NATIONAL QUALITY FORUM
Driving measurable health improvements together

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through National Quality Forum's (NQF) Consensus Development Process (CDP). The information submitted by the measure developers/stewards is included after the *Brief Measure Information* and *Preliminary Analysis* sections.

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

Brief Measure Information

NQF #: 3735

Corresponding Measures:

Measure Title: CVD Risk Follow Up Measure- Proportion of patients with a positive CVD risk assessment who receive follow-up care

Measure Steward: University of California, Irvine

sp.02. Brief Description of Measure: All pregnant and postpartum patients need to be systematically assessed for cardiovascular disease (CVD). Once identified as being at risk for CVD follow-up cardiac tests and consultations are scheduled. UCI implemented and tested a standardized CVD risk assessment algorithm that can be integrated into the EHR system and provides an immediate triage of patients as low and high risk for CVD. This measure assesses the rate of pregnant and postpartum patients who are determined to be at risk for CVD) using a standardized risk assessment who received appropriate follow-up in the form of cardiology consultations and tests. The unit of measurement is the individual patient, and the population is comprised of patients who have an outpatient or inpatient prenatal or postpartum visit at a clinic or facility. This includes pregnant and postpartum emancipated minors. The measure can be calculated at the hospital system level or clinic site level. A hospital system includes Labor and Delivery, outpatient care in a hospital or at affiliated clinics, and private providers contracted with the hospital for delivery. The denominator of the measure is comprised of all patients seen for prenatal or postpartum care at the measurement entity (hospital system or individual clinic sites) who were identified as "at risk for CVD" in a standardized CVD risk assessment such as the California Maternal Quality Care Collaborative (CMQCC) CVD risk assessment algorithm. The aim is to have 100 percent of patients with a positive risk assessment receiving follow-up as recommended by American College of Obstetricians and Gynecologists (ACOG) guidelines. Implementation of the tool in three hospital systems showed a wide variation in the rate of follow-up between 38.4% and 70.8% of patients identified as high risk for CVD. The measure can be calculated annually.

1b.01. Developer Rationale: The implementation of metrics on the adherence to standards signals to clinicians the endorsement of management and value of the metric as a signal to provide quality care to obstetric patients. The measure is easy to understand and can be calculated on the unit or even individual clinician performance (if the patient is assigned to one clinician) in addition to systemwide feedback. The measure allows for the identification of low-performing sites or clinicians and to address modifiable gaps in diagnostic excellence.
CVD is the leading cause of maternal mortality during pregnancy and the postpartum period in the United States. Most pregnant and postpartum patients who die of CVD do not have a known diagnosis of CVD.

Diagnosis of CVD in pregnancy may be challenging as signs and symptoms of normal pregnancy mimic those of CVD which may be missed by the health care providers.
Our CVD quality measures target the childbearing age population who may be at high risk of CVD and access the health care system for maternity services. The universal implementation of a standardized tool to measure the clinic or facility's performance in following up on patients who are identified to be at risk will provide a uniform assessment of patients' CVD risk regardless of the clinicians' diagnostic skills and experience. Information about the follow-up rates will inform quality improvement strategies.
The CVD measure during pregnancy/postpartum care is bound to increase education and awareness in this population and will empower patients to seek early medical care if new signs and symptoms that may be suggestive of CVD develop. Our measure may have implications for long-term health outcomes with improvements in the CVD risk factor profile in the future. The use of a standardized CVD measure to risk stratify pregnant and postpartum patients may improve the timely identification of CVD, thereby decreasing maternal morbidity and/or mortality.
Additionally, the training for the use of the tool and reporting of the clinic's performance compared to other clinic sites in the same hospital network has raised awareness of the importance of CVD risk assessment among clinicians, even in clinics serving primarily low-risk patients. The ease of use (less than 1 minute to do the CVD risk assessment) of the tool allows obstetricians to systematically identify patients who are at risk for CVD and need follow-up on more thorough monitoring during the pregnancy. Additionally, we have anecdotal evidence that administration of the tool and providing patients with a risk score has improved patient awareness of the immediate and lifetime risk of developing CVD, which drives changes in health behavior. We have not seen any evidence that the follow-up of patients who were deemed at high risk for CVD led to inappropriate use of resources.

sp.12. Numerator Statement: Patients who were identified to be at risk for CVD and received follow-up care within 60 days of the risk.

sp.14. Denominator Statement: Pregnant and postpartum patients who have been identified to be at risk for cardiovascular disease (CVD) during the measurement period. Patients who were screened for CVD and had a pregnancy loss or stillbirth will remain in the cohort.

sp.16. Denominator Exclusions: We will exclude patients who discontinued care (no additional visit within 60 days after the risk assessment).

Measure Type: Process

sp.28. Data Source: Electronic Health Records; Paper Medical Records

sp.07. Level of Analysis: Clinician: Group/Practice

Preliminary Analysis: New Measure

Criteria 1: Importance to Measure and Report

1a. Evidence

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

The developer provides the following description for this measure:

- This is a new process measure at the clinician group practice level that measures the rate of follow-up care within 60 days of risk assessment for pregnant or postpartum patients who were identified to be at risk for cardiovascular disease (CVD).
- The developer provides a logic model that demonstrates the importance of follow-up care for pregnant and postpartum patients who may have cardiovascular disease, identified using a risk assessment. The risk assessment assists providers in distinguishing between signs and symptoms of cardiac disease and those of normal pregnant and postpartum patients who may have cardiovascular disease.
- Follow-up care for those identified at risk leads to increased patient awareness, behavior change and ultimately change in maternal mortality and birth outcomes.

The developer provides the following evidence for this measure:

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

Summary:

- The California Maternal Quality Care Collaborative (CMQCC) Cardiovascular Disease in Pregnancy and Postpartum Task Force developed the risk assessment algorithm based on risk factors, symptoms, vital sign abnormalities, and physical examination findings commonly identified in patients who die of various types of cardiovascular disease.
    - The literature establishes that CVD is the leading cause of maternal mortality in the United States and California. CVD accounts for >33% of all pregnancy-related deaths in the US and 25% of pregnancy deaths in CA.
    - The developer cites evidence that the risk assessment was able to accurately identify pregnant or postpartum patients at risk for CVD.
        - The authors assessed the triage algorithm retrospectively on 64 CVD related deaths in CA for 2002-2006. They found that the use of the algorithm would have identified 56 of the 64 cases (88%) of CVD. The proportion of cases increased to 93% when they restricted it to the 60 cases of patients who were symptomatic or had sufficient documentation.
        - A prospective cohort study of obstetrical patients from April 2018 to July 2019 at academic medical centers in CA and NY was conducted with 846 patients. The overall risk assessed positive rate was 8% (5% in CA, 19% in NY). CVD was confirmed in 30% with positive risk assessments with complete follow-up.
- Evidence was not provided on follow-up visits leading to desired health outcomes. While evidence is not presented on follow-up visits, it is likely appropriate to have a follow-up visit for a high risk pregnant or postpartum patient.

Exception to evidence

- N/A

Questions for the Standing Committee:

- *What is the relationship between this measure and patient outcomes?*
- *Is the evidence directly applicable to the process of care being measured?*

Guidance From the Evidence Algorithm

Evidence was not provided on follow-up visits leading to desired health outcomes.

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

RATIONALE: The evidence presented did not address the measure for follow-up visits. While there is empirical evidence to establish that the risk assessment can identify CVD patients, there is no evidence presented on the follow-up visits and their impact on an outcome.

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

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

- CVD risk assessment follow-up rates were provided from 15 measured entities with 312 patients from 9/1/2021 to 2/28/20222.
- Mean performance was 59% and interquartile range (IQR) was 50%.
- The 10th decile performance was 22% and the 90th decile performance was 90%.

Disparities

- The developer provided rates of follow-up for the 312 patients segmented by age, race, ethnicity, insurance status, and timing.
- Differences were shown by age (20-29 [50.0%], 40+ [35.5%], race (Black [40.5%] and White [47.6%]), race/ethnicity (Non-Hispanic White [43.9%], Non-Hispanic Black [40.0%], and Hispanic [45.9%]), insurance status (private [47.5%%] and public [38.7%]), and timing (prenatal [48.6%], postpartum [49.3%]).

Questions for the Standing Committee:

- *Is there a gap in care that warrants a national performance measure?*

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

Criteria 2: Scientific Acceptability of Measure Properties

Complex measure evaluated by the Scientific Methods Panel (SMP)? ☐ **Yes** ☒ **No**
Evaluators: NQF Staff

2a. Reliability: Specifications and Testing

2a1. Specifications require the measure, as specified, to produce consistent (i.e., reliable) and credible (i.e., valid) results about the quality of care when implemented.

2a2. Reliability testing demonstrates whether the measure data elements are repeatable and producing the same results a high proportion of the time when assessed in the same population in the same time period, and/or whether the measure score is precise enough to distinguish differences in performance across providers.

Specifications:

- The developer notes that two facilities are excluded from the signal to noise analysis for reliability testing but are not mentioned in the exclusions. Therefore, it is unclear if these facility types should be

systematically excluded. The developer states that these facilities were excluded for the purpose of parameter estimation.

- The exclusion of patients who discontinued care is not clear.

Reliability Testing:

- Reliability testing conducted at the Patient/Encounter Level:
  - The developer evaluated the degree of agreement between charts extracted from the EHR with manual abstraction using a kappa statistic.
  - The developer evaluated the agreement rates for electrocardiogram (EKG) or echo. The developer found 100% agreement at two facilities (University of California, Irvine [UCI], and University of California, San Diego [UCSD]) received follow-up.
    - However, the developer did not provide testing results for all key data elements (e.g., numerator, denominator, exclusions).
- Reliability testing conducted at the Accountable Entity Level:
  - The developer excluded facilities with a large sample size (N>75th percentile +1.5*IQR) from the analysis. The rationale for this exclusion is unclear, as the developer states that facilities were excluded for the purpose of parameter estimation
  - The median reliability of 0.356 (minimum 0.181; maximum 0.356), implies most of the variability in the measure is attributable to measurement error. The developers concludes that the large measurement error is due to the small denominator in two of the sites (under 20), which results in an unstable rate.

Questions for the Standing Committee regarding reliability:

- *Do you have any concerns that the measure cannot be consistently implemented (i.e., are the measure specifications adequate)?*
- *Are the testing results sufficient to demonstrate reliability?*
- *If large facilities are excluded from the SNR, should that be an exclusion as the measure is not tested as specified?*

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

RATIONALE: The developer found 100% agreement rates for the follow-up visits; however, all key data elements (numerator, denominator, exclusions) did not have any reliability testing. Therefore, looking at the accountable entity level results, the median reliability result is low <0.4.

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

2b2. Validity testing should demonstrate that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

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

Validity Testing

- Validity testing conducted at the Patient/Encounter Level:
  - Agreement between charts extracted from the electronic health record (EHR) were compared with manual abstraction. The developer reviewed 1,399 patient charts over an 18-month

period at the UCI network. The kappa value was equal to 1.0. Of the 1,399 patients, 29 were identified to be at high risk of a cardiovascular event and 20 of those received follow-up testing with EKG or echo within 60 days of the risk assessment.

- The developer conducted a Kappa analysis between automated extracted EHR data and manual review. From October 14, 2020 – May 28, 2022, they reviewed 2,540 UCI charts, 82 UCSD charts and 45 UTENN charts. They checked the presence of follow-up tests and follow-up MFM and cardiology visits in 2,667 charts.
- For the 20 high risk patients with follow up visits, the developer reported that "all critical data elements have a kappa value equal to 1.0."
- However, the developer did not provide testing results for all key data elements (e.g., numerator, denominator, exclusions).

Exclusions

- Patients who discontinued care (no additional visit within 60 days after the risk assessment), are excluded from the denominator.
- The developer stated that it did not perform statistical tests because the exclusions are necessary for the measure to be clinically valid. Per NQF's validity criterion, all threats to validity that are relevant should be empirically assessed. This was not done.

Risk Adjustment

- The measure is not risk-adjusted or stratified. The developer argued that because Black race is one of the variables that contributes to the CVD risk score it is not necessary to include it in a risk adjustment model.

Meaningful Differences

- Rates of follow-up visits for high risk patients (n=29) was 33.3% (25th percentile), 50% (median), and 75% (75th percentile). A Pearson chi-square test indicates that measure rates in different clinics are significantly different (p=0.0013). It is unclear from the submission how many clinics are represented to achieve statistical significance.
- The rate for follow-up of positive CVD risk assessment in pregnant and postpartum patients was 65.7% at UCI, 70.8% at UCSD, and 38.4% at UTENN.

Missing Data

- There is no missing data for this measure. If there is no documentation of a follow-up test in the chart, the procedure is considered not to be done.

Comparability

- The measure only uses one set of specifications for this measure.

Questions for the Standing Committee regarding validity:

- *Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk adjustment approach, etc.)?*
- *The clinical rationale and clarity for exclusion of patients who discontinue care should be discussed by the Standing Committee.*

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

RATIONALE: The developer reported that "all critical data elements have a kappa value equal to 1.0." However, the developer did not provide testing results for all key data elements (e.g., numerator, denominator, exclusions). Additionally, the developer did not empirically assess exclusions to the measure.

Criterion 3. Feasibility

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

- The developer notes all data elements are in defined fields in the electronic claims.
- The developer highlighted that IT departments are prioritizing the transition to telehealth services resulting in delays in processing reliance agreements for IRB approval.
- The developer stated they are exploring the use of tests and labs as follow-up procedures rather than office visits due to office visits being meaningless if the patient was already considered high-risk.

Questions for the Standing Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Are the required data elements available in electronic form (e.g., EHR or other electronic sources)?*
- *Is the data collection strategy ready to be put into operational use?*

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

Criterion 4: Use and Usability

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

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

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

Current uses of the measure

Publicly reported?	☐ Yes	☒ No		
Current use in an accountability program?	☐ Yes	☒ No	☐ UNCLEAR	
Planned use in an accountability program?	☒ Yes	☐ No	☐ NA	

Accountability program details

- The measure was submitted to CMS in April 2022 for the public reporting program for hospital outpatient and inpatient quality reporting programs.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: (1) Those being measured have been given performance results or data, as well as assistance with interpreting

the measure results and data; (2) Those being measured, and other users have been given an opportunity to provide feedback on the measure performance or implementation; and (3) This feedback has been considered when changes are incorporated into the measure.

Feedback on the measure provided by those being measured or others

- The developer provided summaries to clinical sites about their performance and reviewed with the clinicians their performance over time and in comparison to other sites.
- The measures were reviewed with the co-investigators at each site and semi-structured interviews were conducted with five clinicians at each site. Overall, clinicians appreciated the ability to monitor their performance and get a benchmark of their peer's performance.

***Questions for the Standing Committee**:*

- *How have (or can) the performance results be used to further the goal of high quality, efficient healthcare?*
- *How has the measure been vetted in real-world settings by those being measured or others?*

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

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

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

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

Improvement results

- Improvement results are not available because they only have baseline results.

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

Unexpected findings (positive or negative) during implementation

- Patients and clinicians commented on improved patient awareness of the immediate and lifetime risk of developing CVD.
- The consistent use of the tool has raised awareness of the importance of CVD risk assessment among obstetricians.

Potential harms

- None noted.

Additional Feedback:

- At the time of this preliminary analysis development, MAP recommendations for this measure were not available

***Questions for the Standing Committee**:*

- *How can the performance results be used to further the goal of high quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

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

Criterion 5: Related and Competing Measures

Related Measures

- NQF #0607 Pregnant women that had syphilis screening
- NQF #1927 Cardiovascular Health Screening for People with Schizophrenia or Bipolar Disorder Who Are Prescribed Antipsychotic Medications

Harmonization

- The developer states that these measures are harmonized to the extent possible, noting that the syphilis measure has the same target population, but the focus is different. For the cardiovascular health screening measure, the developer states that this measure does not focus on the obstetric population and is a secondary preventive measure (assessing treatment of individuals already identified with cardiovascular health issues).

Criteria 1: Importance to Measure and Report

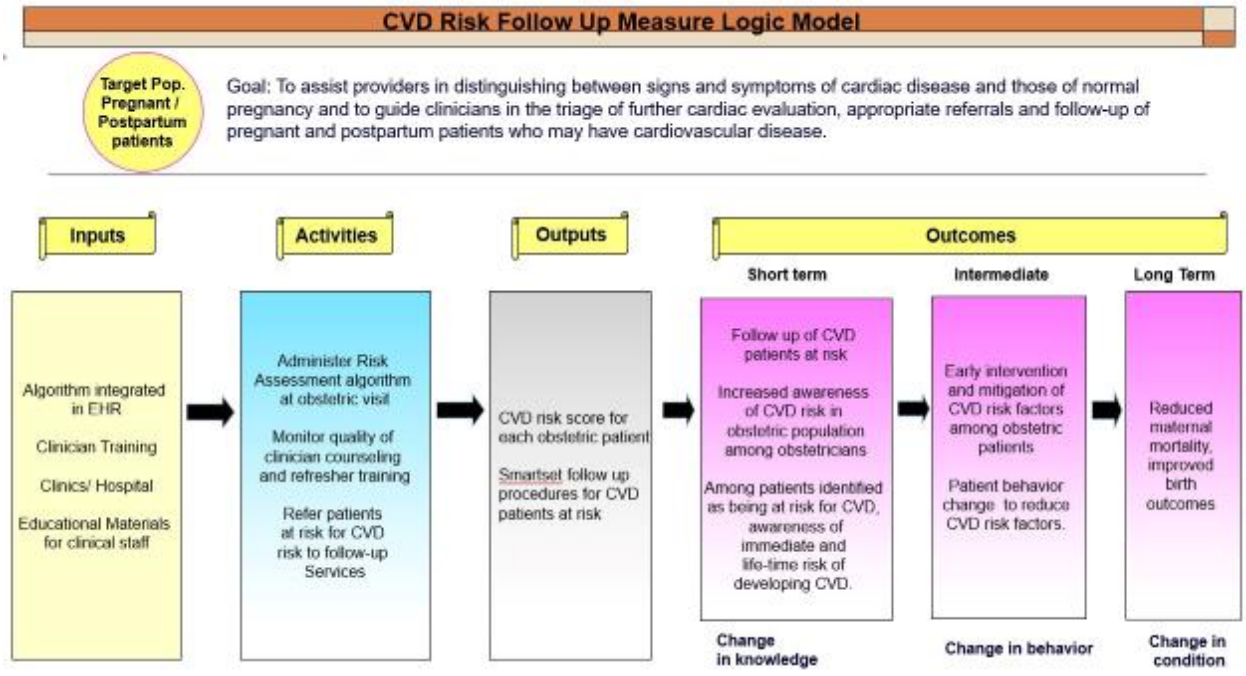
1a. Evidence

1a.01. Provide a logic model.

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

[Response Begins]

The following is the logic model of our FollowUp measure that describes the process and outcome of our measure.



[Response Ends]

1a.02. Select the type of source for the systematic review of the body of evidence that supports the performance measure.

A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data.

[Response Begins]

Other (specify)

[Other (specify) Please Explain]

The CMQCC Cardiovascular Disease in Pregnancy and Postpartum Task Force was charged with developing a toolkit that includes an overview of clinical assessment and management strategies based on risk factors and presenting signs and symptoms. The key components of the Toolkit include an algorithm developed to guide stratification and initial evaluation of symptomatic or high-risk pregnant or postpartum patients.

The goal of the algorithm is to assist providers in distinguishing between signs and symptoms of cardiac disease and those of normal pregnancy and to guide clinicians in the triage of further cardiac evaluation, appropriate referrals, and follow-up of pregnant and postpartum patients who may have cardiovascular disease. Drawing from the literature and analysis of cardiovascular deaths reviewed in the California Pregnancy Associated Mortality Review (CA-PAMR), the authors created this algorithm based on risk factors, symptoms, vital sign abnormalities, and physical examination findings commonly identified in patients who die of various types of cardiovascular disease.

[Response Ends]

If the evidence is not based on a systematic review, skip to the end of the section and do not complete the repeatable question group below. If you wish to include more than one systematic review, add additional tables by clicking "Add" after the final question in the group.

Evidence - Systematic Reviews Table (Repeatable)

Group 1 - Evidence - Systematic Reviews Table

1a.03. Provide the title, author, date, citation (including page number) and URL for the systematic review.

[Response Begins]

N/A

[Response Ends]

1a.04. Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the systematic review.

[Response Begins]

N/A

[Response Ends]

1a.05. Provide the grade assigned to the evidence associated with the recommendation, and include the definition of the grade.

[Response Begins]

N/A

[Response Ends]

1a.06. Provide all other grades and definitions from the evidence grading system.

[Response Begins]

N/A

[Response Ends]

1a.07. Provide the grade assigned to the recommendation, with definition of the grade.

[Response Begins]

N/A

[Response Ends]

1a.08. Provide all other grades and definitions from the recommendation grading system.

[Response Begins]

N/A

[Response Ends]

1a.09. Detail the quantity (how many studies) and quality (the type of studies) of the evidence.

[Response Begins]

N/A

[Response Ends]

1a.10. Provide the estimates of benefit, and consistency across studies.

[Response Begins]

N/A

[Response Ends]

1a.11. Indicate what, if any, harms were identified in the study.

[Response Begins]

N/A

[Response Ends]

1a.12. Identify any new studies conducted since the systematic review, and indicate whether the new studies change the conclusions from the systematic review.

[Response Begins]

N/A

[Response Ends]

1a.13. If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, describe the evidence on which you are basing the performance measure.

[Response Begins]

Empirical Evidence

[Response Ends]

1a.14. Briefly synthesize the evidence that supports the measure.

[Response Begins]

Cardiovascular disease (CVD) is the leading cause of maternal mortality in the United States, accounting for over one-third of all pregnancy-related deaths.[1] *Peripartum cardiomyopathy (PPCM) constitutes the largest group among CVD-related deaths.* **24%** of ALL CVD pregnancy-related deaths (and 31% of cardiomyopathy deaths) were determined to be **potentially preventable.** [2]

CVD also accounts for many folds higher maternal morbidity, a longer length of hospital stays, intensive care unit (ICU) admissions, and future pregnancy risks.[3]

Racial/ethnic disparities in pregnancy-related mortality have also been well established.[4, 5], African American patients exhibit 3-12 times higher mortality [1, 6, 7] as they are more likely to have pre-existing CVD,[3] hypertensive disorders of pregnancy [3, 5] and peripartum cardiomyopathy (PPCM) [5,8] when compared to patients from other racial /ethnic groups.

Timely diagnosis of CVD is critical; however, this may be challenging due to:

1. Pregnancy is a state of hemodynamic stress that may lead to signs and symptoms that are very similar to those of CVD, such as shortness of breath, fatigue, and swelling.[9]
2. Healthcare providers generally do not suspect CVD when evaluating pregnant or postpartum patients with symptoms that may signify an underlying diagnosis of CVD.

There is a need to establish a standardized CVD risk assessment tool to triage pregnant and postpartum patients and provide standardized options of appropriate follow-up. This population-wide risk assessment is likely to reduce CVD-related morbidity and mortality, particularly among African American patients. The proposed measure will monitor follow-up to universal cardiovascular risk assessment in all pregnant patients at their first encounter with an obstetrics provider.[10] The tool facilitates clinicians to evaluate pregnant or postpartum patients presenting with symptoms such as shortness of breath, cough, or excessive fatigue in the context of risk factors, vital sign abnormalities, and abnormal physical examination findings.[1] Use of this measure improves the accurate diagnosis of heart failure rather than attributing symptoms of persistent cough and shortness of breath and bilateral infiltrates on chest X-ray to pneumonia or pregnancy-related.

[Response Ends]

1a.15. Detail the process used to identify the evidence.

[Response Begins]

Cardiovascular disease (CVD) is the leading cause of maternal mortality in the United States and California. CVD accounts for >33% of all pregnancy-related deaths in the US and 25% of pregnancy-related deaths in CA (2002-2006). Data from the California Pregnancy Associated Mortality Review (CA-PAMR)[1] of deaths occurring from 2002-2006 show the following:

- Only a small fraction of these patients had a known diagnosis of cardiovascular disease prior to death.[2]
- Most patients who died had presented with symptoms either during pregnancy or after childbirth.
- A significantly higher proportion of patients sustain short- and long-term morbidity due to undiagnosed or delayed diagnosis of cardiovascular disease, as evidenced by the fact that one of every three intensive care admissions in pregnancy and postpartum period is related to cardiac disease.[3,4]
- 25% of these deaths may have been prevented if heart disease was diagnosed earlier.[2,3,5]

Pregnant and postpartum patients who die from CVD represent the most extreme consequence of missed or delayed recognition of CVD. Accordingly, any triage algorithm should be able to detect the most serious cases and not return a 'false negative' assessment in a patient with underlying CVD. To assess how well the triage algorithm would have identified pregnant and postpartum patients with the most need of further work-up, we compared the 64 cardiovascular disease deaths identified by CA-PAMR for 2002-2006, using the seven critical risks and abnormalities, including heart rate, systolic blood pressure, respiration rate, oxygen saturation, tachypnea, cough, and wheezing. We found that the use of the algorithm would have identified 56 out of 64 (88%) cases of CVD.[1] The proportion of patients identified increased to 93% when we restricted comparison to the 60 cases of patients who were symptomatic or had sufficient documentation with which to compare to the algorithm.[1]

To address these issues, CMQCC together with the California Department of Public Health: Maternal, Child and Adolescent Health Division published the *Improving Health Care Response to Cardiovascular Disease in Pregnancy and Postpartum* Toolkit in 2017.[2] The California Maternal Quality Care Collaborative (CMQCC) developed a CVD risk assessment algorithm, that guides stratification and initial clinical evaluation of symptomatic or high-risk pregnant or postpartum patients. The toolkit includes a risk assessment algorithm, which guides the stratification and initial

evaluation of symptomatic or high-risk pregnant or postpartum patients. The algorithm risk stratifies patients using 18 parameters including patient's history, abnormal symptoms, vital signs, and physical examination findings to identify patients who warrant further cardiac workup. The CMQCC Cardiovascular disease in pregnancy toolkit also includes resources for providers, infographics for patients on signs and symptoms of CVD, future CVD risk and long-term health issues, contraception options, and planning a pregnancy with known CVD. The toolkit also includes a discussion on racial and ethnic disparities in CVD prevention and diagnosis.

The Alliance for Innovation on Maternal Health Cardiac Conditions in Obstetrical Care includes the CMQCC CVD Assessment Algorithm for Pregnant and Postpartum Patients in the Cardiac Conditions in Obstetrical Care Bundle (COCC).[6] In the bundle, cardiac conditions refer to disorders of the cardiovascular system which may impact maternal health. Such disorders may include congenital heart disease or acquired heart disease, including but not limited to cardiac valve disorders, cardiomyopathies, arrhythmias, coronary artery disease, pulmonary hypertension, and aortic dissection despite limitations, recognized as an emerging best practice and an important tool for assessing symptoms and risk in a standardized way.

The American College of Obstetricians and Gynecologists (ACOG) recently endorsed the California (CA) cardiovascular disease (CVD) risk assessment algorithm for pregnant and postpartum patients. The aim is to prospectively determine risk-assessed-positive and true-positive rates of CVD among patients across two populations.

For the initial implementation, a prospective cohort study of obstetrical patients from April 2018 to July 2019 at academic medical centers in CA and New York (NY) was conducted.[7] There were 846 patients who had a risk assessment. There was an attempt to complete a risk assessment for all patients at least once during their pregnancy care (prenatal or postpartum). Patients who had a positive risk assessment ("Red Flags," >3–4 moderate risk factors, abnormal physical examination, and persistent symptoms) underwent further testing. The primary outcome was the risk assessed-positive rate. Secondary outcomes included the true-positive rate and the strength of each moderate factor in predicting a positive CVD risk assessment.

The overall risk assessed-positive rate was 8% (5% in CA vs. 19% in NY). The sites differed in ethnicity, that is, African American patients (2.7% in CA vs. 35% in NY, $p < 0.01$) and substance use (2.7 vs. 5.6%, $p < 0.04$). The true-positive rate was 1.5% at both sites. The percentage of risk assessed-positive patients who did not complete follow-up studies was higher in NY (70%) than in CA (27%). CVD was confirmed in 30% with positive risk assessments with complete follow-up.[7] Combinations of moderate factors were the main driver of risk assessment-positive rates in both populations. This is the first data describing the performance of the CVD risk assessment algorithm in the general obstetric population. Factors, such as the proportion of African American patients affect the likelihood of a positive risk assessment. The CVD risk assessment algorithm highlights patients at higher lifetime risk of CVD and may identify a group that could be targeted for more direct care transitions postpartum. Data may be used to design a larger validation study.

[Response Ends]

1a.16. Provide the citation(s) for the evidence.

[Response Begins]

References for 1a.14):

1. Creanga AA, Syverson C, Seed K, Callaghan WM. Pregnancy-Related Mortality in the United States, 2011–2013. *Obstet Gynecol*. 2017;130(2):366-373. doi:10.1097/AOG.0000000000002114.
2. Hameed AB, Foster E, Main EK, Khandelwal A, Lawton ES. Cardiovascular Disease Assessment in Pregnant and Postpartum Women | California Maternal Quality Care Collaborative. Cardiovascular Disease in Pregnancy Toolkit. Published November 2017. Accessed June 14, 2019. https://www.cmqcc.org/resource/cardiovascular-disease-assessment-pregnant-and-postpartum-women.
3. Fraser A, Nelson SM, Macdonald-Wallis C, et al. Associations of pregnancy complications with calculated cardiovascular disease risk and cardiovascular risk factors in middle age: the Avon Longitudinal Study of Parents and Children. *Circulation*. 2012;125(11):1367-1380. doi:10.1161/CIRCULATIONAHA.111.044784

4. Say L, Pattinson RC, Gülmezoglu AM. WHO systematic review of maternal morbidity and mortality: the prevalence of severe acute maternal morbidity (near miss). *Reprod Health*. 2004;1(1). doi:10.1186/1742-4755-1-3.
5. Petersen EE, Davis NL, Goodman D, et al. Racial/Ethnic Disparities in Pregnancy-Related Deaths — United States, 2007–2016. *MMWR Morb Mortal Wkly Rep*. 2019; 68:762–765. doi:10.15585/mmwr.mm6835a3.
6. Hameed AB, Lawton ES, McCain CL, et al. Pregnancy-related cardiovascular deaths in California: beyond peripartum cardiomyopathy. *Am J Obstet Gynecol*. 2015;213(3):379.e1-379.e10. doi:10.1016/j.ajog.2015.05.008.
7. Boyd LC, Johnson T, Langston A, Mulready-Ward C, Peña J, Wilcox WC. Pregnancy-Associated Mortality Report. New York City, 2006-2010. New York City: New York City Department of Health and Mental Hygiene. Accessed June 14, 2019. https://www1.nyc.gov/assets/doh/downloads/pdf/ms/pregnancy-associated-mortality-report.pdf.
8. Gentry MB, Dias JK, Luis A, Patel R, Thornton J, Reed GL. African-American Women Have a Higher Risk for Developing Peripartum Cardiomyopathy. *J Am Coll Cardiol*. 2010;55(7):654-659. doi:10.1016/j.jacc.2009.09.043.
9. Hunter S, Robson SC. Adaptation of the maternal heart in pregnancy. *Br Heart J*. 1992;68(6):540-543. doi:10.1136/hrt.68.12.540.
10. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2010;122(25):e584-636. doi:10.1161/CIR.0b013e3182051b4c.
11. Abbas AE, Lester SJ, Connolly H. Pregnancy and the cardiovascular system. *Int J Cardiol*. 2005;98(2):179-189. doi:10.1016/j.ijcard.2003.10.028.

References for 1a.15):

1. Hameed A, Main EK, Lawton E, Morton CH, CA-PAMR Committee. Validation Study of Cardiovascular Disease Assessment Algorithm for Pregnant and Postpartum Women among Pregnancy-Related Cardiovascular Deaths in California, 2002-2006. In: University CMQCCaS, ed. 2015; Palo Alto, CA.
2. Hameed AB, Lawton ES, McCain CL, et al. Pregnancy-related cardiovascular deaths in California: beyond peripartum cardiomyopathy. *Am J Obstet Gynecol*. 2015;213(3):379.e1-379.e10. doi:10.1016/j.ajog.2015.05.008.
3. Kuklina E, Callaghan W. Chronic heart disease and severe obstetric morbidity among hospitalisations for pregnancy in the USA: 1995-2006. *BJOG*. 2011;118(3):345-352. doi:10.1111/j.1471-0528.2010.02743.x
4. Small MJ, James AH, Kershaw T, Thames B, Gunatilake R, Brown H. Near-miss maternal mortality: cardiac dysfunction as the principal cause of obstetric intensive care unit admissions. *Obstet Gynecol*. 2012;119(2 Pt 1):250-255. doi:10.1097/AOG.0b013e31824265c7
5. Gunderson EP, Croen LA, Chiang V, Yoshida CK, Walton D, Go AS. Epidemiology of peripartum cardiomyopathy: incidence, predictors, and outcomes. *Obstet Gynecol*. 2011;118(3):583-591. doi:10.1097/AOG.0b013e318229e6de
6. Alliance for Innovation on Maternal Health. Cardiac Conditions in Obstetrical Care Bundle. Accessed May 9, 2022. https://safehealthcareforeverywoman.org/wp-content/uploads/FINAL-U2_AIM_Bundle_CCOC.pdf.
7. Blumenthal EA, Crosland BA, Senderoff D, et al. California Cardiovascular Screening Tool: Findings from Initial Implementation. *AJP Rep*. 2020;10(4):e362-e368. doi:10.1055/s-0040-1718382

[Response Ends]

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

1b.01. Briefly explain the rationale for this measure.

Explain how the measure will improve the quality of care, and list the benefits or improvements in quality envisioned by use of this measure.

[Response Begins]

The implementation of metrics on the adherence to standards signals to clinicians the endorsement of management and value of the metric as a signal to provide quality care to obstetric patients. The measure is easy to understand and can be calculated on the unit or even individual clinician performance (if the patient is assigned to one clinician) in addition to systemwide feedback. The measure allows for the identification of low-performing sites or clinicians and to address modifiable gaps in diagnostic excellence.

CVD is the leading cause of maternal mortality during pregnancy and the postpartum period in the United States. Most pregnant and postpartum patients who die of CVD do not have a known diagnosis of CVD. Diagnosis of CVD in pregnancy may be challenging as signs and symptoms of normal pregnancy mimic those of CVD which may be missed by the health care providers.

Our CVD quality measures target the childbearing age population who may be at high risk of CVD and access the health care system for maternity services. The universal implementation of a standardized tool to measure the clinic or facility's performance in following up on patients who are identified to be at risk will provide a uniform assessment of patients' CVD risk regardless of the clinicians' diagnostic skills and experience. Information about the follow-up rates will inform quality improvement strategies.

The CVD measure during pregnancy/postpartum care is bound to increase education and awareness in this population and will empower patients to seek early medical care if new signs and symptoms that may be suggestive of CVD develop. Our measure may have implications for long-term health outcomes with improvements in the CVD risk factor profile in the future. The use of a standardized CVD measure to risk stratify pregnant and postpartum patients may improve the timely identification of CVD, thereby decreasing maternal morbidity and/or mortality.

Additionally, the training for the use of the tool and reporting of the clinic's performance compared to other clinic sites in the same hospital network has raised awareness of the importance of CVD risk assessment among clinicians, even in clinics serving primarily low-risk patients. The ease of use (less than 1 minute to do the CVD risk assessment) of the tool allows obstetricians to systematically identify patients who are at risk for CVD and need follow-up on more thorough monitoring during the pregnancy. Additionally, we have anecdotal evidence that administration of the tool and providing patients with a risk score has improved patient awareness of the immediate and lifetime risk of developing CVD, which drives changes in health behavior. We have not seen any evidence that the follow-up of patients who were deemed at high risk for CVD led to inappropriate use of resources.

[Response Ends]

1b.02. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.

Include mean, std dev, min, max, interquartile range, and scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]

Performance scores on the measure at the specified level of analysis, September 2020 - February 2022

Level of Analysis	Performance Score
Min	0.0%
Max	100.0%
Mean	59.0%
SD	29.1%
Median	64.7%

Level of Analysis	Performance Score
IQR	50.0%
Score by Decile	*
10th Pctl	22.0%
20th Pctl	31.7%
30th Pctl	42.7%
40th Pctl	50.0%
50th Pctl	64.7%
60th Pctl	75.0%
70th Pctl	81.3%
80th Pctl	85.4%
90th Pctl	90.0%

* *Cell intentionally left empty*

[Response Ends]

1b.03. If no or limited performance data on the measure as specified is reported above, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement. Include citations.

[Response Begins]

N/A

[Response Ends]

1b.04. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]

Proportion of pregnant/postpartum patients at three hospital networks percent of patients who received a follow-up test within 60 days after a positive CVD risk score by demographic variables, the timing of assessment, and three six-month time periods, September 2020 – February 2022

Population Group	Total *n*	*Column %*	Received follow-up test *n*	*Row %*	Chi square test or Fisher's exact test p-value**
Overall	312	100.0%	145	46.5%	*
Health network	*	*	*	*	<.0001
UCI	35	11.2%	23	65.7%	*
UCSD	48	15.4%	34	70.8%	*
UTENN	229	73.4%	88	38.4%	*
Age group	*	*	*	*	0.0326
<20	7	2.2%	0	0.0%	*
20-29	132	42.3%	66	50.0%	*
30-39	142	45.5%	68	47.9%	*
40+	31	9.9%	11	35.5%	*
Race	*	*	*	*	0.0068
Black	121	38.8%	49	40.5%	*
White	124	39.7%	59	47.6%	*
AAPI	12	3.8%	11	91.7%	*
Others	55	17.6%	26	47.3%	*
Ethnicity	*	*	*	*	0.2433
Hispanic	61	19.6%	28	45.9%	*
Non-Hispanic	244	78.2%	116	47.5%	*
Unknown	7	2.2%	1	93.3%	*
Race/Ethnicity	*	*	*	*	0.0012
Non-Hispanic White	98	31.4%	43	43.9%	*
Non-Hispanic Black	120	38.5%	48	40.0%	*
Hispanic	61	19.6%	28	45.9%	*
AAPI	12	3.8%	11	91.7%	*
Others/unknown	21	6.7%	15	71.4%	*
Insurance	*	*	*	*	0.0004
Public (Medicaid, Military, government)	111	35.6%	43	38.7%	*

Population Group	Total *n*	*Column %*	Received follow-up test *n*	*Row %*	Chi square test or Fisher's exact test p-value**
Private (Commercial, Managed care)	179	57.4%	85	47.5%	*
Self-pay	4	1.3%	1	25.0%	*
Unknown	18	5.8%	16	88.9%	*
Timing	*	*	*	*	<.0001
Prenatal	172	55.1%	76	44.2%	*
Postpartum	140	44.9%	69	49.3%	*
Period	*	*	*	*	0.4609
09/01/20 - 02/28/21	74	23.7%	36	48.6%	*
03/01/21- 08/31/21	119	38.1%	59	49.6%	*
09/01/21- 02/28/22	119	38.1%	50	42.0%	*

** Chi-square test or Fisher's exact test testing different distribution of screening status by social-demographic category.

**Cells intentionally left empty,*

[Response Ends]

1b.05. If no or limited data on disparities from the measure as specified is reported above, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in above.

[Response Begins]

N/A

[Response Ends]

1c. Composite – Quality Construct and Rationale

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

sp.01. Provide the measure title.

Measure titles should be concise yet convey who and what is being measured (see What Good Looks Like).

[Response Begins]

CVD Risk Follow-up Measure - Proportion of patients with a positive CVD risk assessment who receive follow-up care

[Response Ends]

sp.02. Provide a brief description of the measure.

Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).

[Response Begins]

All pregnant and postpartum patients need to be systematically assessed for cardiovascular disease (CVD). Once identified as being at risk for CVD follow-up cardiac tests and consultations are scheduled. UCI implemented and tested a standardized CVD risk assessment algorithm that can be integrated into the EHR system and provides an immediate triage of patients as low and high risk for CVD.

This measure assesses the rate of pregnant and postpartum patients who are determined to be at risk for CVD)using a standardized risk assessment who received appropriate follow-up in the form of cardiology consultations and tests.

The unit of measurement is the individual patient, and the population is comprised of patients who have an outpatient or inpatient prenatal or postpartum visit at a clinic or facility. This includes pregnant and postpartum emancipated minors. The measure can be calculated at the hospital system level or clinic site level. A hospital system includes Labor and Delivery, outpatient care in a hospital or at affiliated clinics, and private providers contracted with the hospital for delivery. The denominator of the measure is comprised of all patients seen for prenatal or postpartum care at the measurement entity (hospital system or individual clinic sites) who were identified as "at risk for CVD" in a standardized CVD risk assessment such as the California Maternal Quality Care Collaborative (CMQCC) CVD risk assessment algorithm.

The aim is to have 100 percent of patients with a positive risk assessment receiving follow-up as recommended by American College of Obstetricians and Gynecologists (ACOG) guidelines. Implementation of the tool in three hospital systems showed a wide variation in the rate of follow-up between 38.4% and 70.8% of patients identified as high risk for CVD. The measure can be calculated annually.

[Response Ends]

sp.04. Check all the clinical condition/topic areas that apply to your measure, below.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Surgery: General*

[Response Begins]

Cardiovascular

Perinatal Health

Perinatal Health: Labor and Delivery

Perinatal Health: Post-Partum Care

Perinatal Health: Prenatal Care

[Response Ends]

sp.05. Check all the non-condition specific measure domain areas that apply to your measure, below.

[Response Begins]

Disparities Sensitive

Health and Functional Status: Total Health

Screening

[Response Ends]

sp.06. Select one or more target population categories.

Select only those target populations which can be stratified in the reporting of the measure's result.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]

Adults (Age >= 18)

Children (Age < 18)

Women

[Response Ends]

sp.07. Select the levels of analysis that apply to your measure.

Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

Clinician: Group/Practice

[Response Ends]

sp.08. Indicate the care settings that apply to your measure.

Check ONLY the settings for which the measure is SPECIFIED and TESTED.

[Response Begins]

Inpatient/Hospital

Outpatient Services

[Response Ends]

sp.09. Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials.

Do not enter a URL linking to a home page or to general information. If no URL is available, indicate "none available".

[Response Begins]

https://sites.uci.edu/cvdriskassessmentmeasures/implementation/

[Response Ends]

sp.12. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

Attach an excel or csv file; if this poses an issue, contact staff. Provide descriptors for any codes. Use one file with multiple worksheets, if needed.

[Response Begins]

Available in attached Excel or csv file

[Response Ends]

Attachment: 3735_3731_Data Dictionary For IRB 2020-5693-508.xlsx

Attachment: 3735_3731_CPT-ICD 10 Code Book-508.xlsx

sp.13. State the numerator.

Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome).

DO NOT include the rationale for the measure.

[Response Begins]

Patients who were identified to be at risk for CVD and received follow-up care within 60 days of the risk assessment.

[Response Ends]

sp.14. Provide details needed to calculate the numerator.

All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets.

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]

Patients receiving any follow-up care directed by the algorithm within 60 days of positive risk assessment (positive CVD algorithm calculated risk and signed by the clinician). Follow-up care (i.e., CVD testing) is identified if the patient has one or more ICD codes in their medical chart during the data abstraction time. These codes need to be listed in the data abstraction with the date of input. If CVD testing falls on a date after the risk assessment algorithm was completed but within the 60-day window, it is determined as a follow-up to the risk assessment. See excel attachment "CPT- ICD 10 Code Book".

[Response Ends]

sp.15. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

Pregnant and postpartum patients who have been identified to be at risk for cardiovascular disease (CVD) during the measurement period. Patients who were screened for CVD and had a pregnancy loss or stillbirth will remain in the cohort.

[Response Ends]

sp.16. Provide details needed to calculate the denominator.

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

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

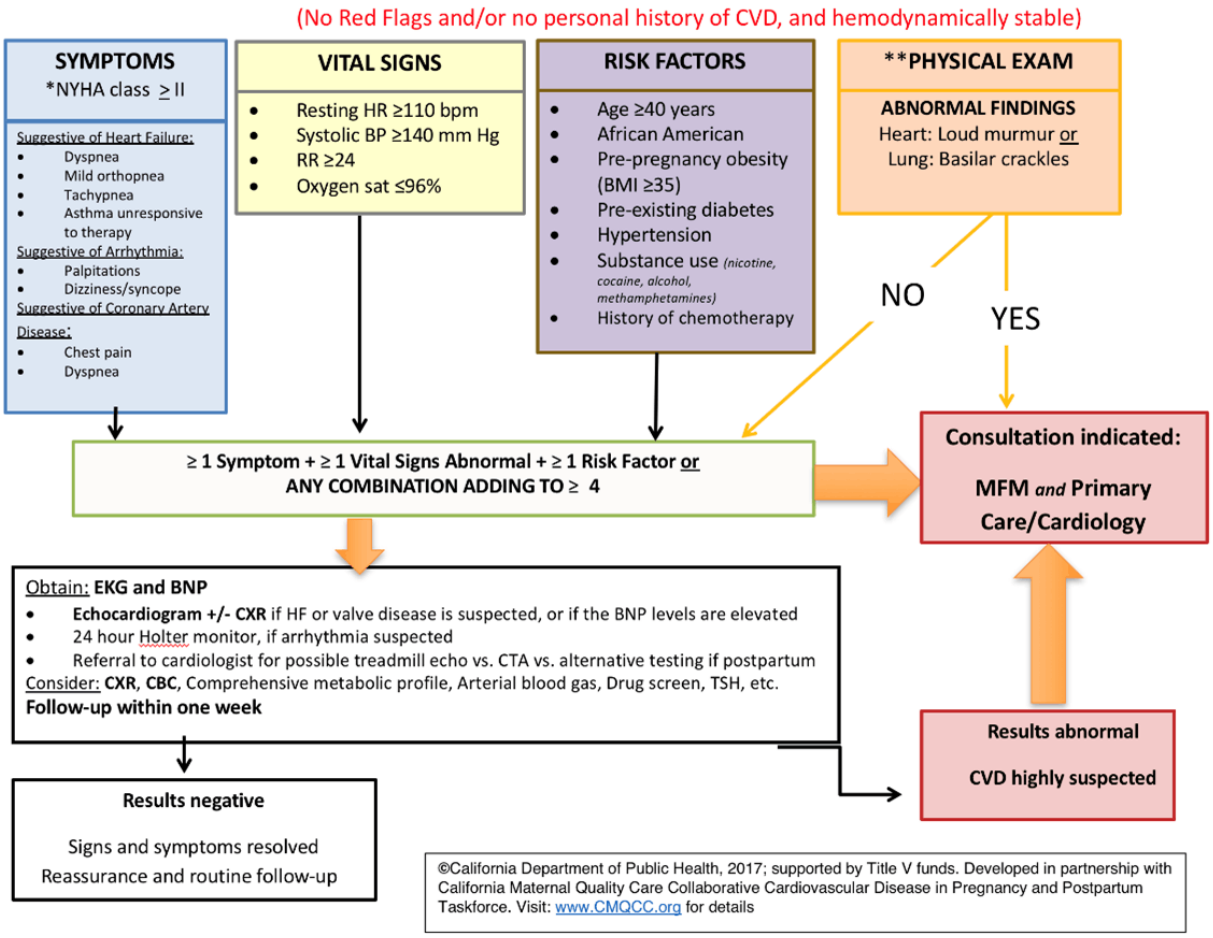
[Response Begins]

The CVD algorithm consists of 18 variables and is integrated in the electronic medical record (EMR), including demographics, vital signs, and any risk factors. The algorithm is administered at the first visit of women receive an obstetric visit (prenatal, labor and delivery, postpartum). It is repeated at presentation of clinical symptoms during the pregnancy. Most of the data are automatically pulled from the medical record (age, race, vital signs, symptoms). See the excel attachment "CPT – ICD 10 Code Book" for the full list of CVD confirmation CPT codes. Completion of the algorithm takes about 30 seconds.

Individual CVD risk scores will be calculated automatically once the algorithm is completed and will be part of the patient's medical record.

Patients are categorized as at risk for CVD and those that are not. Patients who have a positive risk assessment for CVD risk: Have >1 Symptom + >1 vital sign+ >1 Risk factor or ANY COMBINATION ADDING TO >4 (see CVD algorithm figure). Patients are considered to have a positive risk assessment if the CVD algorithm has calculated a positive risk score and is signed by a clinician.

Cardiovascular Disease Algorithm Risk Assessment Toolkit



The figure depicts the CVD Risk Assessment Algorithm toolkit that was integrated into our EHR. The arrows depict the direction of flow based on the risk factors and the output results based on the selection. A combination of each of the different factors can trigger an output of risks or no risks, which results in further follow-up.

The denominator of this measure consists of patients who are categorized to be at risk for CVD. Their patient chart is flagged with a banner and smartset orders with recommendations for follow up (labs/imaging/consults). See Codebook Group G – Cardiovascular Follow-up Visits.

The numerator consists of patients who completed the follow-up visit within 60 days after the risk assessment. The clinic IT system can provide regular updates of the CVD follow-up measures (quarterly, yearly) by clinic site, unit, or the complete hospital network to the medical director.

Medical and demographic data on the patients allow us to calculate the measure for subgroups and identify the need for targeted interventions. IT can extract clinical data on the cohort to identify subgroups in need of targeted interventions.

[Response Ends]

sp.17. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

We will exclude patients who discontinued care (no additional visit within 60 days after the risk assessment).

[Response Ends]

sp.18. Provide details needed to calculate the denominator exclusions.

All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]

We will exclude patients who discontinued care (no additional visit or medical procedure within 60 days after the risk assessment documented in the electronic health record).

[Response Ends]

sp.19. Provide all information required to stratify the measure results, if necessary.

Include the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate. Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format in the Data Dictionary field.

[Response Begins]

N/A

[Response Ends]

sp.20. Is this measure adjusted for socioeconomic status (SES)?

[Response Begins]

No

[Response Ends]

sp.21. Select the risk adjustment type.

Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.

[Response Begins]

No risk adjustment or risk stratification

[Response Ends]

sp.22. Select the most relevant type of score.

Attachment: If available, please provide a sample report.

[Response Begins]

Rate/proportion

[Response Ends]

sp.23. Select the appropriate interpretation of the measure score.

Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score

[Response Begins]

Better quality = Higher score

[Response Ends]

sp.24. Diagram or describe the calculation of the measure score as an ordered sequence of steps.

Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.

[Response Begins]

Cohort: All patients who have an outpatient or inpatient visit for pregnancy, labor and delivery, or postpartum care receive a risk assessment for CVD at the first encounter with the health care system or clinic. The algorithm can be calculated for clinician group/practices and individual clinicians regardless of their patient volume.

Inclusion criteria/denominator: The denominator for this measure is patients who receive a score to be at risk for CVD. In that case, the algorithm provides the clinician with a set of potential referrals for tests and cardiovascular consults (a sidebar with a Smartset in the electronic health record or a handout).

Numerator: The patient is counted to have had CVD follow-up if there is documentation of one or more ICD codes in the medical chart during the data abstraction time period. These codes need to be listed in the data abstraction with the date of input. If CVD testing falls on a date after the risk assessment algorithm is completed, it is determined as follow-up. See the attached document "CPT-ICD 10 Code Book" for the full list of CVD Testing codes.

The follow-up recommended orders include EKG and BNP:

- Echocardiogram +/- CXR if HF or valve disease is suspected, or if the BNP levels are elevated
- B-Type Natriuretic Peptide (BNP)
- Visit diagnosis: heart disease during pregnancy, antepartum [O99.419, I51.9]
- Follow-up within 1 week for Cardiovascular risk assessment testing results

Other orders the algorithm offers for the clinician to consider:

- Consult/Referral to Maternal Fetal Medicine
- Consult/Referral to a Cardiology Clinic
- Consult/Referral to Internal Medicine
- Complete 2D ECHO with Image Enhancement Agent if necessary
- Holter monitor
- Thyroid cascade
- CBD w/Diff
- Comprehensive Metabolic Panel

CVD confirmation is identified if the patient has one or more ICD codes in their medical chart during the data abstraction time period. If CVD confirmation falls on a date prior to CVD algorithm use with a patient who has completed the algorithm, it is considered an exclusion and does not require a CVD risk assessment.

The data on individual patients can be aggregated by the EHR reporting system and be requested by the medical director on a regular basis for the site and quality improvement activities.

Time and Period of Data: Depending on the patient volume, the measure can be calculated on an annual or a quarterly basis for public reporting purposes.

Data extraction for public reporting purposes: The IT department extracts the number of eligible patients (Medical Record Number, visit date, denominator) and the number of patients who received a risk assessment (Date r isk assessment was completed, numerator). Additional data for stratification can be clinic site, clinician, race/ethnicity of mother, insurance,

gestational age, date of birth of infant (to identify whether the assessment was completed during pregnancy or postpartum.

[Response Ends]

sp.27. If measure testing is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

Examples of samples used for testing:

- *Testing may be conducted on a sample of the accountable entities (e.g., hospital, physician). The analytic unit specified for the particular measure (e.g., physician, hospital, home health agency) determines the sampling strategy for scientific acceptability testing.*
- *The sample should represent the variety of entities whose performance will be measured. The 2010 Measure Testing Task Force recognized that the samples used for reliability and validity testing often have limited generalizability because measured entities volunteer to participate. Ideally, however, all types of entities whose performance will be measured should be included in reliability and validity testing.*
- *The sample should include adequate numbers of units of measurement and adequate numbers of patients to answer the specific reliability or validity question with the chosen statistical method.*
- *When possible, units of measurement and patients within units should be randomly selected.*

[Response Begins]

N/A

[Response Ends]

sp.30. Select only the data sources for which the measure is specified.

[Response Begins]

Electronic Health Records

Paper Medical Records

[Response Ends]

sp.31. Identify the specific data source or data collection instrument.

For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.

[Response Begins]

The California Maternal Quality Care Collaborative (CMQCC) developed a CVD risk assessment algorithm guides stratification and initial evaluation of symptomatic or high-risk pregnant or postpartum patients. This is currently (August 2022) the only standardized tool to identify pregnant and postpartum patients who are at risk for CVD. The acceptability of the measure is further strengthened by the support it has received from ACOG, and its inclusion in the CVD bundle by the Alliance for Innovation for Maternal Health.

The algorithm is integrated into EPIC and Cerner electronic health records and all data can be retrieved from the EHR.

[Response Ends]

sp.32. Provide the data collection instrument.

[Response Begins]

Available at measure-specific web page URL identified in sp.09

[Response Ends]

Attachment: 3735 _3735_CVD Risk Assessment Algorithm-508.pdf

Attachment: 3735 _CVD Risk Assessment Algorithm.png

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate fields in the Scientific Acceptability sections of the Measure Submission Form.

- Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.
- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact NQF staff with any questions. Check for resources at the Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the 2021 Measure Evaluation Criteria and Guidance.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a. Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure;

AND

If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration
- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful 16 differences in performance;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias.

2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and

2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(if not conducted or results not adequate, justification must be submitted and accepted)

Definitions

Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

Risk factors that influence outcomes should not be specified as exclusions.

With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v.$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous (Year) Submission:

Testing from the previous submission here.

2a.01. Select only the data sources for which the measure is tested.

[Response Begins]

Electronic Health Records

Paper Medical Records

[Response Ends]

2a.02. If an existing dataset was used, identify the specific dataset.

The dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry).

[Response Begins]

Electronic Health Record data in EPIC from the University of California, Irvine, and University of California, San Diego

Electronic Health Record data in Cerner from St. Thomas/University of Tennessee

[Response Ends]

2a.03. Provide the dates of the data used in testing.

Use the following format: "MM-DD-YYYY - MM-DD-YYYY"

[Response Begins]

09-01-2020 - 02-28-2022

[Response Ends]

2a.04. Select the levels of analysis for which the measure is tested.

Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

Clinician: Group/Practice

[Response Ends]

2a.05. List the measured entities included in the testing and analysis (by level of analysis and data source).

Identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample.

[Response Begins]

Proportion of pregnant/postpartum patients at three hospital networks who received follow up test within 60 days after a positive CVD risk score by clinic site, September 2020 – February 2022

Clinical Site	Total *n*	*Column %*	Received Follow-up Test *n*	*Row %*	Type
UCI Health System	*	*	*	*	*
UCI FQHC ANA FAM MED	2	5.7%	1	50.0%	Family Medicine
UCI FQHC ANA OB/GYN	1	2.9%	0	0.0%	Obstetrics & Gynecology
UCI FQHC SA OB/GYN	16	45.7%	12	75.0%	Obstetrics & Gynecology
UCI MAN MFM	10	28.6%	6	60.0%	Maternal- Fetal Medicine
UCI MAN OB/GYN	4	11.4%	3	75.0%	Obstetrics & Gynecology
UCI TUSTIN OB/GYN	2	5.7%	1	50.0%	Obstetrics & Gynecology
UCSD Health System	*	*	*	*	*
CNV WOMENS HEALTH SVCS	6	12.5%	5	83.3%	Women's Health Services
DIR WOMENS HEALTH SVCS	8	16.7%	7	87.5%	Women's Health Services
MOS WOMENS HEALTH SVCS	12	25.0%	9	75.0%	Women's Health Services
UNC WOMENS HEALTH SVCS	2	4.2%	1	50.0%	Women's Health Services
VLJ WOMENS HEALTH SVCS	17	35.4%	11	64.7%	Women's Health Services
VTC WOMENS HEALTH SVCS	3	6.3%	1	33.3%	Women's Health Services
UTENN Health System	*	*	*	*	*
ST Midtown	178	77.7%	76	42.7%	General
ST River Park	10	4.4%	3	30.0%	General
ST Rutherford	41	17.9%	9	22.0%	General

* *Cells intentionally left empty*

[Response Ends]

2a.06. Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.

If there is a minimum case count used for testing, that minimum must be reflected in the specifications.

[Response Begins]

Number and descriptive characteristics of patients who received follow-up test within 60 days after a positive CVD risk score compared to total patients separated by data source, September 2020 – February 2022

Descriptive Characteristics	Total *n*	Column %	Received follow-up test *n*	Row %
UCI Health System	*	*	*	*
Overall	35	100.0%	23	65.7%
Age group	*	*	*	*
20-29	9	25.7%	8	88.9%
30-39	21	60.0%	14	66.7%
40+	5	14.3%	1	20.0%
Race	*	*	*	*
Black	6	17.1%	4	66.7%
White	24	68.6%	14	58.3%
AAPI	4	11.4%	4	100.0%
Others	1	2.9%	1	100.0%
Ethnicity	*	*	*	*
Hispanic	18	51.4%	11	61.1%
Non-Hispanic	17	48.6%	12	70.6%
Race/Ethnicity	*	*	*	*
Non-Hispanic White	7	20.0%	4	57.1%
Non-Hispanic Black	6	17.1%	4	66.7%
Hispanic	18	51.4%	11	61.1%
AAPI	4	11.4%	4	100.0%
Insurance	*	*	*	*
Public (Medicaid, Military, government)	21	60.0%	11	52.4%
Private (Commercial, Managed care)	5	14.3%	4	80.0%
Unknown	9	25.7%	8	88.9%
Timing	*	*	*	*
Prenatal	11	31.4%	6	54.5%
Postpartum	24	68.6%	17	70.8%
Period	*	*	*	*
09/01/20 - 02/28/21	15	42.9%	11	73.3%

Descriptive Characteristics	Total *n*	Column %	Received follow-up test *n*	Row %
03/01/21- 08/31/21	8	22.9%	5	62.5%
09/01/21- 02/28/22	12	34.3%	7	58.3%
UCSD Health System	*	*	*	*
Overall	48	100.0%	34	70.8%
Age group	*	*	*	*
<20	1	2.1%	0	0.0%
20-29	24	50.0%	19	79.2%
30-39	18	37.5%	13	72.2%
40+	5	10.4%	2	40.0%
Race	*	*	*	*
Black	9	18.8%	6	66.7%
White	14	29.2%	9	64.3%
AAPI	5	10.4%	5	100.0%
Others	20	41.7%	14	70.0%
Ethnicity	*	*	*	*
Hispanic	16	33.3%	9	56.3%
Non-Hispanic	31	64.6%	24	77.4%
Unknown	1	2.1%	1	93.3%
Race/Ethnicity	*	*	*	*
Non-Hispanic White	11	22.9%	7	63.6%
Non-Hispanic Black	9	18.8%	6	66.7%
Hispanic	16	33.3%	9	56.3%
AAPI	5	10.4%	5	100.0%
Others/unknown	7	14.6%	7	100.0%
Insurance	*	*	*	*
Public (Medicaid, military, government)	20	41.7%	12	60.0%
Private (Commercial, Managed care)	19	39.6%	14	73.7%
Self pay	9	18.8%	8	88.9%
Timing	*	*	*	*
Prenatal	17	35.4%	10	58.8%
Postpartum	31	64.6%	24	77.4%

Descriptive Characteristics	Total *n*	Column %	Received follow-up test *n*	Row %
Period	*	*	*	*
09/01/20 - 02/28/21	8	16.7%	6	75.0%
03/01/21- 08/31/21	18	37.5%	13	72.2%
09/01/21- 02/28/22	22	45.8%	15	68.2%
UTENN Health System	*	*	*	*
Overall	229	100.0%	88	38.4%
Age group	*	*	*	*
<20	6	2.6%	0	0.0%
20-29	99	43.2%	39	39.4%
30-39	103	45.0%	41	39.8%
40+	21	9.2%	8	38.1%
Race	*	*	*	*
Black	106	46.3%	39	36.8%
White	86	37.6%	36	41.9%
AAPI	3	1.3%	2	66.7%
Others	34	14.8%	11	32.4%
Ethnicity	*	*	*	*
Hispanic	27	11.8%	8	29.6%
Non-Hispanic	196	85.6%	80	40.8%
Unknown	6	2.6%	0	93.3%
Race/Ethnicity	*	*	*	*
Non-Hispanic White	80	34.9%	32	40.0%
Non-Hispanic Black	105	45.9%	38	36.2%
Hispanic	27	11.8%	8	29.6%
AAPI	3	1.3%	2	66.7%
Others/unknown	14	6.1%	8	57.1%
Insurance	*	*	*	*
Public (Medicaid, Military, government)	70	30.6%	20	28.6%
Private (Commercial, Managed care)	155	67.7%	67	43.2%
Self-pay	4	1.7%	1	25.0%
Timing	*	*	*	*

Descriptive Characteristics	Total *n*	Column %	Received follow-up test *n*	Row %
Prenatal	144	62.9%	60	41.7%
Postpartum	85	37.1%	28	32.9%
Period	*	*	*	*
09/01/20 - 02/28/21	51	22.3%	19	37.3%
03/01/21- 08/31/21	93	40.6%	41	44.1%
09/01/21- 02/28/22	85	37.1%	28	32.9%

**Cells intentionally left empty*

[Response Ends]

2a.07. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing.

[Response Begins]

We used the same data sets for empirical validity tests. However, for SNR reliability testing, we used the same data set but excluded ST Midtown (n=178) and ST Rutherf (n=41) for the purpose of parameter estimation.

[Response Ends]

2a.08. List the social risk factors that were available and analyzed.

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

[Response Begins]

We are limited to social risk factors that are included in the electronic medical record. Most of these risk factors are part of the risk assessment tool; these include age over 40 years, black race, substance use, and medical risk factors (obesity, diabetes, hypertension, history of chemotherapy). Stratifications can be performed by insurance status (commercial/private insurance vs. public insurance) as a proxy for socio-economic status and by geographic location (rural/urban, medically underserved area) using zip code. Patient zip codes also allow us to gauge patient community characteristics.

We conducted a logistic regression analysis with age, race/ethnicity, insurance status, and timing. We found statistically significant differences among those social risk factors (see table below).

Logistic Regression Analysis Table

Social Risk Factors	Total *n*	Column %	Had Follow-up Test	Row %	Chi square test p-value**
Age group	*	*	*	*	0.0326
<20	7	2.2%	0	0.0%	*
20-29	132	42.3%	66	50.0%	*

Social Risk Factors	Total *n*	Column %	Had Follow-up Test	Row %	Chi square test p-value**
30-39	142	45.5%	68	47.9%	*
40+	31	9.9%	11	35.5%	*
Race	*	*	*	*	0.0068
Black	121	38.8%	49	40.5%	*
White	124	39.7%	59	47.6%	*
AAPI	12	3.8%	11	91.7%	*
Others	55	17.6%	26	47.3%	*
Ethnicity	*	*	*	*	0.2433
Hispanic	61	19.6%	28	45.9%	*
Non-Hispanic	244	78.2%	116	47.5%	*
Unknown	7	2.2%	1	93.3%	*
Race/Ethnicity	*	*	*	*	0.0012
Non-Hispanic White	98	31.4%	43	43.9%	*
Non-Hispanic Black	120	38.5%	48	40.0%	*
Hispanic	61	19.6%	28	45.9%	*
AAPI	12	3.8%	11	91.7%	*
Others/unknown	21	6.7%	15	71.4%	*
Insurance	*	*	*	*	0.0004
Public (Medicaid, government)	111	35.6%	43	38.7%	*
Private (Commercial, Managed care)	179	57.4%	85	47.5%	*
Self pay	4	1.3%	1	25.0%	*
Unknown	18	5.8%	16	88.9%	*
Timing	*	*	*	*	<.0001
Prenatal	172	55.1%	76	44.2%	*
Postpartum	140	44.9%	69	49.3%	*

** Chi square test testing different distribution of screening status by social-demographic category

**Cells intentionally left empty*

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.09 check patient or encounter-level data; in 2a.010 enter "see validity testing section of data elements"; and enter "N/A" for 2a.11 and 2a.12.

2a.09. Select the level of reliability testing conducted.

Choose one or both levels.

[Response Begins]

Patient or Encounter-Level (e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements)

Accountable Entity Level (e.g., signal-to-noise analysis)

[Response Ends]

2a.10. For each level of reliability testing checked above, describe the method of reliability testing and what it tests.

Describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used.

[Response Begins]

Patient Encounter Level- Agreement between charts extracted from the EHR with manual abstraction. The statistic used to assess agreement was Kappa. As most of the patients received, EKG or echo as follow up procedure, we assessed this as primary outcome. We reviewed 1,399 patients that underwent CVD risk stratification using the CMQCC algorithm over an 18-month period at a the UCI network. We reviewed the rate of abnormal EKG or echo, defined as abnormal cardiac structure and/or function, among patients who were determined to be at increased risk for CVD. Of 29 patients identified to be at increased risk, 20 received follow-up testing with EKG or echo within 60 days of the risk assessment. Over half (65%) of the patients were found to have underlying EKG or echo abnormality. Abnormal cardiovascular testing results included findings such as sinus tachycardia (HR > 100 bpm), conduction delays, Wolff Parkinson-White syndrome, left ventricular hypertrophy/diastolic dysfunction, and chamber dilation. Using these results as a surrogate for CVD, the CMQCC risk assessment tool identified 13 cases of previously undiagnosed cardiovascular dysfunction in the study population.

Signal to Noise- We reviewed the follow-up codes in our sample and noticed that all follow-up office visits with MFM and cardiologists also had at least one of the follow-up tests. Hence, we used follow-up tests for the SNR analysis. We will review going forward whether office visit codes without any follow-up tests (EKG, echocardiogram, etc) should be considered follow-up for patients who were identified to be likely to have or be at risk of developing CVD."

We eliminate clinics with a relatively large sample size (Denominator, or n) that could have a disproportionate influence. I excluded if n>75th percentile+1.5*(interquartile range).

Across 15 clinics.

25th percentile: 2

75th percentile: 16

interquartile range: 14

75th percentile+1.5*(interquartile range)=37

- ST Midtown (n=178), and ST Rutherf (n=41) are removed for the purpose of the parameter estimation.

Use empirical Bayes shrinkage with n^2 weighting to estimate the signal and noise variances as outlined in Section 5. of Morris[1]:

$$\hat{A} = \sigma^2_{provider-to-provider}$$

$$s_i^2 = \sigma^2_{error}$$

Calculate $Reliability = \frac{\sigma^2_{provider-to-provider}}{\sigma^2_{provider-to-provider}+\sigma^2_{error}}$ for each clinic.

[Response Ends]

2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?

For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, NQF Measure Evaluation Criteria).

[Response Begins]

Patient Encounter Level: 1.0

Signal to Noise Analysis:

k 13

A (Signal Variance) 0.0059

SD 0.0768

b-hat (Mean) 0.666

V-bar 0.017777

Min SNR: 0.181

Max SNR:1.0

Median Reliability: 0.356

[Response Ends]

2a.12. Interpret the results, in terms of how they demonstrate reliability.

(In other words, what do the results mean and what are the norms for the test conducted?)

[Response Begins]

Patient Encounter Level: We are confident that we will be able to identify prenatal and postpartum patients. We are confident that we can identify those who have a positive CVD risk assessment score and who are not. We checked the data extracted by the UCI IT department with results from a manual review of a subset of charts. We reviewed any discrepancies to adjust the logic until the data pull was completely consistent with the gold standard.

We examined discrepancies and adjusted the logic until the data pull was fully aligned with the gold standard. Specifically, we noted that 100% of patients with a positive risk assessment at UCI and UCSD received follow-up. Upon further examination, we saw that a large number at UCI who had a negative risk assessment also entered in the numerator as "follow-up procedure received." These were patients who were seen at the MFM clinic for prenatal visits and it was not possible to distinguish between MFM visits due to a positive risk assessment or due to a normal prenatal visit appointment at these clinics. Although this was not an issue for other obstetric clinics, we removed "MFM office visit" from the list of "follow-up procedures." In addition, we noticed that the original IT script did not limit follow-up tests to 60 days after the risk assessment; after adding the restriction that follow-up procedures had to be performed within 60 days after the assessment, the percentage of completed follow-up up decreased at UCI and UCSD. The measure could provide meaningful and actionable data on the percentage of patients who received a CVD risk assessment at each clinic site. In each hospital system, we identified differences by clinician group or practice, ranging from 0% to 87.5%.

SNR: The median reliability of 0.356, implies most of the variability in a measure is attributable to measurement errors. The large measurement error is due to the small denominator in two of the sites (under 20), which results in an unstable rate. However, based on the data from the third hospital network (UTENN) we expect that this measurement would be stronger with additional data and a higher proportion of black patients.

[Response Ends]

2b. Validity

[Response Ends]

2b.01. Select the level of validity testing that was conducted.

[Response Begins]

Patient or Encounter-Level (data element validity must address ALL critical data elements)

[Response Ends]

2b.02. For each level of testing checked above, describe the method of validity testing and what it tests.

Describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used.

[Response Begins]

Patient Or Encounter Level: We conducted a Kappa analysis between automated extracted EHR data and manual reviewer. From October 14, 2020 – May 28, 2022, we reviewed at UCI 2,540 UCI charts, at UCSD 82 charts and at UTENN 45 charts from UTENN as part of the implementation of the algorithm. We checked the presence of follow-up tests and follow-up MFM and cardiology visits in 2,667 charts.

As most of the patients received, EKG or echo as follow-up procedure, we assessed this as the primary outcome. We reviewed 1,399 patients that underwent CVD risk stratification using the CMQCC algorithm over an 18-month period at a the UCI network. We reviewed the rate of abnormal EKG or echo, defined as abnormal cardiac structure and/or function, among patients who were determined to be at increased risk for CVD. Of 29 patients identified to be at increased risk, 20 received follow-up testing with EKG or echo within 60 days of the risk assessment. Over half (65%) of the patients were found to have underlying EKG or echo abnormality. Abnormal cardiovascular testing results included findings such as sinus tachycardia (HR > 100 bpm), conduction delays, Wolff Parkinson-White syndrome, left ventricular hypertrophy/diastolic dysfunction, and chamber dilation. Using these results as a surrogate for CVD, the CMQCC risk assessment tool identified 13 cases of previously undiagnosed cardiovascular dysfunction in the study population.

[Response Ends]

2b.03. Provide the statistical results from validity testing.

Examples may include correlations or t-test results.

[Response Begins]

Patient or Encounter Level: All critical data elements have kappa of 1.0.

[Response Ends]

2b.04. Provide your interpretation of the results in terms of demonstrating validity. (i.e., what do the results mean and what are the norms for the test conducted?)

[Response Begins]

Patient or Encounter Level: Patient or Encounter Level: We are confident to be able to identify prenatal and postpartum patients. We are confident to identify those who received follow-up after a positive CVD risk assessment score and those

who were assessed to be at low risk. We conducted a chart review of the presence of follow-up tests and follow-up MFM and cardiology visits in 2,540 charts.

We examined discrepancies and adjusted the logic as described in 2a.12.

[Response Ends]

2b.05. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified.

Describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided in Importance to Measure and Report: Gap in Care/Disparities.

[Response Begins]

Our analysis revealed considerable variation in the CVD risk assessment rates and follow-up in various clinical settings. Several differences were identified in the successful completion of CVD measures between the clinical sites, which were primarily based on the size of the clinic (number of patients seen), specialty (high-risk, general practice, family medicine), and staff turnover (drop when a new resident cohort starts). These differences were discussed at grand rounds and led to the identification of gaps in the clinical protocol and clinician training. Bivariate analysis was done using the Chi-square test or Fisher's exact test to examine the difference in categorical variables. Subset analysis was done for each entity as well. For this measure, the Q1 was 33.3%, the median=50%, and the Q3=75%, IQR=41.7%. Pearson chi-square test p=0.0013 indicates that measure 2 rates in different clinics are significantly different.

[Response Ends]

2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.

Examples may include number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined.

[Response Begins]

The rate for follow-up of positive CVD risk assessment in pregnant and postpartum patients was 65.7% at UCI, 70.8% at UCSD, and 38.4% at UTENN. There was a significant variation in the CVD risk assessment rates and follow-up in various clinical settings. Several differences were identified in the successful completion of CVD measures between the clinical sites, which were primarily based on the size of the clinic (number of patients seen) and specialty. Clinics that had a hard stop in the EHR ensured a 100 percent completion rate for the risk assessment. However, follow-up of patients who had a positive risk assessment was considerably lower than in the clinics without a hard stop, suggesting that clinicians had not fully adopted the value of the risk assessment tool.

[Response Ends]

2b.07. Provide your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities.

In other words, what do the results mean in terms of statistical and meaningful differences?

[Response Begins]

Clinics that had a hard stop in the EHR ensured a 100 percent completion rate for the risk assessment measure. However, follow-up of patients who had a positive risk assessment was considerably lower than in the clinics without a hard stop, suggesting barriers to care or that clinicians had not fully adopted the value of the risk assessment tool. Furthermore, the percentage of patients undergoing CVD risk assessment and follow-up of those identified as high risk varied by site

specialty, size, and automated vs. manual entry of the algorithm and demonstrated quality gaps within the same hospital system.

[Response Ends]

2b.08. Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.

Describe the steps—do not just name a method; what statistical analysis was used.

[Response Begins]

If there is no follow-up visit or procedure after a positive risk assessment, the missing value will be part of the measure calculation. Missing risk assessments will be analyzed by patient demographics as part of the measure reporting. We, therefore, do not have any missing data that would bias the results.

Cardiovascular Follow-up Procedure

We assessed the following CPT Codes for CVD Tests. A missing value in the following means a procedure was not completed.

Electrocardiogram = 93000, 93005, 93010

Brain natriuretic peptide test (BNP) = 37386

Echocardiogram = 93303. 93304, 93320, 93321, 93325, 93326

Holter monitor = 93224, 93225, 93226, 93227, 93230, 93231, 93233, 93235, 93236, 93237

Complete blood count (CBC) = 85025, 85027, G0306, H0307

Basic metabolic panel = 80048

Comprehensive metabolic panel = 80053

Arterial blood gas = 82803

Drug screen = 80307 (10 panel)

Thyroid stimulating hormone = 84439; 84443

[Response Ends]

2b.09. Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.

For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).

[Response Begins]

As the algorithm is based on data that is stored in the electronic health record system, there is no missing data that could be used in a sensitivity analysis. If there is no documentation of a follow-up test in the chart, the procedure is considered not to be done (e.g., no follow-up performed).

[Response Ends]

2b.10. Provide your interpretation of the results, in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders), and how the specified handling of missing data minimizes bias.

In other words, what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis was conducted, justify the selected approach for missing data.

[Response Begins]

We do not believe that the data is biased as there is no missing data in the dataset.

[Response Ends]

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

2b.11. Indicate whether there is more than one set of specifications for this measure.

[Response Begins]

No, there is only one set of specifications for this measure

[Response Ends]

2b.12. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications.

Describe the steps—do not just name a method. Indicate what statistical analysis was used.

[Response Begins]

[Response Ends]

2b.13. Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.

Examples may include correlation, and/or rank order.

[Response Begins]

[Response Ends]

2b.14. Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.

In other words, what do the results mean and what are the norms for the test conducted.

[Response Begins]

[Response Ends]

2b.15. Indicate whether the measure uses exclusions.

[Response Begins]

Yes, the measure uses exclusions.

[Response Ends]

2b.16. Describe the method of testing exclusions and what was tested.

Describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used?

[Response Begins]

Current data from the EHR was not designed to identify how many patients discontinued care. Upon manual validation of follow-up test results, we identified less than ten patients who discontinued care in our sample. This number was not sufficient to conduct a statistical analysis.

[Response Ends]

2b.17. Provide the statistical results from testing exclusions.

Include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores.

[Response Begins]

We did not perform statistical tests because the exclusions are necessary for the measure to be clinically valid.

[Response Ends]

2b.18. Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.

In other words, the value outweighs the burden of increased data collection and analysis. Note: If patient preference is an exclusion, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion.

[Response Begins]

The reason for the exclusion is that it does not make sense to administer risk assessment with known CVD and that the exclusion of patients who do not stay at the clinic for prenatal care can not necessarily be counted as an indicator of poor clinical care will reduce lost to follow-up.

[Response Ends]

2b.19. Check all methods used to address risk factors.

[Response Begins]

No risk adjustment or stratification

[Response Ends]

2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.

[Response Begins]

[Response Ends]

2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.

[Response Begins]

It is not risk-adjusted or stratified, because Black race is one of the variables that contribute to the risk score.

[Response Ends]

2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.

[Response Begins]

[Response Ends]

2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.

Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10 or other statistical tests; correlation of x or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).

[Response Begins]

[Response Ends]

2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.

[Response Begins]

[Response Ends]

2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.

Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.

[Response Begins]

[Response Ends]

2b.26. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (describe the steps—do not just name a method; what statistical analysis was used). Provide the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter "N/A" for questions about the statistical risk model discrimination and calibration statistics.

Validation testing should be conducted in a data set that is separate from the one used to develop the model.

[Response Begins]

[Response Ends]

2b.27. Provide risk model discrimination statistics.

For example, provide c-statistics or R-squared values.

[Response Begins]

[Response Ends]

2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).

[Response Begins]

N/A

[Response Ends]

2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.

The preferred file format is .png, but most image formats are acceptable.

[Response Begins]

[Response Ends]

2b.30. Provide the results of the risk stratification analysis.

[Response Begins]

[Response Ends]

2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).

In other words, what do the results mean and what are the norms for the test conducted?

[Response Begins]

[Response Ends]

2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.

Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.

[Response Begins]

[Response Ends]

Criterion 3. Feasibility

3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.

[Response Begins]

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)

Coded by someone other than person obtaining original information (e.g., DRG, ICD-10 codes on claims)

Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

[Response Ends]

3.02. Detail to what extent the specified data elements are available electronically in defined fields.

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.

[Response Begins]

ALL data elements are in defined fields in electronic claims

[Response Ends]

3.03. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using data elements not from electronic sources.

[Response Begins]

N/A

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]

N/A

[Response Ends]

3.06. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

[Response Begins]

The main challenge during implementation was caused by the COVID-19 pandemic, which resulted in IT departments having to prioritize the transition to telehealth services. As we gathered data from different health systems, and faced delays in processing reliance agreements for IRB approval for data transmission to UCI and processing of EMR data not included in the CVD algorithm Smart Sets. During the testing and validation of the extracted data, we had to revise data specifications with our IT team. For example, we noticed that the criterion "office visit with MFM" is meaningless if the patient was already considered a high-risk patient and seen at MFM for other reasons (for example, having twins). We are exploring whether we can use only tests and labs as follow-up procedures rather than office visits. In addition, we had to closely monitor the data extraction. For example, at first, IT pulled all follow-up procedures that were done after the risk assessments. We had to make sure that the IT data extractions indicated the time whether a follow-up laboratory test

was performed within 60 days after a positive risk assessment (as outlined in the data dictionary) or after the 60-day period (presumably due to new symptoms). Another challenge of the measure is for health systems that are decentralized, e.g., including private medical offices that may not use the electronic health record software, the same health system as the delivery hospital, to consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

[Response Ends]

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

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

Attach the fee schedule here, if applicable.

[Response Begins]

Programming code is needed to implement the algorithm and the associated measure is freely available from UCI.

[Response Ends]

4a.01. Check all current uses. For each current use checked, please provide:

Criterion 4: Use and Usability

4a. Use

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

[Response Begins]

Quality Improvement (Internal to the specific organization)

[Quality Improvement (Internal to the specific organization) Please Explain]

Our measures were tested at geographically and ethnically diverse hospital networks: UCI (1,500 births per year), UCSD (3,000 births per year), and UTENN (11,000 births per year). The hospital networks UCI and UCSD are located in Southern California and UTENN in Tennessee. They include regional Level 3 birthing centers with the full scope of inpatient and outpatient hospital services and affiliated community and private medical clinics. All hospitals have Obstetrics/Gynecology (OB/GYN) residency training programs, a high volume of Medicaid patients, and a diverse racial/ethnic demographic mixture. The information on the measure is used for staff training at other additional sites that have adopted the measure, such as Albert Einstein College and the University of Missouri. The Saint Luke's Hospital System is a non-profit 11-hospital system affiliated with the University of Missouri-Kansas City School of Medicine. Five hospitals within the system have obstetrical units that service approximately 5000 deliveries per year. Montefiore Medical Center (MMC), an affiliate of Albert Einstein College, provides a full range of services at more than 20 locations in the Bronx and Westchester County. Its diverse patient population includes mainly Latino (40%) and Black populations (30%). There is an anticipated 5000 births per year at MMC.

Current Users:

- University of Tennessee/ St. Thomas Health, Tennessee

- University of California, San Diego/UC San Diego Health
- University of California, Irvine/UC Irvine Health
- Albert Einstein College/Montefiore Medical Center, New York
- University of Missouri, Kansas City/St. Luke's Health System, Kansas City

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Public reporting

Quality Improvement (internal to the specific organization)

[Response Ends]

4a.03. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?

[Response Begins]

We submitted to the CMS for the Measures Under Consideration (MUC) list in April 2022 and applied to the public reporting program for hospital outpatient and inpatient quality reporting programs. If it is accepted, the measure will be publicly reported three years after acceptance.

[Response Ends]

4a.04. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

A credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.

[Response Begins]

We submitted to the CMS to the Measures Under Consideration (MUC) list in April 2022 and applied to the public reporting program for hospital outpatient and inpatient quality reporting programs. If it is accepted, the measure will be publicly reported three years after acceptance. Each clinic site or unit can calculate its own measure manually or with the support of its Information Technology department annually or semi-annually. UCI is working with two additional health care systems (Albert Einstein College/Montefiore Medical Center and St. Luke's Health/University of Missouri, Kansas City) to implement risk assessment at their obstetric sites for the period 2022-2024. UCI is also advising the University of Pennsylvania on a four-year project (2022-2026) that will implement the algorithm at its health care system and evaluate the use of the risk assessment in emergency room departments. Conversations with additional hospital networks are ongoing.

[Response Ends]

4a.05. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

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

[Response Begins]

During 2020-2022, we worked with three sites: the University of California, Irvine (UCI), the University of California, San Diego (UCSD), and the University of Tennessee/St. Thomas Health (UTENN). Each hospital network implemented a risk assessment tool and calculated the quality measure. Individual sites calculate their measures and provide feedback about the completion rate to their staff through e-mails or at staff meetings.

In addition, individual patient data for the measures were uploaded to UCI, and measures for three six-month periods were calculated by clinic site and patient demographics. The calculated measures were shared with UTENN and UCSD and site differences in performance of the clinic sites were addressed by the site investigators. Implementation of the tool in three hospital systems showed a variation in the follow-up among patients identified as at risk. Site medical directors used these measures to identify low-performance sites and to offer quality improvement initiatives (additional staff training, regular feedback on performance, addressing barriers to care and system issues in low follow-up rates)

[Response Ends]

4a.06. Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

[Response Begins]

The respective sites provided initial data for the first 6-month period after integrating the algorithm. Thereafter, we collected an 18-month period and calculated the measure for three six-month periods. We provided summaries to UCI clinic sites about their performance and reviewed with the clinicians their performance over time and in comparison, to the other sites. Similarly, we provided summaries to UCSD and UTENN by clinic site that were shared and reviewed with the individual sites. Graphs and bar charts have been disseminated with each data run showing the overall performance of risk assessment completion in the various clinical sites within each hospital system.

[Response Ends]

4a.07. Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.

[Response Begins]

The measures of the three hospital networks were reviewed with the co-investigators during virtual co-investigator meetings. Each site co-investigators individually contacted medical directors and/or clinicians with low CVD risk assessment rates to identify any implementation barriers. In addition, UCI conducted semi-structured interviews with five clinicians at each site (n=15) in May 2021 to elicit the value of the measure and barriers to follow up of the measure. The aggregate data presented in the measure facilitated the identification of system problems, such as need to obtain insurance approval for procedures, scheduling timely appointments, and patient logistics to keep health appointments (childcare, transportation, taking time off from work). Overall, clinicians appreciated the ability to monitor their performance and get a benchmark of their peer's performance.

[Response Ends]

4a.08. Summarize the feedback obtained from those being measured.

[Response Begins]

Clinicians appreciated the ability to monitor their performance and get a benchmark of their peer's performance. The measure provided insightful discussions at Safety and Quality meetings.

[Response Ends]

4a.09. Summarize the feedback obtained from other users.

[Response Begins]

We did not obtain any other systematic feedback. However, we collaborated with one additional health system on the integration of the tool in their electronic health systems (University of Missouri-Kansas City/St. Luke's Hospital System). The feedback from management and clinicians was consistent with the feedback already reported.

[Response Ends]

4a.10. Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

[Response Begins]

We formed a 14-member Technical Expert Panel (TEP) representing diverse stakeholders (Measure Developers, Clinical Content – Cardiology, Clinical Content – OB/GYN/MFM, Clinical IT, Patient Representatives). TEP members met virtually every 2-3 months and provided input on the individual elements of the algorithm, the integration of the algorithm in the EHR, and discussed additional clinical criteria such as the appropriate BNP cutoff. The TEP members agreed that:

1. There should not be any upper or lower age limit (so adolescent pregnancies and women with IVF are included).
2. Private providers who contract with the hospital for L&D services can be included in the denominator.
3. How to calculate the measure if the algorithm was administered more than once during a pregnancy episode.

Furthermore, as part of the evaluation of the implementation of the CVD algorithm, we obtained feedback from a purposive sample of 15 clinicians who had used the CVD algorithm. The five clinicians at each site included a range of job categories and levels of use of the algorithm (at least two clinicians who had not used the algorithm consistently).

[Response Ends]

4b. Usability

4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included). If no improvement was demonstrated, provide an explanation. If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

[Response Begins]

The measure to follow up of positive CVD risk assessment was successfully implemented at three large hospital networks in California and Tennessee. The rate for follow-up of positive risk assessment was 65.7%, 70.8%, and 38.4%, respectively. The overall CVD risk follow-up measure was 48.1%. The medical directors at each hospital network identified that the lack of follow-up was partly due to system barriers (need to obtain pre-authorization for follow-up tests from certain insurance programs) and are working on addressing those issues. Another major barrier in Tennessee, which has a higher proportion of low-income rural patients, seems to have transportation barriers to attending clinic appointments, highlighting the need to bundle appointments. Finally, medical directors used the data to review whether patients who changed care during pregnancy and postpartum are advised to alert their new provider about their CVD risk and implement potential standard handoff procedures.

[Response Ends]

4b.02. Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.

[Response Begins]

Patients and clinicians commented on improved patient awareness of the immediate and lifetime risk of developing CVD that drives changes in health behavior. While for most patients the result of the risk assessment may not produce any strong emotions, some patients may have had prior high-risk pregnancies (themselves or family/friends) or overall anxiety about the birth outcome or may have anxiety or other mental health problems that are exacerbated by being labeled "at risk." Physicians need to be altered in the training that the standard explanation of conveying risk may not suffice for all patients. We found that the use of echocardiograms and EKG in patients not at risk was not justified.

[Response Ends]

4b.03. Explain any unexpected benefits realized from implementation of this measure.

[Response Begins]

The consistent use of the tool has raised awareness of the importance of CVD risk assessment among obstetricians. Training of clinicians on how to counsel patients about their CVD risk and address potential concerns to avoid negative emotional reactions related to CVD risk with patients. Ob providers are required to document heart and lung examinations which may not have happened in some patients if it was not for CVD risk assessment.

[Response Ends]

Criterion 5: Related and Competing Measures

5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

(Can search and select measures.)

[Response Begins]

1927: Cardiovascular Health Screening for People With Schizophrenia or Bipolar Disorder Who Are Prescribed Antipsychotic Medications

0607: Pregnant women that had syphilis screening.

[Response Ends]

5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus or target population).

(Can search and select measures.)

[Response Begins]

[Response Ends]

5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

[Response Begins]

N/A

[Response Ends]

5.04. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

[Response Begins]

Yes

[Response Ends]

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

[Response Begins]

The syphilis screening measure has the same target group of pregnant patients. However, their focus is not on follow-up of risk assessments for CVD. The CVD health screening measure that addresses cardiovascular health is not focused on the obstetric population and is a secondary preventive measure (assessing treatment of individuals already identified with cardiovascular health issues).

[Response Ends]

5.06. Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.

Provide analyses when possible.

[Response Begins]

Our measure is used for standardized identification of individuals with suspected disease, or suspected high risk for disease. A CVD risk assessment distinguishes patients with a high probability of disease by analyzing several variables indicated by the algorithm.

For cardiovascular risk assessment and follow-up in pregnant and postpartum women, a reliable clinical screening approach that monitors the hospital and clinician performance is lacking. Timely identification of women at risk of CVD and follow-up may improve maternal health outcomes, i.e., maternal morbidity and mortality and lifetime onset of CVD.

[Response Ends]