



Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

Brief Measure Information

NQF #: 3154

Corresponding Measures:

Measure Title: Informed Participation

Measure Steward:

sp.02. Brief Description of Measure: Improved measurement of the continuity of insurance coverage in the Medicaid and CHIP population is needed to help maximize insurance continuity and coverage for vulnerable children. To further this goal, the AHRQ-CMS CHIPRA PQMP Center of Excellence at the Children's Hospital of Philadelphia developed the metric Informed Coverage. The metric is designed to more accurately measure coverage among children enrolled in Medicaid or CHIP at the state level and overcome the current inability in the Medicaid Analytic eXtract (MAX) dataset to determine whether a child disenrolled from Medicaid and CHIP due to loss of eligibility (such as due to parental income increase or the acquisition of employer-sponsored insurance, a "good" reason) or failure to appropriately re-enroll (a "bad" reason). This measure can help federal and state programs develop strategies to retain children eligible for coverage and minimize gaps that can occur during the renewal process. Informed Coverage assesses the continuity of enrollment of children in publicly financed insurance programs (Medicaid and CHIP), as defined by the ratio of enrolled month to eligible months over an 18 month observation window. Informed Coverage uses a natural experiment based on the random event of appendicitis to "inform" the estimate of coverage in a given state, bounded by two extreme assumptions regarding unknown eligibility information: Coverage Presumed Eligible (PE) and Coverage Presumed Ineligible (PI).

1b.01. Developer Rationale: States are frequently asked to determine public insurance participation rates or measure continuity of enrollment among vulnerable children, both for federal compliance audits and performance-based incentives, and for internal studies concerning vulnerable populations (Patrick et al., 2012; Daly 2003; National Conference of State Legislatures Health Policy Tracking Service, 2003). Participation rates are defined as the fraction of eligible children who are enrolled (Kenney et al., 2009). We developed and validated this administrative claims-based participation metric, "Informed Coverage," using a naturally occurring randomization observed inside each state that dynamically informs assumptions about patterns of eligibility and allows statewide estimates of participation rates using only administrative claims data. This standardized measure can be used by states as a potential indicator of quality and access. The issue of enrollment and retention is a long-standing concern for publically financed insurance programs, and one that states have likely examined using less formal means. Because Medicaid/CHIP enrollees are from low-income families, this measure will benefit vulnerable children it will hold states accountable for retaining children eligible for public coverage. Where data capacity permits, this measure also takes into account children switching from Medicaid and CHIP and vice-versa instead of treating children as disenrolled from public insurance.

sp.12. Numerator Statement: The numerator for Informed Coverage represents the sum (within a state) of months enrolled in Medicaid/CHIP for all children over an 18-month window.

sp.14. Denominator Statement: The sum (within a state) of months eligible for Medicaid/CHIP for all children (0-18 years) over an 18-month window. In addition, months that could be defined as “eligible” are based on known events recorded in the MAX data that would affect eligibility (birth or ageing out).

sp.16. Denominator Exclusions: For the appendicitis calculation, the population is limited to children between the ages of 2 to 16 years old. To determine what is the best assumption to use (either the Appendectomy Coverage Rate (or ACR), PI, or PE) inside each state, we compare the observed appendectomy coverage rate in a state, to the estimated coverage rate that would be calculated in that state with either PI, or PE assumptions.

Measure Type: Outcome

sp.28. Data Source:

Claims

sp.07. Level of Analysis:

Population: Regional and State

IF Endorsement Maintenance – Original Endorsement Date: 2017-07-12 04:00 PM

Most Recent Endorsement Date: 7/12/2017 4:00:39 PM

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

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

sp.03. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?:

1. 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

1ma.01. Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.

[Response Begins]

[Response Ends]

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

2021 Submission:

Updated evidence information here.

2018 Submission:

Evidence from the previous submission here.

1a.01. Provide a logic model.

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

[Response Begins]

[Response Ends]

1a.02. Provide evidence that the target population values the measured outcome, process, or structure and finds it meaningful.

Describe how and from whom input was obtained.

[Response Begins]

[Response Ends]

1a.03. Provide empirical data demonstrating the relationship between the outcome (or PRO) and at least one healthcare structure, process, intervention, or service.

[Response Begins]

[Response Ends]

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

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

[Response Begins]

States are frequently asked to determine public insurance participation rates or measure continuity of enrollment among vulnerable children, both for federal compliance audits and performance-based incentives, and for internal studies concerning vulnerable populations (Patrick et al., 2012; Daly 2003; National Conference of State Legislatures Health Policy Tracking Service, 2003). Participation rates are defined as the fraction of eligible children who are enrolled (Kenney et al., 2009). We developed and validated this administrative claims–based participation metric, “Informed Coverage,” using a naturally occurring randomization observed inside each state that dynamically informs assumptions about patterns of eligibility and allows statewide estimates of participation rates using only administrative claims data. This standardized measure can be used by states as a potential indicator of quality and access. The issue of enrollment and retention is a long-standing concern for publically financed insurance programs, and one that states have likely examined using less formal means. Because Medicaid/CHIP enrollees are from low-income families, this measure will benefit vulnerable children it will hold states accountable for retaining children eligible for public coverage. Where data capacity permits, this measure also takes into account children switching from Medicaid and CHIP and vice-versa instead of treating children as disenrolled from public insurance.

[Response Ends]

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

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

[Response Begins]

The Informed Coverage fraction distributions across states for the time period January 1, 2008 to June 30, 2009 are presented in Table 1 of the Appendix, as are the distributions for the intermediate calculations. The mean state Informed Coverage value was 0.7949 (SD 0.1035). The minimum state Informed Coverage Value was 0.3814, the maximum was 0.9350, and the IQR was 0.7474, 0.8580. The deciles of the state Informed Coverage values from 10% through 90% were as follows: 0.6623, 0.7326, 0.7576, 0.7778, 0.8261, 0.8469, 0.8571, 0.8618, 0.8840. Table 2 of the Appendix provides the number and percentage of children included in the Coverage PE and Coverage PI intermediate calculations (and thus Informed Coverage) by state out of all children present in the MAX dataset for the years 2008 to 2009. Data quality limited analyses to 43 states. Six states failed to have sufficient reporting of managed care claims for the data utilized in the development of this measure and thus were eliminated from analyses for this time period and dataset: Kentucky, Massachusetts, Mississippi, Ohio, Pennsylvania, and West Virginia. Additionally, Maine and the District of Columbia were found to have excessive quality issues in their inpatient records and were likewise eliminated. Table 3 in the Appendix provides all state’s intermediate calculations (Coverage PE, Coverage PI, Appendectomy Coverage Rate) and final Informed Coverage value.

[Response Ends]

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

[Response Begins]

N/A

[Response Ends]

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

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

[Response Begins]

Disparities in continuity of coverage according to ethnicity, geography, insurance type, and special health care need have been observed throughout the literature. Publicly insured children from poorer households are more likely than those from higher income households to have gaps in insurance coverage (Bethell 2011; Angier 2013). Minority children, especially Hispanic children, are more likely to be uninsured, have gaps in coverage, and not have an usual source of care (Federico 2007; Flores 2008; Kogan 2010; Flores 2013; Berdahl 2013). This ethnic disparity is worse for first- and second-generation Latino children (DeCamp 2012). Minorities in Georgia were also found to have lower access to higher quality of healthcare (Ogbuanu 2012a). Rural children are more likely to have longer periods of time without insurance than children in urban settings (Coburn 2002). There is a larger gap between insured and uninsured for children in urban settings than in rural settings, and children in urban settings are less likely to have a usual source of care regardless of insurance status (Ziller 2012). Children with special health care needs are more likely to have public insurance coverage than private insurance coverage but experience more unmet needs (Bethell 2011; Callahan 2007; Okumura 2007). Olson, Tang, and Newacheck (2005) performed a cross-sectional study using National Health Interview Surveys confirming that children with full-year public insurance coverage report a higher prevalence of chronic conditions limiting activities relative to children with full year private insurance coverage (12.3% vs 5.1%). In specific patient populations, such as those with diabetes related complaints, the insurance of children with diabetes was Medicaid rather than private (Park 2012). Children with public insurance had a longer interval between epilepsy seizure onset and referral and subsequent surgery compared to privately insured children (Hauptman 2013).

Race/ethnicity (Appendix Table 4): For these analyses, race and ethnicity was determined based on the race-ethnicity variable reported in the MAX data and classified based on Office of Management and Budget guidelines. White was defined as White, not of Hispanic origin. Black was defined as Black, not of Hispanic origin. For Hispanic, we combined children reported as ‘Hispanic or Latino’ and ‘Hispanic or Latino and one or more races’. Other included American Indian, Alaskan Native, Asian, Pacific Islander and children with missing race/ethnicity. We stratified the 18-month informed coverage fraction by enrollee race/ethnicity. Coverage fractions varied by race within and across states with variations in the race/ethnicity groups with the highest and lowest Informed Coverage values within the state.

Special Health Care Needs (Appendix Table 5): Based on published peer-reviewed literature, we compiled a list of pediatric chronic conditions where each condition was represented in all or most of the papers (Valentine, 2000; Ireys, 1997; Todd, 2006; Fowler, 2001; Neuzil, 2000; Feudtner, 2000; Feudtner, 2001; Seferian, 2006). We stratified the 18-month informed coverage fraction by enrollee chronic condition status. Informed Coverage value were generally higher for children with chronic care needs.

Socioeconomic Status (Appendix Table 6-8): Socioeconomic measures at the individual or census-tract level are not included in the MAX data. Although five digit-zip code-based socioeconomic measures have significant limitations, we performed analyses using three socioeconomic variables (% with high school degree, % with income below federal poverty level, and income level) stratified by quartiles in order to demonstrate that these analyses are feasible (Krieger, 1997). These variables were abstracted from U.S. census 5-digit-zip code-level data and merged with the data. If 9-digit-zip code data were available in the MAX data, these analyses would produce more robust

and meaningful results. As noted in the methods, these analyses were performed for the purposes of demonstrating feasibility and not for the purposes of assessing the significance of associations. Informed Coverage did not vary by much between the highest and lowest quartiles for any measure. The coverage fraction across the poverty quartiles lacked a coherent pattern, although the extremes show that ZIP codes with a lower percentage of enrollees below the FPL had better coverage than ZIP codes with a higher percentage above the FPL. Differences across income quartiles were also small and lacked a coherent pattern across states. When looking at just the extremes the lowest income quartile always had a better coverage than those in the highest quartile. Trends for education were clear: while coverage fractions were generally homogenous and never differed by more than ten percentage points between the most- and least-education quartiles, coverage fractions in every state improved as the high school graduation rate of enrollee ZIP codes fell.

Rurality/Urbanity (Appendix Table 9): A crosswalk was performed between the MAX data using the 2010 Census urban and rural classification (<http://www.census.gov/geo/www/ua/2010urbanruralclass.html>). There are two types of urban areas: urbanized areas have 50,000 or more people residing in that area; urban clusters have at least 2,500 and less than 50,000 people residing in that area. Rural area encompasses all population, housing, and territory not included within an urban area. The Informed Coverage values were drastically lower for enrollees who lacked a geographic status (zip code information missing), compared to any other category. Informed Coverage values were generally similar between urban, rural, and urban cluster areas in each state.

[Response Ends]

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

[Response Begins]

N/A

[Response Ends]

2. 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.

spma.01. Indicate whether there are changes to the specifications since the last updates/submission. If yes, update the specifications in the Measure Specifications section of the Measure Submission Form, and explain your reasoning for the changes below.

[Response Begins]

[Response Ends]

spma.02. Briefly describe any important changes to the measure specifications since the last measure update and provide a rationale.

For annual updates, please explain how the change in specifications affects the measure results. If a material change in specification is identified, data from re-testing of the measure with the new specifications is required for early maintenance review.

For example, specifications may have been updated based on suggestions from a previous NQF CDP review.

[Response Begins]

N/A

[Response Ends]

sp.01. Provide the measure title.

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

[Response Begins]

Informed Participation

[Response Ends]

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

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

[Response Begins]

Improved measurement of the continuity of insurance coverage in the Medicaid and CHIP population is needed to help maximize insurance continuity and coverage for vulnerable children. To further this goal, the AHRQ-CMS CHIPRA PQMP Center of Excellence at the Children's Hospital of Philadelphia developed the metric Informed Coverage. The metric is designed to more accurately measure coverage among children enrolled in Medicaid or CHIP at the state level and overcome the current inability in the Medicaid Analytic eXtract (MAX) dataset to determine whether a child disenrolled from Medicaid and CHIP due to loss of eligibility (such as due to parental income increase or the acquisition of employer-sponsored insurance, a "good" reason) or failure to appropriately re-enroll (a "bad" reason). This measure can help federal and state programs develop strategies to retain children eligible for coverage and minimize gaps that can occur during the renewal process. Informed Coverage assesses

the continuity of enrollment of children in publicly financed insurance programs (Medicaid and CHIP), as defined by the ratio of enrolled month to eligible months over an 18 month observation window. Informed Coverage uses a natural experiment based on the random event of appendicitis to “inform” the estimate of coverage in a given state, bounded by two extreme assumptions regarding unknown eligibility information: Coverage Presumed Eligible (PE) and Coverage Presumed Ineligible (PI).

[Response Ends]

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

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

Please do not select:

- *Surgery: General*

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

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

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]

[Response Ends]

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

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

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

Please do not select:

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

[Response Begins]

Population: Regional and State

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

<http://cor.research.chop.edu/node/26>

[Response Ends]

sp.11. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

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

[Response Begins]

No data dictionary/code table – all information provided in the submission form

[Response Ends]

For the question below: state the outcome being measured. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.12. State the numerator.

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

DO NOT include the rationale for the measure.

[Response Begins]

The numerator for Informed Coverage represents the sum (within a state) of months enrolled in Medicaid/CHIP for all children over an 18-month window.

[Response Ends]

For the question below: describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.13. Provide details needed to calculate the numerator.

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

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

[Response Begins]

The numerator is the summation (within a state) of months enrolled in Medicaid/CHIP for all children (0-18 years) over an 18-month window. A month is considered “covered” if a child has greater than 14 enrolled days in that month or if there is an indicator for S-CHIP coverage for that month. Figures 1 and 2 in the Appendix provide an illustration of Coverage PE and Coverage PI.

To determine what is the best assumption to use (either the Appendectomy Coverage Rate (or ACR), PI, or PE) inside each state, we compare the observed appendectomy coverage rate in a state, to the estimated coverage rate that would be calculated in that state with either PI, or PE assumptions. If $PE < ACR < PI$, we utilize ACR. If $ACR > PI$, we use PI, and if $ACR < PE$ we use PE.

The ACR reflects a natural experiment since appendicitis is a random event, not dependent on healthcare of SES status. Appendicitis is defined using principal diagnosis (ICD-9 CM codes 540-541 Appendicitis; ICD-10 codes K35.2, K35.3, K35.80, K35.89, K37) or procedure (ICD-9 CM 47.0-47.09, 47.2 Appendectomy; ICD-10 codes ODTJ4ZZ, ODTJ0ZZ, ODTJ7ZZ, ODTJ8ZZ, OD9J00Z, OD9J0ZZ, OD9J30Z, OD9J3ZZ, OD9J40Z, OD9J4ZZ, OD9J70Z, OD9J7ZZ, OD9J80Z, OD9J8ZZ). This condition is utilized as it (1) has an acute onset (reflecting a discrete point in time); (2) has an incidence rate that is not influenced by prior care, insurance coverage, or by factors that may influence obtaining coverage, such as socioeconomic status; and, (3) would require hospitalization for all children regardless of insurance status. If a child is hospitalized and generates a bill seen in the Medicaid claims, they must have been eligible for Medicaid. If a child was not enrolled at the time of developing appendicitis, but was eligible, the appendicitis should still be observed because Medicaid and most CHIP programs allow up to three months of retroactive coverage and most states have policies of presumptive eligibility for their public insurance program. By identifying appendicitis hospitalizations and determining whether these children were enrolled prior to their hospitalization, we can utilize the rate of existing enrollment at the specific time point of the event to estimate the participation rate for the state population (number enrolled over number eligible at a given point in time). We determine if a child was enrolled prior to hospitalization using a look-back to their state of enrollment 4 months prior to hospitalization. The numerator for the appendicitis calculation is the number of children with an appendicitis hospitalization during the same 18-month observation window used for the Coverage PE and Coverage PI intermediate calculations, who are enrolled in Medicaid/CHIP four months prior to their inpatient stay.

[Response Ends]

For the question below: state the target population for the outcome. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.14. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

The sum (within a state) of months eligible for Medicaid/CHIP for all children (0-18 years) over an 18-month window. In addition, months that could be defined as “eligible” are based on known events recorded in the MAX data that would affect eligibility (birth or ageing out).

[Response Ends]

For the question below: describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.15. Provide details needed to calculate the denominator.

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

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

[Response Begins]

For the intermediate calculations of “Coverage Presumed Eligible (PE)” and “Coverage Presumed Ineligible (PI)”, the denominator is the summation (within a state) of the months a child is eligible for Medicaid/CHIP over an 18-month observation window. The assumptions used to define a child as “eligible” for Medicaid/CHIP coverage for a given month is specific to which intermediate computation is being calculated. When calculating the intermediate computation of “Coverage Presumed Eligible (PE)”, a child is defined as being eligible based on an 18-month observation, in combination with an 18-month look-back period. If any enrollment is observed in the 18-month look-back period, the child is defined as eligible for the entire 18-month observation window. If there is no evidence of enrollment in the 18-month look-back period, eligibility is defined from the first point of enrollment in the observation window. When calculating the intermediate computation of “Coverage Presumed Ineligible (PI)”, a child is defined as being eligible solely on the 18-month observation window. For Coverage PI, eligibility starts from the first enrolled month during the 18-month observation window.

Again using the point-in-time analysis of appendicitis to calculate the observed participation rate, the denominator for the appendicitis calculation, is the number of children with an appendicitis hospitalization during the same 18-month observation window used for the Coverage PE and Coverage PI intermediate calculations. Appendicitis is defined using principal diagnosis (ICD-9 CM codes 540-541 Appendicitis; ICD-10 codes K35.2, K35.3, K35.80, K35.89, K37) or procedure (ICD-9 CM 47.0-47.09, 47.2 Appendectomy; ICD-10 codes 0DTJ4ZZ, 0DTJ0ZZ, 0DTJ7ZZ, 0DTJ8ZZ, 0D9J00Z, 0D9J0ZZ, 0D9J30Z, 0D9J3ZZ, 0D9J40Z, 0D9J4ZZ, 0D9J70Z, 0D9J7ZZ, 0D9J80Z, 0D9J8ZZ).

Appendicitis was chosen because the aim was to create a population where both enrolled and unenrolled eligible children are identifiable in MAX, we sought a condition that: (1) has an acute onset (reflecting a discrete point in time); (2) has an incidence rate that is not influenced by prior care, insurance coverage, or by factors that may influence obtaining coverage, such as socioeconomic status; and, (3) would require hospitalization for all children, regardless of insurance status. Appendicitis meets these three criteria. Appendicitis has an acute onset which occurs at random and is not influenced by previous care or insurance status; it is not influenced by child or parental characteristics or actions that affect likelihood of coverage; and if children develop appendicitis, they will be hospitalized. If a child is hospitalized and generates a bill seen in the Medicaid claims, they must have been eligible for Medicaid. If a child was not enrolled at the time of developing appendicitis, but was eligible, the appendicitis should still be observed because Medicaid and most CHIP programs allow up to three months of retroactive coverage and most states have policies of presumptive eligibility for their public insurance programs.

[Response Ends]

sp.16. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

For the appendicitis calculation, the population is limited to children between the ages of 2 to 16 years old. To determine what is the best assumption to use (either the Appendectomy Coverage Rate (or ACR), PI, or PE) inside each state, we compare the observed appendectomy coverage rate in a state, to the estimated coverage rate that would be calculated in that state with either PI, or PE assumptions.

[Response Ends]

sp.17. Provide details needed to calculate the denominator exclusions.

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

[Response Begins]

For children who are born within the 18-month window of observation, the total months of eligibility begins from date of birth. Finally, for children who reach the age of 18 before the end of the 18-month window of observation, the total month of eligibility ends with their 18th birthday.

[Response Ends]

sp.18. Provide all information required to stratify the measure results, if necessary.

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

[Response Begins]

Stratification can be performed for Informed Coverage using any desired strata that policymakers choose to study. For example, stratification can be performed within states based on the type of Medicaid and CHIP programs, or by race.

[Response Ends]

sp.19. Select the risk adjustment type.

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

[Response Begins]

No risk adjustment or risk stratification

[Response Ends]

sp.20. Select the most relevant type of score.

Attachment: If available, please provide a sample report.

[Response Begins]

[Response Ends]

sp.21. Select the appropriate interpretation of the measure score.

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

[Response Begins]

Better quality = Higher score

[Response Ends]

sp.22. Diagram or describe the calculation of the measure score as an ordered sequence of steps.

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

[Response Begins]

The following describes the steps for calculating the intermediate computations and their use for the final determination. A minimum of three continuous years of MAX claims data are required. The first 18 months are used for a lookback and the second 18 months are the observation period. The same 18-month observation window is used for all calculations. All calculations are done within a state.

Determine the appendectomy participation rate (APR) Intermediate Calculation:

The prior participation of eligible patients developing appendicitis 4 months prior to developing appendicitis
Step 1- Calculate the denominator for appendectomy participation rate: 1) Identify all children between the ages 2 and 16 at the start of the 18-month observation window; 2) Identify the number of children with an inpatient admission for either a principal diagnosis of appendicitis (ICD-9 CM codes 540-541; ICD-10 codes K35.2, K35.3, K35.80, K35.89, K37) or a principal procedure of appendectomy (ICD-9 CM codes 47.0-47.09, 47.2; ICD-10 codes ODTJ4ZZ, ODTJ0ZZ, ODTJ7ZZ, ODTJ8ZZ, OD9J00Z, OD9J0ZZ, OD9J30Z, OD9J3ZZ, OD9J40Z, OD9J4ZZ, OD9J70Z, OD9J7ZZ, OD9J80Z, OD9J8ZZ). Step 2- calculate the numerator for appendectomy coverage rate: 1) Identify the total number of children with pre-existing enrollment in Medicaid or CHIP. Pre-existing enrollment is defined as an observed enrollment exactly four months prior to their date of admission. Step 3- Calculate the appendectomy participation rate: compute the percentage of children admitted for appendicitis/appendectomy with pre-existing enrollment in Medicaid or CHIP, defined by enrollment 4 months prior to the admission.

Determination of the Appendectomy Never Participated Rate (ANPR) Intermediate Calculation: The fraction of eligible appendectomy patients who did not have any participation noted at any point 4 or more months prior to developing appendicitis (within the limits of the observation and lookback period data).

Coverage PE Intermediate Calculation:

Step 4- To determine the denominator for Coverage PE (total months of eligibility using the PE approach): 1) identify all children enrolled in Medicaid/CHIP at any point within the 18-month window of observation AND/OR the 18-month look back, excluding those older than 18 at the beginning of the 18-month observation window; 2) Identify all children who are born within the 18-month window of observation – for these children, total months of eligibility begin from date of birth; 3) Identify all children who reach the age of 18 before the end of the 18-month window of observation – for these children, total months of eligibility end with their 18th birthday; 4) Identify all children who DO NOT APPEAR as covered at any point within the 18-month look back period (“covered” defined as at least one day of coverage) – for these children, total months of eligibility begin with their first day of coverage within the 18-month observation window; 5) For all other children who do not represent populations in Steps 1, 2, or 3, total months of eligibility equals all 18 months in the observation window; and 6) The Coverage PE denominator is the summation of total number of eligible months for all children in the eligible population. Step 5- to determine the numerator for Coverage PE (total months of coverage using PE approach): 1) Identify total number of months in the 18 month observation window covered by MAX/CHIP for each child in the eligible

population. A month is considered “covered” if the child has greater than 14 days of enrollment in that month or if there is an indicator for S-CHIP coverage for that month; and 2) The Coverage PE numerator is the summation of total months covered within the 18-month observation window for all children in the eligible population. Step 6- Calculate the Coverage PE intermediate value: compute the percentage of months covered within the 18-month observation window (Coverage PE numerator divided by Coverage PE denominator).

PE adjustment for patients never enrolled (PE’): See appendix for derivation (Figure 3). $PE' = PE * (1 - ANPR)$.

Coverage PI Intermediate Calculation:

Step 7- To determine the denominator for Coverage PI (the total months of eligibility using the PI approach): 1) identify all children enrolled in Medicaid/CHIP at any point within the 18-month window of observation, excluding those children older than 18 at the beginning of the 18-month observation window; 2) Identify all children who are born within the 18-month window of observation – for these children, total months of eligibility begin from date of birth; 3) Identify all children who reach the age of 18 before the end of the 18-month window of observation – for these children, total months of eligibility ends with their 18th birthday; 4) For all other children who do not represent populations in Steps 1, 2, or 3, months of eligibility begins with the first observed enrollment in the observation window and continues for the remainder of the observation window; and 5) The Coverage PI denominator is the summation of the total number of eligible months for all children in the eligible population. Step 8- to determine the numerator for Coverage PI (total months of coverage using PI approach): 1) Identify the total number of months in the 18-month observation window covered by MAX/CHIP for each child in the eligible population. A month is considered “covered” if the child has greater than 14 days of enrollment in that month or if there is an indicator for S-CHIP coverage for that month; and 2) The Coverage PI numerator is the summation of the total months covered within the 18-month observation window for all children in the eligible population. Step 9- Calculate the Coverage PI intermediate value: compute the percentage of months covered within the 18-month observation window (Coverage PI numerator divided by Coverage PI denominator).

Informed Coverage:

Step 10- The Informed Coverage is the weighted mean of the state Coverage PE’ and state Coverage PI values, where the weights are determined by the state appendectomy participation rate. The closer the appendectomy rate is to Coverage PE, the more weight that Coverage PE receives in the informed coverage measure, and the closer the appendectomy rate is to Coverage PI, the more weight that Coverage PI receives in the informed coverage. An illustration of the formula for this calculation is provided in Figure 4 of the Appendix.

[Response Ends]

sp.25. If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

[Response Begins]

N/A

[Response Ends]

sp.28. Select only the data sources for which the measure is specified.

[Response Begins]

Claims

[Response Ends]

sp.29. Identify the specific data source or data collection instrument.

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

[Response Begins]

The Medicaid Analytic eXtract (MAX) claims data are used for this metric.

[Response Ends]

sp.30. Provide the data collection instrument.

[Response Begins]

No data collection instrument provided

[Response Ends]

2ma.01. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing).

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

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

[Response Ends]

2ma.02. Indicate whether additional empirical validity testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing. Include information on all testing conducted (prior testing as well as any new testing).

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

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

[Response Ends]

2ma.03. For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?

[Response Begins]

[Response Ends]

2ma.04. For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.

Please update the Scientific Acceptability: Validity - Other Threats to Validity section.

Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.

[Response Begins]

[Response Ends]

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

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

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

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

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

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

AND

If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the

information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

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

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

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

OR

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

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

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

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

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

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

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

Definitions

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

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

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

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

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

With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of

one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of \$25 in cost for an episode of care (e.g., \$5,000 v. \$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

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

2021 Submission:

Updated testing information here.

2018 Submission:

Testing from the previous submission here.

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

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

Please do not select:

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.07 check patient or encounter-level data; in 2a.08 enter “see validity testing section of data elements”; and enter “N/A” for 2a.09 and 2a.10.

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

Choose one or both levels.

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

Examples may include correlations or t-test results.

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

Examples may include correlation, and/or rank order.

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

2b.27. Provide risk model discrimination statistics.

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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.

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

[Response Begins]

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

[Response Ends]

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

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

[Response Begins]

ALL data elements are in defined fields in a combination of electronic sources

[Response Ends]

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

[Response Begins]

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

Attach the fee schedule here, if applicable.

[Response Begins]

N/A

[Response Ends]

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.

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.

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 demonstrating performance improvement.

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

Name of program and sponsor

URL

Purpose

Geographic area and number and percentage of accountable entities and patients included

Level of measurement and setting

[Response Begins]

Public Reporting

Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

Quality Improvement (Internal to the specific organization)

Not in use

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Public reporting

Quality Improvement (internal to the specific organization)

[Response Ends]

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

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

[Response Begins]

The measure is new and has not had widespread or targeted distribution to audiences of interest yet.

[Response Ends]

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

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

[Response Begins]

Within the calendar year of any endorsement, information on this metric will be provided to the National Association of Medicaid Directors (NAMD). The aim will be for the Measure Developer to give a webinar to the NAMD, as past webinars have included similar topics such as data analytics to support Medicaid reform. In addition, overviews will be provided to each state's Medicaid department via electronic mail.

[Response Ends]

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

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

[Response Begins]

At this time, performance results, data, and assistance with interpretation have not been provided to those being measured or other users during implementation. Medicaid/CHIP Programs of New Jersey, Pennsylvania, and Massachusetts were all integral to concept development of the Informed Coverage metric. These entities were part of a stakeholder group that provided feedback to the Children's Hospital of Philadelphia as part of their work as the AHRQ PQMP CoE. This included reviewing the metric design and specifications.

[Response Ends]

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

[Response Begins]

Over the course of a year, approximately 10 conference calls were held with Medicaid/CHIP programs to present the work on PQMP measures, including work related to Continuity of Insurance, and to incorporate stakeholder input into measure development. Their input was specifically sought for their insight into feasibility and usability. In addition, the following questions drove discussions: 1) How would a coverage metric complement the duration metric in measuring state or plan performance?; 2) What aspects or types of care would suffer with good coverage but poor duration and vice versa?; 3) Given the resources to reform the Medicaid/CHIP system, what changes would you implement to improve coverage? Duration?.

[Response Ends]

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

[Response Begins]

Feedback was obtained via conference calls/presentations with Medicaid/CHIP programs.

[Response Ends]

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

[Response Begins]

The NJ Stakeholder Representative (May 14, 2012 discussion) believed that a system which is insensitive to gaps should not be used at all. They also believed that only the proposed metric should be used. The MA Stakeholder Representative (April 23, 2012 discussion) was in favor of the use of an 18-month observation window and recognized that a state which focused on newly enrolling and re-enrolling disenrollees would not be rewarded or have its success tracked under the current CMS metric.

[Response Ends]

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

[Response Begins]

N/A

[Response Ends]

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

[Response Begins]

Medicaid/CHIP programs provided input into measure development which has been included in development and revisions to create the current version of the measure.

[Response Ends]

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

[Response Begins]

Our aim is that by using Informed Coverage, states can better estimate participation in their Medicaid and CHIP populations and in turn improve access and quality of care for their children.

[Response Ends]

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

[Response Begins]

No unintended consequences have been identified.

[Response Ends]

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

[Response Begins]

This measure has not yet been implemented.

[Response Ends]

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.

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02, if the measures are NQF endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

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

(Can search and select measures.)

[Response Begins]

[Response Ends]

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

(Can search and select measures.)

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

Provide analyses when possible.

[Response Begins]

N/A

[Response Ends]

Appendix

Supplemental materials may be provided in an appendix.:

Contact Information

Measure Steward (Intellectual Property Owner):

Measure Steward Point of Contact:

Measure Developer if different from Measure Steward:

Measure Developer Point(s) of Contact:

Additional Information

1. Provide any supplemental materials, if needed, as an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be collated one file with a table of contents or bookmarks. If material pertains to a specific criterion, that should be indicated.

[Response Begins]

[Response Ends]

2. List the workgroup/panel members' names and organizations.

Describe the members' role in measure development.

[Response Begins]

The following people were part of the Children's Hospital of Philadelphia PQMP Center of Excellence and contributed in conceptualizing the measure: Scott A. Lorch, MD, MSCE, Rose E. Calixte, PhD, Ashley E. Zeigler, BA, Jeanhee Moon, PhD, Christopher B. Forrest, MD, PhD, Susmita Pati, MD, MPH, and Shawna R. Calhoun, MPH. In addition, Russell Localio, PhD, Wei Wang, PhD, Justin Ludwig, MA, and Joseph G. Reiter, MS contributed to the conceptualization of the measure, as well as the statistical design.

Scott A. Lorch, MD, MSCE is associated with Center for Outcomes Research, The Children's Hospital of Philadelphia, Philadelphia, PA; The Departments of Pediatrics, The University of Pennsylvania School of Medicine, Philadelphia, PA; The Leonard Davis Institute of Health Economics, The University of Pennsylvania, Philadelphia, PA; and Center for Perinatal and Pediatric Health Disparities Research, The Children's Hospital of Philadelphia, Philadelphia, PA.

Rose E. Calixte, PhD and Susmita Pati, MD, MPH are associated with Division of Primary Care Pediatrics, Stony Brook University School of Medicine & Stony Brook Children's Hospital, Stony Brook, NY.

Ashley E. Zeigler, BA, Shawna R. Calhoun, MPH, Wei Wang, PhD, Justin Ludwig, MA, and Joseph G. Reiter, MS are associated with Center for Outcomes Research, The Children's Hospital of Philadelphia, Philadelphia, PA.

Jeanhee Moon, PhD is associated with Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA.

Christopher B. Forrest, MD, PhD is associated with The Departments of Pediatrics, The University of Pennsylvania School of Medicine, Philadelphia, PA; The Leonard Davis Institute of Health Economics, The University of Pennsylvania, Philadelphia, PA; and Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA.

Russell Localio, PhD is associated with Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA; and Department of Biostatistics and Epidemiology, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA.

[Response Ends]

3. Indicate the year the measure was first released.

[Response Begins]

[Response Ends]

4. Indicate the month and year of the most recent revision.

[Response Begins]

[Response Ends]

5. Indicate the frequency of review, or an update schedule, for this measure.

[Response Begins]

[Response Ends]

6. Indicate the next scheduled update or review of this measure.

[Response Begins]

[Response Ends]

7. Provide a copyright statement, if applicable. Otherwise, indicate "N/A".

[Response Begins]

[Response Ends]

8. State any disclaimers, if applicable. Otherwise, indicate "N/A".

[Response Begins]

[Response Ends]

9. Provide any additional information or comments, if applicable. Otherwise, indicate "N/A".

[Response Begins]

[Response Ends]