NATIONAL QUALITY FORUM

MEASURE WORKSHEET

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

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

Purple and Blue text represents the responses from measure developers

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 3166

Corresponding Measures:

De.2. Measure Title: Antibiotic Prophylaxis Among Children with Sickle Cell Anemia

Co.1.1. Measure Steward: QMETRIC - University of Michigan

De.3. Brief Description of Measure: The percentage of children ages 3 months to 5 years old with sickle cell anemia (SCA) who were dispensed appropriate antibiotic prophylaxis for at least 300 days within the measurement year.

1b.1. Developer Rationale: Children with SCA are at an increased risk of infection compared with children without the disorder. Daily receipt of antibiotic prophylaxis substantially reduces the risk of infection among these high-risk children. Prior studies indicate that appropriate antibiotic prophylaxis rates are low; however, these reports are limited in their generalizability, as they are usually focused on a single healthcare provider or registry. This measure establishes a claims-based method for identifying appropriate antibiotic prophylaxis among larger and broader populations of children with SCA. The performance scores calculated through this measure will identify areas in need of improvement of antibiotic prophylaxis among children with SCA.

S.4. Numerator Statement: The numerator is the number of children ages 3 months to 5 years old with SCA who were dispensed appropriate antibiotic prophylaxis for at least 300 days within the measurement year.

S.6. Denominator Statement: The denominator is the number of children ages 3 months to 5 years with sickle cell anemia (SCA) within the measurement year.

S.8. Denominator Exclusions: There are no denominator exclusions.

De.1. Measure Type: Process

S.17. Data Source: Claims

S.20. Level of Analysis: Health Plan

IF Endorsement Maintenance – Original Endorsement Date: Jul 12, 2017 **Most Recent Endorsement Date:** Jul 12, 2017

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

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

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? NA

Preliminary Analysis: Maintenance of Endorsement

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

Criteria 1: Importance to Measure and Report

- 1a. Evidence

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

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

The developer provides the following evidence for this measure:

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

Summary of prior review in 2017

- This is a maintenance process measure utilizing claims data at the health plan level to assess the use of antibiotic prophylaxis among children ages 3 months to 5 years with sickle cell anemia (SCA).
- The developer provides a logic model depicting that daily receipt of appropriate antibiotic prophylaxis results in substantial reduction of the incidence of infection among children with SCA.
- During the initial review in 2017, the measure developer provided two key sources of evidence. The first was a systematic evidence review and clinical practice guidelines published by the National Heart, Lung, and Blood Institute: Evidence-Based Management of Sickle Cell Disease in 2014.
- The panel recommended that all children with HbSS be administered oral penicillin prophylaxis (125 mg for age <3 years and 250 mg for age ≥3 years) twice daily until age 5 (Strong Recommendation, Moderate-Quality Evidence) based on evidence from three RCTs and one observational study. The grade assigned was strong recommendation, moderate quality evidence.
- Additionally, a Cochrane Systematic Review, published in 2014, found that prophylactic penicillin significantly reduces risk of pneumococcal infection in children with homozygous sickle cell disease, and is associated with minimal adverse reactions. The review included three trials including data from over 800 children. Evidence included in the review was not graded.

Changes to evidence from last review

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

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

Updates: N/A

Exception to evidence

Not Applicable

Questions for the Committee:

- The developer attests the underlying evidence for the measure has not changed since the last NQF endorsement review. Does the Committee agree the evidence basis for the measure has not changed and there is no need for repeat discussion and vote on Evidence?

Guidance from the Evidence Algorithm

Process measure based on systematic review (Box 3)→ QQC present (Box 4)→ Quantity: moderate; Quality: high; Consistency: high (Box 5b)→ Moderate

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

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

Maintenance measures – increased emphasis on gap and variation

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

- Measure scores for the measure as specified across six states from 2005-2010 supported findings from prior studies indicating that appropriate antibiotic prophylaxis rates are low - ranging from 5.71% (South Carolina, 2009) to 36.11% (Illinois, 2007).
- 2010 performance scores in the six states ranged from 15.6% (Florida) to 27.9% (Texas).

Disparities

- The developer indicated that datasets for the performance scores are derived from Medicaid, therefore disparities by insurance or socioeconomic status were not identified. However the majority of children with sickle cell anemia (approximately 90%) have been enrolled in Medicaid at some point in time.
- The developer cited a study assessing compliance with penicillin prophylaxis for sickle cell disease showing that adherence was significantly greater in patients with private versus public insurance (17/28 [61%] vs. 33/90 [37%], respectively). Variation within insurance types is not captured.

Questions for the Committee:

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

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

- **Committee Pre-evaluation Comments:**

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure.

- This is a process measure examining the proportion of children with Sickle Cell who receive antibiotic prophylaxis for at least 300 days in a calendar year. there's been no change in the level of evidence (moderate).
- No new data. Evidence applies to this process measure, with studies showing NNT ~11 to prevent 1 infection in the age group.
- Evidence is updated and applies directly to process being measured.
- Evidence to support measure focus is sound

- This is a maintenance process measure. The goal is reduction of serious pneumococcal infections by daily antibiotic administration in children with sickle cell anemia up to 5 years of age. There was a literature review as part of the clinical guideline development (2014) and another for the Cochran Review (2014). One reference (dated 2016) is listed that postdates the guideline development and the Cochrane review.
- Evidence has not changed since 2017 according to the developer and is moderate in quality.
- Evidence is strong for this measure. Clinical practice guideline supports the measure (National Heart, Lung and Blood Institute). Cochrane Systematic Review found the measure topic significantly reduced risk of pneumococcal infection in children with SCD.
- Process measure, evidence rating is moderate

1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?

- The performance on this measure is generally low. The measure includes all kids aged 3 mos. - 5 years who have had 3 claims for SCD. It does not define a time period over which these claims occur. For example, children who enter care/are diagnosed >65 days into the year will always fail the measure. Disparities were not measured but the developers note a study indicating adherence with PCN prophylaxis is greater in those with private insurance when compared to public insurance, suggesting disparities. It's important to note, however, that this measure is assessing prescription of abx prophylaxis via claims; it is not assessing adherence to therapy.
- Continued performance gap in most current data (at best ~27% of children meet measure in some populations). No data collected regarding disparities in the datasets, however previous studies indicate that privately insured children have higher rates of antibiotic use than those with public insurance.
- Performance data provided that demonstrates a gap.
- yes, there is still a gap
- The goal percentage for daily antibiotic administration in children with sickle cell anemia up to 5 y/o is not mentioned, but can be presumed to be close to 100%. The highest annual rate achieved in any of the cooperating states in any one year was 35%. Medicaid data is used. There is no analysis by populations subgroups. Studies are cited that showed that children with sickle cell anemia and private health insurance received more prescriptions and had better antibiotic dosage adherence than children with sickle cell anemia on public health insurance.
- The performance over time is highly variable and generally low. The question is what factors are contributing to these data for the Medicaid population. For example, do the children go on and off MA and therefore paying for the drugs becomes an issue or it is that compliance for the 300 days each year more the problem. Just holding the providers accountable without understanding the underlying causal factors doesn't advance our quality.
- Studies indicate a gap in compliance with penicillin prophylaxis between private and public insurance (private adherence significantly greater). There is a gap in care.
- A performance gap exists

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

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

- Reliability

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

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

- Validity

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

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

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Staff
Staff Review
Evaluation Summary:
Reliability

- A separate method of reliability testing was not provided by the developer since empirical validity testing was conducted with Medicaid Analytic eXtract (MAX) data for six state Medicaid programs provided by the Centers for Medicare & Medicaid Services (CMS) (2005-2012).
- For some measure types, separate reliability testing of the data elements is not required if empirical validity testing of the data elements is conducted (and results are adequate). Data element validity was used to support the reliability of this measure.

Validity

- The developer conducted data element testing using both ICD-9-CM and ICD-10-CM diagnosis codes.
    - ICD-9-CM denominator was validated via comparison with the gold standard of newborn screening results for the state of Michigan for children enrolled in Michigan Medicaid in 2010 and 2011 with at least one SCD-related healthcare claim within their enrollment year(s).
        - Results: In 2010, a case definition of three Hb SS (282.61, 282.62) claims within the year was 91.4% sensitive and 80% specific in identifying children with sickle cell anemia (Hb SS) (PPV: 80.4%; NPV: 91.3%). These results were replicated with nearly identical precision among the study population in 2011. In comparison, using a case definition of at least one Hb SS claim or at least two Hb SS claims to identify the study population resulted in substantially less specificity.
    - ICD-10-CM denominator was validated through a four step process: (a) developing candidate SCA case definitions; (b) identifying a test population for the definitions; (c) testing the accuracy of candidate case definitions; and (d) testing case definitions within an independent population.
        - Results: The case definition of ≥1 outpatient visit with a SCA-related (D5700, D5701, D5702) or D571 code had the highest area under the ROC curve, with a sensitivity of 95% and specificity of 92%. The same definition also had the highest performance in New York Medicaid (n=2,454), with a sensitivity of 94% and specificity of 86%. There were no clinically relevant changes in the sensitivities and specificities of the definitions when limited to ages 1 through 4.

- Results from both ICD-9-CM and ICD-10-CM diagnosis codes indicate that children with sickle cell anemia can be identified with a high level of accuracy in administrative data.
- The numerator (children with sickle cell anemia that have at least 300 days of filled antibiotics within the year) was validated by a comprehensive examination of National Drug Codes (NDCs) to identify antibiotics using clinical expertise and RxNorm, an assessment of various categorizations of antibiotic prophylaxis using MAX data from 2005-2012, and through a comparison with the gold standard of medical charts.
    - Results: 28,881 NDCs for antibiotic prophylaxis within administrative claims were identified. The average number of days of filled prescriptions for antibiotic prophylaxis was as follows: 162 days of penicillin (SD = 117; median: 160), 164 days of penicillin or erythromycin (SD = 117; median: 160), 178 days of penicillin, erythromycin, or amoxicillin (SD = 113; median: 180), and 193 days of any S pneumoniae antibiotic (SD =116; median: 194). There was 100% agreement between the two trained abstractors for an inter-rater reliability kappa score of 1.00. Given this evidence, the developer asserts the validity of administrative claims in assessing antibiotic prescriptions is very high.

- The developer tested empirical validity by comparing the MAX data for the state of Michigan (obtained from CMS) with the gold standard of Michigan Medicaid data (obtained directly from Michigan's claims data warehouse) for the same time period (2007-2010). Rates of antibiotic prophylaxis using each source of data were calculated and compared using z-tests for two proportions; for these tests, the null hypothesis was that the rate in each year would be the same in both Michigan Medicaid data and MAX data.
    - Results: Compared with the gold standard of Michigan Medicaid data, MAX data has a very high degree of validity. When antibiotic prophylaxis was assessed for the same state (Michigan) from these two data sources for the same time period (2007-2010), no differences in rates were observed (all p-values >0.05).
- Face validity was established by a panel of 14 national experts and advocates for families of children with SCD convened by the Quality Measurement, Evaluation, Testing, Review, and Implementation Consortium (QMETRIC).
    - Results: The QMETRIC expert panel concluded that this measure has a very high degree of face validity through a detailed review of concepts and metrics considered to be essential to effective SCD management and treatment. Concepts and draft measures were rated by this group for their relative importance. This measure was among the most highly rated, receiving an average score of 8.5 (with 9 as the highest possible score).

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?

Questions for the Committee regarding validity:

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

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

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

- **Committee Pre-evaluation Comments:**

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other

specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?

- no risk adjustment. no concerns about implementation.
- Data elements are clearly defined. Consistent implementation is highly likely.
- Data elements are clear. No concerns.
- reliability - specifications are appropriate
- The denominator is children 3 months to 5 years of age continuously enrolled in Medicaid during the measurement year and having at least one outpatient visit coded with a sickle cell anemia diagnostic code. There are no denominator exclusions. The numerator is children in the denominator set for whom appropriate prophylactic antibiotics have been dispensed for at least 300 days in the measurement year. Sampling is not involved. The process seems straightforward.
- Specifications seem comprehensive.
- No concerns with reliability of the measure. All data elements are available and clearly defined.
- No concerns

2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?

- no
- No
- No concerns.
- no
- In Section 2b1.3 of the Testing Form (Validity Testing) there is mention of testing inter-rater reliability by abstractors of 18 charts for antibiotic prophylaxis. On page 37 of the PCCI Measure Worksheet it is noted that reliability testing was done using Medicaid Analytic Extract (MAX) data, but no details are provided.
- It appears that the measure stands up well in all comparisons.
- No concerns with reliability.
- No concerns

2b1. Validity -Testing: Do you have any concerns with the testing results?

- no
- No
- No concerns
- no
- Michigan Medicaid data was compared to Medicaid Analytic Extract (MAX) data with a high degree of agreement. Using both ICD-9 and ICD-10, identification of antibiotic usage for children with sickle cell anemia corresponded with the medical record 88% of the time or better.
- It appears the to be very valid and supported by multiple societies.
- Empirical validity and face validity testing completed. The measure should be able to be consistently implemented. No concerns with measure exclusions.
- No concerns

2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment) 2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure? 2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale

provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?

- the measure does deviate from the guidelines in that it allows for alternate abx for kids who are PCN allergic. This is a reasonable deviation from the guidelines and avoids inappropriate exclusion of some individuals.
- n/a
- Exclusions are consistent with evidence.
- no major concerns
- There were no exclusions or risk adjustments/stratification.
- At present, I see no issues with exclusions or risk adjustment.
- No identified threats to validity.
- No concerns

2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data) 2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?

- developer reports that missing data did not introduce bias.
- No
- No threats to validity.
- no major concerns
- The proportion of children with at least 300 days of antibiotics ranged across states from 6% to nearly 40%. Only one set of specifications was used. Medicaid claims for sickle cell patients seen at three major medical centers in Michigan were matched with their Michigan Medicaid administrative claims data showing 97% concurrence, a low level of missing or conflicting data.
- No issues with threats to validity.
- No identified threats to validity. The results should identify meaningful differences in quality. No concerns about missing data for this measure.
- No concerns

Criterion 3. Feasibility

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

3. **Feasibility** is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.
    - The data elements required for the measure are routinely generated and used during care delivery and all data elements used in the measure are in defined fields in electronic claims.
    - The primary information needed for this measure includes basic demographics, diagnosis codes, and procedure codes and dates. These data are widely available, although obtaining them often requires a restricted-use data agreement.
    - For multiple-state comparisons, MAX data are available from CMS. When the measure is used at the single-state level, state health departments can use their own Medicaid data.

- Acquisition of data directly from states requires the cooperation of those jurisdictions, as well as modification of the statistical programming code developed for MAX files to correctly function using the unique structure of the data files obtained from each state.

Questions for the Committee:

- Is the data collection strategy ready to be put into operational use?

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

- **Committee Pre-evaluation Comments:**

Criteria 3: Feasibility

3. Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?

- no concerns.
- no
- Data elements are routinely generated during care delivery. No concerns
- no major concerns
- All data elements are in defined fields in electronic claims. Data collection is straightforward and has been utilized in six states.
- Continued measure seems to be feasible.
- The data elements for this measure are routinely generated and collected. The data collection strategy is in use.
- No concerns

Criterion 4: Usability and Use

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

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

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

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

Planned use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details

- The developer noted that the State of Michigan Medicaid Program will be implementing this measure for purposes of building a health plan collaborative to improve the care of children with sickle cell anemia in Michigan.
- The developer noted that program implementation was delayed due to COVID-19 and anticipates reporting to occur in 2021.

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

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

Feedback on the measure by those being measured or others

- Specifications and technical assistance were provided to Michigan Medicaid health plans and the State of Michigan Medicaid Program.
- The developer noted that feedback was sought from health plan data teams after implementation and that health plans were able to implement the measure without issue and obtain performance scores thus allowing them to identify opportunities for improvement.
- Using feedback obtained, the developer modified the measure to reflect the validation of the denominator using ICD-10-CM diagnosis codes.

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

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

Improvement results

- There has been no improvement demonstrated.
- The developer shared that a national, strategic effort for accountability, such as inclusion in the Child Core Set, would drive improvement.
- This measure has been recommended for inclusion in the Core Set for multiple years.

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

Unexpected findings (positive or negative) during implementation

- No unintended negative consequences to individuals or populations were identified during testing.
- This measure has been considered for inclusion in the CMS Core Set.

Potential harms

- None identified

***Questions for the 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

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

4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided? 4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?

- used in MI medicated program; considered for CMS core set. feedback sought from others.
- Performance results have been provided to Medicaid plans in Michigan. Feedback regarding the measure purportedly sought and adjustments made.
- Plans to be used in State of Michigan but delay due to COVID
- no major concerns
- Health plans have been able to implement the measure without issue and obtain performance scores. This allowed health plans to identify opportunities for improvement. Use by the Michigan Medicaid Program is planned for 2021, with reporting to the public to follow. All users have indicated that the measure is feasible and usable. The measure development team has provided feedback and clarification when necessary to all users. The only changes in this maintenance measure is adapting the measure to ICD-10-CM.
- Meets the requirements for Use.
- Feedback and technical assistance were provided to the Michigan Medicaid Program by the measure developer. The developer sought feedback from the health plan data teams after implementation. The developer used the feedback to modify the measure to reflect the validation of the denominator using ICD-10-CM diagnosis codes.
- Not currently in use, planned use

4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations? 4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.

- can be used to improve abx prophylaxis in high risk population. benefits outweigh harms. no unintended consequences noted.
- There are ideas to create a collaborative to increase the reach of this measure and help plans to report data and hopefully improve performance. Harms include adverse events with long term antibiotic use, but outweighed by reduction of serious infection in this population.
- No improvement. Proposed to be in child core measure set. No harms.
- no major concerns
- No improvements have been demonstrated. The developers believe that a national and strategic effort for accountability would drive improvement. There is no other analysis for the lack of observed improvement. No unintended negative consequences to individuals or populations were identified.
- Appears to meet the requirements for Usability.
- The measure can be used to improve the quality of healthcare. There were no identified or unexpected harms from the measure cited. It has been vetted in real-world settings.
- No evidence related to unintended consequences

Criterion 5: Related and Competing Measures

Related or competing measures

2797 : Transcranial Doppler Ultrasonography Screening Among Children with Sickle Cell Anemia

Harmonization

- The developer states the measure specifications have been harmonized.
- Different age categories are included in the measures. For example, antibiotic prophylaxis is recommended by NHLBI for ages 0 until 5; TCD screening from ages 2 until 16.
- The numerator specifications differ across each measure.

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

5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?

- no significant concerns here.
- Age ranges for inclusion are different with this measure and measure 2797, but appropriately so.
- No competing or directly related measures.
- no
- There is another measure for children with sickle cell anemia - #2797 : Transcranial Doppler Ultrasonography Screening Among Children with Sickle Cell Anemia. Another measure is proposed in this NQF Fall 2020 cycle - Hydroxyurea Use Among Children with Sickle Cell Anemia. Different age categories and numerator specifications justify the multiple measures.
- Harmonized with other related measures as much as possible.
- Related or competing measure: 2797: Transcranial Doppler Ultrasonography Screening Among Children with Sickle Cell Anemia. The measure developer states that the measures have been harmonized. The numerator specifications differ across these two measures.
- No concerns

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 01/21/2021

- No NQF Members have submitted support/non-support choices as of this date.
- No Public or NQF Member comments submitted as of this date.

- Scientific Acceptability: Preliminary Analysis Form

Measure Number: 3166

Measure Title: Antibiotic Prophylaxis among Children with Sickle Cell Anemia

Type of measure:

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

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

Data Source:

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

Level of Analysis:

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

Measure is:

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

RELIABILITY: SPECIFICATIONS

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

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

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

2. **Briefly summarize any concerns about the measure specifications.**

   No concerns

RELIABILITY: TESTING

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

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

   ☒ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

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

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

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

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

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

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

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

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

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

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

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

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

Precise specifications (Box 1) → Empirical reliability testing (Box 2) → Patient level data validity (Box 3) → (Box 9 of validity algorithm) → Appropriate method to assess data elements (Box 10) → Moderate certainty that the data used in the measure are valid (Box 11a) → Highest possible rating is Moderate.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

Submission document: Testing attachment, section 2b2.

N/A – no exclusions

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

Submission document: Testing attachment, section 2b4.

Logistic regression was used to estimate the associations between each state health plan and at least 300 days of antibiotic prophylaxis, with Illinois used as the reference category. Results indicate that the measure can successfully distinguish differences in performance across state health plans.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
Submission document: Testing attachment, section 2b5.
N/A

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

Submission document: Testing attachment, section 2b6.

The developer tested the completeness and accuracy of Medicaid claims for sickle cell patients seen at three major medical centers in Michigan. For this comparison, children with sickle cell anemia who were enrolled within Michigan Medicaid were matched with their Michigan Medicaid administrative claims data (n=34). The developer attests that missing data is unlikely to bias their performance results, particularly as any missing data would be expected to be non-differential across entities.

16. **Risk Adjustment**

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

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

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:**

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

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

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

16d.**Risk adjustment summary:**

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

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

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

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration)
☐ Yes ☐ No

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

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

VALIDITY: TESTING

17. **Validity testing level: ☐ Measure score ☐ Data element ☒ Both**

18. **Method of establishing validity of the measure score:**

☒ Face validity

☒ Empirical validity testing of the measure score

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

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

Submission document: Testing attachment, section 2b2.2

- The developer conducted data element testing using both ICD-9-CM and ICD-10-CM diagnosis codes.
    - ICD-9-CM denominator was validated via comparison with the gold standard of newborn screening results for the state of Michigan for children enrolled in Michigan Medicaid in 2010 and 2011 with at least one SCD-related healthcare claim within their enrollment year(s).
    - ICD-10-CM denominator was validated through a four step process: (a) developing candidate SCA case definitions; (b) identifying a test population for the definitions; (c) testing the accuracy of candidate case definitions; and (d) testing case definitions within an independent population.
    - The numerator (children with sickle cell anemia that have at least 300 days of filled antibiotics within the year) was validated by a comprehensive examination of National Drug Codes (NDCs) to identify antibiotics using clinical expertise and RxNorm, an assessment of various categorizations of antibiotic prophylaxis using MAX data from 2005-2012, and through a comparison with the gold standard of medical charts
- The developer tested empirical validity by comparing the MAX data for the state of Michigan (obtained from CMS) with the gold standard of Michigan Medicaid data (obtained directly from Michigan's claims data warehouse) for the same time period (2007-2010). Rates of antibiotic prophylaxis using each source of data were calculated and compared using z-tests for two proportions; for these tests, the null hypothesis was that the rate in each year would be the same in both Michigan Medicaid data and MAX data.
- Face validity was established by a panel of national experts and advocates for families of children with SCD convened by the Quality Measurement, Evaluation, Testing, Review, and Implementation Consortium (QMETRIC).

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

Submission document: Testing attachment, section 2b2.3

- **Data element testing:**
    - ICD-9-CM denominator validation: In 2010, a case definition of three Hb SS (282.61, 282.62) claims within the year was 91.4% sensitive and 80% specific in identifying children with sickle cell anemia (Hb SS) (PPV: 80.4%; NPV: 91.3%). These results were replicated with nearly identical precision among the study population in 2011.

- ICD-10-CM denominator validation: measures of performance varied across 23 case definitions, with sensitivities from 0.02-0.97 and specificities from 0.88-1.0. The case definition of ≥1 outpatient visit with a SCA-related (D5700, D5701, D5702) or D571 code had the highest area under the ROC curve, with a sensitivity of 95% and specificity of 92%. The same definition also had the highest performance in New York Medicaid (n=2,454), with a sensitivity of 94% and specificity of 86%. There were no clinically relevant changes in the sensitivities and specificities of the definitions when limited to ages 1 through 4.
- Numerator validation:
  - Comprehensive examination of National Drug Codes (NDCs): the developer identified 28,881 NDCs for antibiotic prophylaxis within administrative claims.
  - Assessment of categorizations of antibiotic prophylaxis: the average number of days of filled prescriptions for antibiotic prophylaxis was as follows: 162 days of penicillin (SD = 117; median: 160), 164 days of penicillin or erythromycin (SD = 117; median: 160), 178 days of penicillin, erythromycin, or amoxicillin (SD = 113; median: 180), and 193 days of any S pneumoniae antibiotic (SD =116; median: 194).
  - Gold standard of medical charts comparison: 34 children with sickle cell anemia who were enrolled within Michigan Medicaid were successfully matched with their Michigan Medicaid administrative claims data. Among these children, 25 cases (76%) had a Medicaid administrative claim for an antibiotic prescription filled within 29 days of the prescription date in the medical record. An additional seven cases (21%) had an antibiotic prescription filled between 30 but less than 90 days following the prescription date. One case (3%) had an antibiotic prescription claim filled 90 or more days after the prescription date in the medical record. Eighteen charts were also chosen for calculation of inter-rater reliability; the two trained abstractors had 100% agreement with each other for abstracting antibiotic prescriptions from the medical records, resulting in a kappa of 1.00.

- **Empirical Validity Testing:**
  - The comparison of rates of appropriate antibiotic prophylaxis from the gold standard of Michigan Medicaid data compared with MAX data can be seen in Table 2, which illustrates that the number of children who were dispensed at least 300 days of antibiotics among children with sickle cell anemia ranged from 14 to 23 in the claims acquired directly from the Medicaid data warehouse, versus a range of 7 to 20 from MAX data for the same time period.
  - Table 3 reports the z-scores and p-values from the two-sample z-tests comparing the proportion of children who were dispensed at least 300 days of antibiotic prophylaxis each year between Michigan Medicaid and MAX data.
  - Figure 1 illustrates the performance scores observed between the Michigan Medicaid data from the state warehouse and MAX data from CMS for each overlapping year noted, respectively: 12.3% versus 14.3% (2007); 19.5% versus 17.8% (2008); 14.8% versus 7.3% (2009), and 15.6% versus 21.1% (2010).
- **Face Validity:** The QMETRIC expert panel concluded that this measure has a very high degree of face validity through a detailed review of concepts and metrics considered to be essential to effective SCD management and treatment.

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

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

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

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

NOTE that data element validation from the literature is acceptable.

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

☒ **Yes**

☐ **No**

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

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

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

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

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

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

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

Box 1: Potential threats to validity assessed Box 2: Empirical validity testing conducted using the measure as specified à Box 5: Testing conducted at the measure score level à Box 6: Testing method described and appropriate à Box 7b: High certainty or confidence that the performance measure scores are a valid indicator of quality à HIGH

ADDITIONAL RECOMMENDATIONS

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

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

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

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

QMETRIC_SCA_Antibiotic_Prophylaxis_NQF_evidence_attachment_508_Compliant.docx

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

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

No

- 1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 3166

Measure Title: Antibiotic Prophylaxis Among Children with Sickle Cell Anemia

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here:

Date of Submission: 11/19/2020

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

Outcome

☐ Outcome:

☐Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (*A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)*

☐ Intermediate clinical outcome (*e.g., lab value*):

☒ Process: Antibiotic Prophylaxis Among Children with Sickle Cell Anemia

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

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

Daily receipt of appropriate antibiotic prophylaxis results in substantial reduction of the incidence of infection among children with sickle cell anemia (SCA).

1a.3 Value and Meaningfulness: **IF** this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

N/A

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

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

X Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

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

☐ Other

Systematic Review

Evidence

Source of Systematic Review:

- Title
- Author
- Date
- Citation, including page number
- URL

National Heart, Lung, and Blood Institute. Evidence-Based Management of Sickle Cell Disease: Expert Panel Report, 2014;13-14. https://www.nhlbi.nih.gov/health-pro/guidelines/sickle-cell-disease-guidelines.

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

RECOMMENDATIONS

1. Administer oral penicillin prophylaxis (125 mg for age <3 years and 250 mg for age ≥3 years) twice daily until age 5 in all children with HbSS.

   (Strong Recommendation, Moderate-Quality Evidence)

2. Discontinue prophylactic penicillin in children with HbSS at age 5 unless they have had a splenectomy or invasive pneumococcal infection. When discontinuing penicillin prophylaxis at age 5, it is important

to assure that the child has completed the recommended pneumococcal vaccination series, and if not, complete the series immediately.

(Weak Recommendation, Moderate-Quality Evidence)

3. Consider withholding penicillin prophylaxis from children with HbSC disease and HbSβ^0-thalassemia unless they have had a splenectomy.

   (Weak Recommendation, Low-Quality Evidence)

4. Assure that people of all ages with SCD have been vaccinated against *Streptococcus pneumoniae*. *

   (Strong Recommendation, Moderate-Quality Evidence)

5. Remind people with SCD, their families, and caregivers to seek immediate medical attention whenever fever (temperature greater than 101.3°F or 38.5°C) occurs, due to the risk for severe bacterial infections.

   (Consensus-Panel Expertise)

*Refer to the Immunization section of this chapter for comprehensive information on immunizations.

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade

Strong Recommendation, Moderate-Quality Evidence

- **Grade of Recommendation:** Strong Recommendation, Moderate-Quality evidence
- **Clarity of Risk/Benefit:** Benefits clearly outweigh harms or burdens, or vice versa
- **Quality of Supporting Evidence:** Evidence from RCTs with important limitations (inconsistent results, methodological flaws, indirect or imprecise evidence), or unusually strong evidence from unbiased observational studies
- **Implications:** Recommendation can apply to most patients in most circumstances. Further research (if performed) is likely to have an impact on our confidence in the estimate of effect and may change the estimate.

Provide all other grades and definitions from the evidence grading system

Strong Recommendation, High-Quality Evidence

- **Clarity of Risk/Benefit:** Benefits clearly outweigh harms or burdens, or vice versa
- **Quality of Supporting Evidence:** Consistent evidence from well-performed RCTs or exceptionally strong evidence from unbiased observational studies*
- **Implications:** Recommendation can apply to most patients in most circumstances. Further research is very unlikely to change our confidence in the estimate of effect.

Strong Recommendation, Moderate-Quality Evidence

- **Clarity of Risk/Benefit:** Benefits clearly outweigh harms or burdens, or vice versa
- **Quality of Supporting Evidence:** Evidence from RCTs with important limitations (inconsistent results, methodological flaws, indirect or imprecise evidence), or unusually strong evidence from unbiased observational studies
- **Implications:** Recommendation can apply to most patients in most circumstances. Further research (if performed) is likely to have an impact on our confidence in the estimate of effect and may change the estimate.

Strong Recommendation, Low-Quality Evidence

- **Clarity of Risk/Benefit:** Benefits clearly outweigh harms or burdens, or vice versa
- **Quality of Supporting Evidence:** Evidence for at least one critical outcome from observational studies, from RCTs with serious flaws, or indirect evidence

- **Implications:** Recommendation may change when higher quality evidence becomes available. Further research (if performed) is likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Strong Recommendation, Very Low-Quality Evidence

- **Clarity of Risk/Benefit:** Benefits clearly outweigh harms or burdens, or vice versa
- **Quality of Supporting Evidence:** Evidence for at least one of the critical outcomes from unsystematic clinical observations or very indirect evidence
- **Implications:** Recommendation may change when higher quality evidence becomes available; any estimate of effect, for at least one critical outcome, is very uncertain.

Weak Recommendation, High-Quality Evidence

- **Clarity of Risk/Benefit:** Benefits closely balanced with harms and burdens
- **Quality of Supporting Evidence:** Consistent evidence from well-performed RCTs or exceptionally strong evidence from unbiased observational studies
- **Implications:** The best action may differ depending on circumstances or patient or societal values. Further research is very unlikely to change our confidence in the estimate of effect.

Weak Recommendation, Moderate-Quality Evidence

- **Clarity of Risk/Benefit:** Benefits closely balanced with harms and burdens
- **Quality of Supporting Evidence:** Evidence from RCTs with important limitations (inconsistent results, methodological flaws, indirect or imprecise evidence), or unusually strong evidence from unbiased observational studies
- **Implications:** Alternative approaches likely to be better for some patients under some circumstances. Further research (if performed) is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Weak Recommendation, Low-Quality Evidence

- **Clarity of Risk/Benefit:** Uncertainty in the estimates of benefits, harms, and burdens; benefits may be closely balanced with harms and burdens
- **Quality of Supporting Evidence:** Evidence for at least one critical outcome from observational studies, from RCTs with serious flaws, or indirect evidence
- **Implications:** Other alternatives may be equally reasonable. Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Weak Recommendation, Very Low-Quality Evidence

- **Clarity of Risk/Benefit:** Major uncertainty in the estimates of benefits, harms, and burdens; benefits may or may not be balanced with harms and burdens
- **Quality of Supporting Evidence:** Evidence for at least one critical outcome from unsystematic clinical observations or very indirect evidence
- **Implications:** Other alternatives may be equally reasonable. Any estimate of effect, for at least one critical outcome, is very uncertain.

Source: Reprinted with permission of the American Thoracic Society. Copyright 2012 American Thoracic Society. Shunemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, Ernst A, Fahy BF, Gould MK, Horan KL, Krishnan JA, Manthous CA, Maurer JR, McNicholas WT, Oxman AD, Rubenfeld G, Turino GM, Guyatt G; ATS Documents Development and Implementation Committee. An official ATS statement; grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med*. 2006 Sep 1:174(5):605-14. Official Journal of the American Thoracic Society.[29]

* Exceptionally strong evidence from unbiased observational studies includes: (1) evidence from studies that yield estimates of the treatment effect that are large and consistent; (2) evidence in which all potential biases may be working to underestimate an apparent treatment effect, and therefore, the actual treatment effect is likely to be larger than that suggested by the study data; and (3) evidence in which a dose-response gradient exists

Grade assigned to the **recommendation** with definition of the grade

Strong Recommendation; for definition, see evidence grading system above.

Provide all other grades and definitions from the recommendation grading system

See above; included within evidence grading system above

Body of evidence:

- Quantity – how many studies?
- Quality – what type of studies?

Three RCTs and one observational study were included.

As stated within the NHLBI Clinical Guidelines: "The three RCTs [Gaston et al., 1986; Faletta et al., 1995; John et al., 1984 {see References at the end of this documents}] were of moderate methodological quality and compared penicillin to no prophylaxis. The initiation of penicillin prophylaxis was associated with a significant reduction in the risk for developing serious pneumococcal infections (2/105 vs. 13/110) and a nonsignificant reduction in mortality (0/105 deaths vs. 3/110 deaths; very low-quality evidence due to severe imprecision). A single trial evaluated the consequences of discontinuing penicillin prophylaxis; it suggested that prophylaxis in children who have not had a prior severe pneumococcal infection or a splenectomy may be discontinued at age 5. Children who continued penicillin had a nonsignificant reduction in systemic pneumococcal infections; there was no effect on mortality. The observational study [Nkouwap et al., 1999] compared penicillin to spiramycin and demonstrated that penicillin was superior. However, the penicillin group had a higher rate of pneumococcal vaccination, confounding the effect of antibiotics and making strong conclusions difficult. The quality of evidence is very low due to severe imprecision (i.e., small number of events) and methodological limitations. Evidence is lacking in children with genotypes other than SS, even though many clinicians prescribe prophylactic penicillin for them both before and after age 5."

Please note that this measure specification deviates slightly from the NHLBI recommendations in two main areas: child age and type of antibiotic dispensed. Instead of including children from birth, this measure includes children starting at 3 months of age. This is to account for any lag in identification of the sickle cell disease status of the child at the State level. Although NHLBI guidelines specifically recommend penicillin for antibiotic prophylaxis, some children may have or be suspected to have penicillin sensitivity. The American Academy of Pediatrics Section on Hematology/Oncology and Committee on Genetics suggests an alternative for children who are allergic to penicillin: "Erythromycin prophylaxis may be used as an alternative for children with suspected or proven penicillin allergy." (American Academy of Pediatrics 2002 and 2016). Therefore, we have included a broader definition of appropriate antibiotics in this measure. These modifications to the guidelines are intended to avoid underestimation of the proportion of children with SCA who are protected against pneumococcal infection.

Estimates of benefit and consistency across studies

The four studies detailed within the NHLBI evidence tables (Gaston et al., 1986; Faletta et al., 1995; John et al., 1984; Nkouwap et al., 1999) included a total of 951 children under the age of 5 years. The majority of children (94%) were sickle cell subtype Hb SS (sickle cell anemia); 5% were Hb SC, and 1% were Hb S beta zero thalassemia. All studies investigating the relationship between antibiotic prophylaxis and infection indicated a net benefit to antibiotic prophylaxis among children with sickle cell anemia. The primary study (Gaston et al., 1986) enrolled children with Hb SS starting at 3-6 months of age into the Prophylactic Penicillin Study (PROPS) to test the protective effect of regular, daily administration of oral penicillin against incidence of documented septicemia due to s. pneumonia in children under the age of 3 years. Children were randomized into study groups; participants were allocated to either receive 125mg of penicillin, twice daily (n=105 children), or a

placebo (n=110 children), for a mean of 15 months. Reduction of septicemia was found in children who received the penicillin, as an 84% reduction in the incidence of infection was observed in the group treated with penicillin compared with the group given placebo (13 of 110 patients vs. 2 of 105; P = 0.0025).

What harms were identified?

Gaston et al. (1986) reported no adverse effects throughout the course of the study. One child in Falletta et al. (1995) experienced nausea and vomiting on a higher dose of penicillin than explored in the PROPS trial. Given the severity of infection and the potential for increased mortality due to pneumococcal infection among these high risk children, the benefit outweighs the potential harm for this recommendation.

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?

N/A

Systematic Review

Evidence

Source of Systematic Review:

- Title
- Author
- Date
- Citation, including page number
- URL

Hirst C, Owusu-Ofori S. Prophylactic antibiotics for preventing pneumococcal infection in children with sickle cell disease. Cochrane Database of Systematic Reviews 2014, Issue 11. Art. No.: CD003427.

DOI: 10.1002/14651858.CD003427.pub3.

http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003427.pub3/abstract;jsessionid=03720C55FB06B08BFB3D12EBCEDF7614.f04t02

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

“Prophylactic penicillin significantly reduces risk of pneumococcal infection in children with homozygous sickle cell disease, and is associated with minimal adverse reactions.”

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade

N/A

Provide all other grades and definitions from the evidence grading system

N/A

Grade assigned to **recommendation** with definition of the grade

N/A

Provide all other grades and definitions from the recommendation grading system

N/A

Body of evidence:

- Quantity – how many studies?
- Quality – what type of studies?

Three trials were included in the review.

As stated within the Cochrane Review: “Methodological quality was assessed based on a method described by Schulz et al. (1995).

“In the John trial participants were randomised, but no details were given of the method of randomisation (John et al., 1984). The publication reports that the group allocation was changed due to the protocol for injected penicillin prophylaxis groups being inconvenient to some families who lived at remote addresses, or due to age of participants at recruitment so that the duration of penicillin treatment would have been too short to assess. Sixteen participants (6.6%) were therefore reassigned to groups which did not receive penicillin prophylaxis. The groups were uneven, with significantly more participants in the penicillin groups (143 in penicillin group compared to 99 in control group). Full baseline data for participant characteristics were not given. The trial was not blinded. Intention-to-treat analysis was undertaken after participants were reassigned. There were 25 withdrawals, 20 due to splenectomy, four due to emigration and one because of recurrent meningitis. The participants discontinued penicillin at the age of three years. However, they continued to be analysed in the groups to which they were randomised. Therefore, although there were no cases of pneumococcal infection amongst participants taking penicillin, there were seven cases in the penicillin assigned groups, all occurring after discontinuation of the drug.

“In the first PROPS trial, a central coordinating center generated a blocked randomisation sequence, and directed participant entry assignment over the telephone (PROPS, 1986 [Gaston et al., 1986; Gaston and Verter, 1990]). Sealed envelopes were also held at the clinical centres in case the central office could not be reached, to maintain allocation concealment. The participants and centre personnel were blinded to allocation, and placebo tablets looked almost identical to penicillin. A sample size calculation was performed based on an estimated 50% reduction in risk of infection, and as a result 219 participants were recruited from 23 centres throughout the USA. Four participants subsequently withdrew due to revisions of diagnosis of genotype; these patients had no severe infections but were not included in subsequent analyses. The baseline characteristics of the children in each group, including history of palpable spleen or infection, were similar. The trial was terminated early due to extreme results. Because of this, there is a possibility that the reported results may be over-estimated.

“In a further PROPS trial, randomisation was by permuted block method, stratified by clinical site and years of previous penicillin use (PROPS II, 1995 [Faletta et al., 1995; Bjornson et al., 1996; Woods et al., 1997]). It was unclear whether allocation concealment had been performed. Identical placebo tablets were used to maintain double blinding of the participants and centre personnel. A sample size calculation ensured that the trial was powered to show a three-fold increase in infection in the placebo group, based on an estimated incidence of 4% in the penicillin group. Four hundred participants were subsequently recruited from 18 centres in the USA. The characteristics of participants in each group were similar at baseline. Four children died after randomisation, but other withdrawals are not reported, and it is unclear whether an intention-to-treat analysis was undertaken.”

Estimates of benefit and consistency across studies

“Three trials with over 800 children are included in the review. All three trials showed a reduced rate of infection in children with sickle cell disease receiving penicillin preventatively. Two trials looked at whether treatment was effective. The third trial followed on from one of the early trials and looked at when it was safe to stop treatment. …

“All of the included trials showed a reduced incidence of infection in children with sickle cell disease (SS or S beta zero thalassemia) receiving prophylactic penicillin. In trials which investigated initiation of penicillin on risk of pneumococcal infection, the odds ratio was 0.37 (95% CI 0.16 to 0.86), while for withdrawal the odds ratio was 0.49 (95% CI 0.09 to 2.71).”

What harms were identified?

“Adverse drug effects were rare and minor. However, there were problems with children keeping to the treatment schedule and with the development of antibiotic resistance.”

Note: The American Academy of Pediatrics Section on Hematology/Oncology and Committee on Genetics suggests an alternative for children who are allergic to penicillin: "Erythromycin prophylaxis may be used as an alternative for children with suspected or proven penicillin allergy" (American Academy of Pediatrics 2002 and 2016). Therefore, we have included a broader definition of antibiotoic prophylaxis to avoid an underestimation of the proportion of children with sickle cell anemia who are protected against pneumococcal infection.

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?

N/A

_____

1a.4 OTHER SOURCE OF EVIDENCE

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

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

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

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

References

American Academy of Pediatrics Section on Hematology/Oncology and Committee on Genetics. Health supervision for children with sickle cell disease. *Pediatrics* 2002; 109(3):526-535.

American Academy of Pediatrics. AAP publications reaffirmed or retired. Pediatrics 2016; 137(5):e20160592.

Bjornson AB, Falletta JM, Verter JI, Buchanan GR, Miller ST, Pegelow CH, et al. Serotype-specific immunoglobulin G antibody responses to pneumococcal vaccine in children with sickle cell anemia: effects of continued penicillin prophylaxis. *J Pediatr*1996;129(6):828-835.

Falletta JM, Woods GM, Verter JI, Buchanan GR, Pegelow CH, Iyer RV, et al. Discontinuing penicillin prophylaxis in children with sickle cell anemia. Prophylactic Penicillin Study II. *J Pediatr* 1995; 127(5):685-690.

Gaston MH, Verter JI, Woods G, Pegelow C, Kelleher J, Presbury G, et al. Prophylaxis with oral penicillin in children with sickle cell anemia. A randomized trial. *N Engl J Med* 1986; 314(25):1593-1599.

Gaston MH, Verter J. Sickle cell anaemia trial. *Stat Med* 1990; **9**(1-2):45-51.

John AB, Ramlal A, Jackson H, Maude GH, Sharma AW, Serjeant GR. Prevention of pneumococcal infection in children with homozygous sickle cell disease. *Br Med J (Clin Res Ed)* 1984; 288(6430):1567-1570.

Nkouwap I, Diara JP, Noyon I, Etienne-Julan M, Merault L. Is there any alternative to oral penicillin in antibioprophylaxis for children with sickle cell disease? [French] Y a-t-il une alternative a la penicilline orale dans l'antibioprophylaxie chez les enfants drepanocytaires? *Med Mal Infect*. 1999; 29(2):111-116.

Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; 273(5):408–412.

Woods GM, Jorgensen JH, Waclawiw MA, Reid C, Wang W, Pegelow CH, et al. Influence of penicillin prophylaxis on antimicrobial resistance in nasopharyngeal S. pneumoniae among children with sickle cell anemia. The Ancillary Nasopharyngeal Culture Study of Prophylactic Penicillin Study II. *J Pediatr Hematol Oncol* 1997; **19**(4):327-333.

- 1b. Performance Gap

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

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

- Disparities in care across population groups.

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

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

Children with SCA are at an increased risk of infection compared with children without the disorder. Daily receipt of antibiotic prophylaxis substantially reduces the risk of infection among these high-risk children. Prior studies indicate that appropriate antibiotic prophylaxis rates are low; however, these reports are limited in their generalizability, as they are usually focused on a single healthcare provider or registry. This measure establishes a claims-based method for identifying appropriate antibiotic prophylaxis among larger and broader populations of children with SCA. The performance scores calculated through this measure will identify areas in need of improvement of antibiotic prophylaxis among children with SCA.

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

Rates of appropriate antibiotic prophylaxis dispensed for at least 300 days within the measurement year for children with sickle cell anemia in Michigan and New York State Medicaid Programs (2011-2018).

Michigan

(Year: Numerator / Denominator = Rate)

2011: 14/85 = 16.47%

2012: 11/83 = 13.25%

2013: 10/85 = 11.76%

2014: 15/73 = 20.55%

2015: 11/75 = 14.67%

2016: 7/110 = 6.36%

2017: 8/91 = 8.79%

2018: 8/95 = 8.42%

New York State

(Year: Numerator / Denominator = Rate)

2011: 68/290 = 23.45%

2012: 68/309 = 22.01%

2013: 64/324 = 19.75%

2014: 60/299 = 20.07%

2015: 74/345 = 21.45%

2016: 72/373 = 19.30%

2017: 69/338 = 20.41%

2018: 63/347 = 18.16%

Rates of appropriate antibiotic prophylaxis dispensed for at least 300 days within the measurement year for children with sickle cell anemia in six states (2005-2010).

Florida

(Year: Numerator/Denominator = Rate)

2005: 27/176 = 15.34%

2006: 21/169 = 12.43%

2007: 23/125 = 18.40%

2008: 18/155 = 11.61%

2009: 30/238 = 12.61%

2010: 44/282 = 15.60%

Illinois

(Year: Numerator / Denominator = Rate)

2005: 19/55 = 34.55%

2006: 24/77 = 31.17%

2007: 26/72 = 36.11%

2008: 21/72 = 29.17%

2009: 22/96 = 22.92%

2010: 13/75 = 17.33%

Louisiana

(Year: Numerator / Denominator = Rate)

2005: 16/104 = 15.38%

2006: 12/89 = 13.48%

2007: 36/126 = 28.57%

2008: 32/128 = 25.00%

2009: 27/120 = 22.50%

2010: 19/120 = 15.83%

Michigan

(Year: Numerator / Denominator = Rate)

2005: 9/42 = 21.43%

2006: 8/56 = 14.29%

2007: 10/70 = 14.29%

2008: 13/73 = 17.81%

2009: 7/96 = 7.29%

2010: 20/95 = 21.05%

South Carolina

(Year: Numerator / Denominator = Rate)

2005: 12/51 = 23.53%

2006: 9/52 = 17.31%

2007: 12/60 = 20.00%

2008: 2/34 = 5.88%

2009: 2/35 = 5.71%

2010: 11/41 = 26.83%

Texas

(Year: Numerator / Denominator = Rate)

2005: 28/84 = 33.33%

2006: 20/77 = 25.97%

2007: 30/100 = 30.00%

2008: 26/122 = 21.31%

2009: 31/123 = 25.20%

2010: 38/136 = 27.94%

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

NA

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

The datasets used for performance scores is Medicaid; therefore, there are no disparities identified by insurance or socioeconomic status. However, the majority of children with sickle cell anemia (approximately 90%) have been enrolled in Medicaid at some point in time.

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

Several studies have pointed to disparities in prophylactic medication use among patients with public versus private insurance. In a study of children with sickle cell disease on Medicaid in Washington state and Tennessee, 10.3% of patients with public insurance received no penicillin or macrolide antibiotic or trimethoprim-sulfamethoxazole during a 365-day period while only 21.5% received more than 270 days of medication in a year. Median duration of prescriptions was 10 days (Sox et al., 2003). In a 10-year retrospective cohort study of 407 infants enrolled in the Tennessee Medicaid program, 60% of infants with sickle cell disease did not have recommended prophylactic antibiotic prescriptions filled within the recommended period (i.e., the first 12 weeks of life) (Warren et al., 2010). A study assessing compliance with penicillin prophylaxis for sickle cell disease found that adherence was significantly greater in patients with private versus public insurance (17/28 [61%] vs. 33/90 [37%], respectively) (Teach et al., 1998.) Clearly, however, significant room for improvement remains, despite type of insurance.

References

Sox CM, Cooper WO, Koepsell TD, DiGiuseppe DL, Christakis DA. Provision of pneumococcal prophylaxis for publicly insured children with sickle cell disease. JAMA 2003; 290(8):1057-1061.

Teach SJ, Lillis KA, Grossi M. Compliance with penicillin prophylaxis in patients with sickle cell disease. Arch Pediatr Adolesc Med 1998; 152(3):274-278.

Warren MD, Arbogast PG, Dudley JA, et al. Adherence to prophylactic antibiotic guidelines among Medicaid infants with sickle cell disease. Arch Pediatr Adolesc Med 2010; 164(3):298-299.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

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

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

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

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

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

https://chear.org/research/projects/qmetric#phase2

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

This is not an eMeasure **Attachment:**
Updated_Sept_8_2020_Antibiotics_NQF_testing_attachment_508_compliant-637413804252354620.docx

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

Attachment**:** SCA_Antibiotic_Measure_Appendix_Tables_20180501.xlsx

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

No, this is not an instrument-based measure **Attachment:**
Updated_Sept_8_2020_Antibiotics_NQF_testing_attachment_508_compliant-637413804342512136.docx

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

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

Yes

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

The denominator has been validated using ICD-10-CM diagnosis codes.

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

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

The numerator is the number of children ages 3 months to 5 years old with SCA who were dispensed appropriate antibiotic prophylaxis for at least 300 days within the measurement year.

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

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

Cases from target population with target process (appropriate antibiotic prophylaxis dispensed for at least 300 days within the calendar year): Antibiotic prophylaxis is defined as at least 300 days covered within the measurement year, which is the summed total of the number of days' supply of antibiotics dispensed within the measurement year (see National Drug Codes (NDC) Table 1).

NOTE: Although NHLBI guidelines specifically recommend penicillin for antibiotic prophylaxis, some children may have or be suspected to have penicillin sensitivity. The American Academy of Pediatrics Section on Hematology/Oncology and Committee on Genetics suggests an alternative for children who are allergic to penicillin: "Erythromycin prophylaxis may be used as an alternative for children with suspected or proven penicillin allergy" (Citation: American Academy of Pediatrics Section on Hematology/Oncology and Committee on Genetics (Pediatrics 2002; 109(3):526-535; Reaffirmed in 2016). Providers may also choose to prescribe amoxicillin. Therefore, we have included a broader definition of antibiotic prophylaxis than penicillin in this measure (penicillin, erythromycin, amoxicillin). This is intended to avoid underestimation of the proportion of children with SCA who are protected against pneumococcal infection.

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

The denominator is the number of children ages 3 months to 5 years with sickle cell anemia (SCA) within the measurement year.

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

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

For calculation of measure using ICD-9: Children with SCA are identified through the presence of at least three separate healthcare encounters related to SCA within the measurement year (ICD-9 codes 282.61, 282.62). Children ages 3 months to 5 years are included within the target population (i.e., must not have a 5th birthday within the measurement year). Children must be continuously enrolled within the health plan in which claims are available and must have no other form of health insurance for the entire measurement year.

For calculation of measure using ICD-10: Children with SCA are identified through the presence of at least one outpatient visit with an ICD-10 diagnosis code of D57.1, D57.00, D57.01 or D57.02. Children ages 3 months to 5 years are included within the target population (i.e., must not have a 5th birthday within the measurement year). Children must be continuously enrolled within the health plan in which claims are available and must have no other form of health insurance for the entire measurement year.

Note: Children with SCA are included starting at 3 months of age to account for any lag in identification and confirmation of the sickle cell disease status of the child.

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

There are no denominator exclusions.

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

NA

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

NA

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

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

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

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

1. Identify the denominator: Determine the eligible population using administrative claims. The eligible population is all individuals who satisfy all specified criteria, including age, continuous enrollment, and benefit requirements within the measurement year.
2. Identify the numerator: Identify numerator events using administrative claims for all individuals in the eligible population (denominator) within the measurement year.
3. Calculate the rate: (numerator/denominator).

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

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

This measure does not involve sampling; all SCA cases meeting the inclusion criteria are used in the calculation.

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

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

NA

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

If other, please describe in S.18.

Claims

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

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

NA

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

No data collection instrument provided

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

Health Plan

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

Other

If other: Any setting represented with prescription medication claims data

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

NA

2. Validity – See attached Measure Testing Submission Form

Updated_Sept_8_2020_Antibiotics_NQF_testing_attachment_508_compliant-637413808900417031.docx

2.1 For maintenance of endorsement

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

No

2.2 For maintenance of endorsement

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

Yes

2.3 For maintenance of endorsement

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

No - This measure is not risk-adjusted

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

Measure Number (*if previously endorsed*)**:** 3166
Measure Title: Antibiotic Prophylaxis among Children with Sickle Cell Anemia
Date of Submission: 8/3/2020

Type of Measure:

Measure	Measure (continued)
☐ Outcome (*including PRO-PM*)	☐ Composite – *STOP – use composite testing form*

Measure	Measure (continued)
☐ Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	- -

- - cell intentionally left blank

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

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

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

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.17***)	**Measure Tested with Data From:**
☐ abstracted from paper record	☒ abstracted from paper record
☒ claims	☒ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☒ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other:	☒ other: Newborn Screening

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

- Michigan Medicaid administrative claims data provided by the Michigan Department of Health and Human Services (MDHHS)
- New York State Medicaid administrative data provided by the New York State Department of Health
- Medicaid Analytic eXtract (MAX) administrative claims data for six state Medicaid programs provided by the Centers for Medicare & Medicaid Services (CMS)
- Michigan Newborn Screening Results
- New York State Newborn Screening Results
- Medical record data from Children's Hospital of Michigan (CHM), Detroit, Michigan; Hurley Medical Center (HMC), Flint, Michigan; and University of Michigan Health System (UMHS), Ann Arbor, Michigan

1.3. What are the dates of the data used in testing? Michigan Medicaid data 2007-2011, 2016; New York State Medicaid data: 2016; MAX data: 2005-2012; Michigan Newborn Screening: 1987-2014; New York State Newborn screening: 2006-2013; Medical record data: 2012

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

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

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

- Michigan Medicaid: all Medicaid claims for children ages 3 months through 5 years continuously enrolled for at least one year in Michigan Medicaid in 2007-2010; all Medicaid claims for children ages 1 through 17 continuously enrolled for at least one year in 2010, 2011, or 2016.
- New York State Medicaid: all Medicaid claims for children ages 1 through 17 continuously enrolled in New York Medicaid in 2016.
- MAX data: all Medicaid claims reported to CMS for children ages 3 months through 5 years continuously enrolled for at least one year from 2005-2012 within one of the six following state Medicaid programs with moderate to high prevalence of sickle cell anemia: Florida, Illinois, Louisiana, Michigan, South Carolina, and Texas.
- Michigan NBS data consisted of all births within the state of Michigan from 1987-2014.
- New York State NBS data consisted of all sickle cell births within the state of New York from 2006-2013.
- Medical record data were obtained from three hospitals in 2012: CHM, HMC, and UMHS. These three large medical centers are located in urban areas in Michigan reflective of the residence of the vast majority (~70%) of children with sickle cell anemia living in Michigan:
    - CHM is a tertiary medical center located in Detroit, Michigan.
    - HMC is a tertiary medical center located in Flint, Michigan.
    - UMHS is an academic medical center located in Ann Arbor, Michigan

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

- Michigan Medicaid data for ICD-9-CM denominator validation (2010 and 2011) included 865 children in 2010 and 863 children in 2011. These children were included if continuously enrolled in either 2010 or 2011, with at least one sickle cell disease-related administrative claim, and a newborn screening result available from 1987-2010.
- Michigan Medicaid data for ICD-10-CM denominator validation (2016 only) included 1,597 children with at least one D57x diagnosis code in 2016.

- Michigan Medicaid data for performance score validation included 114 children in 2007, 118 children in 2008, 149 children in 2009, and 141 children in 2010. These children were included if continuously enrolled in Michigan Medicaid for at least one year from 2007-2010, and met the ICD-9-CM case definition for sickle cell anemia.
- New York State Medicaid data included 2,454 children with at least one D57x diagnosis code in 2016.
- MAX data (2005-2012) included 2,821 children enrolled for a total of 5,014 person-years. The breakdown of person-years by state is as follows: Florida: 1,619; Louisiana: 855; Texas: 897; Illinois: 622; Michigan: 580; South Carolina: 441. Children were included if meeting the denominator criteria for sickle cell anemia.
- Michigan NBS data for ICD-10-CM validation included a total of 1,457 children that had both a D57x diagnosis code in 2016 and a Newborn Screening Result available from 1997-2014.
- New York State NBS data included 2,454 children with a D57x diagnosis code in 2016 and born between 2006-2013.
- Medical record data included 34 children with sickle cell anemia ages 3 months to 5 years who were enrolled in Michigan Medicaid and received care at one of three sickle cell centers in Michigan (CHM, HMC, UMHS) during 2012.

Across all datasets, the majority of children were Black, with a similar proportion of males and females.

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

- Reliability testing data: MAX
- Validity testing data: Michigan Medicaid, MAX, Michigan Newborn Screening, New York State Medicaid, New York State Newborn Screening, and medical records
- Identification of meaningful differences in performance: MAX data

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

The data do not include data on patient-level social risk factors; however, all children included in the data were enrolled in Medicaid, which is a proxy for low socioeconomic status.

2a2. RELIABILITY TESTING

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

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

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

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

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

See section 2b2 for validity testing of data elements.

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

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

_____

2b1. VALIDITY TESTING

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

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

☒ **Performance measure score**

☒ **Empirical validity testing**

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

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

Critical Data Elements

Denominator: The denominator (children with sickle cell anemia) has been validated using both ICD-9-CM and ICD-10-CM diagnosis codes. The process for each separate validation (ICD-9-CM and ICD-10-CM) is detailed below:

ICD-9-CM denominator validation: The accuracy of the case definition (at least three claims for sickle cell anemia [Hb SS] within the measurement year) to identify children with sickle cell anemia was assessed through comparison with the gold standard of newborn screening results for the state of Michigan for children enrolled in Michigan Medicaid in 2010 and 2011 with at least one SCD-related healthcare claim within their enrollment year(s). The area under the receiver operating characteristic (ROC) curve, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated for the case definition. As a comparison, these values were also calculated for those with a minimum of at least one or two Hb SS claims within each year.

ICD-10-CM denominator validation: We developed, tested, and validated the performance of ICD-10-CM case definitions for SCA through the following 4-step process: (a) developing candidate SCA case definitions; (b) identifying a test population for the definitions; (c) testing the accuracy of candidate case definitions; and (d) testing case definitions within an independent population. Briefly, using specific SCA-related (D5700, D5701, D5702) and non-specific (D571) diagnosis codes, 23 SCA case definitions were applied to Michigan Medicaid claims (2016) to identify children with SCA. Measures of performance (sensitivity, specificity, area under ROC curve) were calculated using newborn screening results as the gold standard. A parallel analysis was conducted using New York State Medicaid claims and newborn screening data. We also performed a sensitivity analysis to understand the performance of this case definition for children ages 1 through 4, the target age group for this performance measure.

Additional details regarding the methodology of validating the ICD-10-CM definition can be found at: Reeves SL, Madden B, Wu M, et al. Performance of ICD-10-CM diagnosis codes for identifying children with Sickle Cell Anemia. Health Serv Res. 2020;55(2):310-317. doi:10.1111/1475-6773.13257

Numerator: The numerator consists of children with sickle cell anemia that have at least 300 days of filled antibiotics within the year. We used three methods to validate our numerator.

(1) We performed a comprehensive examination of National Drug Codes (NDCs) to identify antibiotics using clinical expertise and RxNorm. RxNorm is a normalized naming system for all drugs produced and maintained by the National Library of Medicine, and is the gold standard for identifying current and historical medications contained in administrative data. We then identified all filled prescriptions for children in the denominator using MAX data from 2005-2012. We then identified all NDCs that were not related to antibiotic prophylaxis

based on our comprehensive list. We compared this list to RxNorm to manually assess that all remaining NDCs were not related to antibiotic prophylaxis.
(2) We assessed various categorizations of antibiotic prophylaxis using MAX data from 2005-2012. We calculated and compared the days' supply using three definitions: 1) penicillin only; 2) penicillin + erythromycin; 3) penicillin + erythromycin + amoxicillin; and 4) any antibiotic protective against invasive pneumococcal infection. We calculated mean days' supply and standard deviation for each categorization.
(3) The accuracy of administrative claims in identifying antibiotic prophylaxis was assessed through comparison with the gold standard of medical charts. An audit was conducted by trained medical record abstractors to compare administrative claims data with corresponding medical records data. Medical records were abstracted for all children meeting the measure specification criteria. Each medical record was assessed for the presence of a prescription for antibiotics and compared with administrative claims to determine if the prescription was filled and the difference between the date of prescription and the prescription fill. In addition, the reliability of the data element abstracted from the medical chart was assessed by identifying a subset of the charts to be re-abstracted by another trained medical record abstractor; the results of the two abstractors were compared using percent agreement and kappa.

Empirical Validity Testing of Performance Measure

Although a state would typically have direct access to its own Medicaid data, it is unlikely that a state would have similar access to other states' data for comparison. However, CMS develops and maintains standardized Medicaid Analytic eXtract (MAX) data for public use using administrative claims submitted by each state Medicaid program. The MAX data are the only national, person-level administrative claims dataset available for the Medicaid program. As a consequence, MAX data, rather than data acquired directly from individual Medicaid programs, are likely to be used to perform cross-state comparisons of antibiotic prophylaxis among children with sickle cell anemia. Since states submit their Medicaid data to CMS for conversion into the MAX datasets, a state's own Medicaid data can be considered the authoritative source for administrative claims. Our empirical validity testing of this performance measure compared the MAX data for the state of Michigan (obtained from CMS) with the gold standard of Michigan Medicaid data (obtained directly from Michigan's claims data warehouse) for the same time period (2007-2010). Rates of antibiotic prophylaxis using each source of data were calculated and compared using z-tests for two proportions; for these tests, the null hypothesis was that the rate in each year would be the same in both Michigan Medicaid data and MAX data.

Face Validity of Performance Measure Score

The face validity of this measure was established by a panel of national experts and advocates for families of children with SCD convened by the Quality Measurement, Evaluation, Testing, Review, and Implementation Consortium (QMETRIC). The QMETRIC expert panel included nationally recognized experts in SCD, representing hematology, pediatrics, and SCD family advocacy. In addition, measure validity was considered by experts in state Medicaid program operations, health plan quality measurement, health informatics, and health care quality measurement. In total, the QMETRIC SCD panel included 14 experts, providing a comprehensive perspective on SCD management and the measurement of quality metrics for states and health plans. The expert panel assessed whether the performance of the measure would result in improved quality of care for children with sickle cell disease. Specifically, in respect to antibiotic prophylaxis, the panel weighed evidence to determine if the performance outlined in the measure would improve the quality of care provided to patients. The voting process to prioritize the measure was based on the ability of the measure to distinguish good from poor quality.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)
Denominator: The denominator has been validated using both ICD-9-CM and ICD-10-CM diagnosis codes. The results for each separate validation (ICD-9-CM and ICD-10-CM) is detailed below:

ICD-9-CM denominator validation: In 2010, a case definition of **three Hb SS (282.61, 282.62) claims within the year** was 91.4% sensitive and 80% specific in identifying children with sickle cell anemia (Hb SS) (PPV: 80.4%; NPV: 91.3%). These results were replicated with nearly identical precision among the study population in 2011. In comparison, using a case definition of at least one Hb SS claim or at least two Hb SS claims to identify the study population resulted in substantially less specificity.

ICD-10-CM denominator validation: Across the 23 case definitions, measures of performance varied, with sensitivities from 0.02-0.97 and specificities from 0.88-1.0. The case definition of **≥1 outpatient visit with a SCA-related (D5700, D5701, D5702) or D571 code** had the highest area under the ROC curve, with a sensitivity of 95% and specificity of 92%. The same definition also had the highest performance in New York Medicaid (n=2,454), with a sensitivity of 94% and specificity of 86%. There were no clinically relevant changes in the sensitivities and specificities of the definitions when limited to ages 1 through 4.

Numerator:

(1) Through the methodology described above, we identified 28,881 NDCs for antibiotic prophylaxis within administrative claims.

(2) The average number of days of filled prescriptions for antibiotic prophylaxis was as follows: 162 days of penicillin (SD = 117; median: 160), 164 days of penicillin or erythromycin (SD = 117; median: 160), 178 days of penicillin, erythromycin, or amoxicillin (SD = 113; median: 180), and 193 days of any *S pneumoniae* antibiotic (SD =116; median: 194).

(3) For this comparison, 34 children with sickle cell anemia who were enrolled within Michigan Medicaid were successfully matched with their Michigan Medicaid administrative claims data. Among these children, 25 cases (76%) had a Medicaid administrative claim for an antibiotic prescription filled within 29 days of the prescription date in the medical record. An additional seven cases (21%) had an antibiotic prescription filled between 30 but less than 90 days following the prescription date. One case (3%) had an antibiotic prescription claim filled 90 or more days after the prescription date in the medical record. Eighteen charts were also chosen for calculation of inter-rater reliability; the two trained abstractors had 100% agreement with each other for abstracting antibiotic prescriptions from the medical records, resulting in a kappa of 1.00.

Empirical Validity Testing of Performance Measure

The comparison of rates of appropriate antibiotic prophylaxis from the gold standard of Michigan Medicaid data compared with MAX data can be seen in Table 2. This illustrates that the number of children who were dispensed at least 300 days of antibiotics among children with sickle cell anemia ranged from 14 to 23 in the claims acquired directly from the Medicaid data warehouse, versus a range of 7 to 20 from MAX data for the same time period. Table 3 reports the z-scores and p-values from the two-sample z-tests comparing the proportion of children who were dispensed at least 300 days of antibiotic prophylaxis each year between Michigan Medicaid and MAX data.

Table 2: Comparison of appropriate antibiotic prophylaxis by source of Medicaid claims data for the state of Michigan, 2007-2010

Source	Rate Components	2007	2008	2009	2010
Michigan Medicaid data (1)	Numerator	14	23	22	22
Michigan Medicaid data (2)	Denominator	114	118	149	141
Michigan Medicaid data (3)	Percentage	12.28%	19.49%	14.77%	15.60%
MAX data (1)	Numerator	10	13	7	20

Source	Rate Components	2007	2008	2009	2010
MAX data (2)	Denominator	70	73	96	95
MAX data (3)	Percentage	14.29%	17.81%	7.29%	21.05%

Table 3: Comparison of performance score by source of Medicaid claims data, Michigan

Measure	2007	2008	2009	2010
z-score	0.3921	-0.2890	-1.7678	1.0735
p-value	0.6965	0.7718	0.0767	0.2846

Figure 1: Comparison of appropriate antibiotic prophylaxis by source of Medicaid claims data for the state of Michigan, 2007-2010

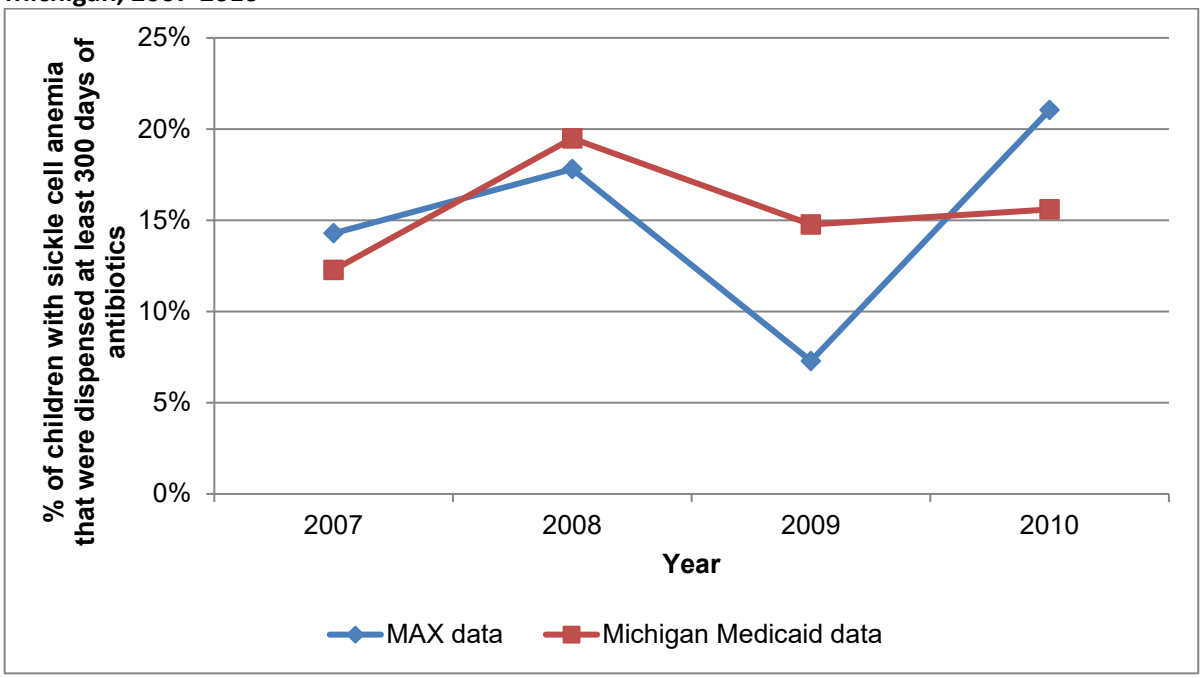

Figure 1 illustrates the performance scores observed between the Michigan Medicaid data from the state warehouse and MAX data from CMS for each overlapping year noted, respectively: 12.3% versus 14.3% (2007); 19.5% versus 17.8% (2008); 14.8% versus 7.3% (2009), and 15.6% versus 21.1% (2010).

Face Validity of Performance Measure Score

The QMETRIC expert panel concluded that this measure has a very high degree of face validity through a detailed review of concepts and metrics considered to be essential to effective SCD management and treatment. Concepts and draft measures were rated by this group for their relative importance. This measure was among the most highly rated, receiving an average score of 8.5 (with 9 as the highest possible score). In addition, the expert panel concluded that the performance outlined in this measure would improve the quality of care provided to patients, and the measure would be able to distinguish good from poor quality.

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

Critical Data Elements

Denominator: Using both ICD-9-CM and ICD-10-CM diagnosis codes, children with sickle cell anemia can be identified with a high level of accuracy in administrative data.

Numerator: Our comprehensive approach to identifying NDCs related to antibiotic prophylaxis, as well as categorization of the NDCs, results in a high level of validity for identifying children within the numerator. The majority (33/34, 97%) of prescribed antibiotics in the medical record were reflected in administrative claims. Further, 30/34 (88%) of those who were prescribed an antibiotic in the medical record were dispensed their antibiotic within 90 days of the prescription. Given this evidence, we believe the validity of administrative claims in assessing antibiotic prescriptions is very high.

Empirical Validity Testing of Performance Measure

Our results suggest that, compared with the gold standard of Michigan Medicaid data, MAX data has a very high degree of validity. When antibiotic prophylaxis was assessed for the same state (Michigan) from these two data sources for the same time period (2007-2010), no differences in rates were observed (all p-values >0.05). Therefore, our results suggest that compared with Michigan Medicaid data, MAX data is highly valid.

Face Validity of Performance Measure Score

Given the high rating of the QMETRIC expert panel, we feel this measure has a very high degree of face validity.

_____

2b2. EXCLUSIONS ANALYSIS

NA ☒ no exclusions — *skip to section 2b4*

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

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

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

_____

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b3.1. What method of controlling for differences in case mix is used?

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with risk factors**

☐ **Stratification by risk categories**

☐ **Other,**

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

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

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g.,*

potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

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

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

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

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

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

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

If stratified, skip to 2b3.9

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

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

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

2b3.9. Results of Risk Stratification Analysis:

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

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

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

We used MAX data to determine if differences at the state health plan level could be assessed using this measure. We calculated proportion of children dispensed at least 300 days of antibiotic prophylaxis for each state health plan from 2005-2012. Logistic regression was used to estimate the associations between each state health plan and at least 300 days of antibiotic prophylaxis, with Illinois used as the reference category. Generalized estimating equation (GEE) models with robust standard errors were used to account for the correlation among children. Odds ratios with 95% confidence intervals were used to assess the final associations. For all models, regression diagnostics were performed to assess normality of error variances; healthcare utilization, year, and demographics were adjusted for within the model.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g.,

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

The proportion of children with at least 300 days of antibiotics ranged across states from 6% to nearly 40%. Compared to Illinois, Florida, Louisiana, Michigan, and South Carolina had lower odds of children receiving at least 300 days of antibiotic prophylaxis (Figure 2).

Figure 2. Multivariable model predicting receipt of at least 300 days of antibiotic prophylaxis among children with sickle cell anemia (n = 5014 person-years).

Variable	OR	Lower 95% CI	Upper 95% CI
Type of visit			
ED	1.02	0.99	1.04
SCD-related outpatient	1.01*	1.01	1.02
Well child	1.08*	1.03	1.14
State			
Florida	0.51*	0.39	0.68
Illinois	1.00	Reference	Reference
Louisiana	0.57*	0.41	0.78
Michigan	0.60*	0.42	0.85
South Carolina	0.62*	0.43	0.89
Texas	1.01	0.76	1.35
Year			
2005	1.00	Reference	Reference
2006	0.92	0.69	1.23
2007	1.21	0.91	1.60
2008	0.98	0.72	1.33
2009	0.80	0.59	1.09
2010	0.99	0.73	1.33
2011	0.98	0.73	1.31
2012	0.85	0.63	1.15

Study consisted of 2821 children. SCD, sickle cell disease.

*$P<.05$.

Citation for Figure 2: Reeves SL, Tribble AC, Madden B, Freed GL, Dombkowski KJ. Antibiotic Prophylaxis for Children With Sickle Cell Anemia. Pediatrics. 2018;141(3):e20172182. doi:10.1542/peds.2017-2182

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

This measure can successfully distinguish differences in performance across state health plans.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note*: This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to*

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

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

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

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

We tested the completeness and accuracy of Medicaid claims for sickle cell patients seen at three major medical centers in Michigan. For this comparison, children with sickle cell anemia who were enrolled within Michigan Medicaid were matched with their Michigan Medicaid administrative claims data (n=34).

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

Of these children, the vast majority of prescribed antibiotics in the medical record were reflected in Medicaid claims (33/34, 97%). We found that the preponderance of those who were prescribed an antibiotic in the medical record had a corresponding Medicaid claims for a dispensed antibiotic within 30 days of the prescription (76%). An additional seven cases (21%) had an antibiotic prescription filled between 30 but less than 90 days following the prescription date. We noted only one case (3%) that did not have an antibiotic prescription claim filled within 90 days of a prescription date.

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

Given this evidence, we believe that missing data is unlikely to bias our performance results, particularly as any missing data would be expected to be non-differential across entities.

3. Feasibility

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

3a. Byproduct of Care Processes

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

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

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

If other:

3b. Electronic Sources

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

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

ALL data elements are in defined fields in electronic claims

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

NA

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

Attachment:

3c. Data Collection Strategy

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

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

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

This measure was tested using Medicaid administrative claims data. The primary information needed for this measure includes basic demographics, diagnosis codes, and procedure codes and dates. These data are widely available, although obtaining them often requires a restricted-use data agreement. For multiple-state comparisons, MAX data are available from CMS. When the measure is used at the single-state level, state health departments can use their own Medicaid data. Acquisition of data directly from states requires the cooperation of those jurisdictions, as well as modification of the statistical programming code developed for MAX files to correctly function using the unique structure of the data files obtained from each state.

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

NA

4. Usability and Use

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

4a. Accountability and Transparency

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

4.1. Current and Planned Use

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

Specific Plan for Use	Current Use (for current use provide URL)
Payment Program Quality Improvement (external benchmarking to organizations) Quality Improvement (Internal to the specific organization)	- -

- - cell intentionally left blank

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

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

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

The State of Michigan Medicaid Program will be implementing this measure for purposes of building a health plan collaborative to improve the care of children with sickle cell anemia in Michigan. This program is currently being implemented; we anticipate reporting to occur in 2021.

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

The health plan collaborative is currently being implemented; we anticipate reporting to occur in 2021. Program implementation was delayed due to COVID-19.

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

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

Specifications and technical assistance were provided to Michigan Medicaid health plans and the State of Michigan Medicaid Program.

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

The measure development team has provided feedback and clarification when necessary to all users.

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

Describe how feedback was obtained.

We spoke to the health plan data teams after implementation.

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

Health plans were able to implement the measure without issue and obtain performance scores. This allowed health plans to identify opportunities for improvement.

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

All users have indicated that the measure is feasible and usable.

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

The measure has only been modified to reflect the validation of the denominator using ICD-10-CM diagnosis codes.

Improvement

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

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

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

There has been no improvement demonstrated. A national, strategic effort for accountability, such as inclusion in the Child Core Set, would drive improvement. This measure has been recommended for inclusion in the Core Set for multiple years.

4b2. Unintended Consequences

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

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

No unintended negative consequences to individuals or populations were identified during testing.

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

This measure has been considered for inclusion in the CMS Core Set.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria **and** there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

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

Yes

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

2797 : Transcranial Doppler Ultrasonography Screening Among Children with Sickle Cell Anemia

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

Hydroxyurea Use Among Children with Sickle Cell Anemia - Regents of the University of Michigan - Being submitted for initial NQF review (Fall 2020 cycle)

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):

Are the measure specifications harmonized to the extent possible?

Yes

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

Different age categories are included in the measures. For example, antibiotic prophylaxis is recommended by NHLBI for ages 0 until 5; TCD screening from ages 2 until 16; and hydroxyurea beginning at 9 months of age. Further, the numerators are identifying different events (antibiotics, hydroxyurea, TCD); therefore, the numerator specifications differ across each measure.

5b. Competing Measures

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

OR

Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment **Attachment:** UM_Antibiotic_Prophylaxis_Measure_Steward_Agreement.pdf

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): QMETRIC - University of Michigan

Co.2 Point of Contact: Sarah, Reeves, sleasure@med.umich.edu, 734-615-9755-

Co.3 Measure Developer if different from Measure Steward: QMETRIC – University of Michigan

Co.4 Point of Contact: Gary, Freed, gfreed@med.umich.edu, 734-647-3610-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

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

The face validity of this measure was established by a national panel of experts and advocates for families of children with SCD convened by the Quality Measurement, Evaluation, Testing, Review, and Implementation Consortium (QMETRIC) at the University of Michigan. The QMETRIC Representative Panel included nationally recognized experts in SCD, representing hematology, pediatrics, and SCD family advocacy. The QMETRIC Feasibility Panel included experts in state Medicaid program operations, health plan quality measurement, health informatics, and health care quality measurement. In total, the QMETRIC SCD panels included 14 experts, providing a comprehensive perspective on SCD management and the measurement of quality metrics for states and health plans.

The QMETRIC expert panels concluded that this measure has a very high degree of face validity through a detailed review of concepts and metrics considered to be essential to effective SCD management and treatment. Concepts and draft measures were rated by this group for their relative importance. This measure was among the most highly rated, receiving an average score of 8.5 (with 9 as the highest possible score).

Sickle Cell Disease Representative Panel:

Samir Ballas, MD, Professor, Division of Hematology, Thomas Jefferson University, Philadelphia, PA

Mary E. Brown, President and Chief Executive Officer, Sickle Cell Disease Association, Los Angeles, CA

George Buchanan, MD, Pediatric Hematologist, University of Texas Southwest Medical Center at Dallas, TX

Peter Lane, MD, Pediatric Hematologist-Oncologist, Children's Healthcare of Atlanta Pediatric Hospital, Atlanta, GA

Suzette Oyeku, MD, Assistant Professor of Pediatrics, Albert Einstein College, Bronx, NY

Lynnie Reid, Parent Representative, Boston, MA

Elliott Vichinsky, MD, Pediatric Hematology-Oncology, Children's Hospital and Research Center, Oakland, CA

Winfred Wang, MD, Hematologist, St. Jude Children's Hospital, Memphis, TN

Sickle Cell Disease Feasibility Panel:

Cathy Call, BSN, MSC, Senior Policy Analyst and Director for Health Quality Research, Altarum Institute, Alexandria, VA

J. Mitchell Harris, PhD, Director Research and Statistics, Children's Hospital Association, (formerly NACHRI), Alexandria, VA

Kevin Johnson, MD, MS, Professor and Vice Chair of Biomedical Informatics, Vanderbilt University, Nashville, TN

Don Lighter, MD, MBA, FAAP, FACHE, Director, The Institute for Health Quality Research and Education, Knoxville, TN

Sue Moran, BSN, MPH, Director of the Bureau of Medicaid Program Operations and Quality Assurance, Michigan Department of Community Health, Lansing, MI

Joseph Singer, MD, Vice President Clinical Affairs, HealthCore, Inc., Wilmington, DE

C. Jason Wang, MD, PhD, Associate Professor of Pediatrics, Stanford School of Medicine, Stanford, CA

QMETRIC Investigators:

Kevin J. Dombkowski, DrPH, MS, Research Professor of Pediatrics, School of Medicine, University of Michigan, Ann Arbor, MI

Gary L. Freed, MD, MPH, Professor of Pediatrics, School of Medicine and Professor of Health Management and Policy, School of Public Health, University of Michigan, Ann Arbor, MI (principal investigator)

Sarah Reeves, PhD, MPH, Assistant Professor of Pediatrics, School of Medicine, University of Michigan, Ann Arbor, MI

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2017

Ad.3 Month and Year of most recent revision: 05, 2018

Ad.4 What is your frequency for review/update of this measure? 1 year

Ad.5 When is the next scheduled review/update for this measure? 05, 2021

Ad.6 Copyright statement:

Ad.7 Disclaimers: This work was funded by the Agency for Healthcare Research and Quality (AHRQ) and the Centers for Medicare & Medicaid Services (CMS) under the CHIPRA Pediatric Quality Measures Program Centers of Excellence grant number U18 HS020516 and contract number HHSP233201600166A. AHRQ, in accordance to CHIPRA 42 U.S.C. Section 1139A(b), and consistent with AHRQ´s mandate to disseminate research results, 42 U.S.C. Section 299c-3, has a worldwide irrevocable license to use and permit others to use products and materials from the grant for government purposes, which may include making the materials available for verification or replication by other researchers and making them available to the health care community and the public, if such distribution would significantly increase access to a product and thereby produce substantial or valuable public health benefits. The Measures can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the measures require a license agreement between the user and the Quality Measurement, Evaluation, Testing, Review and Implementation Consortium (QMETRIC) at the University of Michigan (U-M). Neither QMETRIC/U-M nor their members shall be responsible for any use of the Measures. QMETRIC/U-M makes no representations, warranties or endorsement about the quality of any organization or physician that uses or reports performance measures, and QMETRIC/U-M has no liability to anyone who relies on such measures. The QMETRIC performance measures and specifications are not clinical guidelines and do not establish a standard of medical care.

This statement is signed by Gary L. Freed, MD, MPH, who, as the principal investigator of QMETRIC, is authorized to act for any holder of copyright on the submitted measure.

Gary L. Freed, MD, MPH

Percy and Mary Murphy Professor of Pediatrics, School of Medicine

Professor of Health Management and Policy, School of Public Health

Principal Investigator, QMETRIC

Child Health and Evaluation Research (CHEAR) Unit

Division of General Pediatrics

University of Michigan Hospital and Health Systems

Ann Arbor, MI 48109-5456

Ad.8 Additional Information/Comments: