and updates to Joint Commission standards, policies, and other requirements for all Joint Commission-accredited and -certified organizations and healthcare settings.

Name of program and sponsor: Hospital Inpatient Quality Reporting, Hospital Compare; Centers for Medicare & Medicaid Services

- Purpose: A public website that provides information to help consumers decide where to obtain healthcare and encourages hospitals to improve the quality of care they provide.
- Geographic area and number and percentage of accountable entities and patients included: Nationwide; 4500+ medicare-certified hospitals (2019)

Name of program and sponsor: Perinatal Care Certification-The Joint Commission

- Purpose: A certification program that recognizes hospitals that have achieved integrated, coordinated, patient-centered care for clinically uncomplicated pregnancies and births.
- Geographic area and number and percentage of accountable entities and patients included: Nationwide; 65 Joint Commission-accredited hospitals (2018)

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

Not Applicable

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

Not Applicable

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

The Joint Commission provides accredited healthcare organizations feedback reports for the measures submitted. The results are shared with organizations on a quarterly and/or annual basis depending on the reporting cycle of the measure. In addition, the Joint Commission has launched a new program called Continuous Customer Engagement (CCE) to assist organization in improving the quality of the performance measures. CCE includes enhanced dashboards with QI tools embedded into the dashboard, as well as focused and targeted solutions to assist organizations with gaps in the performance of their measures. The initial outreach to organizations utilizes an email process for hospital contact related to their measure rates and analysis. Response is provided in a timely manner either by email or directly by phone. Additionally, the data is available publicly through The Joint Commission Quality Check website. Individual hospital data for each rolling yearly time period is viewable and can be downloaded from this website.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

The Joint Commission is committed to provided valuable and actionable feedback to accredited organizations submitted the performance measurement data. The Joint Commission aggregates the Patient level data is aggregated at the hospital level quarterly. The hospital Performance Measure Report and Quality Check website are updated either quarterly or annually to reflect organization results, as well as National

Benchmarks. A user guide to the Performance Measure Report is posted on the Joint Commission website. Quality Check includes yearly and quarterly hospital rates, state and national averages, and the top 10 percentile at the national and state level.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

The Joint Commission utilizes an automated feedback system with access available to the measured entities and the vendors contracted by measured entities. A clinical lead is responsible for each individual measure set. The system is monitored daily and response is provided typically within 8 business hours. If queries cannot be managed via written response, arrangements are made to address any issues or concerns via phone. In addition, the Joint Commission developed dashboards as part of an ongoing project to provide continuous customer engagement. The Joint Commission analyzes aggregate performance in each of measure and identifies the measures for which the greatest opportunities for improvement exist among accredited hospitals. Based on those findings, an educational webinar series that address the high-opportunity topics is developed. All accredited hospitals have access to the educational webinar series. Organizations with high opportunity for improvement are particularly encouraged to participate. The dashboard report—posted in the Resources and Tools section of an accredited hospital's secure Joint Commission Connect® extranet site—is representative of each organization's relative performance on each of the selected measures. For each measure, the dashboard shows that organization's performance compared to national, state, and Joint Commission–accredited organization averages. The dashboard is not a score-able element on survey, but rather, a tool to facilitate discussion about ongoing quality improvement work. For example, surveyors may ask an organization how it addresses the subset of performance measures in the report and what action(s) the organization is taking to improve processes.

4a2.2.2. Summarize the feedback obtained from those being measured.

The Joint Commission provides several venues for the organizations being measured to provide feedback. Questions on the measures are most likely to come through the clinical and data receipt mailboxes provided on all communications. In addition, the Joint Commission has advisory committees for the Hospital Accreditation Program, which meet on a quarterly basis, and have the opportunity to provide feedback on the measures being collected.

Most statistical questions on this measure were regarding how this measure was to be publicly reported in 2020. There was strong support for the public reporting of this measure from multiple stakeholders.

Queries submitted via the automated feedback system have decreased significantly for the early elective delivery measure in the past three years.

Correction

The statement "Queries submitted via the automated feedback system have decreased significantly for the measure in the past three years." does apply for PC-01

4a2.2.3. Summarize the feedback obtained from other users

Same as above in 4a2.2.2.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

Note: All feedback is tracked and considered. If upon analysis there are trends noted giving cause for updates, this is reviewed by the measure workgroup to confirm the need for revision. Additionally, The Joint Commission engages a Technical Advisory Panel (TAP) for review and/or approval of updates which may require their additional subject matter expertise. All measure specifications are reviewed twice a year and updates are made as needed based on feedback from the measure users, input from the TAP, changes in the guidelines, or changes in clinical practice.

Minor modifications have been made to this measure based upon feedback received.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

Not Applicable

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

Unintended Consequence:

Cases with prior uterine surgery were inappropriately failing the measure.

Mitigating Action:

The measure rate calculation algorithm was revised to include a check prior to a cesarean birth via a new data element (Prior Uterine Surgery) created to enable cases with prior uterine surgery to remain in the denominator population and pass the measure.

Unintended Consequence:

Patients who did not receive prenatal care were inappropriately included in the measure denominator, as the gestational age data element was abstracted as unable to be determined (UTD).

Mitigating Action:

In order to avoid penalizing hospitals, cases with UTD were removed from the measure population.

Unintended Consequence:

Some hospitals have reported higher rates due to small denominator populations as a result of sampling.

Mitigating Action:

Vital Records reports, delivery logs and clinical information systems were added as acceptable data sources to help hospitals identify all cases with 37 and 38 weeks gestation, so that 100% of these cases could be reviewed to increase the denominator population size.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

Not applicable

5. Comparison to Related or 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.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

No

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications harmonized to the extent possible?

Yes

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

The measures are completely harmonized to the extent possible, given the fact that the data source for #0469e is the electronic clinical quality measure record.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. If this measure conceptually addresses 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.)

Not Applicable

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested

information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Available at measure-specific web page URL identified in S.1 **Attachment:**

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): The Joint Commission

Co.2 Point of Contact: JohnMarc, Alban, jalban@jointcommission.org, 630-792-5304-

Co.3 Measure Developer if different from Measure Steward: The Joint Commission

Co.4 Point of Contact: Tricia, Elliott, TElliott2@jointcommission.org, 630-792-5643-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

Michael Ross, MD, MPH (Chair)

Harbor-UCLA Medical Center

Torrance, CA

Martin McCaffrey, MD

UNC North Carolina Children's Hospital

Chapel Hill, NC

Debra Bingham PhD, RN, FAAN

Institute of Perinatal Quality Improvement

Washington, DC

James Christmas, MD

HCA Clinical Services Group

Elizabeth Rochin, PhD, RN, NE-BC

National Perinatal Information Center

Providence, RI

Cathy Ivory, PhD, RNC-OB, RN-BC, FAAN

Indiana University Health

Indianapolis, IN

Joseph Kunisch, PhD, RN-BC, CPHQ

Memorial Hermann Healthcare System

Houston, TX

B. Dale Magee, MD, MS

Shrewbury, MA

Elliott Main, MD

Stanford University

Mill Valley, CA

Susan Matney, PhD, RNC-OB

Intermountain Healthcare

Salt Lake City, UT

Elizabeth O'Neil-Greiner, RN, MHA

BJC Healthcare

St. Louis, MI

Patrick Romano, MD, MPH

University of California Davis Health

Sacramento, CA

Mark Tomlinson, MD

Providence Health System

Portland, OR

Brooke Villarreal, DNP, MSN, RN-BC

HCA Healthcare

Nashville, TN

The technical advisory panel (TAP) members determined priority areas that could be evaluated to improve care related to perinatal care during the development timeframe. After implementation, minor revisions, acknowledged by TAP representatives, were made to improve clarity. Hospital feedback will be reviewed during the reliability testing phase of the project to assist the TAP in making the final measure recommendations.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2010

Ad.3 Month and Year of most recent revision: 10, 2015

Ad.4 What is your frequency for review/update of this measure? Biannual

Ad.5 When is the next scheduled review/update for this measure? 01, 2020

Ad.6 Copyright statement: No royalty or use fee is required for copying or reprinting this manual, but the following are required as a condition of usage: 1) disclosure that the Specifications Manual is periodically updated, and that the version being copied or reprinted may not be up-to-date when used unless the copier or printer has verified the version to be up-to-date and affirms that, and 2) users participating in Joint Commission accreditation, including vendors, are required to update their software and associated documentation based on the published manual production timelines.

Ad.7 Disclaimers:

Ad.8 Additional Information/Comments: