Severe Obstetric Complications Electronic Clinical Quality Measure (eCQM) Methodology Report

Version 2.0

December 2022

Prepared by:

Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation

This material was prepared by Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (YNHHSC/CORE), under contract to the Centers for Medicare & Medicaid Services (CMS), an agency of the U.S. Department of Health and Human Services. The contents presented do not necessarily reflect CMS policy.

Table of Contents

List of Tables

Yale New Haven Health Services Corporation - Center for Outcomes Research and Evaluation (YNHHSC/CORE) Project Team

Kathleen M.B. Balestracci, PhD, MSW* – Division Lead, Hospital Research and Development
Valery A. Danilack, MPH, PhD* – Project Lead
Grace Glennon, MS, RD – Project Lead
Andrea Barbo Barthel, MS – Lead Analyst
Tahmidul Islam, PhD* – Analyst
Sarah Attanasio, MS – Project Coordinator
Nicole Walton, BS – Research Support
Emma Turchick, MPH – Research Support
Kerry McDowell, MS, MPhil, MSEd – Project Manager
Lisa G. Suter, MD* – Project Director and Contract Director, Quality Measurement Program
Harlan M. Krumholz, MD, SM* - Principal Investigator
*Yale School of Medicine

The Joint Commission Project Team

JohnMarc Alban, MS, RN, CPHMS – Associate Director
Susan Yendro, MSN, RN – Project Associate Director, Project Team Lead
Stephen Schmaltz, MPH, PhD – Associate Director and Senior Biostatistician
Sheila Aguilar, MBA – Senior Research Associate, Project Coordinator
Jennifer Hurlburt, DNP, MSN, BSN – Associate Director, Department of Standards and Survey Methods, Subject Matter Expert
Christine Walas, MSN, RN– Managing Project Director, Project Clinical Lead
Marilyn Parenzan, MBA, RHIA, CPHQ – Associate Project Director, Project eCQM Lead

Acknowledgements

We would like to acknowledge the support and input from former CORE Project Team members: Rajvi Shah, Fior Rodriguez, Rachelle Zribi, and Jacqueline Grady; and former Joint Commission Project Team members: Tricia Elliott and Mia Nievera. The authors also gratefully acknowledge the individuals that provided feedback, experiences, and expert input on the development of this measure: our Technical Expert Panel (TEP) members; our Patient Working Group; our clinical experts, Elliott Main, MD, Katherine Nash, MD, and Mona Sharifi, MD, MPH; the Person and Family Engagement team; Ada Ezennia, MPH, and James Poyer, MS, MBS, formerly of the Centers for Medicare & Medicaid Services (CMS); and Annese Abdullah-Mclaughlin and Raquel Myers (our Contracting Officer Representative), of CMS. (Please note that individual acknowledgements for TEP and Patient Working members can be found in Appendix A.)

Disclaimer: The views, thoughts, and opinions expressed in this report belong solely to the author, and not necessarily to any contributors or consultants, including the Patient Working Group and TEP members and their affiliated organizations. Acknowledgement of input does not imply endorsement of the methodology and policy decisions.

Executive Summary

Measure Background

The United States experiences higher rates of maternal morbidity and mortality than most other developed countries, and rates have trended upward in recent decades.[1] There is national interest across maternal health advocacy organizations, payors, and the public to evaluate hospital performance and improve maternal morbidity and mortality rates, and a need to provide timely and accurate data to inform hospital improvement efforts and patient decision-making. The broad availability of electronic health record (EHR) data presents an opportunity to measure maternal complication rates that cannot be fully measured using claims data alone.

The Centers for Medicare & Medicaid Services (CMS) contracted with Yale New Haven Health Services Corporation - Center for Outcomes Research and Evaluation (CORE) to support The Joint Commission (TJC) in the development of an EHR-based outcome measure of maternal morbidity and mortality. The goal for this measure is to assess the occurrence of specific severe obstetric complications in the hospital setting by using a methodology that reliably allows comparison across hospitals. Reduction in maternal complications will reduce maternal death and disability and improve maternal quality of life. The Severe Obstetric Complications electronic clinical quality measure (eCQM) is expected to inform hospital efforts to improve maternal health outcomes and thus reduce the costs associated with adverse health outcomes. We sought to keep measure specifications harmonized with other Joint Commission perinatal measures (for cohort alignment) and with the Center for Disease Control and Prevention's (CDC's) 21 indicators of severe maternal morbidity (SMM) (for measure outcome alignment) to minimize burden and for broad applicability across hospitals.

Measure Development

This report describes our approach to the development of the Severe Obstetrics Complications eCQM. We vetted measure decisions through multiple stakeholder groups, including a Technical Expert Panel (TEP), clinical expert consultants, and a Patient Working Group. In this report, we outline the approach to development, and provide detailed measure specifications for this eCQM. We describe the process and results of testing of this eCQM, which was conducted in three phases, across multiple hospitals with a variety of EHR systems.

Measure Specifications Summary

Data Sources: The Severe Obstetric Complications eCQM primarily uses electronic health record data, and data from other electronic clinical systems, depending on hospital site workflows, to define all components of the measure, including the measure denominator, measure numerator, risk adjustment variables, and candidate stratification variables.

Measure Cohort: The measure cohort for this eCQM is drawn from the initial patient population (IPP), defined as all inpatient hospitalizations for patients greater than or equal to eight years and less than 65 years of age who undergo a delivery procedure with a discharge date during the measurement period.

The measure cohort, or denominator, is further defined as patients in the IPP who are greater than or equal to 20 weeks, zero days gestation at the time of delivery. Patients are excluded if they have a confirmed diagnosis of COVID with a COVID-related respiratory condition or if they have a confirmed diagnosis of COVID and undergo a COVID-related respiratory procedure.

Measure Outcome: The measure outcome for this eCQM, severe obstetric complications during the delivery hospitalization, is based on the Centers for Disease Control and Prevention (CDC) definition of severe maternal morbidity (SMM), consisting of 21 indicators of SMM defined using International Classification of Diseases, Tenth Revisions (ICD-10) diagnosis and procedure codes. The numerator also includes patients who expire (die) during the inpatient encounter.

A second measure outcome is defined as severe obstetric complications (as defined above) excluding delivery hospitalizations for which blood transfusion is the only numerator event. Blood transfusions, generally in response to excessive bleeding around delivery, account for the greatest proportion of patients identified as having an obstetric complication, but patients for whom this is the only identified numerator event may represent a less severe outcome experience.

Risk Adjustment: The Severe Obstetric Complications eCQM is a risk-adjusted measure. Candidate risk variables of SMM or maternal mortality for consideration in the measure risk adjustment model were identified in literature and with input from clinical experts. Following the identification of risk-adjustment variables, a risk model was developed for both outcomes. The same variables are included in the risk models for severe obstetric complications and severe obstetric complications excluding blood transfusion-only encounters; however, due to very low prevalence of a few risk variables in the risk model of severe obstetric complication excluding transfusion-only encounters, some risk factors have been grouped. Complications that arise during the hospitalization are not used in risk adjustment.

Measure Testing: Alpha and Beta testing conducted for this eCQM included feasibility, reliability and validity testing of data elements and measure outcomes. For Alpha testing, virtual EHR walkthroughs were conducted with nine healthcare sites consisting of 27 individual hospitals, representing three different EHR systems. Alpha testing included assessment of clinical and documentation workflows compared to measure intent, assessment of data element availability and accuracy, and assessment of use of data standards.

For Stage 1 Beta testing, the Measure Authoring Tool (MAT) specifications were tested using data from eight healthcare sites and 25 hospitals, representing three different EHR systems, to further establish the feasibility and validity of each of the data elements as well as the validity of the outcome. Data were pulled for delivery hospital encounters discharged from January 1, 2020, to December 31, 2020**.** The accuracy of the data extracted from the EHR using the MAT specifications was assessed by comparing the data values to those identified in the medical record during clinical medical record review.

For Stage 2 Beta testing, data from five additional hospitals representing two EHR systems were recruited to test the measure specifications and measure logic, to further assess the feasibility of data elements required for the measure calculation, and to adjudicate the presence of conditions indicative of severe obstetric complication in the medical record. Data were pulled for delivery hospital encounters

discharged from January 1, 2019, to December 31, 2020, for four of the five hospitals; one of the five hospitals provided data for February 1, 2020 through June 30, 2021**.**

Measure Testing: Alpha testing revealed high data element feasibility, with a rate of 98% for final measure specifications, and Stage 1 Beta testing revealed an overall data element agreement rate of 90.4%. Final measure score results for all 30 Beta testing hospitals (Stage 1 and Stage 2 Beta testing) indicated high reliability, with a median site reliability score of 0.958 (range: 0.792 – 0.996) for the outcome measuring any severe obstetric complication and 0.918 (range: 0.652 – 0.992) for the outcome measuring severe obstetric complications excluding blood transfusion-only cases. Reliability increased when tested among hospitals with at least 200 deliveries in the measurement period: a median site reliability score of 0.968 (range: 0.860 – 0.996) for any severe obstetric complication and 0.937 (range: 0.751 – 0.992) for severe obstetric complications excluding blood transfusion-only cases. Likewise, clinical adjudication of EHR data to identify encounters with severe obstetric complications using medical chart review during Stage 1 Beta testing (six sites) and during Stage 2 Beta testing (five hospitals) revealed high validity: positive predictive value (PPV) was 94.74% (Stage 1) and 98.91% (Stage 2), and negative predictive value (NPV) was 100% (Stage 1) and 95.53% (Stage 2). The measure outcome agreement rates and kappa scores indicate overall 91.2% agreement with a kappa score 0.881, indicating very good agreement.

1. Measure Introduction

1.1 Measure Overview

The Centers for Medicare & Medicaid Services (CMS) contracted with Yale New Haven Health Services Corporation - Center for Outcomes Research and Evaluation (CORE) to support The Joint Commission in the development of an electronic health record (EHR)-based outcome measure of maternal morbidity and mortality. This measure, the Severe Obstetric Complications electronic clinical quality measure (eCQM), reflects a collaborative effort, from finalization of measure specifications through measure testing and completion.

The United States experiences higher rates of maternal morbidity and mortality than most other developed countries. These rates have continued to trend upward in recent decades.[1] Research indicates that the overall rate of severe maternal morbidity (SMM) has increased by almost 200% between 1993 and 2014 to 144 per 10,000 delivery hospitalizations[1], with more than 25,000 women per year experiencing obstetric complications.[2] Recent maternal mortality data from 2018 reveal that 658 women died from maternal causes, resulting in a rate of 17.4 deaths per 100,000 live births, with 77% of the deaths attributed to direct obstetric causes like hemorrhage, preeclampsia, obstetric embolism, and other complications.[3] This has prompted national health experts and organizations to prioritize quality improvement strategies to mitigate risk of adverse outcomes among maternal populations. The U.S. Department of Health & Human Services (HHS) has also called for action to improve maternal health and outcomes and outlines seven actions for healthcare professionals, including participating in quality improvement and safety initiatives.[4] There are currently only a small number of quality measures focused on maternal health, and those implemented at the national level are mostly process measures

and limited in scope. While these existing measures aim to promote coordination of care and standardize health care processes, maternal health outcome measures are sorely needed. Measures that are focused on maternal health outcomes will address the patient safety priority area under the Meaningful Measures 2.0 framework, and likewise will use EHR data to address interoperability, another meaningful measure area for assessing quality of health care.[5]

Our goal was to develop a reliable outcome-based eCQM to evaluate hospital-level quality of maternal care for patients who were hospitalized for delivery. This measure uses EHR data captured during the delivery hospitalization for an all-payer population. Utilizing EHR data for quality improvement and measurement efforts has several advantages compared to claims data alone, because the data tend to be clinically rich and produced in real time.[6] The Severe Obstetric Complications eCQM is the first hospital quality measure of maternal morbidity and mortality developed for national reporting.

This methodology report includes comprehensive information on the measure development approach, specifications, and testing results of the Severe Obstetric Complications eCQM. CORE convened a Technical Expert Panel (TEP) and a Patient Working Group and consulted with a clinical expert to provide input and expertise throughout the development of this eCQM.

1.2 Key Terminology

Key terms utilized throughout this report include the following:

- Severe Maternal Morbidity (SMM) – "unexpected outcomes of labor and delivery that result in significant short- or long-term consequences to a woman's health" (American College of Obstetricians and Gynecologists and the Society for Maternal-Fetal Medicine).[7] Note the outcome for this measure includes both SMM and maternal mortality occurring during the delivery hospitalization.
- Maternal Mortality – defined as the death of a pregnant woman "irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes" (from the World Health Organization definition).[8] Many definitions also include a time period relative to pregnancy (e.g., within 42 days, within one year).[8,9] For the purposes of this measure, we focus on death that occurs during the delivery hospitalization.
- Healthcare Disparity – defined as "differences in the quality of care that are not due to access-related factors or clinical needs, preferences, and appropriateness of interventions" (National Quality Forum).[10]
- Electronic Clinical Quality Measure (eCQM) – A measure that "uses data electronically extracted from electronic health records (EHRs) and/or health information technology systems to measure the quality of health care provided" (The Office of the National Coordinator for Health Information Technology [ONC]).[11]
- Electronic Health Record (EHR) – "A digital version of a patient's paper chart. EHRs are real-time, patient-centered records that make information available instantly and securely to authorized users" (ONC).[12]

Other terminology used to describe specifications for the Severe Obstetric Complications eCQM can be found in the Glossary of this report (Appendix B).

1.3 Severe Obstetric Complications as a Measure of Quality

1.3.1 Importance

Maternal morbidity and mortality pose serious health threats to pregnant women in the United States, where rates have been on the rise in comparison to other developed nations.[13] Recent data indicate a rate of 17.4 maternal deaths per 100,000 live births[3], and SMM occurring in 144 out of 10,000 delivery hospitalizations.[1] Hemorrhage, hypertensive disorders of pregnancy (HDP), sepsis/infection, cardiovascular conditions, cardiomyopathy, embolism, and mental health conditions have been identified as overall leading causes of peripartum death.[14] Nearly 16% of pregnancy-related deaths can be attributed to cardiovascular conditions.[15] The Centers for Disease Control and Prevention (CDC) report significant increases in SMM events since 1993.[16] The CDC specifically defines SMM by 21 indicators, defined by International Classification of Diseases, Tenth Revision (ICD-10) diagnosis and procedure codes.[17] The top SMM indicators include blood transfusions, which occurred in 122.3 per 10,000 delivery hospitalizations in 2014 and resulted in a substantial 399% rate increase from 1993 to 2014. Acute renal failure, another identified SMM indicator, has steadily increased over the years, with a 300% rate increase from 1993 to 2014. Other events identified among the CDC's SMM indicators with increasing rates over this period include adult respiratory distress syndrome with a rate increase of 205%; cardiac arrest, fibrillation, or conversion of cardiac rhythm with a 175% rate increase; and shock with a 173% increase.[16] The consequences of maternal morbidity are well documented; not only are these conditions leading causes of pregnancy-related death, but often lead to further pregnancy complications and other SMM conditions.[18,19]

The costs associated with delivery complications are high. Investigators evaluating costs for women with a live inpatient birth in 2013 calculated a 37% increase in delivery hospitalization costs for women experiencing SMM over those without SMM among commercially insured women ($20,380 versus $14,840), and a 47% increase in delivery costs for women experiencing SMM over those without SMM among women insured with Medicaid ($10,134 versus $6,894).[20] The differential in costs was even higher in two studies using the Agency for Healthcare Research and Quality's (AHRQ's) Healthcare Cost and Utilization Project (HCUP) National Inpatient Sample. These studies, one using 2011 to 2012 data[21] and the other using 2012 to 2014 data[22], calculated average risk-adjusted hospital costs (not including physician costs) for SMM during delivery hospitalizations at over two times greater for patients with any SMM compared to patients with no SMM, 5.5 times the cost if the patient had two or more SMM events[22], and over 10 times the cost with five or more SMM events[21]. Costs are incurred due to the treatment required by obstetric complications and the impact on hospital lengths of stay; Premier's Bundle of Joy™ Report (2019) found that women with SMM delivering vaginally have hospital stays that are 70% longer than women with vaginal deliveries experiencing no SMM, and costs that are almost 80 percent higher.[23]

Lastly, there are considerable racial and ethnic disparities in maternal outcomes. Historically marginalized women of color are at a significantly higher risk for developing severe maternal complications compared to non-Hispanic White women.[24] Non-Hispanic Black women are three to four times more likely to die from pregnancy-related causes than non-Hispanic White women.[25] Non-Hispanic American Indian/Alaska Native (AI/AN) women have the second highest pregnancy-related mortality ratio compared to non-Hispanic White, Asian/Pacific Islander, and Hispanic women.[25] Non-Hispanic Black women experience higher mortality from cardiomyopathy and cardiovascular conditions, while AI/AN women have an increased risk of death due to hemorrhage and hypertensive disorders.[26] Based on SMM defined using the 21 indicators identified by the CDC, Black, Hispanic, Asian/Pacific Islander, and AI/AN women had 2.1, 1.3, 1.2, and 1.7 times higher rates of severe morbidity, respectively, compared with White women in data from seven states.[27]

1.3.2 Performance and Preventability

The high maternal mortality and morbidity rates in the United States present unique opportunities for large-scale quality measurement and improvement activities. Statistics on preventability vary but suggest that a considerable proportion of maternal mortality and morbidity events could be prevented. A 2019 report from 14 maternal mortality review committees conducting a thorough review of pregnancy-related deaths determined that 65.8% were preventable (data from 14 U.S. Maternal Mortality Review Committees, 2008-2017).[14] Additionally, a study that examined preventability of pregnancy-related death, women with near-miss morbidity, and those with severe morbidity found that 40.5% of deaths, 45.5% of near miss morbidity, and 16.7% of other severe morbidities were preventable.[28] Study investigators identified areas of focus for preventability of morbidity and mortality that included assessment/point of entry to care, diagnosis and recognition of high risk, referral to experts, treatment, management hierarchy, education, communication, policies and procedures, documentation, and discharge.

Although there are limited measures to assess variability among hospitals, rates in the United States are higher than all other developed countries, presenting opportunity for improvement. USA Today's database of childbirth complication rates at maternity hospitals, with data from 1,027 hospitals in 13 states from 2014-2017, showed marked variation in median rates of childbirth complications. Using the CDC definition of SMM, the US median rate was 1.4%, whereas the highest hospital rate was 12.2%. [29] This variability may reflect similar trends for maternal complications.

Maternal morbidity has garnered much national attention, with a broad range of SMM events and outcomes that can be examined, many of which are closely associated with mortality.[15,30] Several initiatives have shown promise in reducing maternal morbidity events. For example, following the inception of the California Maternal Quality Care Collaborative (CMQCC), which focused on metrics and toolkits to improve maternal outcomes, the maternal mortality rate in California declined by 55% between 2006 and 2013.[31] The CMQCC obstetric hemorrhage collaborative resulted in a 20.8% reduction in SMM in California hospitals compared with the 1.2% reduction in SMM among nonparticipating hospitals.[30] The state of California has established a successful framework for assessing

and improving quality of maternal care, and outcomes suggest great potential for nationally reducing maternal care complications.

1.3.3 Measurement Gap

National evaluation of hospitals' performance on maternal morbidity and mortality is limited because there are currently no maternal morbidity or obstetric complications outcome measures in national reporting programs. Current quality measures related to pregnancy and maternal health proposed for or in public reporting programs are largely process measures (e.g., Maternity Care: Post-partum Follow Up and Care Coordination) and outcome measures related to delivery type (e.g., PC-01 Elective Delivery).

There are numerous state agencies, private and/or non-profit organizations, and collaboratives that have spearheaded maternal health and quality improvement initiatives. For instance, the Alliance for Innovation in Maternal Health (AIM) developed evidence-based patient safety bundles to address leading causes of SMM, like obstetric hemorrhage and hypertension. The CDC Perinatal Collaboratives also support various state-based efforts to promote high quality maternal care. The CMQCC created the Maternal Data Center (MDC) for hospitals with Labor and Delivery units in California, Oregon, and Washington. The MDC is an online tool that receives patient discharge data on maternity care services, links these data to birth certificate or clinical data, and provides clinicians with perinatal performance data for supporting quality improvement.[32] The MDC allows hospital performance regional and statewide comparisons. Overall, such quality metrics do not currently cater to a national population because there is extensive variation and timing delays in the widespread adoption and implementation of safety protocols in obstetric care across states.[30,33] Moreover, data examining the nationwide implementation of these resources are not widely available.[30,34] Therefore, the development of a obstetric complications outcome measure addresses a national measurement gap and can build on learning from existing maternal health initiatives and measures.

1.3.4 Feasibility and Usability

State and national initiatives to measure, track, and reduce maternal morbidity and mortality have produced encouraging results. The Severe Obstetric Complications eCQM could expand these improvements in care, outcomes, and cost savings at a national level. This eCQM will provide hospitals with benchmarking and actionable data to inform their quality improvement efforts; the use of EHR data will provide them with the potential to repurpose the data and measure logic for internal quality control using real-time feedback to further mitigate harm to patients. Additionally, the eCQM can provide information that allows patients to compare hospitals' performance to aid in their decision-making when choosing care.

Although efforts may require hospitals to initially invest resources to support measure reporting, we anticipate that such investments will help them more fully utilize their EHRs to improve care for pregnant patients, which is a shared goal among stakeholders. Using EHR data instead of administrative data allows for more patient-centric, potentially real-time measure results to support hospital quality improvement efforts.[6,35,36]

However, using data from the EHR is only the first step to securing accurate and reliable data for measuring severe obstetric complications. The quality of our measure results depends on the reliability of the data extracted from structured fields in the EHR. To reduce hospital burden, we aimed to build a measure based on data in structured fields that are consistently captured during clinical care. We did not use data that might have required natural language processing prior to measure calculation. The only data manipulation required is calculation of gestational age at delivery. During measure testing, we tested the feasibility and validity of data elements required to determine the measure cohort, the outcome, and risk adjustment. Additionally, we adjudicated outcomes to ensure that the electronically specified definition correlated with the actual occurrence of a severe obstetric complication, according to clinical adjudication of the medical record.

Our goal was to build an eCQM that does not require changes in clinical workflow and for which the electronic specifications are easy to understand and implement.

1.4 Measure Use

This is a *de novo* eCQM intended to measure inpatient acute care hospital quality and performance related to severe obstetric complications and death during the delivery hospitalization. The measure is intended to be used alongside existing perinatal process of care quality measures and existing quality improvement efforts focused on reducing maternal morbidity and mortality.

1.5 Approach to Measure Development

The goal of the Severe Obstetric Complications eCQM is to assess prevalence of SMM and mortality during hospital delivery encounters for an all-payer population based on EHR data. We began by assessing the critical drivers of maternal morbidity and mortality, health disparities, and risk adjustment variables through an environmental scan and literature review (ES/LR). We then drafted Measure Authoring Tool (MAT) specifications, value sets, and a testing plan. To develop preliminary specifications, we built on prior published specifications when available. It is important to note that a standard and consistent definition for maternal morbidity and mortality is currently lacking; existing definitions vary in scope and in the time frame during which SMM or maternal death is captured.[8,9,15] For this measure, measure specifications are modeled after the nationally available and adopted CDC definition for SMM, which encompasses "unexpected outcomes of labor and delivery that result in significant short- or long-term consequences to a woman's health".[7] We also solicited input from clinicians and a diverse group of stakeholders throughout the development process; specifications were developed with input from a TEP and Patient Working Group. Our goal was to ensure usability by keeping specifications feasible and straightforward.

Development testing included Alpha testing and two stages of Beta testing. Alpha testing consisted of virtual EHR walkthroughs with recruited hospitals to assess feasibility of the data elements necessary to define the measure specifications. Beta testing consisted of testing the measure specifications in the MAT to further establish the feasibility and validity of each of the data elements as well as validity of the severe obstetric complications composite outcome. The accuracy of the data extracted from the EHR, and the identification of severe obstetric complications were assessed through medical record

abstraction. In Stage 1 Beta testing we examined data element feasibility issues and numerator validation; findings informed updates to the measure specifications. In Stage 2 Beta testing, conducted with data from additional hospitals, we tested the measure specifications and further validated measure results.

1.5.1 Information Gathering

CORE initially conducted an ES/LR on maternal morbidity and mortality to inform the development of a maternal health eCQM, and subsequently conducted focused literature reviews on three common maternal morbidity events often associated with mortality: obstetric hemorrhage, maternal hypertension and preeclampsia, and maternal infection and sepsis.

In parallel, TJC identified through work on the Unexpected Complications in Term Newborn measure that there was need for a similar measure for maternal care. A broad environmental scan and literature review was conducted on the topic of maternal complications.

These literature reviews served to gather evidence on the prevalence, health consequences, and evidence of preventability of various maternal morbidity events and how they might be measured based on clinical research, prior measurement efforts, and clinical guidelines. Methods to measure maternal morbidity outcomes through extraction of data from the EHR and through chart review for clinical adjudication were explored. These reviews informed eCQM specification considerations for measurement of severe obstetric complications.

The environmental scans served to identify existing related or competing quality measures addressing maternal morbidity and mortality overall and measures specific to obstetric complications. An online scan of both pre-specified websites and search engines was conducted to identify existing quality measures related to maternal morbidity outcomes using electronic and other medical record systems, cross-checked against maternal health measure inventories provided by CMS. Websites were searched using keywords for pregnancy and maternity complications in combination with keywords reflecting the 21 SMM indicators used by the CDC to operationally define SMM. We supplemented this search via Google search engine using the following keywords: maternal morbidity and mortality measure, maternal morbidity measure, maternal mortality measure.

Ultimately, these literature reviews and environmental scans, in addition to discussions with key stakeholders led by TJC, led us to leverage the existing CDC indicators[17] and The American College of Obstetricians and Gynecologists' (ACOG) detailed list of ICD-10 codes to identify SMM[37] as a foundation for measure development.

In addition, literature revealed the importance of risk adjustment for this patient population. Literature was used to identify common risk factors for SMM[38-41] and risk prediction for SMM to help identify potential risk variables for this eCQM[30,42] through the EHR.

1.5.2 Expert and Stakeholder Input

Expert and stakeholder input for the development of this measure was sought from a TEP, a Patient Working Group, and ongoing consultation with Dr. Elliott K. Main, the Medical Director at CMQCC, Clinical Professor of Obstetrics and Gynecology at Stanford University and a nationally recognized expert and leader in maternal health outcomes measurement. The TEP was composed of 17 members (16 members initially, with an additional member replacing a departing member in 2021), including several individuals who had served on TJC's Technical Advisory Panel supporting the development of their perinatal care measures. Members brought expertise in quality improvement, electronic capture of medical information, healthcare disparities, obstetrics and gynecology, and the patient perspective. TEP members were nominated or nominated themselves to participate in this stakeholder group. The members were engaged during key development milestones.

The first TEP meeting was held in person in February 2020 in Baltimore, MD, during which TEP members provided input on draft measure specifications for the measure cohort, outcome, and risk adjustment. The second TEP meeting was held via a web-based webinar in July 2021, during which TEP members provided input on Alpha testing and feasibility results, initial Beta testing results, and proposed updated measure specifications. At the third TEP meeting, a web-based webinar held in November 2021, TEP members provided input on the risk adjustment model, measure scores, and further testing results.

To gain targeted input from the patient and caregiver perspective, a Patient Working Group was recruited through collaboration with Rainmakers Strategic Solutions LLC. The Patient Working Group was composed of seven members, including patients and caregivers with diverse experiences and perspectives. The first Patient Working Group meeting was held in August 2020 via web-based webinar during which Patient Working Group members provided input on initial measure specifications for the measure cohort, outcome, and risk adjustment. The second meeting was held in July 2021 via web-based webinar, at which Patient Working Group members provided input on measure specification updates, Alpha testing and feasibility results, and initial Beta testing results. At the third meeting, a web-based webinar held in November 2021, Patient Working Group members provided input on the risk adjustment model, measure scores, and further testing results.

Dr. Main served as a clinical expert consultant, providing ongoing consultation for this work throughout measure development and testing. His clinical expertise and evidence he provided from prior research informed the development and evolution of the measure specifications.

2. Methods

2.1 Overview

The Severe Obstetric Complications eCQM captures SMM events and in-hospital mortality extracted from the EHR to assess quality of maternal care in the hospital setting for an all-payer population. The measure identifies ICD-10 codes consistent with CDC's 21 SMM indicators, as well as death, to define the outcome. The initial patient population was built upon existing specifications from the PC-01 Elective Delivery and PC-02 Cesarean Birth eCQMs[43] developed by TJC. Measure specification definitions, including risk variable decisions, were informed by published research,[30,42] expert clinical input from Dr. Main and members of the TEP, and valuable patient experience narratives from Patient

Working Group members. We partnered with hospitals and qualified vendors to evaluate feasibility, reliability, and validity of clinical data and measure logic.

Many of the data elements within the measure specifications are defined by ICD-10 diagnosis and procedure codes (value sets for the numerator, denominator, and risk adjustment are listed in Appendix C). Additional work has been done to map Systematized Nomenclature of Medicine (SNOMED) codes consistent with delivery encounters, the CDC's 21 SMM indicators, and risk variables in the measure specifications. Although SNOMED codes are available for clinical data capture in the EHR, we found that hospitals participating in the testing of this measure chose to submit ICD-10 codes rather than SNOMED codes for almost all data elements. Therefore, at this time, only select data elements using SNOMED codes are included in the measure logic, as follows: blood transfusion, COVID-19 infection, delivery procedures, patient expired, ED visit and OB triage, inpatient encounter, and observation services. SNOMED codes representing numerator events and risk factors have been captured in value sets, and inclusion of these SNOMED codes in measure specifications will be considered during reevaluation. We believe that including both ICD-10 and SNOMED codes to define these data elements in the future will allow for inclusivity and flexibility for this measure.

Measure testing included Alpha and two stages of Beta testing, described below:

- Alpha Testing: Alpha testing was conducted via virtual EHR walkthroughs with recruited hospitals to confirm preliminary feasibility of documentation and data elements necessary to define the measure. Alpha testing was conducted in three different EHR systems.
- Beta Testing: Testing of the MAT output was conducted with recruited hospitals. The MAT output describes the measure logic and value sets associated with each required data element; testing was conducted to further establish the feasibility and validity of each of the data elements as well as the validity of the Severe Obstetric Complications eCQM outcomes. In Stage 1 Beta testing, conducted with 8 health systems consisting of 25 hospitals, we determined the accuracy of the data extracted from the EHR using the MAT specifications by comparing the data values to values identified through medical record review. Results informed updates to the measure specifications, including the removal of trauma codes initially identified for denominator exclusion and numerator definitions initially considered in addition to the CDC 21 SMM indicators. In Stage 2 Beta testing, five additional hospitals were recruited, and the measure specifications and measure logic were tested. We confirmed the accuracy of the outcome through clinical medical record review. Beta Testing was conducted in three different EHR systems.

2.2 Data Sources

The Severe Obstetric Complications eCQM primarily uses electronic health record data and data from other electronic clinical systems, depending on hospital site workflows, to define all components of the measure, including the measure denominator, measure numerator, risk adjustment variables, and candidate stratification variables.

For Alpha testing, virtual EHR walkthroughs were conducted with nine healthcare sites consisting of 27 individual hospitals, representing three different EHR systems, including Epic, Cerner, and Meditech. The EHR walkthroughs included EHR experts, report writers, and clinical leads to assess feasibility of the data elements necessary to define the measure specifications. Alpha testing included assessment of clinical and documentation workflows compared to measure intent, assessment of data element availability and accuracy, and assessment of use of data standards. A feasibility scorecard was completed for each healthcare site.

For Stage 1 Beta testing, the MAT specifications were tested using data from eight healthcare sites and 25 hospitals, representing Epic, Cerner, and Meditech EHR systems, to further establish the feasibility and validity of each of the data elements as well as the validity of the outcome. Data were pulled for delivery hospital encounters discharged from January 1, 2020, to December 31, 2020**.** The accuracy of the data extracted from the EHR using the MAT specifications was assessed by comparing the data values to those identified in the medical record during clinical medical record review.

For Stage 2 Beta testing, data from five additional hospitals using Epic and Meditech EHR systems were pulled to test the measure specifications and measure logic, to further assess the feasibility of data elements required for the measure calculation, and to adjudicate the presence of conditions indicative of severe obstetric complications in the medical record. For four of the five hospitals, data were pulled for delivery hospital encounters discharged from January 1, 2019, to December 31, 2020**.** However, one of the five hospitals pulled data from February 1, 2020, to June 30, 2021.

When data from Stage 1 and Stage 2 hospitals were combined for select analyses of all 30 hospitals, data from January 1, 2020 to December 31, 2020 were used for all hospitals.

2.2.1 Limitations

While rates of maternal morbidity and mortality have continued to trend upward in the U.S. in recent decades[1], severe maternal morbidity is a relatively rare outcome. As defined with 22 numerator definitions (21 SMM indicators as identified by the CDC and mortality), SMM requires a substantial sample size for testing. For this reason, eight sites representing 25 hospitals were included for initial Beta testing, and an additional five hospitals were identified for subsequent Beta testing. As testing results have revealed low frequencies for some of the numerator definitions, future testing in reevaluation will be important for assessing measure specifications.

As noted in Section 2.1, only select data elements using SNOMED codes are included in the measure logic; hospitals participating in the testing submitted ICD-10 codes rather than SNOMED codes for almost all data elements, precluding testing of many of the data elements using these codes. When SNOMED codes are more readily used across healthcare settings, an update to the measure specifications to implement SNOMED code value sets and timing logic can be tested for future implementation.

2.2.2 Missing Data

We developed this eCQM with the intent to use variables expected to be consistently obtained for the target population, available in a structured field, and captured as part of standard clinical workflow. During Alpha testing, data elements were evaluated for feasibility and availability; two data elements were removed from measure specifications when several test sites were unable to accurately capture them (timestamp for procedure performed, and lab result for PaO2/FiO2). All other data elements were assessed to be feasible and available.

Many of the data elements used in the Severe Obstetric Complications eCQM are defined with ICD-10 diagnosis or procedure codes (for example, severe maternal mortality numerator events and risk adjustment variables). None of these data elements are considered missing when absent, since the absence of a given code implies absence of the corresponding condition.

For data elements representing vital signs and lab results, it is clinically acceptable that certain vital signs and labs were not performed for certain patients. However, vital sign and lab result fields with more than 20% missing were not considered as potential risk adjustment variables. For vital sign and lab results included as risk adjustment variables, values were categorized and a separate category for missing data was included. No imputation was performed.

2.2.3 Generalizability

Hospital recruitment for participation in testing was aimed at gathering test data from a diversity of settings and multiple EHR systems. The 28 hospitals used in Alpha and Stage 1 Beta testing (27 represented in Alpha testing, 25 represented in Stage 1 Beta testing) across 10 healthcare sites were located in 11 states. Twenty-five hospitals were urban and three were rural. Three were operated by not-for-profit faith-based organizations, 24 were other not-for-profit hospitals, and one was government (county) owned. Three of the 28 hospitals were primarily obstetrics and gynecology hospitals. Total births per hospital per year ranged from 150 to 8,800, with four hospitals with fewer than 500 births, 6 hospitals with 500-999 births, 11 hospitals with 1000-4999 births, and four hospitals with greater than 5000 births (three hospitals did not report these data). Three EHR systems were utilized across these hospitals: Epic, Meditech, and Cerner (see Section 3.2).

The five hospitals used in Stage 2 Beta testing were located in three states. Four hospitals were urban, and one was rural. Two of the five hospitals were teaching hospitals and had obstetrics and gynecology residencies. Total births per year ranged from 400 to 3,600, with one hospital with fewer than 500 births, two hospitals with 500-999 births, and two hospitals with 1,000-4,999 births. Two EHR systems were utilizes across these hospitals: Epic and Meditech (see Section 3.2).

However, given that this was neither a national nor a randomized sample, we recommend further testing in reevaluation to monitor measure specifications and update them as needed.

2.3 Measure Cohort (Denominator)

The measure cohort for this eCQM is drawn from the initial patient population (IPP), defined as all inpatient hospitalizations for patients greater than or equal to eight years and less than 65 years of age who undergo a delivery procedure with a discharge date during the measurement period. The measure

cohort, or denominator, is further defined as patients in the IPP who are greater than or equal to 20 weeks, zero days gestation at the time of delivery. The IPP is defined using delivery procedure codes from the EHR, and the measure denominator is further defined by gestation at the time of delivery. Patients with COVID-19 and respiratory complications are excluded from the denominator.

2.3.1 Inclusion Criteria

The measure includes all delivery hospitalizations for live births and stillbirths with ≥ 20 weeks 0 days gestation completed at delivery for patients greater than or equal to eight years and less than 65 years of age. The measure does not include delivery hospitalizations for patients with gestation less than 20 weeks.

Gestational age is defined by either measure logic calculating an estimated gestation age (EGA) or by EGA identified in a discrete field in the EHR. The EGA calculation in the measure logic uses the American College of Obstetricians and Gynecologists ReVITALize guidelines.[44] Gestational age = (280-(EDD minus Reference Date))/7 where the Estimated Due Date (EDD) is defined as: last menstrual period if confirmed by early ultrasound or no ultrasound performed, or early ultrasound if no known last menstrual period or the ultrasound is not consistent with last menstrual period, or known date of fertilization (e.g., assisted reproductive technology). The "Reference Date" is the date for which gestational age is being calculated. For purposes of this eCQM, "Reference Date" is the identified "Date of Delivery."

Rationale: This measure intends to include still and live births for patients of childbearing age. Patients delivering at less than 20 weeks' gestation represent miscarriages, where the fetus is not viable to survive.[45] The influence of the quality of hospital quality on miscarriages is less, making this population less applicable for hospital-based quality and performance improvement.

2.3.2 Exclusion Criteria

Following completion of Alpha and Beta testing, a denominator exclusion for patients with a COVID-19 diagnosis at admission who also had at least one diagnosis code for respiratory distress or a procedure code for a respiratory procedure was evaluated. Although rare, cases with this exclusion were found during analysis. The COVID-19 exclusion was added to ensure patients with this condition who are symptomatic with respiratory conditions would not be counted as a numerator case for hospitals.

Analyses outlining the impact of the exclusion on the measure results are found in Section 3.7. All other results presented are without the COVID-19 denominator exclusion applied.

Rationale: The evidence base for COVID-19 and related variants is rapidly growing and changing. Available studies suggest that symptomatic pregnant women with COVID-19 are at increased risk of more severe illness compared with nonpregnant peers.[46] Treatment protocols are being developed and tested and the measure should not include these patients while preventability of these complications is unknown.

2.4 Measure Outcome (Numerator)

The measure outcome (numerator) for this eCQM is based on the CDC definition of SMM (21 indicators) and uses ICD-10 to define diagnoses and procedures that are indicative of an SMM. ICD-10 codes are used for billing in hospitals and therefore are generally widely available and offer stability over time.[15] The numerator also includes patients who expire (die) during the inpatient delivery encounter.

The measure numerator is defined as the number of inpatient delivery hospitalizations in the denominator for patients who experience any of the following numerator events. Note that only diagnoses not present on admission are considered a numerator event.

- Severe maternal morbidity diagnoses and procedures
    - Acute myocardial infarction
    - Aortic aneurysm
    - Cardiac arrest/ventricular fibrillation
    - Heart failure/arrest during procedure or surgery
    - Disseminated intravascular coagulation
    - Shock
    - Acute renal failure
    - Adult respiratory distress syndrome
    - Pulmonary edema/Acute heart failure[1]
    - Sepsis
    - Air and thrombotic embolism
    - Amniotic fluid embolism
    - Eclampsia
    - Severe anesthesia complications
    - Puerperal cerebrovascular disorder
    - Sickle cell disease with crisis
    - Blood transfusion
    - Conversion of cardiac rhythm
    - Hysterectomy
    - Temporary tracheostomy
    - Ventilation
- Patients who expire (die) during the inpatient encounter

[1] CDC utilizes 21 indicators for defining SMM, but for the purposes of this measure's outcome, one of the indicators (Pulmonary edema/Acute heart failure) is defined using two distinct value sets. It is listed here as one indicator, but the value sets identify these as two distinct diagnoses. Likewise, the Measure Authoring Tool (MAT) header that supports this eCQM identifies these two diagnoses separately.

In addition to severe obstetric complications as defined above, the measure includes an additional outcome: severe obstetric complications excluding delivery hospitalizations for which blood transfusion is the only numerator event.

Rationale: We chose to align the severe obstetric complications outcome with the 21 diagnoses and procedures widely accepted as SMM as defined by CDC. Stakeholders supported alignment to ensure comparability of rates with other maternal morbidity reporting. We included death in this measure outcome because this critical outcome may occur in the absence of one of the defined severe obstetric complication events. We requested feedback from TEP and Patient Working Group members on these specifications, which, along with clinical input and testing, helped inform key decisions for the measure outcome definition.

In development, four additional numerator events were included for consideration in the measure outcome: 1) intensive care unit (ICU) stay > 12 hours during the delivery hospitalization, 2) platelet count < 100 10*3/uL, 3) serum creatinine >= 2 mg/dL, and 4) PaO2 < 60 mmHg. These four candidate numerator definitions were not included in the numerator after clinical adjudication during Stage 1 Beta testing revealed that: patients with ICU stay and patients with creatinine >= 2 mg/dL generally also met other numerator definitions; platelet count <100 10*3/uL alone did not identify severe obstetric complications; and PaO2 is not administered consistently in this population and is burdensome for providers to map in the EHR. In addition, specific concerns about hospitals who may not have ICUs and differential use of these units for patient care supported removal of this indicator in the numerator.

The second severe obstetric complications outcome excluding patients for whom blood transfusion is the only numerator event addresses input from some clinical experts about the level of severity blood transfusions represent. Blood transfusions generally performed in response to excessive bleeding around delivery, account for the greatest proportion of patients identified as having an obstetric complication, but patients for whom this is the only identified numerator event may represent a less severe outcome experience. This second outcome requires that patients who experience a blood transfusion during the delivery hospitalization also experience at least one other numerator event to be counted as having a severe obstetric complication.

2.5 Attribution

This Severe Obstetric Complications eCQM was developed as a hospital-level measure, with outcomes attributable to acute care settings, because deliveries most commonly occur in the acute inpatient setting.

2.6 Risk Adjustment

The goal of risk adjustment is to account for patient-level factors that are clinically relevant, have strong relationships with the outcome, and are outside of the control of the reporting entity, without obscuring important quality differences. Risk factors can increase (or decrease) the likelihood that a patient experiences a certain outcome.

Risk adjustment for case mix differences among hospitals is based on clinical status of the patient and other patient characteristics at the time of admission. Only conditions or comorbidities that convey information about the patient at the time of the admission are included in risk adjustment, determined by present on admission indicators. Complications that arise during the hospitalization are not used in risk adjustment.

We identified candidate risk variables predictive of SMM for consideration in the measure risk adjustment model by utilizing literature and research findings, including a study on a comorbidity scoring system by Leonard et al.[42], the NQF Maternal Morbidity and Mortality Environmental Scan[15], and our initial ES/LR findings on specific drivers of severe obstetric complications and maternal mortality. In addition, we identified candidate risk variables from the list of Hospital Core Clinical Data Elements[47], a set of 21 clinical variables from EHRs routinely collected for use in risk-adjusted hospital-level outcome measures; these data represent the first set of vital signs and basic laboratory tests collected from patients, reflecting a patient's clinical status as they present to the hospital prior to or during the process of inpatient admission. We sought input from the clinical expert consultant and other clinical experts to select data elements from this list that are routinely collected from patients admitted for a delivery hospitalization and could provide valuable clinical information regarding patient risk. We also solicited input from clinicians, patients, and other experts on the TEP who identified for consideration numerous risk-adjustment variables at the patient and hospital levels. These included, but were not limited to, prior pregnancy history, housing instability, and availability of specialists and trauma care in hospitals. The team acknowledged and carefully considered recommendations from the TEP and Patient Working Group for selection of candidate risk-adjustment variables.

Following the identification of risk-adjustment variables, a risk model was developed for the severe obstetric complications and severe obstetric complications excluding blood transfusion-only encounters. The risk model was developed and tested with data from the healthcare sites included in Stage 1 Beta testing; delivery hospitalizations were randomly divided in a 70/30 split for a development dataset and a validation dataset. We included variables in the model that were identified *a priori* as being clinically associated with severe obstetric complications. Risk variables were not excluded from the final model based on statistical considerations; we determined that sample size limitations of testing data should not prevent inclusion of risk variables with an evidence base that we expect could be statistically significant in data available from national measure reporting. However, three adjustments from the plan to include all variables were employed. First, vital sign and laboratory risk variables were removed from the model that had greater than 20% missing values. Second, due to a lack of variation across encounters, temperature and respiratory rate were not included in the final model. Third, due to very low prevalence of a few risk variables in the risk model of severe obstetric complications excluding transfusion-only encounters, Human Immunodeficiency Virus (HIV) was combined with autoimmune disease and obstetric venous thromboembolism (VTE) was combined with long-term anticoagulant medication use for the model of severe obstetric complications excluding transfusion-only encounters only. Otherwise, the same risk variables were included in the risk models for severe obstetric complications and severe obstetric complications excluding blood transfusion-only encounters.

The following variables were included in the final risk model:

- Patient demographics: maternal age (derived from birthdate)
- Preexisting conditions and pregnancy characteristics defined by ICD-10 codes:
    - Anemia
    - Asthma
    - Autoimmune disease
    - Bariatric surgery
    - Bleeding disorder
    - Body Mass Index (BMI) >= 40
    - Cardiac disease
    - Gastrointestinal disease
    - Gestational diabetes
    - Human Immunodeficiency Virus (HIV)
    - Hypertension
    - Mental health disorder
    - Multiple pregnancy
    - Neuromuscular disease
    - Obstetric venous thromboembolism (VTE)
    - Other pre-eclampsia
    - Placental accreta spectrum
    - Placental abruption
    - Placenta previa
    - Preexisting diabetes
    - Preterm birth
    - Previous cesarean
    - Pulmonary hypertension
    - Renal disease
    - Severe pre-eclampsia
    - Substance abuse
    - Thyrotoxicosis
- Laboratory tests and vital signs upon hospital arrival [first resulted value within 24 hours prior to initial encounter (earliest between inpatient admission, emergency department/obstetric triage, observation stay) and before delivery]: Hematocrit, White blood cell (WBC) count, Heart rate, Systolic blood pressure
- Long-term anticoagulant medication use
- Social Risk Factors: economic/housing instability

Given the changes made to the initial list of risk factors due to missing data and a small sample for one of the outcomes, we recommend reevaluating the components of the risk model in a larger dataset.

2.6.1 Social Determinants of Health

Our goal in selecting risk factors for adjustment was to develop a parsimonious model that included clinically relevant variables strongly associated with a severe obstetric complication. Social risk factors were considered dependent on the availability of information in the EHR. As noted above, economic/housing instability was included in the model, and was chosen due to support in research literature for its inclusion and availability in the EHR. [42]

Because of the stark differences in maternal outcomes by race/ethnicity as demonstrated in the literature, race/ethnicity were examined as stratification variables rather than risk variables. It was determined that illumination of outcome disparities by race/ethnicity, rather than adjustment of outcomes by race/ethnicity, would best inform stakeholders and patients and be most impactful in incentivizing improvements in the quality and equity of maternal care.

2.7 Statistical Approach to Model Development

Risk model performance was assessed by examining the model performance (C-statistics), model calibration (lack of fit), and model discrimination in terms of predictivity (range of observed outcomes among deciles of predicted outcomes). We calculated the model estimates as well as the coefficients, and adjusted odds ratios with 95% Confidence Intervals (CIs) for risk-adjustment variables. Calibration plots were created to assess the agreement between observed severe obstetric complication rates and the risk predicted by the risk model.

With the list of risk variables identified for the risk model, we estimated the hospital-specific risk standardized obstetric complications rate (RSOCR) using a hierarchical logistic regression model (hierarchical model). This strategy accounts for within-hospital correlation of the observed outcome among patients and accommodates the assumption that underlying differences in the quality of care across hospitals lead to systematic differences in patient outcomes. This approach models the log odds of a severe obstetric complication as a function of patient demographics and clinically relevant comorbidities with a random intercept for the hospital-specific effect.

2.8 Calculation of Measure Score

Hospital-level measure scores were calculated as a standardized proportion of the number of delivery hospitalizations for patients who experience a severe obstetric complication, as defined by the numerator, by the total number of delivery hospitalizations in the denominator during the measurement period. The hospital specific RSOCRs were calculated as the ratio of a hospital's "predicted" number of delivery hospitalizations with a severe obstetric complication to "expected" number of delivery hospitalizations with a severe obstetric complication, multiplied by the overall observed rate of delivery hospitalizations with a severe obstetric complication. The expected number of delivery hospitalizations with a complication for each hospital (denominator) was estimated using its patient mix and the average hospital-specific intercept (i.e., the average intercept among all hospitals in the sample). The predicted number of delivery hospitalizations with a complication for each hospital (numerator) was estimated given the same patient mix but an estimated hospital-specific intercept. Operationally, the expected number of delivery hospitalizations with a complication for each hospital was obtained by summing the expected complications for all delivering patients in the hospital. The

expected complications outcome for each delivering patient was calculated via the hierarchical model, which applies the estimated regression coefficients to the observed patient characteristics and adds the average of the hospital-specific intercept. The predicted number of delivery hospitalizations with a complication for each hospital was calculated by summing the predicted complications for all delivering patients in the hospital. The predicted complications outcome for each delivering patient was calculated through the hierarchical model, which applies the estimated regression coefficients to the patient characteristics observed and adds the hospital-specific intercept.

More specifically, we used a hierarchical model to account for the natural clustering of observations within hospitals. The model employs a logit link function to link the risk factors to the outcome with a hospital-specific random effect:

Let Y_{ij} denote the outcome (equal to one if the delivery encounter has a severe obstetric complication, zero otherwise) for patient *i* at hospital *j*; $\boldsymbol{Z}_{ij}$ denotes a set of risk factors for patients i at hospital j; and n_j is the number of delivery admissions to hospital j. We assume the outcome is related linearly to the covariates via a logit function:

Logistic Regression Model

$$\boldsymbol{logit}\big(\mathbf{Prob}(\boldsymbol{Y_{ij}} = \mathbf{1})\big) = \boldsymbol{\alpha} + \boldsymbol{\beta Z_{ij}} \qquad (1)$$

and $\boldsymbol{Z_{ij}} = (\boldsymbol{Z_{1ij}}, \boldsymbol{Z_{2ij}}, \dots, \boldsymbol{Z_{pij}})$ is a set of $\boldsymbol{p}$ patient-specific covariates.

To account for the natural clustering of observations within hospitals, we estimate a hierarchical logistic regression model that links the risk factors to the same outcomes and a hospital-specific random effect.

Hierarchical Logistic Regression Model

$$\boldsymbol{logit}\big(\mathbf{Prob}(\boldsymbol{Y_{ij}} = \mathbf{1})\big) = \boldsymbol{\alpha_j} + \boldsymbol{\beta Z_{ij}} \qquad (2)$$

$$\text{where } \alpha_j = \mu + \omega_j;\ \omega_j \sim N(0, \tau^2) \qquad (3)$$

Where $\alpha_j$represents the hospital-specific intercept, $\boldsymbol{Z}_{ij}$is defined as above, μ is the adjusted average intercept over all hospitals in the sample, ω_j is the hospital-specific intercept deviation from μ, and τ^2 is the between-hospital variance component. This model separates within-hospital variation from between-hospital variation. Both the hierarchical logistic regression model and the logistic regression model are estimated using the SAS software system (GLIMMIX and LOGISTIC procedures, respectively).

The ratio of a hospital's "predicted" number of delivery hospitalizations with a severe obstetric complication to "expected" number of delivery hospitalizations with a severe obstetric complication, referred to as the standardized risk ratio (SRR), is calculated as follows:

After obtaining the estimates of the hierarchical logistic regression model parameters $\hat{\mu}, \{\widehat{\alpha_1}, \widehat{\alpha_2}, \dots, \widehat{\alpha_I}\}, \hat{\beta}$,and $\widehat{\tau^2}$, we calculate a standardized risk ratio (SRR), $\widehat{s_j}$, for each site by computing the ratio of the number of predicted SMM events to the number of expected SMM events.

Specifically, we calculate:

Predicted Value: $\widehat{p_{ij}} = logit^{-1}(\hat{\alpha}_j + \hat{\beta}Z_{ij}) = \frac{exp(\hat{\alpha}_j+\hat{\beta}Z_{ij})}{1+ exp(\hat{\alpha}_j+\hat{\beta}Z_{ij})}$

Expected Value: $\widehat{e_{ij}} = logit^{-1}(\hat{\mu}_j + \hat{\beta}Z_{ij}) = \frac{exp(\hat{\mu}_j+\hat{\beta}Z_{ij})}{1+ exp(\hat{\mu}_j+\hat{\beta}Z_{ij})}$

Standardized Risk Ratio (SRR): $\hat{s}_j = \frac{\sum_{i=1}^{n_j} \widehat{p_{ij}}}{\sum_{i=1}^{n_j} \widehat{e_{ij}}}$

The risk-standardized obstetric complication rate is calculated by multiplying the SRR by the national observed severe obstetric complications rate, $\bar{y}$. For testing, this rate was the observed severe obstetric complications rate across all testing hospitals.

Risk-Standardized Obstetric Complications Rate: $\widehat{RSOCR} = SRR_j \times \bar{y}$

For measure reporting, we report the measure scores as a rate per 10,000 delivery hospitalizations. Measure scores for Stage 1 and Stage 2 Beta testing is reported in Section 3.4 Measure Results.

Measure scores calculated during Stage 1 Beta testing health sites and hospitals are provided in Appendix D.

2.9 Measure Testing

2.9.1 Data Element Reliability

Data element reliability and feasibility were assessed in nine sites with a total of 27 hospitals with virtual EHR walkthrough sessions conducted with each site. The site shared their screen while navigating through their EHR system as the measure data elements, specifications, and clinical workflows were discussed. Using the NQF's eCQM Feasibility Scorecard template, a scorecard was completed for each site during this time. The feasibility scorecard results were analyzed for each site and aggregated across all sites. Each data element score was examined within each of the domains

2.9.2 Measure Score Reliability

During measure testing, we assessed measure score reliability, which is the degree to which repeated measurements of the same entity agree with each other. We estimated the measure score reliability using a signal-to-noise ratio to assess the values according to conventional standards.[48] We used signal-to-noise reliability to assess how well the measure can distinguish the performance of one hospital from another. The signal is the proportion of the variability in measured performance that can be explained by real differences in performance. Scores can range from zero to one. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real difference in performance.[49,50] We used the formula presented by Adams and colleagues (2010) to calculate site-level reliability for hospital systems, and separately, individual hospitals.[50] The site-to-site variance is estimated from the hierarchical logistic regression model, n is equal to each site volume, and the site error variance is estimated using the variance of the

logistic distribution ($\frac{\pi^2}{3}$). We examined reliability using two minimum case volume thresholds for assessment of the impact on measure score reliability.

Measure score reliability for Stage 1 and Stage 2 Beta testing is reported in Section 3.5.2 Measure Score Reliability.

Initial measure score reliability testing during Stage 1 Beta testing health sites and hospitals are provided in Appendix D.

2.9.3 Data Element Validity

For this measure, the determination of both outcomes and risk factors involves many data elements from hospital EHR systems. We first ensured that the critical data elements were complete by examining:

- Distribution and availability of the data elements, and
- Variation of distribution and completeness of data elements across different hospitals and EHR systems.

In Stage 1 Beta testing, a statistically representative sample of the electronically submitted inpatient encounters from six sites (one site with 10 hospitals, the other five representing individual hospitals) was selected for re-abstraction for reliability testing and clinical adjudication. The minimum number of denominator cases per measured entity for adjudication was established to achieve sufficient measure score reliability and was determined to be 30 to 36 sampled cases were examined per site. This includes 30 to 36 charts at each of the five individual hospitals and three-to-four charts for each of the ten hospitals in the system. 100% of the sites met the minimum denominator requirement.

During the virtual visits, site staff shared their screen, navigated through the electronic health records of the sampled patients while Joint Commission staff manually re-abstracted each data element. To determine reliability and validity, re-abstraction findings were compared with the original electronic data submission and any disagreements were adjudicated with reasons for discrepancies noted.

2.9.4 Measure Score Validity

For Stage 1 Beta testing, measure score validity was assessed clinically adjudicating numerator and denominator encounters as identified by EHR data. Each component of the measure was validated and considered to have 'agreement' if the EHR and chart abstracted data both identified the encounter as appropriately belonging in the measure numerator, or in the denominator only. We calculated the positive predictive value (PPV), measure sensitivity, specificity, agreement, and negative predictive value (NPV).

For Stage 2 Beta testing, data from five additional hospitals representing two electronic health record (EHR) systems were recruited to test the measure specifications and measure logic, to further assess the feasibility of collecting data elements required for measure calculation, and to adjudicate the identification in EHR data of hospital delivery encounters as those with and those without severe

obstetric complications/numerator events with review of these delivery encounters in the medical record for conditions indicative of severe obstetric complications.

Sample size calculations were performed to identify the number of numerator encounters and the number of denominator-only encounters for clinical adjudication based on factors including alpha level of 0.05, margin of error at various sample sizes, target PPV (for numerator encounter adjudication), and target NPV (for denominator-only encounter adjudication).

A total of 275 numerator encounters were adjudicated, of which 136 were identified as those with severe obstetric complications in addition to or without transfusion and 139 were transfusion-only encounters. All severe obstetric complication numerator encounters (136) from the five hospitals were adjudicated. For transfusion-only numerator encounters, at least 10 transfusion-only numerator encounters were adjudicated per hospital. For hospitals with more than 10 transfusion-only numerator encounters, the following algorithm was used: in hospitals with fewer than 2,000 encounters, 50% of transfusion-only cases were adjudicated. For hospitals with ≥2,000 encounters, 34% of transfusion-only cases were adjudicated to make up the remainder of the planned adjudication of 139 transfusion-only numerator encounters. Note that 394 separate numerator events were adjudicated among the 275 delivery encounters with at least one qualifying numerator event.

For the denominator, 179 denominator-only encounters were adjudicated out of the 17,855 overall denominator encounters. Denominator-only encounters with high-risk conditions for severe obstetric complications were selected in order to maximize the likelihood of identifying false negatives. High-risk conditions included: ICU stay > 12 hours; prolonged length of stay > 2 days after delivery for vaginal delivery or > 4 days after delivery for cesarean delivery; platelet count < 100 10^3 uL; and patients with placental accreta spectrum, placental previa, renal disease, or pulmonary hypertension at hospital admission. For hospitals with fewer than 2,000 encounters, all denominator-only high-risk cases were adjudicated. For hospitals with ≥2,000 encounters, all cases of ICU >12, placenta accreta, and pulmonary HTN were adjudicated while 82% of placenta previa-only and renal disease-only cases were adjudicated.

These selected numerator and denominator only encounters were investigated by clinical adjudicators by reviewing medical records from qualifying delivery hospitalizations, with a focus on labor and delivery documentation and discharge summaries. We calculated PPV and NPV, as well as estimates of sensitivity and specificity given the parameters of clinical adjudication.

Definitions for calculated measures of validity are as follows:

- PPV: describes the probability that a patient with a positive result (numerator case) in the EHR data also was a positive result in the abstracted medical record data, as confirmed by a clinical adjudicator.
- NPV: describes the probability that a patient with a negative result (not in the numerator) in EHR data also was a negative result in the abstracted medical record, confirmed by the clinical adjudicator.
- Sensitivity: describes the probability that a patient with a positive result in the abstracted medical record data was also a positive result in the EHR data.

- Specificity: describes the probability that a patient with a negative result in the abstracted medical record data was also a negative result in the EHR data.
- Agreement: defined as the amount of remaining agreement between the maternal morbidity outcomes based on EHR and the maternal morbidity outcomes based on the abstracted medical record after the agreement by chance is factored in, measured by a Kappa statistic with values closer to one reflecting higher agreement.

2.9.5 Face Validity

To systematically assess face validity, we surveyed the TEP and Patient Working Group. We asked TEP members to rate a series of five statements using a six-point scale (1=Strongly Agree, 2=Moderately Agree, 3=Somewhat Agree, 4=Somewhat Disagree, 5= Moderately Disagree, and 6=Strongly Disagree) to assess the importance, reliability and validity, feasibility, and usability of the measure, as well as the ability of the measure to help distinguish better and worse quality of care at hospitals. Using the same six-point scale, the Patient Working Group members were asked to rate two of the five statements (Statement 1 and Statement 5) to assess the importance of the measure and the ability of the measure to help distinguish better and worse quality of care at hospitals.

The statements read as follows:

- Statement 1: The severe obstetric morbidity and mortality captured by the Severe Obstetric Complications eCQM is an important health outcome to measure because it is an area with room for improvement.
- Statement 2: The Severe Obstetric Complications eCQM will produce reliable and valid hospital measurement of severe obstetric morbidity and mortality rates across hospitals.
- Statement 3: The Severe Obstetric Complications eCQM is feasible to implement because required data are routinely collected as part of clinical care and are extractable from electronic health records.
- Statement 4: Hospitals can use the Severe Obstetric Complications eCQM performance results for performance improvement.
- Statement 5: The risk standardized rate of severe obstetric morbidity and mortality events obtained from the Severe Obstetric Complications eCQM as specified is a critical component (that is, necessary but not all-inclusive) of defining and comparing quality of obstetric care between hospitals.

3. Results

3.1 Measure Cohort

Table 1a provides information on the number of delivery encounters and patient demographic characteristics within Stage 1 Beta testing sites and across the eight sites. A corresponding description of

the number of delivery encounters and patient demographic characteristics in five Stage 2 Beta testing hospitals is provided in Table 1b.

Table 1a. Patient Characteristics of Delivery Encounters (8 Sites, Stage 1 Beta Testing)

Characteristics		Site #1	Site #2	Site #3	Site #5	Site #6	Site #7	Site #9	Site #10	Across Sites
		N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Number of encounters		18,070	7,196	7,955[a]	6,139	3,359	4,369[b]	3,918	9,178[c]	60,184
Average Maternal Age in Years [Mean (STD)]		30 (6.0)	31 (6.0)	29 (6.0)	29 (6.0)	33 (5.0)	32 (5.0)	32 (5.0)	31 (5.0)	30 (6.0)
Maternal Age in Years	**<18**	111 (0.6)	39 (0.5)	78 (1.0)	51 (0.8)	1 (0.0)	2 (0.0)	10 (0.3)	52 (0.6)	344 (0.6)
	18-<25	3,158 (17.5)	1,130 (15.7)	1,822 (22.9)	1,530 (24.9)	145 (4.3)	391 (8.9)	356 (9.1)	1,255 (13.7)	9,787 (16.3)
	25-<30	4,917 (27.2)	1,791 (24.9)	2,416 (30.4)	1,885 (30.7)	490 (14.6)	959 (22.0)	860 (21.9)	2,194 (23.9)	15,512 (25.8)
	30-<35	5,908 (32.7)	2,413 (33.5)	2,223 (27.9)	1,708 (27.8)	1,417 (42.2)	1,622 (37.1)	1,542 (39.4)	3,404 (37.1)	20,237 (33.6)
	35-<40	3,161 (17.5)	1,458 (20.3)	1,177 (14.8)	800 (13.0)	1,007 (30.0)	1,118 (25.6)	914 (23.3)	1,864 (20.3)	11,499 (19.1)
	40-<45	749 (4.1)	341 (4.7)	223 (2.8)	153 (2.5)	277 (8.2)	263 (6.0)	215 (5.5)	387 (4.2)	2,608 (4.3)
	45-<50	60 (0.3)	21 (0.3)	15 (0.2)	12 (0.2)	19 (0.6)	13 (0.3)	19 (0.5)	18 (0.2)	177 (0.3)
	>=50	6 (0)	3 (0.0)	0 (0.0)	0 (0.0)	3 (0.1)	1 (0.0)	2 (0.1)	4 (0.0)	19 (0.0)
Race/ Ethnicity	**Hispanic**	2,468 (13.7)	2,110 (29.3)	734 (9.2)	485 (7.9)	497 (14.8)	1,739 (39.8)	163 (4.2)	235 (2.6)	8,431 (14.0)
	Non-Hispanic, Black/African American	4,084 (22.6)	606 (8.4)	2,971 (37.3)	952 (15.5)	89 (2.6)	254 (5.8)	1,307 (33.4)	1,590 (17.3)	11,853 (19.7)
	Non-Hispanic, Asian/Pacific Islander	743 (4.1)	117 (1.6)	157 (2.0)	66 (1.1)	364 (10.8)	703 (16.1)	250 (6.4)	532 (5.8)	2,932 (4.9)
	Non-Hispanic, White	9,322 (51.6)	3,658 (50.8)	3,940 (49.5)	4,507 (73.4)	2,307 (68.7)	1,648 (37.7)	2,077 (53.0)	5,912 (64.4)	33,371 (55.4)
	Non-Hispanic, Other	651 (3.6)	633 (8.8)	135 (1.7)	58 (0.9)	35 (1.0)	17 (0.4)	112 (2.9)	40 (0.4)	1,681 (2.8)
	Declined/Unknown	802 (4.4)	72 (1.0)	18 (0.2)	71 (1.2)	67 (2.0)	8 (0.2)	9 (0.2)	869 (9.5)	1,916 (3.2)
Primary Payer	**Medicare**	50 (0.3)	12 (0.2)	27 (0.3)	36 (0.6)	7 (0.2)	1 (0.0)	6 (0.2)	84 (0.9)	223 (0.4)
	Medicaid	5,857 (32.4)	305 (4.2)	3,790 (47.6)	2,600 (42.4)	97 (2.9)	408 (9.3)	10 (0.3)	3,154 (34.4)	16,221 (27.0)
	Private Insurance	11,170 (61.8)	6,863 (95.4)	4,119 (51.8)	3,482 (56.7)	3,230 (96.2)	3,869 (88.6)	3,894 (99.4)	4,439 (48.4)	41,066 (68.2)
	Self-pay or Uninsured	0 (0.0)	15 (0.2)	19 (0.2)	21 (0.3)	15 (0.4)	0 (0.0)	8 (0.2)	71 (0.8)	149 (0.2)
	Other	993 (5.5)	0 (0.0)	0 (0.0)	0 (0.0)	10 (0.3)	86 (2.0)	0 (0.0)	1,429 (15.6)	2,518 (4.2)

Characteristics		Site #1	Site #2	Site #3	Site #5	Site #6	Site #7	Site #9	Site #10	Across Sites
		N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
	Unknown	0 (0.0)	1 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	5 (0.1)	0 (0.0)	1 (0.0)	7 (0.0)

[a] *Site 3: 7,949 unique patients had 7,955 encounters*
[b] *Site 7: 4,367 unique patients had 4,369 encounters*
[c] *Site 10: 9,173 unique patients had 9,178 encounters*

Table 1b. Patient Characteristics of Delivery Encounters (5 Hospitals, Stage 2 Beta Testing)

Characteristics		Measure Cohort					
		Hospital A	Hospital B	Hospital C	Hospital D	Hospital E	Across Sites
		N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Number of encounters		1,610	6,932	7,101	1,724	488	17,855
Average Maternal Age in Years [Mean (STD)]		27.84	28.90	31.90	29.62	26.42	29.60
Maternal Age in Years	**<18**	23 (1.4)	104 (1.5)	12 (0.2)	11 (0.6)	10 (2.0)	160 (0.9)
	18-<25	433 (26.9)	1,672(24.1)	460 (6.5)	301 (17.5)	173 (35.5)	3,039 (17.0)
	25-<30	554 (34.4))	1,929 (27.8)	1,582 (22.3)	565 (32.8)	183 (37.5)	4,813 (27.0)
	30-<35	398 (24.7)	1,907 (27.5)	2,918 (41.1)	503 (29.2)	87 (17.8)	5,813 (32.6)
	35-<40	177 (11.0)	1,039 (15.0)	1,784 (25.1)	287 (16.6)	33 (6.8)	3,320 (18.6)
	40-<45	25 (1.6)	264 (3.8)	331 (4.7)	56 (3.2)	1 (0.2)	677 (3.8)
	45-<50	0 (0.0)	17 (0.2)	14 (0.2)	0 (0.0)	1 (0.2)	32 (0.2)
	>=50	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.1)	0 (0.0)	1 (0.0)
Ethnicity	**Hispanic or Latino**	24 (1.5)	524 (7.6)	130 (1.8)	58 (3.4)	1 (0.2)	737 (4.1)
	Not Hispanic or Latino	1,577 (98.0)	6,383 (92.1)	6,937 (97.7)	1,664 (96.5)	487 (99.8)	17,048 (95.5)
	Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Unknown	9 (0.6)	25 (0.4)	34 (0.5)	2 (0.1)	0 (0.0)	70 (0.4)
Race	**Black/African American**	249 (15.5)	3,483 (50.2)	935 (13.2)	98 (5.7)	14 (2.9)	4,779 (26.8)
	White	1,328 (82.5)	2,870 (41.4)	5,718 (80.5)	1,528 (88.6)	457 (93.6)	11,901 (66.7)
	Asian	6 (0.4)	256 (3.7)	331 (4.7)	58 (3.4)	0 (0.0)	651 (3.6)
	American Indian or Alaska Native	3 (0.2)	21 (0.3)	7 (0.1)	4 (0.2)	0 (0.0)	35 (0.2)
	Native Hawaiian or Other Pacific Islander	1 (0.1)	20 (0.3)	15 (0.2)	10 (0.6)	0 (0.0)	46 (0.3)
	Other	7 (0.4)	129 (1.9)	43 (0.6)	12 (0.7)	17 (3.5)	208 (1.2)
	Unknown	16 (1.0)	153 (2.2)	52 (0.7)	14 (0.8)	0 (0.0)	235 (1.3)
Primary Payer	**Medicare**	2 (0.1)	10 (0.1)	3 (0.0)	0 (0.0)	1 (0.2)	16 (0.0)
	Medicaid	0 (0.0)	4 (0.1)	0 (0.0)	1 (0.1)	239 (49.0)	244 (1.4)

Characteristics		Measure Cohort					
		Hospital A	Hospital B	Hospital C	Hospital D	Hospital E	Across Sites
		N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
	Private Insurance	825 (51.2)	3,309 (47.7)	6,260 (88.2)	1,320 (76.6)	204 (41.8)	11,918 (66.7)
	Self-pay or Uninsured	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	5 (1.0)	5 (0.0)
	Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	39 (8.0)	39 (0.0)
	Unknown	783 (48.6)	3,609 (52.1)	838 (11.8)	403 (23.4)	0 (0.0)	5,633 (31.5)

Table 2 shows observed (unadjusted) frequencies for each defined severe obstetric complication among patients in Stage 1 and Stage 2 Beta testing. Singular numerator events are not mutually exclusive; delivery encounters in which multiple numerator events occurred are included in the frequency for each numerator event experienced.

Table 2. Observed (Unadjusted) Frequencies for Numerator Events (Stage 1 and Stage 2 Beta Testing)

Numerator Events	Stage 1 Beta Testing		Stage 2 Beta Testing	
	Total N	%	Total N	%
Delivery Encounters	60,184	*	17,855	*
Delivery encounter with any of the 21 CDC numerator events **or** mortality	1,466	2.44	502	2.81
Delivery encounter with **blood transfusion-only (encounter has no other numerator events)**	1,164	1.93	366	2.05
Delivery encounter with any of the 21 CDC numerator events or mortality but excluding **blood transfusion-only encounters**	302	0.50	136	0.76
Delivery encounter with **mortality**	3	< 0.01	3	0.02
Delivery encounter with **acute heart failure**	6	0.01	5	0.03
Delivery encounter with **acute myocardial infarction**	0	0.00	1	< 0.01
Delivery encounter with **aortic aneurysm**	0	0.00	0	0.00
Delivery encounter with **cardiac arrest/ventricular fibrillation**	2	< 0.01	4	0.02
Delivery encounter with **heart failure/arrest during procedure or surgery**	0	0.00	0	0.00
Delivery encounter with **disseminated intravascular coagulation**	71	0.12	21	0.12
Delivery encounter with **shock**	33	0.05	29	0.16
Delivery encounter with **acute renal failure**	94	0.16	66	0.37
Delivery encounter with **adult respiratory distress syndrome**	31	0.05	18	0.10
Delivery encounter with **pulmonary edema**	18	0.03	9	0.05
Delivery encounter with **sepsis**	31	0.05	11	0.06
Delivery encounter with **air and thrombotic embolism**	7	0.01	1	0.01
Delivery encounter with **amniotic fluid embolism**	1	< 0.01	2	0.01
Delivery encounter with **eclampsia**	10	0.02	3	0.02
Delivery encounter with **severe anesthesia complications**	3	< 0.01	0	0.00
Delivery encounter with **puerperal cerebrovascular disease**	1	< 0.01	1	< 0.01
Delivery encounter with **conversion of cardiac rhythm**	4	0.01	2	< 0.01
Delivery encounter with **hysterectomy**	57	0.09	13	0.07
Delivery encounter with **temporary tracheostomy**	0	0.00	1	0.01
Delivery encounter with **ventilation**	26	0.04	13	0.07
Delivery encounter with **sickle cell disease with crisis**	0	0.00	1	0.01
Delivery encounter with **blood transfusion**	1,295	2.15	417	2.34

** Cell intentionally left empty*

3.2 Attribution

Table 3a and Table 3b provide health care system specific characteristics for each of the sites included in measure testing. In Table 3a, identification of whether a site was included in Alpha testing, Stage 1 Beta testing, and Stage 1 Beta clinical adjudication for reliability and validity testing is provided. Nine sites were included in Alpha testing (Sites 1 – 9), eight sites were included in Stage 1 Beta testing (Sites 1 – 3, 5 – 7, 9 – 10), and six sites were included in clinical adjudication (Sites 1 – 3, 6, 7, 9). Table 3b includes the characteristics for the five Stage 2 Beta testing hospitals.

Table 3a. Site Characteristics (10 sites, Alpha Testing and Stage 1 Beta Testing)

Site ID	# Of Hospitals	Geography (Urban, Suburban, Rural)	# Total Beds[a]	# Of Births[a]	Teaching Program in OB/GYN	Obstetric unit care level	NICU Level	Clinical EHR Software	Included in Alpha Testing	Included in Stage 1 Beta Testing	Included in Stage 1 Beta Clinical Adjudication
Site 1	10	Urban	1,800 (range 36 - 740)	16,350 + (range 450 – 5,550)[b]	No	*(Information not provided)*	Level 2 Level 3 Level 4	Epic	Yes	Yes	Yes
Site 2	1	Urban	250	8,800	No	Services all serious illnesses & abnormalities	Level 4	Cerner/ Siemens	Yes	Yes	Yes
Site 3	1	Urban	250	8,300	No	Services all serious illnesses & abnormalities	Level 3	Meditech	Yes	Yes	Yes
Site 4[c]	2	Urban	450	2,900	No	Services uncomplicated maternity & newborn cases	Level 2 Level 3	Cerner	Yes	No	No
Site 5	9	6 Urban 3 Rural	1,650 (range 35 - 595)	9,300 + (range 150– 3,400)[b]	No	*2 hospitals* = Services all serious illnesses & abnormalities *2 hospitals* = Services uncomplicated & most complicated cases *3 hospitals* = Services uncomplicated maternity & newborn cases *2 hospitals* = (Information not provided)	Level 3 (1 central NICU for all hospitals)	Epic	Yes	Yes	No
Site 6	1	Urban	450	3,300	No	Services all serious illnesses & abnormalities	Level 3	Meditech	Yes	Yes	Yes
Site 7	1	Urban	550	4,650	Yes	Services uncomplicated & most complicated cases	Level 3	Epic	Yes	Yes	Yes
Site 8[d]	1	Urban	650	2,450	Yes	Services all serious illnesses & abnormalities	Level 3	Epic	Yes	No	No
Site 9	1	Urban	400	3,850	No	Services all serious illnesses & abnormalities	Level 3	Epic	Yes	Yes	Yes
Site 10[e]	1	Urban	300	8,800	Yes	Services all serious illnesses & abnormalities	Level 3	Cerner	No	Yes	No

[a] The number of total beds and number of births have been rounded to the nearest 50 to maintain confidentiality of the hospitals

[b] Not all hospitals within this site reported number of births, so total births across site is higher than indicated

[c] Test Site 4 declined continued participation after Alpha Testing

[d] Data from Test Site 8 was not available in time for Beta Testing

[e] Test Site 10 joined after Alpha Testing

Table 3b. Test Site Characteristics (5 Hospitals, Stage 2 Beta Testing)

Site ID	Geography (Urban, Suburban, Rural)	# Total Beds[a]	# Of Births[a]	Teaching Program in OB/GYN	Obstetric unit care level	NICU Level	Clinical EHR Software
Hospital A	Urban	150	800	No	Level I	Level I	Epic
Hospital B	Urban	1,250	3,600	Yes	Level IV	Level II	Epic
Hospital C	Urban	450	3,600	Yes	Level IV	Level II	Epic
Hospital D	Urban	50	900	No	Level I	Level I	Epic
Hospital E	Rural	100	400	No	Level I	NA	Meditech

[a] *The number of total beds and number of births have been rounded to the nearest 50 to maintain confidentiality of the hospitals.*

3.3 Risk Model and Model Performance Results

Table 4 provides frequencies, adjusted odds ratios (ORs) and 95% confidence intervals (CIs) of the demographic and clinical variables in the risk models for any severe obstetric complications and severe obstetric complications excluding blood-transfusion only encounters developed and tested using data from Stage 1 Beta testing. The risk model was developed and tested with data from the test sites included in Stage 1 Beta testing; 60,184 delivery hospitalizations were randomly divided in a 70/30 split for a development dataset and a validation dataset. The same risk variables were included in the model for severe obstetric complications and the model for severe obstetric complications excluding blood transfusion-only encounters; however, due to the impact of very low prevalence of a few risk variables in the model of severe obstetric complication excluding transfusion-only encounters, Human Immunodeficiency Virus (HIV) was combined with autoimmune disease, and obstetric venous thromboembolism (VTE) was combined with long-term anticoagulant medication use.

Table 4. Risk Variables with Frequencies and Adjusted Odds Ratio for Risk Model in Stage 1 Beta Testing Development and Validation Samples for Both Severe Obstetric Complication Outcomes

	Frequencies		Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Stage 1 Development Sample N = 42,129 n (%)	Stage 1 Validation Sample N = 18,055 n (%)	Stage 1 Development Dataset	Stage 1 Validation Dataset	Stage 1 Development Dataset	Stage 1 Validation Dataset
Maternal Age in Years	*	*	*	*	*	*
<20	1,097 (2.6)	476 (2.6)	REF	REF	REF	REF
20-<25	5,945 (14.1)	2,613 (14.5)	0.92 (0.63, 1.34)	1.34 (0.73, 2.45)	1.06 (0.36, 3.09)	0.94 (0.20, 4.40)
25-<30	11,028 (26.2)	4,484 (24.8)	0.78 (0.54, 1.13)	0.98 (0.54, 1.78)	1.29 (0.46, 3.64)	1.18 (0.27, 5.27)
30-<35	14,088 (33.4)	6,149 (34.1)	0.77 (0.53, 1.11)	0.91 (0.50, 1.66)	1.31 (0.46, 3.67)	1.16 (0.26, 5.10)
35-<40	8,029 (19.1)	3,470 (19.2)	0.81 (0.55, 1.18)	0.89 (0.48, 1.67)	0.99 (0.34, 2.89)	1.24 (0.27, 5.66)
>=40	1,942 (4.6)	863 (4.8)	1.36 (0.89, 2.08)	1.33 (0.66, 2.70)	2.12 (0.69, 6.55)	1.67 (0.32, 8.84)
Anemia	8,016 (19.0)	3,450 (19.1)	1.70 (1.48, 1.96)	1.91 (1.54, 2.37)	1.25 (0.89, 1.76)	2.04 (1.25, 3.35)
Asthma	3,587 (8.5)	1,512 (8.4)	1.27 (1.04, 1.55)	1.11 (0.80, 1.53)	2.09 (1.45, 3.02)	1.70 (0.92, 3.14)
BMI >= 40	2,551 (6.1)	1,115 (6.2)	1.11 (0.87, 1.42)	0.72 (0.46, 1.12)	1.82 (1.15, 2.88)	0.76 (0.31, 1.84)
Bariatric Surgery	318 (0.8)	127 (0.7)	1.13 (0.64, 1.98)	0.28 (0.04, 2.01)	1.13 (0.34, 3.80)	**
Bleeding Disorder	1,238 (2.9)	530 (2.9)	2.17 (1.66, 2.83)	1.89 (1.23, 2.91)	3.00 (1.82, 4.96)	1.84 (0.77, 4.40)
Cardiac Disease	673 (1.6)	266 (1.5)	1.54 (1.07, 2.21)	1.79 (1.00, 3.18)	2.42 (1.31, 4.48)	4.33 (1.86, 10.10)
Economic Housing Instability	41 (0.1)	21 (0.1)	2.74 (0.96, 7.85)	**	9.47 (2.61, 34.31)	**
Gastrointestinal Disease	658 (1.6)	309 (1.7)	1.01 (0.63, 1.62)	1.85 (1.10, 3.14)	0.62 (0.21, 1.89)	1.97 (0.68, 5.71)
Gestational Diabetes	3,988 (9.5)	1,805 (10.0)	1.08 (0.88, 1.34)	0.91 (0.65, 1.29)	1.39 (0.92, 2.11)	1.52 (0.82, 2.84)
Hypertension	1,816 (4.3)	797 (4.4)	0.98 (0.75, 1.28)	1.22 (0.81, 1.83)	0.72 (0.40, 1.28)	0.96 (0.43, 2.14)
Mental Health Disorder	6,086 (14.4)	2,667 (14.8)	1.19 (1.01, 1.41)	1.39 (1.09, 1.79)	1.15 (0.80, 1.64)	1.68 (1.00, 2.81)
Multiple Pregnancy	832 (2.0)	346 (1.9%)	2.10 (1.56, 2.82)	2.05 (1.29, 3.26)	1.75 (0.92, 3.33)	0.95 (0.29, 3.19)
Neuromuscular	214 (0.5)	89 (0.5)	0.95 (0.43, 2.11)	0.94 (0.23, 3.89)	1.42 (0.33, 6.05)	**
Other Preeclampsia	4,278 (10.2)	1,747 (9.7)	1.42 (1.17, 1.73)	1.15 (0.82, 1.62)	1.38 (0.88, 2.17)	1.64 (0.81, 3.33)

	Frequencies		Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Stage 1 Development Sample N = 42,129 n (%)	Stage 1 Validation Sample N = 18,055 n (%)	Stage 1 Development Dataset	Stage 1 Validation Dataset	Stage 1 Development Dataset	Stage 1 Validation Dataset
Placenta Previa	179 (0.4)	92 (0.5)	4.84 (3.01, 7.76)	2.10 (0.88, 5.05)	1.17 (0.41, 3.31)	2.04 (0.48, 8.74)
Placental Abruption	402 (1.0)	146 (0.8)	3.53 (2.51, 4.96)	4.02 (2.33, 6.91)	2.15 (0.97, 4.78)	3.62 (1.20, 10.88)
Placental Accreta Spectrum	47 (0.1)	19 (0.1)	45.36 (21.71, 94.78)	67.99 (22.88, 201.98)	171.79 (77.38, 381.39)	195.71 (60.74, 630.63)
Preexisting Diabetes	637 (1.5)	266 (1.5)	1.43 (0.98, 2.09)	2.10 (1.24, 3.55)	1.85 (0.95, 3.60)	1.85 (0.69, 4.97)
Preterm Birth	2,893 (6.9)	1,204 (6.7)	1.41 (1.16, 1.72)	1.59 (1.17, 2.16)	2.32 (1.57, 3.44)	2.21 (1.25, 3.93)
Previous Cesarean	7,201 (17.1)	3,055 (16.9)	1.22 (1.04, 1.44)	1.45 (1.13, 1.85)	1.08 (0.75, 1.56)	1.24 (0.72, 2.12)
Pulmonary Hypertension	18 (0.0)	5 (0.0)	0.69 (0.12, 3.99)	6.24 (0.59, 66.34)	3.56 (0.63, 20.05)	5.72 (0.39, 83.53)
Renal Disease	110 (0.3)	36 (0.2)	3.34 (1.90, 5.87)	1.27 (0.37, 4.40)	3.66 (1.48, 9.07)	2.61 (0.45, 15.07)
Severe Preeclampsia	1,615 (3.8)	722 (4.0)	2.35 (1.82, 3.03)	2.95 (2.04, 4.27)	3.48 (2.11, 5.77)	4.94 (2.51, 9.74)
Substance Abuse	2,799 (6.6)	1,249 (6.9)	1.12 (0.89, 1.39)	0.97 (0.69, 1.38)	1.34 (0.84, 2.13)	0.98 (0.45, 2.11)
Thyrotoxicosis	150 (0.4)	62 (0.3)	0.60 (0.19, 1.94)	**	1.09 (0.15, 8.03)	**
Autoimmune Disease	108 (0.3)	49 (0.3)	2.60 (1.22, 5.53)	1.40 (0.38, 5.14)	NA[a]	NA[a]
HIV	53 (0.1)	18 (0.1)	1.47 (0.45, 4.86)	3.20 (0.70, 14.68)	NA[a]	NA[a]
Grouped: Autoimmune Disease or HIV	160 (0.4)	67 (0.4)	NA[b]	NA[b]	1.80 (0.41, 7.90)	1.13 (0.12, 10.40)
Long Term Anticoagulant Use	136 (0.3)	45 (0.2)	1.27 (0.60, 2.69)	1.87 (0.49, 7.15)	NA[a]	NA[a]
Obstetrical VTE	37 (0.1)	15 (0.1)	0.71 (0.13, 3.91)	**	NA[a]	NA[a]
Grouped: Long Term Anticoagulant Use or Obstetrical VTE	167 (0.4)	57 (0.3)	NA[b]	NA[b]	0.83 (0.21, 3.23)	1.80 (0.20, 16.55)
Vitals - Heart Rate	*	*	*	*	*	*
Result <110	35,700 (84.7)	15,245 (84.4)	REF	REF	REF	REF
Result >=110	3,859 (9.2)	1,748 (9.7)	1.21 (0.99, 1.49)	1.33 (0.99, 1.79)	1.58 (1.04, 2.40)	1.03 (0.52, 2.05)
Missing	2,570 (6.1)	1,062 (5.9)	2.05 (0.97, 4.34)	2.70 (0.84, 8.65)	2.15 (0.38, 12.09)	0.75 (0.03, 17.48)

	Frequencies		Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Stage 1 Development Sample N = 42,129 n (%)	Stage 1 Validation Sample N = 18,055 n (%)	Stage 1 Development Dataset	Stage 1 Validation Dataset	Stage 1 Development Dataset	Stage 1 Validation Dataset
Vitals - Systolic BP	*	*	*	*	*	*
Result <140	33,382 (79.2)	14,295 (79.2)	REF	REF	REF	REF
Result >=140 & <160	5,050 (12.0)	2,225 (12.3)	1.16 (0.95, 1.40)	1.01 (0.75, 1.38)	0.83 (0.54, 1.29)	1.16 (0.62, 2.17)
Result >=160	1,178 (2.8)	486 (2.7)	1.31 (0.96, 1.78)	0.77 (0.45, 1.32)	0.42 (0.20, 0.90)	1.08 (0.43, 2.68)
Missing	2,519 (6.0)	1,049 (5.8)	0.76 (0.34, 1.67)	0.60 (0.18, 2.04)	0.82 (0.13, 5.04)	1.87 (0.08, 43.64)
Labs - Hematocrit	*	*	*	*	*	*
Result <33	7,929 (18.8)	3,415 (18.9)	2.54 (2.20, 2.93)	2.89 (2.32, 3.61)	1.22 (0.86, 1.73)	0.95 (0.54, 1.67)
Result >=33	28,911 (68.6)	12,382 (68.6)	REF	REF	REF	REF
Missing	5,289 (12.6)	2,258 (12.5)	1.14 (0.83, 1.56)	1.24 (0.76, 2.03)	0.71 (0.34, 1.49)	1.08 (0.39, 2.94)
Labs - WBC	*	*	*	*	*	*
Result <14	29,472 (70.0)	12,627 (69.9)	REF	REF	REF	REF
Result >=14	4,879 (11.6)	2,131 (11.8)	1.15 (0.95, 1.40)	1.20 (0.89, 1.62)	1.47 (0.99, 2.18)	1.37 (0.73, 2.58)
Missing	7,778 (18.5)	3,297 (18.3)	0.65 (0.50, 0.84)	0.58 (0.38, 0.86)	0.62 (0.34, 1.14)	0.81 (0.33, 1.98)

** Cell intentionally left empty*

*** Odds ratios not estimable due to small sample size and distribution of values across variables*

[a]NA indicates odds ratios not calculated because variables were combined into a composite

[b]NA indicates odds ratios not calculated because variables appeared in model individually and were not combined into a composite

Table 5 shows model performance statistics for the model of any severe obstetric complications and for the model of severe obstetric complications excluding blood transfusion-only encounters, developed and validated using Stage 1 Beta testing. The calculated C-statistic for the risk model for any severe obstetric complications was 0.74 using the development dataset and 0.75 using the same beta coefficients from the development dataset model applied to the validation dataset. The calculated C-statistic for the risk model for severe obstetric complications excluding blood transfusion-only encounters was 0.77 using the development dataset and 0.73 using the validation dataset. For both versions of the measure, the C-statistics indicate good model discrimination.

The calibration indices (γ0, γ1) used to assess the risk model for any severe obstetric complications in the validation dataset are (0.14, 1.05) and for the severe obstetric complications excluding blood transfusion-only encounters in the validation dataset are (0.15, 1.02). The calibration values which are consistently close to 0 at one end and close to 1 at the other end indicate good calibration of the model. If the γ0 in the model performance using validation data is substantially far from zero and the γ1 is substantially far from 1, there is potential evidence of over-fitting.

Predictive ability displays the percent of cases with severe obstetric complications in the (lowest, highest) decile of predicted risk. With both the development and validation datasets, both models show a reasonable range between the lowest decile and highest decile of predicted ability, given the low prevalence of the outcome. Overall, these diagnostic results demonstrate the risk-adjustment model adequately controls for differences in patient characteristics.

Table 5. Model Performance Statistics for Risk Models for Both Severe Obstetric Complication Outcomes (Stage 1 Beta Testing)

Model Performance Statistic	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Development Dataset	Validation Dataset	Development Dataset	Validation Dataset
C-statistic	0.74 (0.72,0.76)	0.75 (0.72,0.77)	0.77 (0.73,0.81)	0.73 (0.67,0.79)
Calibration (γ0, γ1)	(0.00,1.00)	(0.14,1.05)	(0.00,1.00)	(0.15,1.02)
Predictive ability	(0.56%,9.56%)	(0.43%,10.04%)	(0.19%,2.58%)	(0.17%,2.49%)

A calibration plot is a graphical tool to assess the agreement between the number of observed severe obstetric complications and the number predicted by the risk model. The average observed and predicted rates are plotted by decile groups of predicted probabilities. A calibration curve closer to the perfect calibration (diagonal) line reflects excellent model fit to the data. 95% confidence intervals of the observed rates are also provided in the plot.

In the calibration (risk decile) plots (Figures 1 and 3 for Any Severe Obstetric Complications, and Figures 2 and 4 for Severe Obstetric Complications Excluding Blood Transfusion-Only Encounters), we visualize the agreement between observed severe obstetric complication rates and the risk predicted by the risk model. The plots show reasonable agreement between the observed severe obstetric complication rates with the model predictions demonstrating that the risk-adjustment model adequately fits the patient characteristics (case mix) data.

Figure 1. Calibration for Any Severe Obstetric Complication Rates by Predicted Risk Deciles in the Development Dataset (Stage 1 Beta Testing)

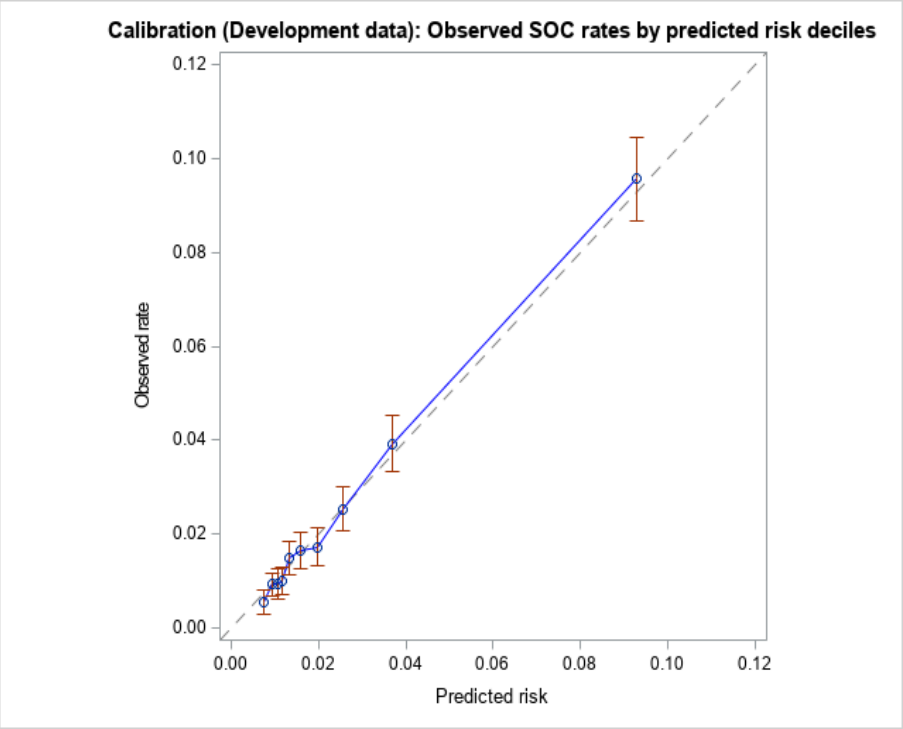


Figure 2. Calibration for Severe Obstetric Complications Excluding Blood Transfusion-Only Encounter Rates by Predicted Risk Deciles in the Development Dataset (Stage 1 Beta Testing)

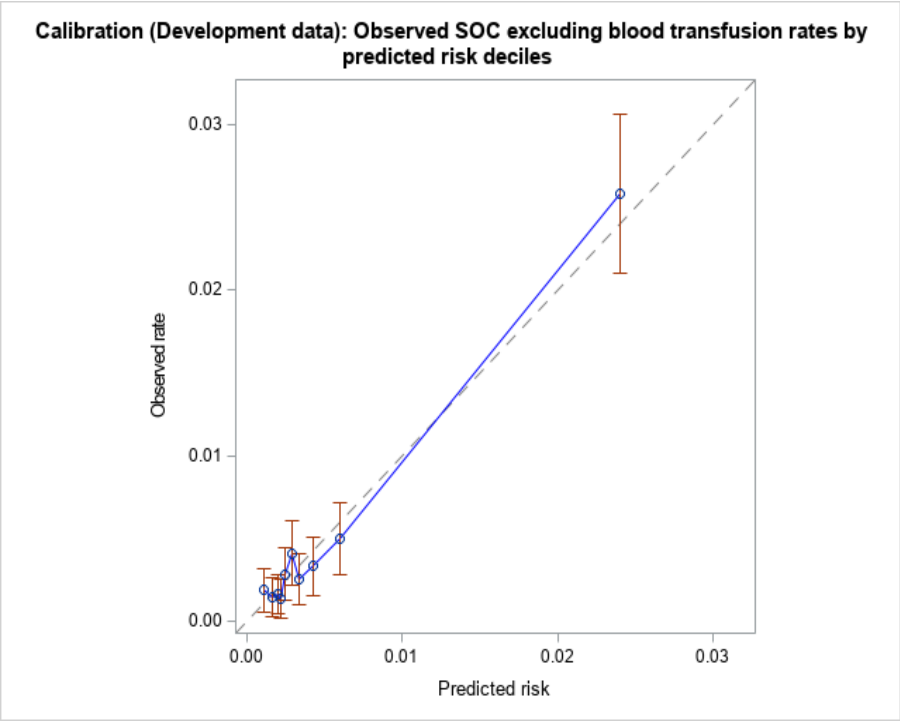


Figure 3. Calibration for Any Severe Obstetric Complication Rates by Predicted Risk Deciles in the Validation Dataset (Stage 1 Beta Testing)

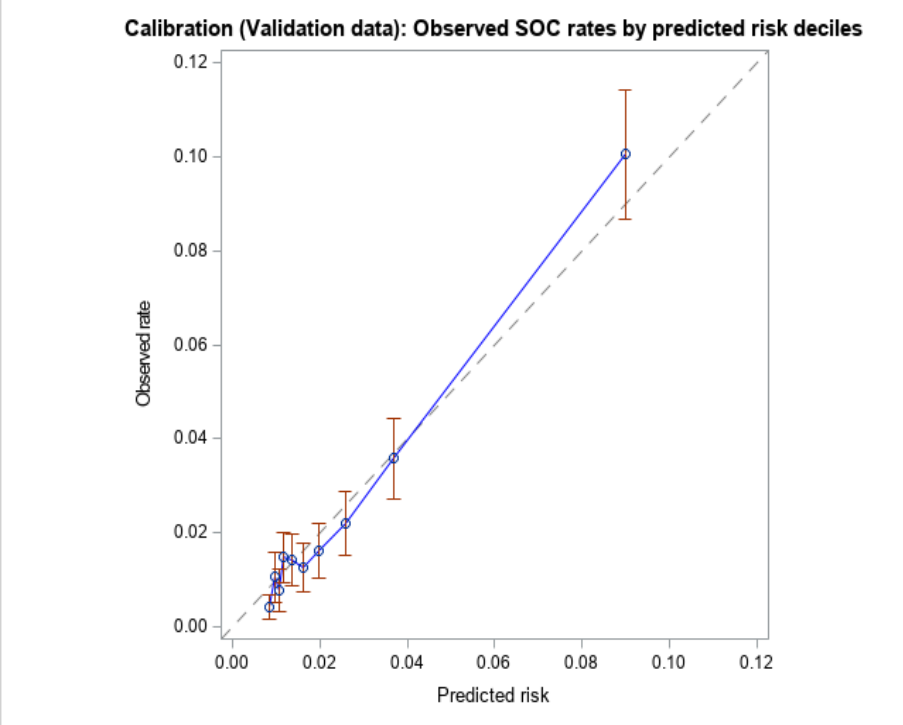


Figure 4. Calibration for Severe Obstetric Complications Excluding Blood Transfusion-Only Encounter Rates by Predicted Risk Deciles in the Validation Dataset (Stage 1 Beta Testing)

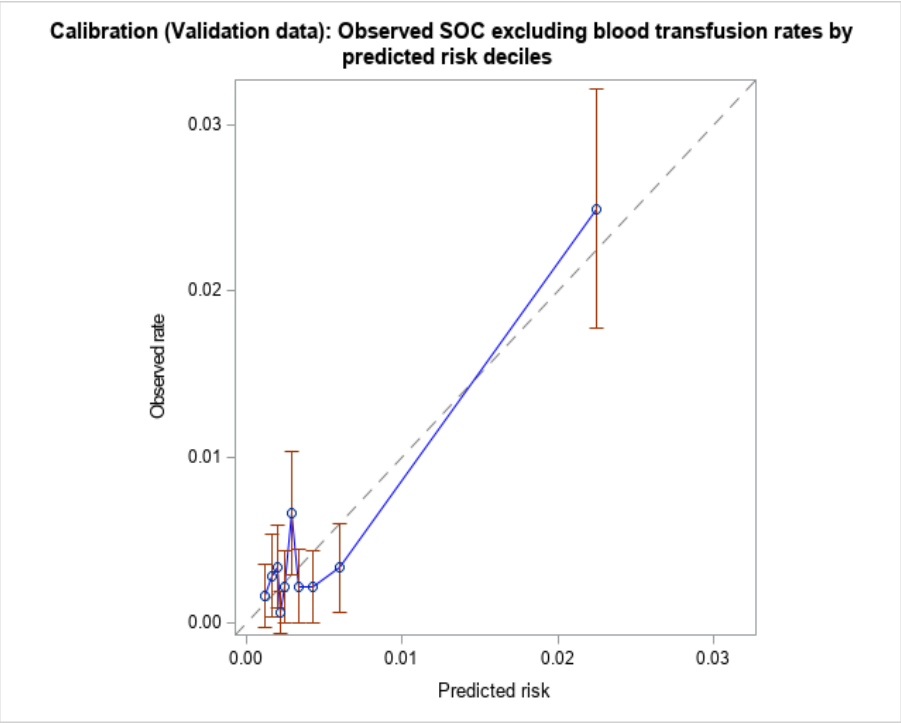


Table 6 shows the risk variables with frequencies and adjusted odds ratio for the risk model in stage 1 beta testing development sample, and the validation sample in stage 2. The outcomes for both stages were grouped as any severe obstetric complication, and for severe obstetric complications excluding blood transfusions.

Table 6. Risk Variables with Frequencies and Adjusted Odds Ratio for Risk Model in Stage 1 Beta Testing Development Sample and Stage 2 Beta Testing Validation Sample for Both Severe Obstetric Complication Outcomes

Variable	Frequencies		Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Stage 1 Development Sample N = 42,129 n (%)	Stage 2 Validation Sample N = 17,855 n (%)	Stage 1 Development Dataset	Stage 2 Validation Dataset	Stage 1 Development Dataset	Stage 2 Validation Dataset
Maternal Age in Years	*	*	*	*	*	*
<20	1,097 (2.6)	783 (4.4)	REF	REF	REF	REF
20-<25	5,945 (14.1)	2,767 (15.5)	0.92 (0.63, 1.34)	0.66 (0.43, 0.99)	1.06 (0.36, 3.09)	0.67 (0.27, 1.67)
25-<30	11,028 (26.2)	5,027 (28.2)	0.78 (0.54, 1.13)	0.61 (0.41, 0.90)	1.29 (0.46, 3.64)	0.90 (0.38, 2.12)
30-<35	14,088 (33.4)	5,749 (32.2)	0.77 (0.53, 1.11)	0.51 (0.34, 0.76)	1.31 (0.46, 3.67)	0.71 (0.29, 1.70)
35-<40	8,029 (19.1)	2,953 (16.5)	0.81 (0.55, 1.18)	0.82 (0.54, 1.25)	0.99 (0.34, 2.89)	0.91 (0.37, 2.28)
>=40	1,942 (4.6)	576 (3.2)	1.36 (0.89, 2.08)	0.78 (0.44, 1.40)	2.12 (0.69, 6.55)	1.43 (0.50, 4.12)
Anemia	8,016 (19.0)	2,271 (12.7)	1.70 (1.48, 1.96)	2.17 (1.75, 2.69)	1.25 (0.89, 1.76)	1.14 (0.72, 1.80)
Asthma	3,587 (8.5)	1,955 (10.9)	1.27 (1.04, 1.55)	1.39 (1.08, 1.77)	2.09 (1.45, 3.02)	1.96 (1.29, 2.98)
BMI >=40	2,551 (6.1)	1,539 (8.6)	1.11 (0.87, 1.42)	0.91 (0.67, 1.23)	1.82 (1.15, 2.88)	1.20 (0.73, 1.96)
Bariatric Surgery	318 (0.8)	84 (0.5)	1.13 (0.64, 1.98)	1.05 (0.39, 2.81)	1.13 (0.34, 3.80)	0.42 (0.05, 3.54)
Bleeding Disorder	1,238 (2.9)	590 (3.3)	2.17 (1.66, 2.83)	1.37 (0.89, 2.12)	3.00 (1.82, 4.96)	2.77 (1.48, 5.19)
Cardiac Disease	673 (1.6)	536 (3.0)	1.54 (1.07, 2.21)	1.83 (1.30, 2.56)	2.42 (1.31, 4.48)	3.22 (1.98, 5.23)
Economic Housing Instability	41 (0.1)	47 (0.3)	2.74 (0.96, 7.85)	0.98 (0.27, 3.62)	9.47 (2.61, 34.31)	**
Gastrointestinal Disease	658 (1.6)	340 (1.9)	1.01 (0.63, 1.62)	1.44 (0.85, 2.43)	0.62 (0.21, 1.89)	1.46 (0.57, 3.76)
Gestational Diabetes	3,988 (9.5)	1,195 (6.7)	1.08 (0.88, 1.34)	0.59 (0.38, 0.92)	1.39 (0.92, 2.11)	0.39 (0.16, 0.95)
Hypertension	1,816 (4.3)	776 (4.3)	0.98 (0.75, 1.28)	1.14 (0.79, 1.64)	0.72 (0.40, 1.28)	1.77 (1.02, 3.10)
Mental Health Disorder	6,086 (14.4)	2,805 (15.7)	1.19 (1.01, 1.41)	1.03 (0.81, 1.30)	1.15 (0.80, 1.64)	1.39 (0.93, 2.08)
Multiple Pregnancy	832 (2.0)	472 (2.6)	2.10 (1.56, 2.82)	1.38 (0.90, 2.11)	1.75 (0.92, 3.33)	0.60 (0.23, 1.54)
Neuromuscular	214 (0.5)	130 (0.7)	0.95 (0.43, 2.11)	0.90 (0.39, 2.10)	1.42 (0.33, 6.05)	0.72 (0.18, 2.85)

Variable	Frequencies		Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Stage 1 Development Sample N = 42,129 n (%)	Stage 2 Validation Sample N = 17,855 n (%)	Stage 1 Development Dataset	Stage 2 Validation Dataset	Stage 1 Development Dataset	Stage 2 Validation Dataset
Other Preeclampsia	4,278 (10.2)	2,126 (11.9)	1.42 (1.17, 1.73)	1.18 (0.88, 1.57)	1.38 (0.88, 2.17)	1.17 (0.64, 2.12)
Placenta Previa	179 (0.4)	114 (0.6)	4.84 (3.01, 7.76)	3.53 (1.88, 6.63)	1.17 (0.41, 3.31)	1.98 (0.60, 6.52)
Placental Abruption	402 (1.0)	184 (1.0)	3.53 (2.51, 4.96)	2.84 (1.71, 4.74)	2.15 (0.97, 4.78)	0.21 (0.03, 1.67)
Placental Accreta Spectrum	47 (0.1)	14 (0.1)	45.36 (21.71, 94.78)	31.32 (9.19, 106.67)	171.79 (77.38, 381.39)	35.01 (9.11, 134.50)
Preexisting Diabetes	637 (1.5)	337 (1.9)	1.43 (0.98, 2.09)	1.52 (0.95, 2.42)	1.85 (0.95, 3.60)	1.45 (0.73, 2.89)
Preterm Birth	2,893 (6.9)	2,253 (12.6)	1.41 (1.16, 1.72)	1.74 (1.36, 2.23)	2.32 (1.57, 3.44)	3.54 (2.29, 5.47)
Previous Cesarean	7,201 (17.1)	3,557 (19.9)	1.22 (1.04, 1.44)	1.73 (1.39, 2.15)	1.08 (0.75, 1.56)	1.30 (0.84, 2.01)
Pulmonary Hypertension	18 (0.0)	16 (0.1)	0.69 (0.12, 3.99)	16.05 (5.38, 47.89)	3.56 (0.63, 20.05)	8.60 (2.10, 35.27)
Renal Disease	110 (0.3)	113 (0.6)	3.34 (1.90, 5.87)	2.54 (1.39, 4.64)	3.66 (1.48, 9.07)	3.70 (1.60, 8.57)
Severe Preeclampsia	1,615 (3.8)	1,038 (5.8)	2.35 (1.82, 3.03)	2.18 (1.59, 2.98)	3.48 (2.11, 5.77)	2.62 (1.56, 4.42)
Substance Abuse	2,799 (6.6)	1,893 (10.6)	1.12 (0.89, 1.39)	0.97 (0.74, 1.27)	1.34 (0.84, 2.13)	0.95 (0.57, 1.56)
Thyrotoxicosis	150 (0.4)	60 (0.3)	0.60 (0.19, 1.94)	0.41 (0.06, 3.06)	1.09 (0.15, 8.03)	1.78 (0.23, 13.76)
Autoimmune Disease	108 (0.3)	82 (0.5)	2.60 (1.22, 5.53)	1.96 (0.84, 4.58)	NA[a]	NA[a]
HIV	53 (0.1)	49 (0.3)	1.47 (0.45, 4.86)	0.40 (0.05, 2.99)	NA[a]	NA[a]
Grouped: Autoimmune Disease or HIV[c]	160 (0.4)	131 (0.7)	NA[b]	NA[b]	1.80 (0.41, 7.90)	0.32 (0.04, 2.67)
Long Term Anticoagulant Use	136 (0.3)	57 (0.3)	1.27 (0.60, 2.69)	1.45 (0.47, 4.42)	NA[a]	NA[a]
Obstetrical VTE	37 (0.1)	5 (0.0)	0.71 (0.13, 3.91)	**	NA[a]	NA[a]
Grouped: Long Term Anticoagulant Use or Obstetrical VTE[c]	167 (0.4)	60 (0.3)	NA[b]	NA[b]	0.83 (0.21, 3.23)	1.26 (0.24, 6.68)
Vitals – Heart Rate	*	*	*	*	*	*
Result <110	35,700 (84.7)	10,601 (59.4)	REF	REF	REF	REF

Variable	Frequencies		Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Stage 1 Development Sample N = 42,129 n (%)	Stage 2 Validation Sample N = 17,855 n (%)	Stage 1 Development Dataset	Stage 2 Validation Dataset	Stage 1 Development Dataset	Stage 2 Validation Dataset
Result >=110	3,859 (9.2)	884 (5.0)	1.21 (0.99, 1.49)	1.89 (1.39, 2.57)	1.58 (1.04, 2.40)	1.90 (1.11, 3.25)
Missing	2,570 (6.1)	6,370 (35.7)	2.05 (0.97, 4.34)	0.74 (0.38, 1.45)	2.15 (0.38, 12.09)	1.44 (0.24, 8.51)
Vitals – Systolic BP	*	*	*	*	*	*
Result <140	33,382 (79.2)	9,168 (51.3)	REF	REF	REF	REF
Result >=140 & <160	5,050 (12.0)	1,610 (9.0)	1.16 (0.95, 1.40)	1.04 (0.77, 1.41)	0.83 (0.54, 1.29)	0.58 (0.33, 1.02)
Result >=160	1,178 (2.8)	486 (2.7)	1.31 (0.96, 1.78)	1.12 (0.73, 1.72)	0.42 (0.20, 0.90)	0.57 (0.28, 1.15)
Missing	2,519 (6.0)	6,591 (36.9)	0.76 (0.34, 1.67)	1.02 (0.53, 1.97)	0.82 (0.13, 5.04)	0.28 (0.05, 1.65)
Labs – Hematocrit	*	*	*	*	*	*
Result <33	7,929 (18.8)	4,323 (24.2)	2.54 (2.20, 2.93)	2.32 (1.90, 2.84)	1.22 (0.86, 1.73)	1.25 (0.84, 1.87)
Result >=33	28,911 (68.6)	13,196 (73.9)	REF	REF	REF	REF
Missing	5,289 (12.6)	336 (1.9)	1.14 (0.83, 1.56)	0.93 (0.42, 2.06)	0.71 (0.34, 1.49)	**
Labs – WBC	*	*	*	*	*	*
Result <14	29,472 (70.0)	15,457 (86.6)	REF	REF	REF	REF
Result >=14	4,879 (11.6)	2,062 (11.5)	1.15 (0.95, 1.40)	1.20 (0.91, 1.59)	1.47 (0.99, 2.18)	0.99 (0.59, 1.66)
Missing	7,778 (18.5)	336 (1.9)	0.65 (0.50, 0.84)	**	0.62 (0.34, 1.14)	**

** Cell intentionally left empty*

*** Odds ratios not estimable due to small sample size and distribution of values across variables*

[a]NA indicates odds ratios not calculated because variables were combined into a composite

[b]NA indicates odds ratios not calculated because variables appeared in model individually and were not combined into a composite

C Due to low prevalence of select risk variables, for the risk model of severe obstetric complication excluding transfusion-only encounters, Human Immunodeficiency Virus (HIV) was combined with autoimmune disease, and obstetric venous thromboembolism (VTE) was combined with long-term anticoagulant medication use.

Table 7 shows model performance statistics for the risk model of any severe obstetric complications and for the model of severe obstetric complications excluding blood transfusion-only encounters, validated using Stage 2 Beta testing data. The calculated C-statistic for the risk model for any severe obstetric complications was 0.77 when using the same risk variables to create a risk model with new beta coefficients in the Stage 2 validation dataset. The calculated C-statistic for the risk model for severe obstetric complications excluding blood transfusion-only encounters was 0.83. For both versions of the measure, the C-statistics indicate good model discrimination.

Both models show a slightly larger range between the lowest decile and highest decile of predicted ability compared to the Stage 1 development dataset, which indicates good model discrimination. Overall, these additional validation diagnostic results further support that the risk-adjustment model adequately controls for differences in patient characteristics.

Table 7. Model Performance Statistics for Risk Model for Both Severe Obstetric Complication Outcomes (Stage 2 Beta Testing)

Model Performance Statistic	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Stage 1 Development Dataset	Stage 2 Validation Dataset	Stage 1 Development Dataset	Stage 2 Validation Dataset
C-statistic	0.74 (0.72,0.76)	0.77 (0.75,0.79)	0.77 (0.73,0.81)	0.83 (0.79,0.87)
Predictive ability	(0.56%,9.56%)	(0.58%,12.66%)	(0.19%,2.58%)	(0.14%,4.54%)

For Stage 2 Beta testing in the calibration (risk decile) plots (Figure 5, Any Severe Obstetric Complications, and Figure 6 for Severe Obstetric Complications Excluding Blood Transfusion-Only Encounters), we visualize the agreement between observed severe obstetric complication rates and the risk predicted by the risk model. The plots show reasonable agreement between the observed severe obstetric complication rates with the model predictions demonstrating that the risk-adjustment model adequately fits the patient characteristics (case mix) data.

Figure 5. Calibration for Severe Obstetric Complications Rates by Predicted Risk Deciles (Stage 2 Beta Testing)

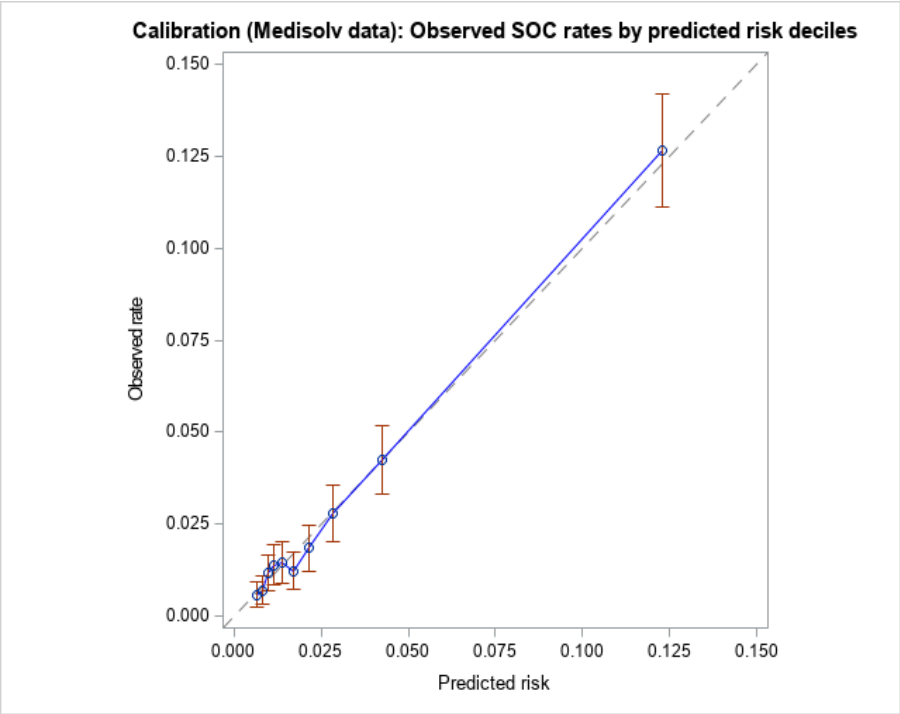
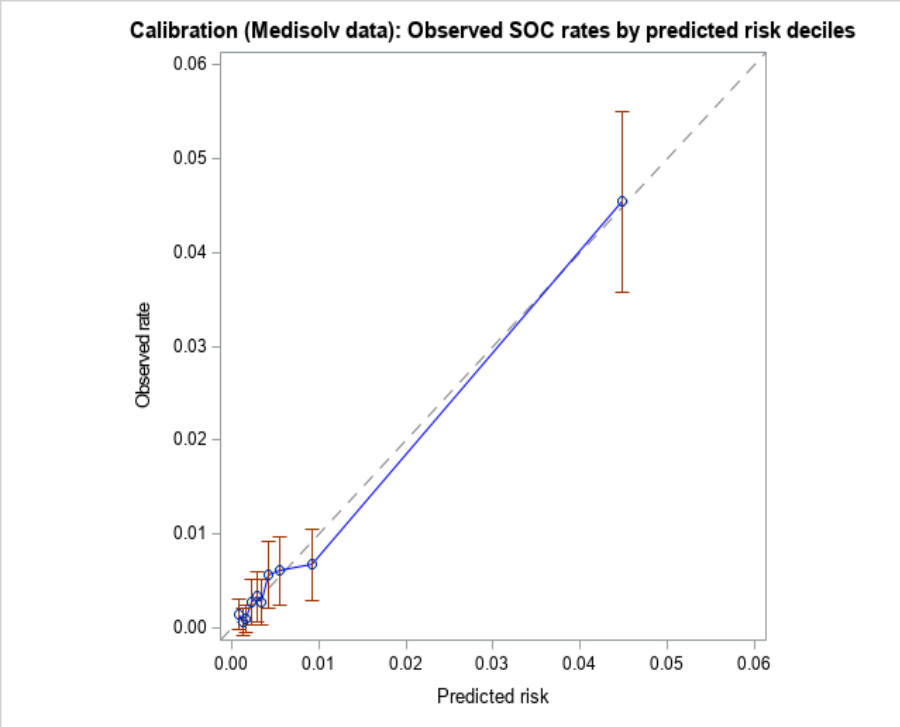


Figure 6. Calibration for Severe Obstetric Complications Excluding Blood Transfusion-Only Encounters Rates by Predicted Risk Deciles (Stage 2 Beta Testing)



After risk model development and validation in Stage 1 Beta testing data, and external validation in Stage 2 Beta testing data, the final analytic dataset used for measure score calculation was created by combining Stage 1 and Stage 2 Beta testing encounters from January 1, 2020, to December 31, 2020. Table 8 shows frequencies and adjusted odds ratios with 95% confidence intervals for the variables in the hierarchical risk model, which accounts for risk factors, the natural clustering of observations within hospitals, and hospital-specific random effects.

Table 8. Risk Variables w/Adjusted Odds Ratio for Risk Models for Both Severe Obstetric Complications Outcomes (30 Hospitals, Stage 1 and Stage 2 Beta Testing)

Variable	Full Sample = 69,018	Adjusted OR (95% CI)	
	N (%)	Any Severe Obstetric Complication(s)	Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters
Maternal Age in Years	*	*	*
<20	1,977 (2.9)	REF	REF
20-<25	9,961 (14.4)	0.96 (0.73, 1.27)	0.74 (0.38, 1.44)
25-<30	17,949 (26.0)	0.81 (0.61, 1.06)	0.96 (0.51, 1.82)
30-<35	23,100 (33.5)	0.78 (0.59, 1.02)	0.95 (0.50, 1.79)
35-<40	12,934 (18.7)	0.83 (0.62, 1.10)	0.86 (0.45, 1.68)
>=40	3,097 (4.5)	1.28 (0.92, 1.77)	1.40 (0.67, 2.90)
Anemia	12,590 (18.2)	1.79 (1.60, 2.00)	1.41 (1.09, 1.81)
Asthma	6,091 (8.8)	1.26 (1.08, 1.46)	2.13 (1.63, 2.78)
BMI >= 40	4,449 (6.4)	1.00 (0.83, 1.21)	1.35 (0.95, 1.91)
Bariatric Surgery	496 (0.7)	0.92 (0.55, 1.52)	0.62 (0.19, 2.03)
Bleeding Disorder	2,085 (3.0)	2.08 (1.69, 2.56)	2.86 (1.98, 4.14)
Cardiac Disease	1,199 (1.7)	1.65 (1.27, 2.13)	2.79 (1.88, 4.14)
Economic Housing Instability	83 (0.1)	1.35 (0.55, 3.34)	2.55 (0.73, 8.85)
Gastrointestinal Disease	1,140 (1.7)	1.23 (0.89, 1.70)	0.97 (0.49, 1.94)
Gestational Diabetes	6,427 (9.3)	0.98 (0.83, 1.17)	1.22 (0.88, 1.69)
Hypertension	3,042 (4.4)	1.03 (0.84, 1.25)	0.85 (0.57, 1.27)
Mental Health Disorder	10,134 (14.7)	1.22 (1.08, 1.38)	1.34 (1.04, 1.73)
Multiple Pregnancy	1,414 (2.0)	1.86 (1.47, 2.36)	1.03 (0.59, 1.78)
Neuromuscular	375 (0.5)	0.85 (0.46, 1.56)	0.75 (0.22, 2.52)
Other Preeclampsia	7,115 (10.3)	1.23 (1.05, 1.44)	1.40 (0.99, 1.96)
Placenta Previa	316 (0.5)	4.40 (3.04, 6.38)	1.55 (0.72, 3.33)
Placental Abruption	636 (0.9)	3.83 (2.95, 4.99)	1.90 (1.02, 3.53)
Placental Accreta Spectrum	73 (0.1)	45.95 (25.69, 82.16)	135.46 (72.98, 251.43)
Preexisting Diabetes	1,067 (1.5)	1.54 (1.17, 2.04)	1.72 (1.08, 2.75)
Preterm Birth	5,162 (7.5)	1.35 (1.15, 1.58)	2.21 (1.65, 2.95)

Variable	Full Sample = 69,018	Adjusted OR (95% CI)	
	N (%)	Any Severe Obstetric Complication(s)	Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters
Previous Cesarean	12,020 (17.4)	1.39 (1.23, 1.57)	1.18 (0.90, 1.54)
Pulmonary Hypertension	30 (0.0)	1.78 (0.56, 5.61)	2.30 (0.59, 9.02)
Renal Disease	198 (0.3)	2.94 (1.91, 4.52)	3.12 (1.63, 5.95)
Severe Preeclampsia	2,850 (4.1)	2.47 (2.04, 2.98)	4.06 (2.87, 5.74)
Substance Abuse	4,993 (7.2)	1.02 (0.86, 1.20)	1.09 (0.77, 1.54)
Thyrotoxicosis	243 (0.4)	0.47 (0.17, 1.29)	1.17 (0.28, 4.82)
Autoimmune Disease	197 (0.3)	2.32 (1.32, 4.08)	NA[a]
HIV	98 (0.1)	1.33 (0.57, 3.11)	NA[a]
Grouped: Autoimmune Disease or HIV[b]	294 (0.4)	NA[c]	1.28 (0.44, 3.71)
Long Term Anticoagulant Use	211 (0.3)	1.04 (0.56, 1.94)	NA[a]
Obstetrical VTE	53 (0.1)	0.60 (0.12, 3.03)	NA[a]
Grouped: Long Term Anticoagulant Use or Obstetrical VTE[c]	255 (0.4)	NA[c]	0.90 (0.32, 2.50)
Vitals – Heart Rate	*	*	*
Result <110	55,742 (80.8)	REF	REF
Result >=110	6,032 (8.7)	1.40 (1.20, 1.63)	1.53 (1.12, 2.09)
Missing	7,244 (10.5)	1.79 (1.02, 3.15)	2.01 (0.59, 6.81)
Vitals – Systolic BP	*	*	*
Result <140	51,817 (75.1)	REF	REF
Result >=140 & <160	8,027 (11.6)	1.16 (1.00, 1.35)	0.92 (0.67, 1.27)
Result >=160	1,882 (2.7)	1.23 (0.96, 1.57)	0.68 (0.42, 1.11)
Missing	7,292 (10.6)	0.65 (0.36, 1.15)	0.52 (0.15, 1.82)
Labs – Hematocrit	*	*	*
Result <33	13,571 (19.7)	2.62 (2.35, 2.93)	1.20 (0.92, 1.55)
Result >=33	47,758 (69.2)	REF	REF
Missing	7,689 (11.1)	1.36 (1.04, 1.78)	0.90 (0.51, 1.59)
Labs – WBC	*	*	*
Result <14	49,785 (72.1)	REF	REF
Result >=14	8,016 (11.6)	1.23 (1.06, 1.43)	1.35 (1.00, 1.82)
Missing	11,217 (16.3)	0.60 (0.47, 0.76)	0.71 (0.43, 1.18)

** Cell intentionally left empty*

[a] NA indicates odds ratios not calculated because variables were combined into a composite

[b]Due to low prevalence of select risk variables, for the risk model of severe obstetric complication excluding transfusion-only encounters, Human Immunodeficiency Virus (HIV) was combined with autoimmune disease, and obstetric venous thromboembolism (VTE) was combined with long-term anticoagulant medication use
[c]NA indicates odds ratios not calculated because variables appeared in model individually and were not combined into a composite

3.3.1 Social Risk Factor Assessment

Table 9 shows the distribution of delivery encounters and unadjusted severe obstetric complication rates by race/ethnicity across all Stage 1 and Stage 2 Beta testing locations. Non-Hispanic Black/African American patients have the highest observed (unadjusted) rates of severe obstetric complications and severe obstetric complications excluding blood transfusion-only encounters. Non-Hispanic patients who declined identifying race or for whom race is unknown have the lowest observed rate of any severe obstetric complications, and non-Hispanic patients of "Other" race have the lowest observed rates of severe obstetric complications excluding blood transfusion-only encounters.

Table 9. Observed Severe Obstetric Complication Rates among Race/Ethnicity Groups (30 Hospitals, Stage 1 and Stage 2 Beta Testing)

Race/Ethnicity	Denominator	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
		Numerator	Observed Outcome Rate Per 10,000	Numerator	Observed Outcome Rate Per 10,000
Unique Encounters	69,018	1,719	249	376	54
Hispanic	8,807	220	250	44	50
Non-Hispanic – Black/African American	14,218	515	362	99	70
Non-Hispanic – Asian/Pacific Islander	3,246	83	256	19	59
Non-Hispanic – White	39,060	816	209	197	50
Non-Hispanic – Other/ Multiple	1,546	33	213	4	26
Non-Hispanic – Declined/unknown	230	4	174	1	43

3.4 Measure Results

Table 10 provides the observed and the risk-standardized rate per 10,000 deliveries for severe obstetric complications and severe obstetric complications excluding blood transfusion-only encounters for each Stage 1 and Stage 2 Beta testing hospital and across all Stage 1 and Stage 2 Beta testing hospitals. The starting point for risk-standardized rates is the observed rate across all observations, which is then adjusted up or down according to the case mix of patients at one hospital relative to other hospitals. Therefore, a hospital with zero or very few cases may have a risk-standardized rate that is close to the average hospital rate and is much higher than the observed rate for that hospital. It is best to compare

the risk-standardized rates to the overall across hospital rate, and not to a given hospital's observed rate.

The variation in severe obstetric complication rates suggests that there could be meaningful differences in delivery of maternal care across Stage 1 and Stage 2 Beta testing hospitals.

Table 10. Observed and Risk-Standardized Severe Obstetric Complication Rates per 10,000 Delivery Hospitalizations (30 Hospitals, Stage 1 and Stage 2 Beta Testing)

Hospital	Delivery Encounters	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
		Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations
Hospital 1.1	496	202	244 (166, 358)	0	51 (32, 76)
Hospital 1.2	3,875	248	290 (222, 383)	52	59 (39, 77)
Hospital 1.3	1,518	158	222 (160, 328)	33	53 (36, 79)
Hospital 1.4	534	412	374 (243, 527)	19	52 (38, 72)
Hospital 1.5	2,383	105	169 (122, 232)	29	52 (38, 72)
Hospital 1.6	5,952	269	293 (226, 383)	54	57 (42, 76)
Hospital 1.7	1,678	244	325 (239, 426)	36	55 (41, 71)
Hospital 1.8	733	164	217(149, 319)	14	52 (35, 74)
Hospital 1.9	608	214	229 (163, 327)	16	51 (35, 74)
Hospital 1.10	293	171	240 (164, 345)	34	54 (38, 76)
Hospital 2	7,196	235	286 (228, 364)	72	70 (52, 91)
Hospital 3	7,955	303	301 (239, 377)	48	53 (40, 68)
Hospital 5.1	292	137	229 (144, 371)	0	52 (34, 77)
Hospital 5.2	224	179	263 (154, 484)	45	55 (35, 87)
Hospital 5.3	139	72	226 (135, 426)	0	54 (31, 86)
Hospital 5.4	347	144	246 (145, 410)	29	55 (38, 91)
Hospital 5.5	799	50	172 (100, 268)	13	52 (36, 73)
Hospital 5.6	163	0	201 (115, 348)	0	53 (35, 84)
Hospital 5.7	560	143	221 (140, 337)	18	53 (32, 81)
Hospital 5.8	3,316	305	287 (223, 380)	66	56 (44, 71)
Hospital 5.9	299	33	191 (113, 306)	33	55 (41, 82)
Hospital 6	3,359	104	166 (123, 232)	27	50 (33, 70)
Hospital 7	4,369	213	285 (222, 378)	41	53 (42, 70)
Hospital 9	3,918	202	342 (243, 477)	26	49 (35, 66)
Hospital 10	9,178	341	330 (263, 426)	81	58 (46, 75)
Hospital A	781	205	254 (174, 378)	13	51 (32, 72)
Hospital B	3,468	464	313 (246, 381)	161	72 (55, 105)
Hospital C	3,394	165	216 (168, 288)	47	54 (43, 72)
Hospital D	862	116	187 (123, 292)	12	50 (32, 71)

Hospital	Delivery Encounters	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
		Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations
Hospital E	329	304	305 (195, 502)	0	52 (36, 76)
Across All Encounters	**69,018**	**249**	*	**54**	*
Average Among Hospitals	*	*	**254**	*	**55**

** Cell intentionally left empty*

3.5 Reliability

3.5.1 Data Element Feasibility

Data element feasibility were assessed with virtual EHR walkthrough sessions conducted with each Alpha testing site. Each data element score was examined within each of four domains: data availability, data accuracy, data standards, and workflow. Subsequent to the fourth EHR Walkthrough, Joint Commission staff determined several of the sites were unable to accurately capture two main data elements: the timestamp for the procedure performed and the laboratory test result of the PaO2/FiO2 ratio. Joint Commission staff proposed to address these feasibility challenges by revising the draft specifications used for Alpha testing to better align with clinical intent and decrease burden for a lab result not commonly calculated in the EHR. Consequently, feasibility scores based on the revised specifications increased to 98%. All other data elements were assessed to be feasible and available.

Table 11 provides the data element feasibility rates prior to and following revision of draft measure specifications during Alpha testing. Feasibility Rate 1 reflects the rate inclusive of the timestamp for the procedure performed and the laboratory test result of the PaO2/FiO2 ratio. Feasibility Rate 2 reflects the rate with the revised specifications, using date only for procedures performed (no timestamp) and laboratory test results of PaO2. Feasibility Rate 1 was 95% overall; Feasibility Rate 2, at 98%, shows a very high rate of data element feasibility.

Among nine sites, overall data availability was 95% with a range of 87% to 97%. Overall data accuracy was 98% with range 94% to 100%. Overall data standards scores were 96% with range 87% to 100%. Overall workflow feasibility was 99% with range 94% to 100% (Table 12).

The Severe Obstetric Complications eCQM is feasible to implement because required data are routinely collected as part of clinical care and are extractable from electronic health records. Feasibility testing revealed that race and ethnicity data elements are routinely collected; however, there is not standardization amongst hospitals.

Table 11. Feasibility Rate (9 Alpha Testing Sites, 75 data elements)

Test Sites	Feasibility Rate 1 Initial	Feasibility Rate 2 Revised
Site 1	97%	97%
Site 2	87%	94%
Site 3	97%	100%
Site 4	97%	97%
Site 5	96%	98%
Site 6	91%	100%
Site 7	97%	100%
Site 8	97%	100%
Site 9	90%	99%
Overall	95%	98%

Table 12. Feasibility Rates by Domain (9 Alpha Testing Sites, 75 data elements)

Test Sites	Data Availability	Data Accuracy	Data Standards	Workflow
Site 1	97%	97%	87%	100%
Site 2	87%	94%	94%	94%
Site 3	97%	100%	100%	100%
Site 4	97%	97%	96%	99%
Site 5	96%	98%	94%	99%
Site 6	91%	100%	100%	100%
Site 7	97%	100%	100%	100%
Site 8	97%	100%	100%	100%
Site 9	90%	99%	96%	100%
Overall	95%	98%	96%	99%

3.5.2 Measure Score Reliability

The signal-to-noise ratio was calculated to assess how well the measure can distinguish the performance of one hospital from another. Results for hospitals in Stage 1 and Stage 2 Beta testing are presented in Table 13.

Signal-to-noise reliability was calculated for the 30 individual hospitals at two volume thresholds; results are provided for hospitals with at least 25 delivery encounters in the year (all hospitals included in testing) and for hospitals with at least 200 delivery encounters in the year (28 of the 30 hospitals included in testing). For hospitals with at least 25 delivery encounters, the mean (SD) reliability score was 0.946 (0.055) for any severe obstetric complication outcome and 0.900 (0.095) for severe obstetric complications excluding blood transfusion-only encounters. The signal-to-noise reliability is higher when included hospitals had at least 200 delivery encounters in a year, rather than 25 delivery encounters, particularly for the second outcome (severe complications excluding blood transfusion-only encounters). The mean (SD) reliability score among hospitals with at least 200 delivery encounters was 0.956 (0.041) for any severe obstetric complication outcome and 0.917 (0.074) for severe obstetric complications excluding blood transfusion-only encounters.

Results at the hospital level, using a threshold for hospital inclusion of at least 25 delivery encounters per hospital, indicate very high reliability for the outcome measuring any severe obstetric complications, and a lower reliability for the outcome measuring severe obstetric complications excluding blood transfusion-only encounters. Setting a minimum threshold of at least 200 delivery encounters per hospital increases reliability, particularly for the severe obstetric complications excluding transfusion-only encounters, which impacts fewer patients and represents a rarer outcome.

Our interpretation of these results is based on standards established by Landis and Koch:[45]

- <0 = Less than chance agreement
- 0 – 0.2 = Slight agreement
- 0.21 – 0.39 = Fair agreement
- 0.4 – 0.59 = Moderate agreement
- 0.6 – 0.79 = Substantial agreement
- 0.8 – 0.99 = Almost Perfect agreement
- 1 = Perfect agreement

Table 13. Summary Statistics of Signal-to-Noise-Reliability of Hospital Measure Scores for Both Severe Obstetric Complication Outcomes (30 Hospitals, Stage 1 and Stage 2 Beta Testing)

Outcome	Volume Threshold (Number of Delivery Encounters per Hospital per year)	# Of Hospitals	Median	Mean (SD)	Minimum	Maximum	Interquartile Range	
							Q1	Q3
Any Severe Obstetric Complication(s)	> 25	30	0.958	0.946 (0.055)	0.792	0.996	0.905	0.995
Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	> 25	30	0.918	0.900 (0.095)	0.652	0.992	0.824	0.991
Any Severe Obstetric Complication(s)	> 200	28	0.968	0.956 (0.041)	0.860	0.996	0.934	0.995
Severed Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	> 200	28	0.937	0.917 (0.074)	0.751	0.992	0.874	0.991

3.6 Validity

3.6.1 Data Element Validity

Data element validity testing was completed for six Stage 1 Beta testing sites, one system of 10 hospitals and five individual hospitals. 30 to 36 sampled cases were examined per test site. Overall, the data element agreement rate for all six sites was 90.4%, indicating excellent agreement (Table 14).

Table 14. Data Element Agreement Rates (6 Sites, Stage 1 Beta Testing)

*	Data Element Name	Site 1 Match/N (Rate)	Site 2 Match/N (Rate)	Site 3 Match/N (Rate)	Site 6 Match/N (Rate)	Site 7 Match/N (Rate)	Site 9 Match/N (Rate)	Total Match/N (Rate)
Demographics	DOB	36/36 (100.0%)	31/31 (100.0%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	204/204 (100.0%)
	ONC Administrative Sex Code	36/36 (100.0%)	31/31 (100.0%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	204/204 (100.0%)
	Race	36/36 (100.0%)	31/31 (100.0%)	35/35 (100.0%)	35/35 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	203/203 (100.0%)
	Ethnicity	36/36 (100.0%)	31/31 (100.0%)	35/35 (100.0%)	35/35 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	203/203 (100.0%)
	Payer	32/36 (88.9%)	31/31 (100.0%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	200/204 (98.0%)
	Admission Source	34/36 (94.4%)	31/31 (100.0%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	202/204 (99.0%)
	Discharge Disposition	35/36 (97.2%)	30/31 (96.8%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	202/204 (99.0%)
Encounter History	Encounter, Performed: Encounter Inpatient	36/36 (100.0%)	31/31 (100.0%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	204/204 (100.0%)
	Admission Date Time (Relevant Period Start Time)	36/36 (100.0%)	30/31 (96.8%)	35/35 (100.0%)	35/36 (97.2%)	1/30 (3.3%)	36/36 (100.0%)	173/204 (84.8%)
	Discharge Date Time (Relevant Period End Time)	36/36 (100.0%)	31/31 (100.0%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	36/36 (100.0%)	204/204 (100.0%)
	Encounter, Performed: Emergency Department Visit	6/6 (100.0%)	12/12 (100.0%)	16/16 (100.0%)	0/0	0/0	3/3 (100.0%)	37/37 (100.0%)
	ED Start Date Time (relevant Period)	6/6 (100.0%)	12/12 (100.0%)	16/16 (100.0%)	0/0	0/0	3/3 (100.0%)	37/37 (100.0%)
	ED End Date Time (relevant Period)	6/6 (100.0%)	12/12 (100.0%)	16/16 (100.0%)	0/0	0/0	3/3 (100.0%)	37/37 (100.0%)
	Encounter, Performed: Preadmission Observation (PreAdmObs) Undelivered Mother	0/0	0/7 (0.0%)	0/0	9/9 (100.0%)	0/29 (0.0%)	36/36 (100.0%)	45/81 (55.6%)
	PreAdmObs Start Date Time (relevant Period)	0/0	0/7 (0.0%)	0/0	9/9 (100.0%)	0/29 (0.0%)	36/36 (100.0%)	45/81 (55.6%)

*	Data Element Name	Site 1	Site 2	Site 3	Site 6	Site 7	Site 9	Total
		Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)
Encounter History	PreAdmObs End Date Time (relevant Period)	0/0	0/7 (0.0%)	0/0	9/9 (100.0%)	0/29 (0.0%)	36/36 (100.0%)	45/81 (55.6%)
	Encounter, Performed: Observation Services	25/27 (92.6%)	1/1 (100.0%)	0/0	1/1 (100.0%)	0/0	36/36 (100.0%)	63/65 (96.9%)
	Obs Start Date Time (relevant Period)	25/27 (92.6%)	1/1 (100.0%)	0/0	1/1 (100.0%)	0/0	36/36 (100.0%)	63/65 (96.9%)
	Obs End Date Time (relevant Period)	25/27 (92.6%)	1/1 (100.0%)	0/0	1/1 (100.0%)	0/0	36/36 (100.0%)	63/65 (96.9%)
	Facility Locations: Intensive Care Unit Code	2/2 (100.0%)	0/0	5/5 (100.0%)	1/1 (100.0%)	0/0	0/0	8/8 (100.0%)
	ICU Start Date Time	2/2 (100.0%)	0/0	5/5 (100.0%)	1/1 (100.0%)	0/0	0/0	8/8 (100.0%)
	ICU End Date Time	2/2 (100.0%)	0/0	5/5 (100.0%)	1/1 (100.0%)	0/0	0/0	8/8 (100.0%)
Dx	Diagnosis POA	245/391 (62.7%)	397/397 (100.0%)	312/312 (100.0%)	208/346 (60.1%)	319/319 (100.0%)	327/328 (99.7%)	1808/2093 (86.4%)
	Diagnosis code	391/391 (100.0%)	397/397 (100.0%)	312/312 (100.0%)	346/346 (100.0%)	319/319 (100.0%)	328/328 (100.0%)	2093/2093 (100.0%)
Procedure	Procedure code & date	103/104 (99.0%)	140/142 (98.6%)	114/115 (99.1%)	103/103 (100.0%)	93/93 (100.0%)	78/79 (98.7%)	631/636 (99.2%)
Blood	Blood Transfusion code	33/33 (100.0%)	27/148 (18.2%)	31/31 (100.0%)	31/31 (100.0%)	10/15 (66.7%)	19/20 (95.0%)	151/278 (54.3%)
	Blood Transfusion start	33/33 (100.0%)	25/141 (17.7%)	31/31 (100.0%)	31/31 (100.0%)	10/15 (66.7%)	19/20 (95.0%)	149/271 (55.0%)
	Blood Transfusion end	25/33 (75.8%)	24/138 (17.4%)	31/31 (100.0%)	30/30 (100.0%)	5/15 (33.3%)	19/20 (95.0%)	134/267 (50.2%)
Delivery Details	Relevant Date Time Assessment, Performed: Date and time of obstetric delivery	35/36 (97.2%)	30/31 (96.8%)	34/35 (97.1%)	36/36 (100.0%)	28/30 (93.3%)	36/36 (100.0%)	199/204 (97.5%)
	Result: Date and time of obstetric delivery	35/36 (97.2%)	30/31 (96.8%)	34/35 (97.1%)	36/36 (100.0%)	28/30 (93.3%)	36/36 (100.0%)	199/204 (97.5%)

*	Data Element Name	Site 1	Site 2	Site 3	Site 6	Site 7	Site 9	Total
		Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)
Delivery Details	Relevant Date Time Assessment, Performed: Delivery date Estimated	34/36 (94.4%)	0/31 (0.0%)	34/35 (97.1%)	36/36 (100.0%)	28/30 (93.3%)	36/36 (100.0%)	168/204 (82.4%)
	Result: Delivery date Estimated	34/36 (94.4%)	30/31 (96.8%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	34/36 (94.4%)	199/204 (97.5%)
	Relevant Date Time Assessment, Performed: Estimated Gestational Age at Delivery	35/36 (97.2%)	24/30 (80.0%)	34/35 (97.1%)	36/36 (100.0%)	28/30 (93.3%)	34/36 (94.4%)	191/203 (94.1%)
	Result: Estimated Gestational Age at Delivery	36/36 (100.0%)	24/31 (77.4%)	35/35 (100.0%)	36/36 (100.0%)	30/30 (100.0%)	34/36 (94.4%)	195/204 (95.6%)
Laboratory Results	Creatinine Result Date Time	0/0	0/0	2/2 (100.0%)	0/0	1/1 (100.0%)	0/0	3/3 (100.0%)
	Creatinine Result	0/0	0/0	2/2 (100.0%)	0/0	1/1 (100.0%)	0/0	3/3 (100.0%)
	PaO2 Result Date Time	0/0	1/1 (100.0%)	0/0	0/0	0/0	2/10 (20.0%)	3/11 (27.3%)
	PaO2 Result	0/0	1/1 (100.0%)	0/0	0/0	0/0	2/10 (20.0%)	3/11 (27.3%)
	Platelet Result Date Time	9/9 (100.0%)	4/4 (100.0%)	3/3 (100.0%)	9/9 (100.0%)	8/9 (88.9%)	2/2 (100.0%)	35/36 (97.2%)
	Platelet Result	9/9 (100.0%)	4/4 (100.0%)	3/3 (100.0%)	9/9 (100.0%)	8/9 (88.9%)	2/2 (100.0%)	35/36 (97.2%)
	Hemoglobin Result Date Time	117/117 (100.0%)	98/99 (99.0%)	89/89 (100.0%)	108/109 (99.1%)	69/92 (75.0%)	50/50 (100.0%)	531/556 (95.5%)
	Hemoglobin Result	117/117 (100.0%)	98/99 (99.0%)	89/89 (100.0%)	108/109 (99.1%)	71/92 (77.2%)	50/50 (100.0%)	533/556 (95.9%)
	Hematocrit Result Date Time	117/117 (100.0%)	97/99 (98.0%)	93/93 (100.0%)	108/109 (99.1%)	70/92 (76.1%)	111/112 (99.1%)	596/622 (95.8%)
	Hematocrit Result	117/117 (100.0%)	98/99 (99.0%)	93/93 (100.0%)	108/109 (99.1%)	70/92 (76.1%)	111/112 (99.1%)	597/622 (96.0%)
	WBC Result Date Time	105/105 (100.0%)	97/99 (98.0%)	92/92 (100.0%)	108/109 (99.1%)	70/92 (76.1%)	49/49 (100.0%)	521/546 (95.4%)
	WBC Result	105/105 (100.0%)	98/99 (99.0%)	92/92 (100.0%)	108/109 (99.1%)	70/92 (76.1%)	49/49 (100.0%)	522/546 (95.6%)
	Glucose Result Date Time	19/19 (100.0%)	16/32 (50.0%)	31/31 (100.0%)	27/28 (96.4%)	1/9 (11.1%)	16/28 (57.1%)	110/147 (74.8%)
Laboratory Results	Glucose Result	19/19 (100.0%)	16/32 (50.0%)	31/31 (100.0%)	27/28 (96.4%)	1/9 (11.1%)	16/28 (57.1%)	110/147 (74.8%)
	Bicarbonate Result Date Time	0/11 (0.0%)	6/6 (100.0%)	27/27 (100.0%)	0/26 (0.0%)	5/6 (83.3%)	14/14 (100.0%)	52/90 (57.8%)

*	Data Element Name	Site 1	Site 2	Site 3	Site 6	Site 7	Site 9	Total
		Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)
	Bicarbonate Result	0/11 (0.0%)	6/6 (100.0%)	27/27 (100.0%)	0/26 (0.0%)	5/6 (83.3%)	14/14 (100.0%)	52/90 (57.8%)
Vital Signs	Relevant Date Time Physical Exam, Performed: Oxygen saturation in Arterial blood by Pulse oximetry (%)	4/35 (11.4%)	19/27 (70.4%)	34/34 (100.0%)	34/34 (100.0%)	29/29 (100.0%)	31/31 (100.0%)	151/190 (79.5%)
	Result: Oxygen saturation	34/35 (97.1%)	21/27 (77.8%)	34/34 (100.0%)	34/34 (100.0%)	29/29 (100.0%)	31/31 (100.0%)	183/190 (96.3%)
	Relevant Date Time Physical Exam, Performed: Heart rate (BPM)	12/36 (33.3%)	19/31 (61.3%)	35/35 (100.0%)	31/35 (88.6%)	28/30 (93.3%)	36/36 (100.0%)	161/203 (79.3%)
	Result: Heart rate	36/36 (100.0%)	23/31 (74.2%)	35/35 (100.0%)	31/35 (88.6%)	28/30 (93.3%)	36/36 (100.0%)	189/203 (93.1%)
	Relevant Date Time Physical Exam, Performed: Systolic blood pressure (mmHg)	12/36 (33.3%)	19/31 (61.3%)	35/35 (100.0%)	31/35 (88.6%)	28/30 (93.3%)	36/36 (100.0%)	161/203 (79.3%)
	Result: Systolic blood pressure	36/36 (100.0%)	23/31 (74.2%)	35/35 (100.0%)	31/35 (88.6%)	28/30 (93.3%)	36/36 (100.0%)	189/203 (93.1%)
	Relevant Date Time Physical Exam, Performed: Respiratory rate (breaths per minute)	10/36 (27.8%)	19/31 (61.3%)	35/35 (100.0%)	22/24 (91.7%)	29/30 (96.7%)	36/36 (100.0%)	151/192 (78.6%)
	Result: Respiratory rate	35/36 (97.2%)	23/31 (74.2%)	35/35 (100.0%)	22/24 (91.7%)	29/30 (96.7%)	36/36 (100.0%)	180/192 (93.8%)

*	Data Element Name	Site 1	Site 2	Site 3	Site 6	Site 7	Site 9	Total
		Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)	Match/N (Rate)
Vital Signs	Relevant Date Time Physical Exam, Performed: Body temperature (degrees Fahrenheit or degrees Celsius)	7/36 (19.4%)	19/31 (61.3%)	35/35 (100.0%)	29/32 (90.6%)	29/30 (96.7%)	36/36 (100.0%)	155/200 (77.5%)
	Result: Body temperature	36/36 (100.0%)	23/31 (74.2%)	35/35 (100.0%)	29/32 (90.6%)	29/30 (96.7%)	36/36 (100.0%)	188/200 (94.0%)
*	**Totals**	**2,447/2,780 (88.0%)**	**2,343/2,900 (80.8%)**	**2,472/2,477 (99.8%)**	**2,369/2,594 (91.3%)**	**1,935/2,243 (86.3%)**	**2,423/2,476 (97.9%)**	**13,989/15,470 (90.4%)**

** Cell intentionally left empty*

3.6.2 Measure Score Validity

Measure score validity was assessed in both Stage 1 and Stage 2 Beta testing.

3.6.2.1 Stage 1 Beta Testing

Measure score validity testing during Stage 1 Beta testing was conducted in the same six sites as data element reliability and validity. Table 15 displays the PPV (agreement rate) for the numerator among delivery encounters clinically adjudicated in Stage 1 Beta testing. The PPV rate was 100% at Sites 1, 2, 3, 6, and 7, and 70% at Site 9, with an overall PPV of 94.74%. In almost all delivery encounters with a numerator event adjudicated, the delivery encounters with a severe obstetric complication in the EHR data were shown to have a severe obstetric complication in the chart abstracted data, indicating strong measure validity. Although we do not always expect perfect agreement, as we expect some degree of human error in entering and matching values, we consider these PPVs to show excellent measure score validity. The absence of a perfect PPV does not threaten validity as we do not expect any systematic error in this small amount of disagreement across hospitals that might bias the measure results.

Table 15. Agreement Statistics for Measure Numerator between EHR Extraction and Manual Chart Abstraction (PPV) (6 Sites, Stage 1 Beta Testing)

Test Sites	# Of Numerator Events Verified by Clinical Adjudication	# Of Numerator Events from EHR	Positive Predictive Value (PPV)
Site 1	20	20	100%
Site 2	16	16	100%
Site 3	20	20	100%
Site 6	20	20	100%
Site 7	18	18	100%
Site 9	14	20	70.00%
Across 6 Sites	108	114	94.74%

Table 16 displays the sensitivity, specificity, and negative predictive value (NPV) calculated using Stage 1 Beta testing adjudication results. Estimated specificity and sensitivity are high. Estimated sensitivity is 100% in all sites and estimated specificity is 100% in Sites 1, 2, 3, 6, and 7, while at 62.5% in Site 9. NPV was 100% in all sites, showing that when EHR data indicated a severe obstetric complication did not occur, 100% of the time the chart abstraction confirmed a harm did not occur.

Table 16. Measure Score Validity Statistics Between EHR Extraction and Manual Chart Abstraction (Sensitivity, Specificity, NPV) (6 Sites, Stage 1 Beta Testing) (N=114)

Test Sites	Sensitivity	Specificity	Negative Predictive Value (NPV)
Site 1	100%	100%	100%
Site 2	100%	100%	100%
Site 3	100%	100%	100%
Site 6	100%	100%	100%
Site 7	100%	100%	100%
Site 9	100%	62.50%	100%
Across 6 Sites	100%	90.48%	100%

Table 17 provides the measure outcomes agreement rates and kappa scores for the Stage 1 Beta testing clinical adjudication sites. These data indicate overall 91.2% agreement with a kappa score of .881, indicating very good agreement.

Table 17. Measure Outcome Agreement Rates (6 Sites, Stage 1 Beta Testing)

Test Site	N	Agreement Rate	kappa
Site 1	36	97.2%	.963
Site 2	31	83.9%	.786
Site 3	35	94.3%	.922
Site 6	36	97.2%	.963
Site 7	30	96.7%	.953
Site 9	36	77.8%	.703
Total	204	91.2%	.881

3.6.2.2 Stage 2 Beta Testing

For clinical adjudication during Stage 2 Beta testing, a total of 275 numerator encounters were adjudicated: 139 encounters that were transfusion-only encounters, and all 136 encounters with severe obstetric complications from the five hospitals that were not transfusion-only numerator encounters. Adjudication of 179 denominator-only encounters out of the 17,855 overall denominator encounters was conducted with patients identified with high-risk conditions for severe obstetric complications in order to maximize the likelihood of identifying false negatives.

3.6.2.3 Numerator

Table 18 and Table 19 provide clinical adjudication results for each severe obstetric complication condition identified in the EHR data (including adjudication of each numerator event identified for encounters with multiple numerator events). In Table 18, the results are provided at the hospital-level; each potential adjudication response is identified in Columns A through F. The overall agreement rate at

the numerator event-level was 98.48% (Table 18). Event-specific positive predictive values (PPV) are shown in Table 19**.**

Table 18. Numerator Adjudication – All Severe Obstetric Complication Numerator Events Adjudicated Individually

Hospital	Numerator Event Verified in Medical Chart Review	Numerator Event Not Verified in Medical Chart			Total	Agreement
		Event Occurred Prior to Delivery Encounter[a]	Event Did Not Occur (Possible Miscoding of Related Event)[b]	No Evidence of Event[c]		
Hospital A	21	0	0	0	21	100.00%
Hospital B	274	3	0	1	278	98.56%
Hospital C	69	0	1	1	71	97.18%
Hospital D	14	0	0	0	14	100.00%
Hospital E	10	0	0	0	10	100.00%
Total	**388**	**3**	**1**	**2**	**394**	**98.48%**

[a] *Condition and/or procedure occurred prior to the delivery hospitalization and was present upon admission*
[b] *Condition and/or procedure did not occur, but there is another related event that may have been miscoded*
[c] *There is no evidence of the procedure and/or condition for this patient*

Table 19. Positive Predictive Value (PPV) – All Numerator Events Adjudicated Individually

Numerator Event	# Of Numerator Events Verified by Clinical Adjudication	# Of Numerator Events from EHR Data	Positive Predictive Value (PPV)
Overall, Across All Five Hospitals	**388**	**394**	**98.48%**
Acute Heart Failure	4	5	80.00%
Acute Myocardial Infarction	1	1	100.00%
Acute Renal Failure	65	66	98.48%
Acute Respiratory Distress Syndrome	18	18	100.00%
Air and Thrombotic Embolism	1	1	100.00%
Amniotic Fluid Embolism	2	2	100.00%
Aortic Aneurysm	NA	0	NA
Blood Transfusion	190	190	100.00%
Cardiac Arrest Ventricular Fibrillation	4	4	100.00%
Conversion of Cardiac Rhythm	1	2	50.00%
Disseminated Intravascular Coagulation	21	21	100.00%
Eclampsia	3	3	100.00%
Heart Failure – Cardiac Arrest	NA	0	NA
Hysterectomy	13	13	100.00%
Mortality	3	3	100.00%
Puerperal Cerebrovascular Disorder	1	1	100.00%
Pulmonary Edema	9	9	100.00%
Sepsis	11	11	100.00%
Severe Anesthesia Complications	NA	0	NA
Shock	27	29	93.10%
Sickle Cell Disease with Crisis	1	1	100.00%
Tracheostomy	1	1	100.00%
Ventilation	12	13	92.31%

Table 20 provides clinical adjudication results at the hospital level for each delivery encounter identified as having at least one severe obstetric complication condition identified in EHR data. At the numerator encounter-level, the overall PPV was 98.91%.

Table 20. Positive Predictive Value (PPV) – Numerator Encounters

Hospital	Number of Numerator Encounters Verified by Clinical Adjudication	Number of Numerator Encounters Clinically Adjudicated (Identified as Numerator Encounters in EHR data)	Positive Predictive Value (PPV)
Hospital A	18	18	100.00%
Hospital B	180	183	98.36%
Hospital C	51	51	100.00%
Hospital D	13	13	100.00%
Hospital E	10	10	100.00%
Total	**272**	**275**	**98.91%**

3.6.2.4 Denominator-Only

Table 21 provides clinical adjudication results at the hospital level for each denominator-only encounter that was adjudicated; the overall negative predictive value was 95.53%. Since denominator-only cases at high risk of complications were selected for adjudication, we would expect the negative predictive value of all denominator eligible delivery encounters to be higher than 95.53%.

Table 21. Negative Predictive Value (NPV) – Denominator-Only Encounters

Hospital	# Of Denominator-Only Encounters Verified by Clinical Adjudication	# Of Denominator-Only Cases Clinically Adjudicated (Identified as Denominator Encounters in EHR data)[a]	Negative Predictive Value (NPV)
Hospital A	7	7	100.00%
Hospital B	100	107	93.46%
Hospital C	53	54	98.15%
Hospital D	10	10	100.00%
Hospital E	1	1	100.00%
Total	**171**	**179**	**95.53%**

[a] *The following severe obstetric complications were found among the 8 encounters identified by EHR data as denominator only encounters:*
Acute renal failure (n=5)
Eclampsia
Seizure, intubation, eclampsia
Pulmonary edema

Table 22 displays the estimated sensitivity and specificity calculated using Stage 2 Beta testing adjudication results. Given that we oversampled high risk denominator-only encounters, we anticipate the true NPV to be higher than what was calculated. If there are a greater proportion of true negatives compared to false negatives in the Stage 2 Beta testing dataset than in the sample we adjudicated, we would expect both the sensitivity and specificity in the Stage 2 Beta testing dataset to be higher than what we found in the adjudication sample. Estimated sensitivity is 100% in Hospitals A, D, and E, while 96.26% and 98.08% for Hospitals B and C, respectively. Estimated specificity is 100% in Hospitals A, C, D, and E, while 97.09% in Hospital B.

Table 22. Measure Score Validity Statistics (Sensitivity, Specificity) (5 Hospitals, Stage 2 Beta Testing; Estimated from 454 Clinically Adjudicated Numerator and Denominator-Only Encounters)

Site ID	Sensitivity	Specificity
Hospital A	100%	100%
Hospital B	96.26%	97.09%
Hospital C	98.08%	100%
Hospital D	100%	100%
Hospital E	100%	100%
Total	97.14%	98.28%

3.6.2.5 Face Validity

Fifteen TEP members and five Patient Working Group members completed the face validity survey. TEP members rated the importance, reliability and validity, feasibility, and usability of the measure, as well as the ability of the measure to help distinguish better and worse quality of care at hospitals. Patient Working Group members rated the importance of the measure as well as the ability of the measure to help distinguish better and worse quality of care at hospitals.

Most TEP members (12/15) and all Patient Working Group members (5/5) strongly agreed that the measure was important. More than half of TEP members (8/15) strongly agreed that the measure was usable; the majority of TEP members moderately or strongly agreed that the measure was reliable and valid (13/15), feasible (10/15), and that it would be able to help distinguish better and worse quality of care at hospitals (11/15). More than half of Patient Working Group (3/5) members strongly agreed that the measure can distinguish quality of care at hospitals (Table 23).

Table 22. Results of Face Validity Survey – Questions and Frequency of Ratings Among TEP (N=15) and Patient Working Group (N=5) Members

Statements - Respondents	Strongly Agree	Moderately Agree	Somewhat Agree	Somewhat Disagree	Moderately Disagree	Strongly Disagree
Statement 1: Importance – TEP	12	3	0	0	0	0
Statement 1: Importance – Patient Working Group	5	0	0	0	0	0
Statement 2: Reliability/Validity – TEP	3	10	2	0	0	0
Statement 3: Feasibility – TEP	6	4	2	2	1	0
Statement 4: Usability – TEP	8	5	2	0	0	0
Statement 5: Quality – TEP	5	6	1	3	0	0
Statement 5: Quality – Patient Working Group	3	2	0	0	0	0

3.7 Denominator Exclusion – COVID

Following completion of Alpha and Beta testing, a denominator exclusion for which persons with a COVID-19 diagnosis at admission who also had at least one diagnosis code for respiratory distress or a respiratory procedure was considered for measure specifications. Measure testing results, including measure scores, reported in earlier sections of this document do not reflect this exclusion. Analyses were conducted to explore this COVID-19 exclusion. Table 24 provides the frequency distribution of COVID exclusions for all hospitals in Stage 1 and Stage 2 Beta Testing. Table 25 provides the impact the COVID exclusion had on measure scores for all hospitals in Stage 1 and Stage 2 Beta Testing.

Table 23. Distribution of COVID Exclusions by Hospital (30 Hospitals, Stage 1 and Stage 2 Beta Testing)

Hospital	Denominator			Any Severe Obstetric Complication(s)			Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters		
	Delivery Encounters	# Of Encounters that would meet the Denominator Exclusions	Denominator Exclusions %	Patients with a Numerator Event	# Of Patients that would be Excluded from Numerator	Exclusions %	Patients with a Numerator Event	# Of Patients that would be Excluded from Numerator	Exclusions %
1.1	496	0	0.00	10	0	0.00	0	0	*
1.2	3,875	5	0.13	96	3	3.13	20	3	15.00
1.3	1,518	1	0.07	24	0	0.00	5	0	0.00
1.4	534	0	0.00	22	0	0.00	1	0	0.00
1.5	2,383	4	0.17	25	1	4.00	7	1	14.29
1.6	5,952	6	0.10	160	2	1.25	32	2	6.25
1.7	1,678	4	0.24	41	0	0.00	6	0	0.00
1.8	733	0	0.00	12	0	0.00	1	0	0.00
1.9	608	0	0.00	13	0	0.00	1	0	0.00
1.10	293	0	0.00	5	0	0.00	1	0	0.00
2	7,196	3	0.04	169	1	0.59	52	1	1.92
3	7,955	6	0.08	241	4	1.66	38	4	10.53
5.1	292	1	0.34	4	0	0.00	0	0	*
5.2	224	0	0.00	4	0	0.00	1	0	0.00
5.3	139	1	0.72	1	0	0.00	0	0	*
5.4	347	0	0.00	5	0	0.00	1	0	0.00
5.5	799	0	0.00	4	0	0.00	1	0	0.00
5.6	163	0	0.00	0	0	*	0	0	*
5.7	560	0	0.00	8	0	0.00	1	0	0.00
5.8	3,316	4	0.12	101	1	0.99	22	1	4.55
5.9	299	0	0.00	1	0	0.00	1	0	0.00
6	3,359	0	0.00	35	0	0.00	9	0	0.00
7	4,369	1	0.02	93	0	0.00	18	0	0.00
9	3,918	1	0.03	79	1	1.27	10	1	10.00

Hospital	Denominator			Any Severe Obstetric Complication(s)			Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters		
	Delivery Encounters	# Of Encounters that would meet the Denominator Exclusions	Denominator Exclusions %	Patients with a Numerator Event	# Of Patients that would be Excluded from Numerator	Exclusions %	Patients with a Numerator Event	# Of Patients that would be Excluded from Numerator	Exclusions %
10	9,178	0	0.00	313	0	0.00	74	0	0.00
Hospital A	781	0	0.00	16	0	0.00	1	0	0.00
Hospital B	3,468	7	0.20	161	5	3.11	56	5	8.93
Hospital C	3,394	0	0.00	56	0	0.00	16	0	0.00
Hospital D	862	1	0.17	10	1	10.00	1	0	0.00
Hospital E	329	0	0.00	10	0	0.00	0	0	0.00
Total	69,018	45	0.065	1,719	19	1.11	376	18	4.79

**Estimate undefined because denominator for calculation is zero*

Table 24. COVID Exclusions: Impact on Measure Scores, by Hospital (30 Hospitals, Stage 1 and Stage 2 Beta Testing)

Hospital	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Without Exclusion	With COVID Denominator Exclusion	Without Exclusion	With COVID Denominator Exclusion
	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations
1.1	244 (166, 358)	241 (162, 360)	51 (32, 76)	48 (31, 77)
1.2	289 (222, 383)	280 (215, 369)	59 (39, 77)	53 (35, 75)
1.3	221 (160, 328)	220 (160, 330)	53 (36, 79)	50 (34, 72)
1.4	374 (243, 527)	375 (241, 538)	52 (38, 72)	49 (36, 69)
1.5	168 (122, 232)	163 (117, 228)	51 (38, 72)	48 (35, 67)
1.6	293 (226, 383)	290 (223, 379)	56 (42, 76)	53 (40, 72)
1.7	325 (239, 426)	324 (234, 427)	55 (41, 71)	52 (38, 71)
1.8	216 (149, 319)	214 (146, 325)	51 (35, 74)	49 (33, 70)
1.9	229 (163, 327)	226 (159, 340)	51 (35, 74)	48 (34, 71)
1.10	239 (164, 345)	237 (165, 354)	54 (38, 76)	51 (40, 73)
2	285 (228, 364)	284 (225, 367)	70 (52, 91)	67 (51, 91)
3	301 (239, 377)	295 (234, 370)	53 (40, 68)	48 (36, 62)
5.1	228 (144, 371)	227 (137, 360)	52 (34, 77)	49 (33, 75)
5.2	262 (154, 484)	261 (158, 455)	55 (35, 87)	52 (37, 84)
5.3	225 (135, 426)	222 (128, 397)	53 (31, 86)	51 (32, 72)
5.4	245 (145, 410)	244 (152, 410)	54 (38, 91)	52 (38, 81)
5.5	171 (100, 268)	169 (93, 265)	52 (36, 73)	50 (33, 70)
5.6	201 (115, 348)	197 (115, 309)	53 (35, 84)	50 (34, 75)
5.7	220 (140, 337)	218 (137, 338)	52 (32, 81)	50 (33, 78)
5.8	287 (223, 380)	289 (226, 370)	56 (44, 71)	54 (42, 68)
5.9	191 (113, 306)	188 (110, 302)	54 (41, 82)	52 (40, 80)
6	165 (123, 232)	164 (121, 230)	49 (33, 70)	47 (33, 68)
7	285 (222, 378)	288 (220, 381)	53 (42, 70)	52 (40, 68)
9	342 (243, 477)	346 (247, 480)	49 (35, 66)	47 (34, 64)
10	329 (263, 426)	331 (261, 427)	58 (46, 75)	56 (45, 72)
Hospital A	253 (174, 378)	249 (170, 379)	51 (32, 72)	48 (33, 72)
Hospital B	312 (246, 381)	306 (240, 376)	72 (55, 105)	68 (51, 98)
Hospital C	215 (168, 288)	213 (165, 286)	54 (43, 72)	51 (41, 67)
Hospital D	187 (123, 292)	176 (116, 271)	49 (32, 71)	47 (32, 67)
Hospital E	304 (195, 502)	301 (189, 493)	52 (36, 76)	49 (36, 77)

Hospital	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
	Without Exclusion	With COVID Denominator Exclusion	Without Exclusion	With COVID Denominator Exclusion
	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations
Average Across Sites	254.071	251.790	54.550	51.972

4. Summary

In this report, we described the development and testing of the Severe Obstetric Complication eCQM, for which the primary goal was to assess the occurrence of specific severe obstetric complications in the hospital setting by using a methodology that reliably allows comparison across hospitals. This measure supports the public health goal of lowering the occurrence of maternal complications to reduce maternal death and disability and improve maternal quality of life. Measure specifications incorporated guidance from CMS, statistical experts, and subject matter experts.

This report documents the successful development and testing of a reliable, valid, and feasible eCQM of severe obstetric complications. The measure used two stages of Beta testing to ensure a reliable, valid and feasible measure; Stage 2 findings reinforce the Stage 1 results.

This measure was developed by The Joint Commission in close collaboration with CORE under contract with CMS. This measure fills a critical gap in maternal health surveillance in the United States and will enable both monitoring of hospital care and health equity for this important patient population.

References

1. Centers for Disease Control and Prevention. Severe Maternal Morbidity in the United States. January 31, 2020; https://www.cdc.gov/reproductivehealth/maternalinfanthealth/severematernalmorbidity.html.
2. U.S. Department of Health & Human Services. HHS Outlines New Plans and a Partnership to Reduce U.S. Pregnancy-related Deaths. 2020; https://www.hhs.gov/about/news/2020/12/03/hhs-outlines-new-plans-to-reduce-us-pregnancy-related-deaths.html.
3. Hoyert DL, Miniño AM. Maternal mortality in the United States: changes in coding, publication, and data release, 2018. 2020.
4. U.S. Department of Health & Human Services. *The Surgeon General's Call to Action to Improve Maternal Health.* 2020.
5. Centers for Medicare & Medicaid Services. Meaningful Measures 2.0: Moving from Measure Reduction to Modernization. 2020; https://www.cms.gov/meaningful-measures-20-moving-measure-reduction-modernization, 2020.
6. OPTUM. *The benefit of using both claims data and electronic medical record data in health care analysis.* 2012.
7. Kilpatrick SK, Ecker JL. Severe maternal morbidity: screening and review. *Am J Obstet Gynecol.* 2016;215(3):B17-22.
8. World Health Organization. Maternal deaths. https://www.who.int/data/gho/indicator-metadata-registry/imr-details/4622, 2020.
9. Centers for Disease Control and Prevention. Preventing Pregnancy-Related Deaths. 2019; https://www.cdc.gov/reproductivehealth/maternal-mortality/preventing-pregnancy-related-deaths.html, 2020.
10. National Quality Forum. *A Roadmap for Promoting Health Equity and Eliminating Disparities: The Four I's for Health Equity.* 2017.
11. eCQI Resource Center. eCQMs. 2020; https://ecqi.healthit.gov/ecqms?qt-tabs_ecqm=0.
12. The Office of the National Coordinator for Health Information Technology. What is an electronic health record (EHR)? 2019; https://www.healthit.gov/faq/what-electronic-health-record-ehr.
13. Kassebaum NJ, Barber RM, Bhutta ZA, et al. Global, regional, and national levels of maternal mortality, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet.* 2016;388(10053):1775-1812.
14. Davis NL, Smoots AN, Goodman DA. Pregnancy-Related Deaths: Data from 14 US Maternal Mortality Review Committees. *Education.* 2019;40(36):8.2.
15. National Quality Forum. *Maternal Morbidity and Mortality Environmental Scan.* 2020.
16. Centers for Disease Control and Prevention. Rates in Severe Morbidity Indicators per 10,000 Delivery Hospitalizations, 1993-2014. 2020; https://www.cdc.gov/reproductivehealth/maternalinfanthealth/smm/rates-severe-morbidity-indicator.htm.
17. Centers for Disease Control and Prevention. How Does CDC Identify Severe Maternal Morbidity? 2019; https://www.cdc.gov/reproductivehealth/maternalinfanthealth/smm/severe-morbidity-ICD.htm.
18. Creanga AA, Berg CJ, Ko JY, et al. Maternal mortality and morbidity in the United States: where are we now? *Journal of women's health.* 2014;23(1):3-9.
19. Shields LE, Wiesner S, Klein C, Pelletreau B, Hedriana HL. Use of maternal early warning trigger tool reduces maternal morbidity. *American journal of obstetrics and gynecology.* 2016;214(4):527. e521-527. e526.

20. Vesco KK, Ferrante S, Chen Y, Rhodes T, Black CM, Allen-Ramey F. Costs of Severe Maternal Morbidity During Pregnancy in US Commercially Insured and Medicaid Populations: An Observational Study. *Maternal and Child Health Journal.* 2020;24(1):30-38.
21. Chen H-Y, Chauhan SP, Blackwell SC. Severe maternal morbidity and hospital cost among hospitalized deliveries in the United States. *American journal of perinatology.* 2018;35(13):1287-1296.
22. Lin C-CC, Hirai AH, Li R, Kuklina EV, Fisher SK. Rural–urban differences in delivery hospitalization costs by severe maternal morbidity status. *Annals of Internal Medicine.* 2020;173(11_Supplement):S59-S62.
23. Premier Inc. *Report 2: The Added Cost of Complications During and After Delivery.* 2019.
24. Leonard SA, Main EK, Scott KA, Profit J, Carmichael SL. Racial and ethnic disparities in severe maternal morbidity prevalence and trends. *Annals of epidemiology.* 2019;33:30-36.
25. Petersen EE, Davis NL, Goodman D, et al. Vital signs: pregnancy-related deaths, United States, 2011–2015, and strategies for prevention, 13 states, 2013–2017. *Morbidity and Mortality Weekly Report.* 2019;68(18):423.
26. Petersen EE, Davis NL, Goodman D, et al. Racial/ethnic disparities in pregnancy-related deaths—United States, 2007–2016. *Morbidity and Mortality Weekly Report.* 2019;68(35):762.
27. Creanga AA, Bateman BT, Kuklina EV, Callaghan WM. Racial and ethnic disparities in severe maternal morbidity: a multistate analysis, 2008-2010. *American journal of obstetrics and gynecology.* 2014;210(5):435. e431-435. e438.
28. Geller SE, Rosenberg D, Cox SM, et al. The continuum of maternal morbidity and mortality: factors associated with severity. *American journal of obstetrics and gynecology.* 2004;191(3):939-944.
29. Deadly Deliveries: Childbirth complication rates at maternity hospitals. https://www.usatoday.com/maternal-mortality-harm-hospital-database/.
30. Main EK. Reducing maternal mortality and severe maternal morbidity through state-based quality improvement initiatives. *Clinical obstetrics and gynecology.* 2018;61(2):319-331.
31. California Maternal Quality Care Collaborative (CMQCC). Who We Are. https://www.cmqcc.org/who-we-are, 2020.
32. California Maternal Quality Care Collaborative (CMQCC). Maternal Data Center. https://www.cmqcc.org/maternal-data-center, 2020.
33. Lenfant C. Clinical research to clinical practice—lost in translation? *New England Journal of Medicine.* 2003;349(9):868-874.
34. Maher-Griffiths C. Maternal Quality Outcomes and Cost. *Critical Care Nursing Clinics.* 2019;31(2):177-193.
35. Jha A, Pronovost P. Toward a safer health care system: the critical need to improve measurement. *Jama.* 2016;315(17):1831-1832.
36. Classen DC, Griffin FA, Berwick DM. Measuring patient safety in real time: an essential method for effectively improving the safety of care. In: American College of Physicians; 2017.
37. Alliance for Innovation on Maternal Health (AIM). AIM SMM Codes List v3-20-2019. In:2019.
38. Gibbins KJ, Einerson BD, Varner MW, Silver RM. Placenta previa and maternal hemorrhagic morbidity. *The Journal of Maternal-Fetal & Neonatal Medicine.* 2018;31(4):494-499.
39. Merriam AA, Wright JD, Siddiq Z, et al. Risk for postpartum hemorrhage, transfusion, and hemorrhage-related morbidity at low, moderate, and high volume hospitals. *The Journal of Maternal-Fetal & Neonatal Medicine.* 2018;31(8):1025-1034.
40. Chandra I, Sun L. Preterm and term preeclampsia: differences in biochemical parameter and pregnancy outcomes. *Postgraduate medicine.* 2018;130(8):703-707.

41. Al-Ostad G, Kezouh A, Spence AR, Abenhaim HA. Incidence and risk factors of sepsis mortality in labor, delivery and after birth: Population-based study in the USA. *Journal of Obstetrics and Gynaecology Research.* 2015;41(8):1201-1206.
42. Leonard SA, Kennedy CJ, Carmichael SL, Lyell DJ, Main EK. An expanded obstetric comorbidity scoring system for predicting severe maternal morbidity. *Obstetrics & Gynecology.* 2020;136(3):440-449.
43. The Joint Commission. Measure Information Form (PC-01 Elective Delivery). 2020; https://manual.jointcommission.org/releases/TJC2021A/MIF0166.html.
44. The American College of Obstetricians and Gynecologists. reVITALize: Obstetrics Data Definitions. 2014; https://www.acog.org/practice-management/health-it-and-clinical-informatics/revitalize-obstetrics-data-definitions.
45. Centers for Disease Control and Prevention. What is Stillbirth? 2020; https://www.cdc.gov/ncbddd/stillbirth/facts.html.
46. Khan DSA, Pirzada AN, Ali A, Salam RA, Das JK, Lassi ZS. The Differences in Clinical Presentation, Management, and Prognosis of Laboratory-Confirmed COVID-19 between Pregnant and Non-Pregnant Women: A Systematic Review and Meta-Analysis. *International Journal of Environmental Research and Public Health.* 2021;18(11):5613.
47. Services CfMM. Core Clinical Data Elements for the Hybrid Hospital-Wide Readmission (HWR) Measure with Claims and Electronic Health Record Data
48. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *biometrics.* 1977:159-174.
49. Yu H, Mehrotra A, Adams J. Reliability of utilization measures for primary care physician profiling. Paper presented at: Healthcare2013.
50. Adams JL, Mehrotra A, Thomas JW, McGlynn EA. Physician cost profiling—reliability and risk of misclassification. *New England Journal of Medicine.* 2010;362(11):1014-1021.

Appendix A: Acknowledgement Details

Clinical Consultant

We would like to acknowledge the expertise from our clinical consultant who has offered invaluable guidance to inform clinical and methodological decisions for the Severe Obstetrics Complications eCQM.

Elliot K. Main, MD

Medical Director, California Maternal Quality Care Collaborative (CMQCC) and Clinical Professor, Obstetrics and Gynecology at Stanford University

Technical Expert Panel

We would like to acknowledge the contributions of our TEP (Table A1). The TEP members brought a diverse range of expertise and provided feedback for consideration in the development of the Severe Obstetric Complications eCQM.

Table A1. Technical Expert Panel Members

Name	Affiliation	Location
Suzanne McMurtry Baird, DNP, RN	Co-Owner and Nursing Director, Clinical Concepts in Obstetrics, LLC	Brentwood, TN
Debra Bingham, DrPH, RN, FAAN	Executive Director, Institute for Perinatal Quality Improvement	Quincy, MA
James Christmas, MD	National Medical Director, Women's and Obstetrics, HCA Healthcare	Nashville, TN
Blair Dudley, MPH	Senior Manager, Transform Maternity Care, Pacific Business Group on Health	Oakland, CA
Tomeka Isaac, MBA	Patient Representative	Denver, NC
Ajshay James	Patient Representative	Houston, TX
Deborah Kilday, MSN, RN	Manager, Performance Partner – Women, infants, and Children, Strategy, Innovation, and Population Health, Premier Healthcare Solutions, Inc.	Woodstock, GA
Joseph Kunisch, PhD, RN-BC Informatics, CPHQ	VP Quality Programs, Harris Health	Houston, TX
David C. Lagrew Jr., MD	Executive Medical Director, Providence Health System	Irvine, CA
Elizabeth O'Neil-Greiner, RN, MHA	Business Process Consultant, BJC Healthcare	St. Louis, MO
Sarosh Rana, MD, MPH	Professor, Department of Obstetrics and Gynecology Section Chief, Maternal Fetal Medicine, University of Chicago	Chicago, IL
Elizabeth Rochin, PhD, RN, NE-BC	President, National Perinatal Information Center	Providence, RI

Name	Affiliation	Location
Michael Ross, MD, MPH	Professor of Obstetrics and Gynecology and Public Health, David Geffen School of Medicine and Fielding School of Public Health, UCLA Investigator, The Lundquist Institute	Los Angeles, CA
Karey M. Sutton, PhD	Director, Health Equity Research Workforce, Association of American Medical Colleges	Washington, DC
Aswita Tan-McGrory, MBA, MSPH	Director, The Disparities Solutions Center, Massachusetts General Hospital Adjunct Faculty, Northeastern University	Boston, MA
Brooke Villarreal, DNP, MSRN, RN-BC	Director, Public Reporting and Outcomes Measurement, HCA Healthcare	Nashville, TN

Patient Working Group

We would also like to thank members of our Patient Working Group (Table A2) for their personal and insightful perspectives on key measure aspects of measure development and decisions.

Table A2. Patient Working Group Members

Patient Expert Name	Location
Leah Bahrencu	Austin, TX
Marianne Drexler	Durham, NC
Nikki Montgomery	Euclid, OH
Katie Silwa	Hagerstown, MD
Molly Firth	Tumwater, WA
Kayleigh Summers	Pottstown, PA
Kim Sandstrom	Ocala, FL

Appendix B: Glossary

Frequent terminology and definitions used in the Severe Obstetric Complications eCQM:

Acute care hospital: A hospital that provides inpatient medical care for surgery and acute medical conditions or injuries. Short-term acute care hospitals provide care for short-term illnesses and conditions. In contrast, long-term acute care hospitals generally treat medically complex patients who require long-stay hospital-level care, which is generally defined as an inpatient length of stay greater than 25 days.

C-Statistic: An indicator of the model's discriminant ability or ability to correctly classify those patients who have and have not experienced the outcome. Potential values range from 0.5, meaning no better than chance, to 1.0, an indication of perfect prediction. Perfect prediction implies patients' outcomes can be predicted completely by their risk factors, and physicians play no role in their patients' outcomes.

Case mix: The particular illness severity and demographic characteristics of patients with encounters/admissions at a given hospital.

Cohort: The encounters used to calculate the measure after inclusion and exclusion criteria have been applied.

Comorbidities: Medical conditions the patient had in addition to their primary reason for admission to the hospital.

Complications: Medical conditions that may have occurred because of care rendered during hospitalization.

Confidence Interval: A CI is a range of values that describes the uncertainty surrounding an estimate. It is indicated by its endpoints; for example, a 95% CI for an odds ratio (OR) noted as "1.09 – 1.15" would indicate that there is 95% confidence that the true OR lies between 1.09 and 1.15.

Outcome: The result of a broad set of healthcare activities that affect patients' well-being. For the Severe Obstetric Complications eCQM, the outcome is the number of inpatient hospitalizations for patients who experience SMM diagnoses not present on admission during a delivery hospitalization.

Risk-adjustment variables: Patient demographics and comorbidities used to standardize rates for differences in case mix across hospitals.

Appendix C: Value Sets for Severe Obstetric Complications eCQM Specifications

Table C1 outlines the Value Sets that are used to define the measure specifications. The Value Set Authoring Center is the authoritative data source for Value Sets and Organizational Object Identifiers (OIDs).

Table C1. Value Set Name and OID for measure numerator, denominator, and risk adjustment

Measure Specification	Value Set Name	Code System	OID
Numerator	**Severe Maternal Morbidity Procedures**	Grouping[a]	2.16.840.1.113762.1.4.1029.256
	Severe Maternal Morbidity Diagnoses	Grouping	2.16.840.1.113762.1.4.1029.255
	Acute Heart Failure	Grouping	2.16.840.1.113762.1.4.1029.351
	Acute Myocardial Infarction	Grouping	2.16.840.1.113883.3.666.5.3011
	Aortic Aneurysm	Grouping	2.16.840.1.113762.1.4.1029.344
	Cardiac Arrest/Ventricular Fibrillation	Grouping	2.16.840.1.113762.1.4.1029.345
	Heart Failure/ Arrest Related to Procedure or Surgery	Grouping	2.16.840.1.113762.1.4.1029.348
	Disseminated Intravascular Coagulation	Grouping	2.16.840.1.113762.1.4.1029.346
	Shock	Grouping	2.16.840.1.113762.1.4.1029.354
	Renal (Acute Renal Failure Grouping)	Grouping	2.16.840.1.113762.1.4.1029.342
	Adult Respiratory Distress Syndrome	Grouping	2.16.840.1.113762.1.4.1029.367
	Pulmonary Edema	Grouping	2.16.840.1.113762.1.4.1029.350
	Sepsis	Grouping	2.16.840.1.113762.1.4.1029.353
	Air and Thrombotic Embolism	Grouping	2.16.840.1.113762.1.4.1029.356
	Amniotic Fluid Embolism	Grouping	2.16.840.1.113762.1.4.1029.343

Measure Specification	Value Set Name	Code System	OID
Numerator	**Eclampsia**	Grouping	2.16.840.1.113762.1.4.1029.347
	Severe Anesthesia Complications	Grouping	2.16.840.1.113762.1.4.1029.352
	Puerperal Cerebrovascular Disorder	Grouping	2.16.840.1.113762.1.4.1029.349
	Sickle Cell Disease with Crisis	Grouping	2.16.840.1.113762.1.4.1029.355
	Blood Transfusion	Grouping	2.16.840.1.113762.1.4.1029.213
	Conversion of Cardiac Rhythm	Grouping	2.16.840.1.113762.1.4.1029.357
	Hysterectomy	Grouping	2.16.840.1.113762.1.4.1029.358
	Tracheostomy	Grouping	2.16.840.1.113762.1.4.1029.359
	Ventilation	Grouping	2.16.840.1.113762.1.4.1029.360
	Hemorrhage	Grouping	2.16.840.1.113762.1.4.1029.258
Denominator	**Delivery Procedures**	Grouping	2.16.840.1.113762.1.4.1045.59
Denominator Exclusions	**COVID**	Grouping	2.16.840.1.113762.1.4.1029.373
	COVID Related Respiratory Illness – Diagnosis Codes	Grouping	2.16.840.1.113762.1.4.1029.376 2.16.840.1.113762.1.4.1029.379
	COVID Related Respiratory Illness – Procedure Codes	Grouping	2.16.840.1.113762.1.4.1029.379
Risk Adjustment	**Anemia**	Grouping	2.16.840.1.113762.1.4.1029.323
	Asthma	Grouping	2.16.840.1.113883.3.117.1.7.1.271
	Autoimmune Disease	Grouping	2.16.840.1.113762.1.4.1029.311
	Bariatric Surgery	Grouping	2.16.840.1.113762.1.4.1029.317
	Bleeding Disorder	Grouping	2.16.840.1.113762.1.4.1029.287

Measure Specification	Value Set Name	Code System	OID
Risk Adjustment	**BMI >=40**	Grouping	2.16.840.1.113762.1.4.1029.290
	Cardiac Disease	Grouping	2.16.840.1.113762.1.4.1029.341
	Gastrointestinal Disease	Grouping	2.16.840.1.113762.1.4.1029.338
	Gestational Diabetes	Grouping	2.16.840.1.113762.1.4.1029.269
	HIV	Grouping	2.16.840.1.113762.1.4.1029.272
	Hypertension	Grouping	2.16.840.1.113762.1.4.1029.332
	Mental Health Disorder	Grouping	2.16.840.1.113762.1.4.1029.314
	Multiple Pregnancy	Grouping	2.16.840.1.113762.1.4.1029.284
	Neuromuscular Disease	Grouping	2.16.840.1.113762.1.4.1029.308
	Obstetric VTE	Grouping	2.16.840.1.113762.1.4.1029.363
	Other Preeclampsia	Grouping	2.16.840.1.113762.1.4.1029.329
	Placental Accreta Spectrum	Grouping	2.16.840.1.113762.1.4.1029.302
	Placental Abruption	Grouping	2.16.840.1.113762.1.4.1029.305
	Placenta Previa	Grouping	2.16.840.1.113762.1.4.1110.37
	Preexisting Diabetes	Grouping	2.16.840.1.113762.1.4.1029.275
	Preterm Birth	Grouping	2.16.840.1.113762.1.4.1029.299
	Previous Cesarean	Grouping	2.16.840.1.113762.1.4.1029.278
	Pulmonary Hypertension	Grouping	2.16.840.1.113762.1.4.1029.281
	Renal Disease	Grouping	2.16.840.1.113762.1.4.1029.335
	Severe Preeclampsia	Grouping	2.16.840.1.113762.1.4.1029.327

Measure Specification	Value Set Name	Code System	OID
Risk Adjustment	**Substance Abuse**	Grouping	2.16.840.1.113762.1.4.1029.320
	Thyrotoxicosis	Grouping	2.16.840.1.113762.1.4.1029.296
	Heart Rate	LOINC	8867-4
	Systolic Blood Pressure	LOINC	8480-6
	Hematocrit	LOINC	2.16.840.1.113762.1.4.1045.114
	White Blood Cells Count Lab Test	LOINC	2.16.840.1.113762.1.4.1045.129
	Long-term Anticoagulant Use	Grouping	2.16.840.1.113762.1.4.1029.366
	Economic Housing Instability	Grouping	2.16.840.1.113762.1.4.1029.292

[a] *Grouping of ICD-10 and SNOMED-CT value sets*

Appendix D. Stage 1 Beta Testing Results

Table D1 provides the observed and the risk-standardized rate per 10,000 deliveries rates for Severe Obstetric Complications and Severe Obstetric Complications Excluding Blood Transfusion-Only encounters for all Stage 1 Beta testing (8 sites).

Table D1. Observed and Risk-Standardized Severe Obstetric Complication Rates Across Test Sites (8 Sites, Stage 1 Beta Testing)

Site ID	Delivery Encounters	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
		Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations
Site 1	18,070	226	241	41	49
Site 2	7,196	235	248	72	55
Site 3	7,955	303	268	48	50
Site 5	6,139	209	223	44	50
Site 6	3,359	104	158	27	48
Site 7	4,369	213	255	41	50
Site 9	3,918	202	299	26	48
Site 10	9,178	341	285	81	51
Across All Encounters	**60,184**	**244**	*	**50**	*
Average Among Hospitals	*	*	**252**	*	**50**

** Cell intentionally left empty*

Table D2provides the unadjusted and the risk-standardized rate per 10,000 deliveries rates for severe obstetric complications and severe obstetric complications excluding blood transfusion-only encounters for each of the 25 individual Stage 1 Beta testing hospitals and across all Stage 1 Beta testing hospitals.

Table D2. Observed and Risk-Standardized Severe Obstetric Complication Rates per 10,000 Delivery Hospitalizations across Hospitals (25 Hospitals, Stage 1 Beta Testing)

Hospital	Delivery Encounters	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
		Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations
Hospital 1.1	496	202	239	0	49
Hospital 1.2	3,875	248	284	52	51
Hospital 1.3	1,518	158	216	33	50

Hospital	Delivery Encounters	Any Severe Obstetric Complication(s)		Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	
		Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations	Observed rate per 10,000 Delivery Hospitalizations	Risk-Standardized Rate per 10,000 Delivery Hospitalizations
Hospital 1.4	534	412	372	19	50
Hospital 1.5	2,383	105	164	29	50
Hospital 1.6	5,952	269	288	54	51
Hospital 1.7	1,678	244	317	36	50
Hospital 1.8	733	164	210	14	50
Hospital 1.9	608	214	223	16	49
Hospital 1.10	293	171	233	34	50
Hospital 2	7,196	235	269	72	55
Hospital 3	7,955	303	293	48	50
Hospital 5.1	292	137	226	0	50
Hospital 5.2	224	179	262	45	50
Hospital 5.3	139	72	221	0	50
Hospital 5.4	347	144	244	29	50
Hospital 5.5	799	50	170	13	50
Hospital 5.6	163	0	196	0	50
Hospital 5.7	560	143	221	18	50
Hospital 5.8	3,316	305	294	66	51
Hospital 5.9	299	33	187	33	50
Hospital 6	3,359	104	156	27	49
Hospital 7	4,369	213	282	41	50
Hospital 9	3,918	202	345	26	49
Hospital 10	9,178	341	313	81	51
Across All Encounters	**60,184**	**244**	*	**50**	*
Average Among Hospitals	*	*	**249**	*	**50**

** Cell intentionally left empty*

Table D3 provides the Signal-to-Noise reliability measures scores for Severe Obstetric Complications and Severe Obstetric Complications Excluding Blood Transfusion-only Encounters for all Stage 1 Beta testing (8 sites).

Table D3. Signal-to-Noise-Reliability, Measure Scores, by Site (8 Sites, Stage 1 Beta Testing)

Outcome	# Hospitals	Median	Mean (SD)	Minimum	Maximum	Interquartile Range	
						Q1	Q3
Any Severe Obstetric Complication(s)	8	0.991	0.990 (0.469)	0.983	0.997	0.986	0.997

Outcome	# Hospitals	Median	Mean (SD)	Minimum	Maximum	Interquartile Range	
						Q1	Q3
Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	8	0.957	0.951 (2.214)	0.918	0.984	0.932	0.984

Table D4 provides the Signal-to-Noise reliability measures scores for Severe Obstetric Complications and Severe Obstetric Complications Excluding Blood Transfusion-only Encounters for all Stage 1 Beta testing (25 hospitals).

Table D4. Signal-to-Noise-Reliability, Measure Scores, by Hospital (25 Hospitals, Stage 1 Beta Testing)

Outcome	Volume Threshold (Number of Delivery Encounters per Hospital per year)	# Pilot Hospitals	Median	Mean (SD)	Minimum	Maximum	Interquartile Range	
							Q1	Q3
Any Severe Obstetric Complication(s)	>25	25	0.960	0.947 (0.055)	0.805	0.996	0.912	0.996
Severe Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	>25	25	0.684	0.695(0.229)	0.274	0.961	0.485	0.956
Any Severe Obstetric Complication(s)	>200	23	0.978	0.958 (0.040)	0.869	0.996	0.936	0.996
Severed Obstetric Complication(s) Excluding Blood Transfusion-Only Encounters	>200	23	0.805	0.730 (0.202)	0.378	0.961	0.574	0.956