NATIONAL QUALITY FORUM

Memo

June 4, 2018

To: Consensus Standards Approval Committee (CSAC)

From: Andrew Anderson, Tara Murphy, Yetunde Ogungbemi, Robyn Nishimi

Re: Prevention and Population Health, Fall 2017 Cycle

CSAC Action Required

The CSAC will review recommendations from the Prevention and Population Health Standing Committee at its June 4-5, 2018 meeting and vote on whether to uphold the recommendations from the Committee.

This memo includes a summary of the project, recommended measures, and themes identified and responses to the public and member comments. The following documents accompany this memo:

1. **Prevention and Population Health, Fall 2017 draft report.** The draft report has been updated to reflect the changes made following the Standing Committee's discussion of public and member comments. The complete draft report and supplemental materials are available on the project webpage.
2. Comment Table. Staff has identified themes within the comments received. This table lists 12 comments received during the post-meeting comment period and the NQF/Standing Committee responses.

Background

Performance measurement is necessary to assess whether healthcare stakeholders effectively use strategies to increase prevention and improve population health. Strengthening measurement of prevention and population health will require joint efforts from communities, public health entities, healthcare providers, and other nonhealthcare stakeholders that influence health outcomes. Growing evidence shows that targeted programs and policies can prevent disease, increase productivity, and yield billions of dollars in savings for the U.S. healthcare system. The United States can reduce the incidence of morbidity and premature mortality by identifying the right measures and implementing evidence-based interventions.

This project seeks to identify and endorse measures that can be used to assess prevention and population health in both healthcare and community settings. It also focuses on the assessment of disparities in health outcomes. NQF's prevention and population health portfolio includes measures that assess the promotion of healthy behaviors, community-level indicators of health, oral health, and primary prevention strategies. In this cycle, NQF reviewed two screening measures and five measures related to pediatric dentistry for maintenance of endorsement.

Draft Report

The Prevention and Population Health Fall 2017 draft report presents the results of the evaluation of seven measures considered under the Consensus Development Process (CDP). Five are recommended for endorsement, and two are not recommended.

The measures were evaluated against the 2017 version of the measure evaluation criteria.

	Maintenance	New	Total
Measures under consideration	7	–	7
Measures recommended for endorsement	5	–	5
Measures not recommended for endorsement	2	–	2
Measures withdrawn from consideration	1	–	1
Reasons for not recommending	Importance – 0 Scientific Acceptability – 2 Overall – 0 Competing Measure – 0	Importance – N/A Scientific Acceptability – N/A Overall – N/A Competing Measure – N/A	

CSAC Action Required

Pursuant to the CDP, the CSAC is asked to consider endorsement of five candidate consensus measures.

Measures Recommended for Endorsement

- 0024 Weight Assessment and Counseling for Nutrition and Physical Activity for Children/Adolescents (WCC) (National Committee for Quality Assurance)

  Overall Suitability for Endorsement: Yes-10; No-5

- 0034 Colorectal Cancer Screening (National Committee for Quality Assurance)

  Overall Suitability for Endorsement: Yes-15; No-0

- 2511 Utilization of Services, Dental Services (American Dental Association on behalf of the Dental Quality Alliance)

  Overall Suitability for Endorsement: Yes-13; No-0

- 2517 Oral Evaluation, Dental Services (American Dental Association on behalf of the Dental Quality Alliance)

  Overall Suitability for Endorsement: Yes-13; No-0

- 2528 Prevention: Topical Fluoride for Children at Elevated Caries Risk, Dental Services (American Dental Association on behalf of the Dental Quality Alliance)

  Overall Suitability for Endorsement: Yes-14; No-0

Measures Not Recommended

(See Appendix B for the Committee's votes and rationale.)

- 2508 Prevention: Sealants for 6-9 Year-Old Children at Elevated Caries Risk (American Dental Association on behalf of the Dental Quality Alliance)
- 2509 Prevention: Sealants for 10-14 Year-Old Children at Elevated Caries Risk (American Dental Association on behalf of the Dental Quality Alliance)

Comments and Their Disposition

NQF received 15 comments from seven organizations (including four member organizations) and individuals pertaining to the draft report and to the measures under consideration.

A table of comments submitted during the comment period, with the responses to each comment and the actions taken by the Standing Committee and measure developers, is posted to the Prevention and Population Health project webpage.

Comments Themes and Committee Responses

Comments about specific measure specifications and rationale were forwarded to the developers, who were invited to respond.

The Standing Committee reviewed all of the submitted comments (general and measure-specific) and developer responses. Committee members focused their discussion on measures or topic areas with the most significant and recurring issues.

Measure-Specific Comments: Recommended Measure

0024 Weight Assessment and Counseling for Nutrition and Physical Activity for Children/Adolescents (WCC) (NCQA)

NQF received two comments on this measure. One comment expressed support for the measure as specified, while the other expressed concern about the lack of reliability and validity testing for the measure when collected through abstraction from the medical record. The commenter also noted that the submission does not include testing results for the testing of exclusions.

Measure Steward/Developer Response:

The reliability and validity testing information represents results for both administrative claims and medical record review (i.e., the measure as specified). The exclusions are identifiable in claims or by medical record abstraction; all HEDIS measures are audited in order to ensure members of the eligible population who are excluded are done so appropriately.

Measure-Specific Comments: Measures Not Recommended

2508 Prevention: Sealants for 6-9 Year-Old Children at Elevated Caries Risk, Dental Services (ADA for DQA)

NQF received five post-evaluation comments on this measure. All five comments raised concern over the measure's specifications, specifically the lack of exclusions for individuals with zero sealable molars. One comment disagreed with the measure's inclusion of individuals with "elevated" risk in the denominator, noting there is evidence that current tools to assess caries risk are not reliable. One comment also requested the creation of implementation guidelines.

Measure Steward/Developer Response:

The developer addressed many of the commenters' concerns regarding clinical exclusions in its memo. Further, the developer responded to comments individually, noting that the current state of science on caries risk assessment and developed guidance on risk categorization found that current caries risk assessment tools share many common elements to assess risk and affirmed that they have dichotomous predictive ability to quantify "low risk" and "elevated risk." Consequently, the developer continues to support the focus of the measure on the priority population of children at elevated risk for developing dental caries.

2509 Prevention: Sealants for 10-14 Year-Old Children at Elevated Caries Risk, Dental Services (ADA for DQA)

NQF received two post-evaluation comments on this measure. One comment agreed with the Committee's concern about the need for an exclusion of patients with previously sealed molars. Additionally, however, the commenter noted support for the measure's endorsement. The second comment called for clarification as to whether the measure is specified exclusively for Medicaid plans or if it also applied to commercial plans with a dental benefit. The commenter also agreed with the Committee's request for clarification on how individuals are classified as "high" or "moderate" risk.

Measure Steward/Developer Response:

The developer responded that this measure is calculated using administrative enrollment and claims data, and it is specified for reporting at the program (e.g., Medicaid or CHIP) or plan (e.g., MCO or DBA) level. Further, the developer responded to the concerns regarding exclusions in their memo for measure #2509. The developer also provided the rationale for the risk assessment levels, noting that testing data found that significant performance gaps existed within the elevated caries risk populations. During initial measure development, it was recognized that the ability to make reliable distinctions between at-risk levels (e.g., between "moderate" and "high" risk) was not well established. Consequently, the measure adopted a clearer distinction of "low" risk and "elevated" risk. (The measure does not require distinguishing "moderate" risk from "high" risk.)

Member Expression of Support

Throughout the 16-week continuous public commenting period, NQF members had the opportunity to express their support ('support' or 'do not support') for each measure submitted

for endorsement consideration to inform the Committee's recommendations. Two NQF members provided their expression of support. Appendix C details the expression of support.

Removal of NQF Endorsement

One measure previously endorsed by NQF has not been re-submitted, and endorsement has been removed.

Measure	Measure Description	Reason for Removal of Endorsement
2020 Adult Current Smoking Prevalence	Percentage of adult (age 18 and older) U.S. population that currently smoke.	Developer is no longer able to support the measure.

Appendix A: CSAC Checklist

The table below lists the key considerations to inform the CSAC's review of the measures submitted for endorsement consideration.

Key Consideration	Yes/No	Notes
Were there any process concerns raised during the CDP project? If so, briefly explain.	No	
Did the Standing Committee receive requests for reconsideration? If so, briefly explain.	Yes	The Standing Committee received a request to reconsider measure *2508 Prevention: Sealants for 6-9 Year-Old Children at Elevated Caries Risk, Dental Services*. The Committee denied the request because they had concerns that the measure does not assess whether a child who has received the sealants has met the recommended guidelines. Instead, the measure assesses whether or not a child has received the recommended sealants in the measurement year. The Committee also had concerns that the measure does not exclude children with previously sealed molars.
Did the Standing Committee overturn any of the Scientific Methods Panel's ratings of Scientific Acceptability? If so, state the measure and why the measure was overturned.	No	
If a recommended measure is a related and/or competing measure, was a rationale provided for the Standing Committee's recommendation? If not, briefly explain.	Yes	Measure *0034 Colorectal Cancer Screening (COL)* is related to 0658 *Appropriate Follow-Up Interval for Normal Colonoscopy In Average Risk Patients* (Minnesota Community Measurement). Measures 0658 and 0034 have similar numerators but different denominators. The two measures are harmonized. However, measure 0658 was not reviewed during this cycle.
Were any measurement gap areas addressed? If so, identify the areas.	No	

Key Consideration	Yes/No	Notes
Are there additional concerns that require CSAC discussion? If so, briefly explain.	No	

Appendix B: Measures Not Recommended for Endorsement

The table below lists the Committee's vote and rationale for measures not recommended for endorsement.

Legend: H = High; M = Moderate; L = Low; I = Insufficient

Measure	Voting Results	Standing Committee Rationale
2508 Prevention: Dental Sealants for 6-9 Year-Old Children at elevated Caries Risk, Dental Services American Dental Association on behalf of the Dental Quality Alliance	**Evidence** Previous Evidence Evaluation Accepted **Gap** Previous Performance Gap Evaluation Accepted **Reliability** Previous Reliability Evaluation Accepted **Validity** H-0; M-5; L-8; I-0	Committee members had concerns that the measure does not assess whether a child who has received the sealants has met the recommended guidelines. Instead, the measure assesses whether or not a child has received the recommended sealants in the measurement year. The Committee also had concerns that the measure does not exclude children with previously sealed molars. The measure failed on the Validity criterion during the February 9 in-person meeting. The Committee denied a request for reconsideration during the April 30 post-comment call.

Measure	Voting Results	Standing Committee Rationale
2509 Prevention: Dental Sealants for 10-14 Year-Old Children at Elevated Caries Risk, Dental Services American Dental Association on behalf of the Dental Quality Alliance	**Evidence** Previous Evidence Evaluation Accepted **Gap** Previous Performance Gap Evaluation Accepted **Reliability** Previous Reliability Evaluation Accepted **Validity** H-0; M-6; L-7; I-0 **Feasibility** Previous Feasibility Evaluation Accepted **Usability and Use** *Use* Pass-13; No Pass-0 *Usability* H-0; M-11; L-0; I-1 **Post Comment Call Vote: Validity** H-0; M-8; L-6; I-1	The Committee did not reach consensus during the February 9 in-person meeting, and the measure failed on the Validity criterion during the April 30 post-comment call. Similar to measure 2508, Committee members had concerns that the measure does not assess whether a child who has received the sealants has met the recommended guidelines. Instead, the measure assesses whether or not a child has received the recommended sealants in the measurement year. The Committee also had concerns that the measure does not exclude children with previously sealed molars.

Appendix C: NQF Member Expression of Support Results

Two NQF members provided their expressions of support. Four of seven measures under consideration received support from NQF members. Results for each measure are provided below.

0024 Weight Assessment and Counseling for Nutrition and Physical Activity for Children/Adolescents (WCC) (National Committee for Quality Assurance)

Member Council	Support	Do Not Support	Total
Health Professional	1	0	1

2508 Prevention: Sealants for 6-9 Year-Old Children at Elevated Caries Risk, Dental Services (American Dental Association on behalf of the Dental Quality Alliance)

Member Council	Support	Do Not Support	Total
Health Plan	0	1	1

2509 Prevention: Sealants for 10-14 Year-Old Children at Elevated Caries Risk, Dental Services (American Dental Association on behalf of the Dental Quality Alliance)

Member Council	Support	Do Not Support	Total
Health Plan	0	1	1

2528 Prevention: Topical Fluoride for Children at Elevated Caries Risk, Dental Services (American Dental Association on behalf of the Dental Quality Alliance)

Member Council	Support	Do Not Support	Total
Health Plan	0	1	1

Appendix D: Details of Measure Evaluation

Recommended Measures

0024 Weight Assessment and Counseling for Nutrition and Physical Activity for Children Adolescents (WCC)

Submission

Description: Percentage of patients 3-17 years of age who had an outpatient visit with a primary care physician (PCP) or an OB/GYN and who had evidence of the following during the measurement year:

- Body mass index (BMI) percentile documentation
- Counseling for nutrition
- Counseling for physical activity

Numerator Statement: Patients who had evidence of the following during the measurement year: a body mass index (BMI) percentile documentation, counseling for nutrition, counseling for physical activity.

Denominator Statement: Patients 3-17 years of age with at least one outpatient visit with a primary care physician (PCP) or OB-GYN during the measurement year.

Exclusions: The measure excludes female patients who have a diagnosis of pregnancy and patients who use hospice services during the measurement year.

Level of Analysis: Health Plan, Integrated Delivery System

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Claims, Electronic Health Records, Paper Medical Records

Measure Steward: National Committee for Quality Assurance

STANDING COMMITTEE MEETING 02/09/2018

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **H-0; M-9; L-1; I-5**; 1b. Performance Gap: **H-4; M-10; L-0; I-0**

Re-votes from Standing Committee Post-Comment Meeting 04/30/2018

1a. Evidence: **H-0; M-3; L-1; I-11**; Insufficient Evidence with Exception to Evidence: **Y: 13; N: 2**

Rationale:

- This maintenance measure focuses on weight assessment and counseling for nutrition and physical activity for patients ages 3-17 years.
- The developer cites the USPSTF's moderate recommendation for childhood and adolescent body mass index (BMI) assessment to support the measure. The USPSTF recommendation states that there is no direct evidence on the benefits of screening; however, given the base of evidence on lifestyle modifications to reduce excess weight, the USPSTF awarded a moderate rating.

- Some Committee members believed the measure's evidence was mostly based on a "consensus recommendation," which is reasonable for continued endorsement.
- Other Committee members questioned the measure's age range, noting that the USPSTF guidelines address patients ages 6 and older. The developer responded that it used supplemental recommendations from AAP's "Bright Futures" to support the measure's inclusion of patients ages 3 and older.
- The Committee also questioned the accuracy of BMI when applied to young children, specifically that the USPSTF recommendations noted the decreased accuracy of BMI for children under six as rationale for excluding that population. A systematic review and meta-analysis of the diagnostic performance of BMI in children and adolescents (ages 4-18 years) published in 2015 concluded that BMI had high specificity, but low sensitivity, to detect excess fat and fails to identify over a quarter of children with excess body fat percentage.
- The Committee discussed the appropriateness of the nutrition and physical activity counseling included in the measure and stated that the literature suggests that more than counseling is needed. The USPSTF recommendation focuses on screening for BMI and recommends referral to comprehensive, intensive behavioral interventions to promote improvement in weigh status. The developer noted that it used supplemental recommendations from AAP's "Bright Future," which supports the inclusion of counseling. The developer also stated that this measure attempts to balance the feasibility of obtaining intervention information with making sure people get at least a minimum amount of counseling.
- Since most effective BMI interventions occur outside the clinician's office, the Committee agreed it would like to see measures that address the decision makers and institutions that can impact children and adolescent BMI (schools, communities, etc.), not just clinicians.
- The developer indicated it is developing a set of measures focused on health promotion and targeted at organizations surrounding wellness and health promotion. The developer also noted challenges defining and assessing whether referrals to more intensive weight management interventions have occurred.
- In addition to reporting performance on the measure, the developer also provides the plan with information about three separate rates for BMI screening, nutrition counseling, and physical activity counseling.
- There are known disparities in the prevalence of obesity among African American females as identified in the systematic review cited by the developer. The Committee urged the developer to improve collection of race/ethnicity data, which is often reported as unknown or refused, and to require reporting of that data to it. The developer did note that while disparities are not reported in the measure, health plans may analyze the disparities data for their own internal understanding and use.
- The Committee questioned whether it is worthwhile for NQF to endorse measures such as this, which focus only on the basic standard of care, but it ultimately agreed this was outside the scope of this particular measure evaluation.
- This measure is calculated at the plan level. Committee members discussed the appropriate attribution of the measure, noting that while plans can support activities that lead to healthy BMI, ultimately the individual clinicians will be held accountable to the measure whether appropriate or not.

- Ultimately, the Committee did not reach consensus on whether the measure meets the Evidence criterion during the measure evaluation meeting. The Committee re-voted on the Evidence criterion during the post-comment call.
- During the post-comment call, the Committee continued to express concerns about the lack of evidence that supporting the counseling specified in the measure. In response to a query from the lead discussant, NQF staff explained the process by which a measure would be eligible for the Exception to Evidence criterion. The Committee voted >60% Insufficient on the Evidence criterion, which prompted a vote on an Exception to the Evidence, which passed (Y-13; N-2).
- The developer provided performance data from HEDIS for 2014, 2015, and 2016 stratified by age group and plan type. The developer characterizes the change in performance from 2014 to 2016 as "slight" improvement (3%-6%) across commercial and Medicaid plans.
- Average performance for the measure is approximately 55%-60%.
- The Committee agreed the measure meets the Performance Gap criterion that a performance gap still exists.

---

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **H-7**; **M-6**; **L-1**; **I-0**; 2b. Validity: **H-0**; **M-8**; **L-7**; **I-0**

Re-votes from Standing Committee Post-Comment Meeting

2b. Validity: **H-0; M-10; L-3; I-2**

<u>Rationale</u>:

- The beta-binomial method was used to assess the ratio of signal to noise, where the signal is the proportion of variability in measured performance that can be explained by real differences in performance and noise is variability that is attributable to error. A reliability score of 1 implies that all the variability is attributable to real differences in performance and a score of 0 implies all variability is attributable to measurement error. The developer states a reliability score of 0.7 is considered "very good."
- Using 246 Medicaid plans and 406 commercial plans in the 2016 HEDIS dataset, the reliability statistics 0.999 in Commercial plans and 0.993 or above in Medicaid plans.
- The Committee agreed that the measure meets the Reliability criterion.
- To assess the construct validity of this measure, the developer used a Pearson correlation test to assess the correlation between this measure and a similar measure of Adult BMI assessment, hypothesizing that plans with a high performance on weight assessment and counseling for nutrition and physical activity also will have high performance on adult BMI assessment. The magnitude of correlation ranges from -1 to +1. A correlation coefficient of 1 indicates a perfect linear dependence between the measures, where increasing values on one measure are associated with increasing values of the second measure. A value of -1 indicates a perfect linear relationship, where increasing values of the first measure are associated with decreasing values of the second. A score of 0 indicates no relationship between the measures.

- The developer reported that Pearson correlation coefficients for commercial plans show a strong positive correlation with of 0.79 and higher. In Medicaid plans, results were moderate, with correlations of 0.64-0.65.
- Recent changes to the NQF measurement criteria require that maintenance measures present empiric validity testing and that face validity is no longer sufficient for continued endorsement. The Committee discussed the relevance of the construct validity hypothesis, noting the weak justification that comparison to a similar adult measure provides. Rather, the Committee recommended using a related plan-level measure, such as recent diagnoses of obesity, diabetes, or metabolic syndrome in the pediatric population or a similar measure of screening (i.e., lead) in the same population.
- Because this measure uses data from administrative claims, paper records, and electronic health records (EHRs), the Committee questioned the measure's sensitivity to data sources across plans. The developer responded that it audits all data and provides clear rules on how to report different data sources; it also requires a HEDIS vendor certification for data collection to further support data integrity across plans.
- The developer also purports a high degree of face validity and described its detailed method of assessment. However, the developer presented no information on the results of the face validity assessment.
- Ultimately, the Committee did not reach consensus on whether the measure meets the Validity criterion during the measure evaluation meeting. The Committee re-voted on the Evidence criterion during the post-comment call.
- The Committee requested that the developer present the results of the data collection auditing during the re-vote on the post-comment call and for its other measures in the future.
- The Committee also requested that the developer provide further construct validity testing, using a more appropriate measure for comparison during the post-comment call.
- During the post-comment call, The Committee questioned the whether the codes that correspond to counseling in the medical record represents the type of counseling intended by the clinical guideline. The codes included in the measure do not specify the level of counseling that is provided, and so the Committee had concerns over whether the counseling that was provided would be considered a comprehensive, intensive behavioral intervention, as the USPSTF guidelines recommend.
- Ultimately, the Committee agreed that NQF measure 0024 is able to measure simple screening and whether counseling took place with a moderate to high degree of validity.

---

3. Feasibility: H-5; M-10; L-0; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c.Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- This measure is calculated using data from administrative claims, paper medical records, and EHRs. Data for this measure are generation or collected by healthcare personnel

during the provision of care, coded by a separate individual, and abstracted from a record by someone other than the person obtaining the original information.
- The Committee noted that this measure is used in multiple programs and had no concerns regarding the measure's feasibility.

---

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-13; No Pass-1;** 4b. Usability: **H-0; M-14; L-1; I-0**

Rationale:

- This measure is used in a variety of accountability programs, health plans, CMS programs and Medicaid CHIP.
- Per the developer, performance data from 2014-2016 have shown modest improvement indicating that either performance or documentation is improving.
- The Committee also noted the lack of an outcome measure that addresses weight in children and adolescents.
- Following the vote, several Committee members noted their concern that the measure is "letting plans off the hook" and incentivizing interventions that show no evidence to support weight reduction. The Committee strongly recommended that during future maintenance review the measure include a component that assesses the adequacy of the counseling or meets more rigorous USPSTF guidelines.

---

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-10**; **N-5**

6. Public and Member Comment

Two comments were submitted on this measure, one pre-evaluation and one post-evaluation. The pre-evaluation comment expressed support for the measure as specified, while the other expressed concern about the lack of reliability and validity testing for the measure when collected through abstraction from the medical record. The commenter also noted that the submission does not include testing results for the testing of exclusions.

Developer Response:

- The reliability and validity testing information represents results for both administrative claims and medical record review (i.e., the measure as specified). The exclusions are identifiable in claims or by medical record abstraction; all HEDIS measures are audited in order to ensure members of the eligible population who are excluded are done so appropriately.

---

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals

0034 Colorectal Cancer Screening (COL)

Submission

Description: The percentage of patients 50–75 years of age who had appropriate screening for colorectal cancer.

Numerator Statement: Patients who received one or more screenings for colorectal cancer according to clinical guidelines.

Denominator Statement: Patients 51–75 years of age

Exclusions: This measure excludes patients with a history of colorectal cancer or total colectomy. The measure also excludes patients who use hospice services or are enrolled in an institutional special needs plan (SNP) or living long-term in an institution any time during the measurement year.

Level of Analysis: Health Plan, Integrated Delivery System

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Claims, Electronic Health Data, Paper Medical Records

Measure Steward: National Committee for Quality Assurance

STANDING COMMITTEE MEETING 02/09/2018

1. Importance to Measure and Report: The measure meets the Importance to Measure criterion

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **H-10**; **M-5**; **L-0**; **I-0**; 1b. Performance Gap: **H-12**; **M-3**; **L-0**; **I-0**

Rationale:

- This maintenance measure focuses on the percentage of patients ages 50-75 years who received appropriate screening for colorectal cancer. The measure aligns with the A-rated 2016 U.S. Preventive Services Task Force (USPSTF) guideline recommending screening for colorectal cancer starting at age 50 and continuing until age 75.
- The evidence for the measure was updated to reflect the 2016 USPSTF guideline, which added two tests in addition to colonoscopy to satisfy the screening requirement.
- It was noted that according to the USPSTF guideline, clinicians should engage patients to make an informed decision about the type of screening they receive. However, patient education in decision making is not addressed in the measure.
- The Committee noted that the measure does not exclude patients with limited life expectancy comorbidities, which is a critical area for stemming overuse of screening. The measure excludes hospice patients and patients in long-term care or skilled nursing facilities, but the Committee noted this does not account for additional, broader limited-

life expectancy. The Committee emphasized the importance of and opportunity to move toward measures of appropriate care and encouraged the developer to consider such a change, while also acknowledging the practical concerns of defining limited life expectancy and the challenges of having known prognoses for many diagnoses.
- The developer provided performance data from HEDIS for 2014-2016 stratified by plan type. The developer stated that during this period, performance rates for the measure have shown slight decline (-1%) across commercial plans and slight improvement (+2%) in Medicare plans.
- Average performance for the measure is approximately 60% in commercial plans and 67% in Medicare plans, which the developer noted demonstrates a persistent performance gap.
- Some Committee members questioned whether a persistent gap in performance indicates a flaw in the measure itself. The developer responded that it is fair to assess the measure construct itself, but that in this case the clinical guideline supports the measure as constructed. Others pointed out that the persistent gap could be attributed to a variety of other forces in the healthcare system besides the measure itself and is not problematic if the gap is supported by evidence from additional sources.
- The developer noted that published literature has identified disparities in the rate of colorectal cancer screening based on race, ethnicity, income, education, and English language proficiency.
- The Committee requested that the developer provide a statistical test for the change in performance scores in future submissions.

---

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **H-8**; **M-7**; **L-0**; **I-0**; 2b. Validity: **H-1**; **M-14**; **L-0**; **I-0**

<u>Rationale</u>:

- The beta-binomial method was used to assess the ratio of signal to noise, where the signal is the proportion of variability in measured performance that can be explained by real differences in performance and noise is variability that is attributable to error. A reliability score of 1 implies that all the variability is attributable to real difference in performance and a score of 0 implies all variability is attributable to measurement error. The developer stated a reliability score of 0.7 is considered "very good."
- Using the 2017 HEDIS dataset, the reliability statistics were 0.997 in commercial plans and 0.988 in Medicare plans.
- The Committee agreed that the measure meets the Reliability criterion.
- To assess the construct validity of this measure, the developer used a Pearson correlation test to assess the correlation between this measure and a similar measure of breast cancer screening, hypothesizing that plans with high performance on colorectal cancer screening will also have high performance on breast cancer screening. The magnitude of correlation ranges from -1 to +1. A correlation coefficient of 1 indicates a perfect linear dependence between the measures, where increasing values on one measure are associated with increasing values of the second measure. A value of -1 indicates a perfect linear relationship, where increasing values of the first measure are

associated with decreasing values of the second. A score of 0 indicates no relationship between the measures.
- The developer reported that Pearson correlation coefficients for commercial plans showed a strong positive correlation of 0.711. In Medicaid plans, results were moderate, with correlation of 0.716.
- The developer purported a high degree of face validity and described its detailed method of assessment. However, the developer presented no information on the results of the face validity assessment.
- The Committee indicated some concern about reliability and validity, given the many places from which data are collected in the medical record. It requested that the developer provide a greater articulation of the data auditing process and results in future submissions.
- The Committee agreed the measure meets the Validity criterion

---

3. Feasibility: H-3; M-12; L-1; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c.Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- This measure is calculated using data from administrative claims, paper medical records, and EHRs. Data for this measure are generated or collected by healthcare personnel during the provision of care, coded by a separate individual, and abstracted from a record by someone other than the person obtaining the original information.
- The Committee noted it had similar objections to the collection of data from multiple sources (administrative claims, paper medical records, and EHRs) as raised in the discussion of measure #0024.
- Ultimately, the Committee agreed the measure meets the criterion for Feasibility.

---

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-15; No Pass-0;** 4b. Usability: **H-6; M-9; L-0; I-0**

Rationale:

- This measure is currently in widespread use in multiple accountability programs including the Medicare Advantage Star Rating Program and the Quality Payment Program.
- The Committee noted that given that the measure numerator has changed to include additional methods of screening, measure trends over time should be assessed with caution in recognition of the expanded specifications.
- The Committee agreed that the measure meets the Use criterion.
- The Committee agreed that the measure meets the Usability criterion.

---

5. Related and Competing Measures

- This measure is related to:
    - 0658 Appropriate Follow-Up Interval for Normal Colonoscopy in Average Risk Patients (American Gastroenterological Association)

Standing Committee Recommendation for Endorsement: **Y-15**; **N-0**

6. Public and Member Comment

NQF received three post-evaluation comments from two organizations about this measure. One comment did not support the Committee's recommendation and expressed concern over the lack of exclusions for patients with limited life expectancy who are not in hospice care. The other comment expressed concern about the lack of reliability and validity testing for the measure when collected through abstraction from the medical record. The commenter also noted that the submission does not include testing results for the testing of exclusions.

Developer Response:

- The reliability and validity testing information represents results for both administrative claims and medical record review (i.e., the measure as specified). The exclusions are identifiable in claims or by medical record abstraction; all HEDIS measures are audited in order to ensure members of the eligible population who are excluded are done so appropriately.
- NCQA recognizes the hospice/living long-term in institutional care exclusion is an important step towards ensuring these patients are removed from measures that require services at an intensity and frequency that may be inappropriate. NCQA continues to assess whether there are additional ways to identify members who should be removed from these types of measures.

---

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals

2511 Utilization of Services, Dental Services

Submission

Description: Percentage of enrolled children under age 21 years who received at least one dental service within the reporting year.

Numerator Statement: Unduplicated number of children under age 21 years who received at least one dental service

Denominator Statement: Unduplicated number of enrolled children under age 21 years

Exclusions: Medicaid/CHIP programs should exclude those individuals who do not qualify for dental benefits. The exclusion criteria should be reported along with the number and percentage of members excluded.

Level of Analysis: Health Plan, Integrated Delivery System

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Claims

Measure Steward: American Dental Association on behalf of the Dental Quality Alliance

---

STANDING COMMITTEE MEETING 02/09/2018

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Previous Evidence Evaluation Accepted**; 1b. Performance Gap: **Previous Performance Gap Evaluation Accepted**

Rationale:

- This maintenance measure captures whether children ages 1-21 years have received any dental services during the reporting year, which assesses access to oral care. The developer provided evidence that there are documented disparities in untreated dental caries and receipt of dental services. The American Academy of Pediatric Dentistry recommends that all children have a dental home established by 12 months of age.
- Committee members acknowledged the strong evidence and guidelines regarding preventive dental care/dental homes for children, but expressed concerns that the specific format of the measure does not directly address the evidence and guidelines. Particularly, "any dental service" could include extractions or other dental procedures for caries that would indicate poor or limited preventive care in previous years.
- Ultimately, the Committee chose to accept the vote on Evidence from its previous evaluation of the measure.
- The developer stated that performance range of 28% to 74% in CY 2010 is indicative of a significant performance gap. The developer did not provide more recent performance data, stating that due to the start-up phase for integration of the measures, most of the entities that have adopted the measure are just beginning implementation, so there are limited data.
- The developer stated that its findings demonstrate disparities by age, geographic location, and race/ethnicity. The developer also evaluated whether the measure could detect disparities (by income, children's health status, Medicaid program type, commercial product line, and preferred language for program communications), and noted disparities were detected for each of these factors, but data on all characteristics were not consistently available for all programs.

---

2. Scientific Acceptability of Measure Properties: The measure meets the Scientific Acceptability criteria

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Previous Reliability Evaluation Accepted;** 2b. Validity: **Previous Validity Evaluation Accepted**

Rationale:

- The developer provided the original testing data as previously evaluated during the measures initial review, which is permitted by NQF's maintenance of endorsement policy. Because the testing had not been updated, the Committee chose to accept its votes from the measure's prior evaluation of Reliability and Validity.

---

3. Feasibility: Previous Feasibility Evaluation Accepted

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c.Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- This measure relies on standard data elements in administrative claims data, which are readily available and can be easily retrieved because they are routinely used for billing and reporting purposes. There have been no reports of feasibility issues with implementing this measure, as provided by the developer.
- Because no feasibility issues had arisen since the previous review, the Committee chose to accept the vote on Feasibility from its previous evaluation of the measure.

---

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-12; No Pass-0**; 4b. Usability: **H-1 M-12; L-0; I-0**

Rationale:

- The measure is used for public reporting and accountability by the Texas HHS Commission: Texas Medicaid/CHIP Pay-for-Quality (P4Q) program.
- In 2016, the Dental Quality Alliance (DQA) expanded its scope of review for its measures by convening participants who shared experiences implementing DQA measures, including any challenges related to the measures specifications and use of the measures in their quality improvement programs. Participants did not have any significant issues related to the clarity or feasibility of implementing the measure.
- The developer also noted that data retrieved to date suggest a trend in improvement over time, although these are initial performance data for one program. It stated that most measure users are just now getting their quality measurement programs underway, and therefore do not have substantial data to provide.
- The Committee voted to approve #2511 on both Use and Usability and Use criterion.

---

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-13**; **N-0**

6. Public and Member Comment

NQF received one comment about this measure. The comment requested clarification on whether the measure is specified for Medicaid plans only or if it is also specified for commercial plans with a dental benefit. The commenter also requested clarification on how this measure differs from the existing NCQA HEDIS measure Annual Dental Visit (NQF #1388), a similar measure that is no longer NQF endorsed but is currently in use in various public reporting programs.

Developer Response:

- Thank you for your comment. Measure 2511 is calculated using administrative enrollment and claims data, and it is specified for reporting at the program (e.g., Medicaid or CHIP) or plan (e.g., MCO or DBA) level. Entities that do not have dental administrative data will not be able to report this measure. Reporting on the measure for a unit (e.g., provider level) or using a data source other than that for which the measure was developed may not be reliable. A primary difference between the DQA measure- Utilization of Services and the NCQA HEDIS Annual Dental Visit (ADV) is the denominator. The enrollment requirement for the proposed measure, Utilization of Dental Services, is 6 months continuous enrollment during the reporting year as opposed to a full-year enrollment, allowing for no more than a single 45-day gap in enrollment (or one month for entities that measure enrollment on a monthly basis) for the ADV. The testing of Utilization of Dental Services included five alternative enrollment requirements, including full-year (allowing for a one-month gap) enrollment. Our results indicated a significant decrease in the percentage of members eligible for the measure when increasing the enrollment requirement from 6 months to one year. (Data from Table 2a2.2-1. in the Measure Testing Form.) The percentage of members included with a full-year enrollment requirement ranged from only 1/3 of members up to just shy of 2/3 of members. In contrast, 2/3 – 4/5 of members were retained in the measure with a six-month enrollment requirement, which also was deemed to provide sufficient time to seek, schedule, and obtain a dental visit.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals

2517 Oral Evaluation, Dental Services

Submission

Description: Percentage of enrolled children under age 21 years who received a comprehensive or periodic oral evaluation within the reporting year.

Numerator Statement: Unduplicated number of enrolled children under age 21 years who received a comprehensive or periodic oral evaluation as a dental service

Denominator Statement: Unduplicated number of enrolled children under age 21 years

Exclusions: Medicaid/CHIP programs should exclude those individuals who do not qualify for dental benefits. The exclusion criteria should be reported along with the number and percentage of members excluded

Level of Analysis: Health Plan, Integrated Delivery System

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Claims

Measure Steward: American Dental Association on behalf of the Dental Quality Alliance

---

STANDING COMMITTEE MEETING 02/09/2018

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Previous Evidence Evaluation Accepted**; 1b. Performance Gap: **Previous Performance Gap Evaluation Accepted**

Rationale:

- This maintenance measure focuses on the percentage of enrolled children under age 21 years who received a comprehensive or periodic oral evaluation within the reporting year.
- This developer cited national guidelines from the American Academy of Pediatric Dentistry (AAPD) and the American Academy of Pediatrics (AAP), which recommend that children receive oral health services by 1 year of age and have regular visits thereafter.
- The Committee supported the need for individuals under the age of 21 to have routine dental visits.
- The developer updated the evidence with a recent Cochrane Review, which was directionally the same as the previous evidence.
- Since the developer did not provide new evidence outside of the Cochrane Review, the Committee chose not to vote again and accepted its previous evaluation for the Evidence criterion.
- The developer provided plan and program level performance information from four sources: Texas Medicaid, Florida CHIP, and Florida Medicaid programs, as well as national commercial data from Dental Service of Massachusetts, Inc. The performance range of 26% to 67% in CY 2010 (year in which data were available for all five programs) indicates variation in topical fluoride application across programs. The developer did not provide more current performance data, stating that due to the start-up phase for integration of the measure, most of the entities that have adopted the measure are just implementing it and so there is limited data reporting.
- The Committee accepted the developer's rationale for the lack of updated performance gap information. Because it had not been updated since its 2014 evaluation of the measure, the Committee chose to accept the vote on Performance Gap from its previous evaluation of the measure.

2. Scientific Acceptability of Measure Properties: The measure meets the Scientific Acceptability criteria

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Previous Reliability Evaluation Accepted**; 2b. Validity: **Previous Validity Evaluation Accepted**

Rationale:

- The developer previously provided empirical reliability testing with statistical tests, but did not update the testing, as permitted by NQF's maintenance of endorsement policy.
- NQF permits data-element level validity testing to suffice for reliability testing. This previous testing focused on assessing the accuracy of dental procedure codes reported in claims data as the data elements that contribute most to computing the measure score. To evaluate data element validity, the developer conducted reviews of dental records and compared them to extracted administrative claim data, which matched more than 86% of the time.
- In addition to the data element-level validity, the developer also submitted results from a systematic assessment of the measure's face validity. The assessment used a modified Delphi method, among other activities, to assess the measure's face validity and found unanimous agreement that the calculated measure scores can be used to evaluate quality of care.
- Ultimately, given that no new testing was provided, the Committee opted to accept its previous evaluation and votes to approve the Reliability and Validity criteria.

3. Feasibility: Previous Feasibility Evaluation Accepted

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c.Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- This measure relies on standard data elements in administrative claims data (e.g., patient ID, patient birthdate, enrollment information, CDT codes, date of service, and provider taxonomy). These data are readily available and can be easily retrieved because they are routinely used for billing and reporting purposes.
- Because no feasibility issues had arisen since the previous review and the specifications remained the same, the Committee chose to accept the vote on Feasibility from its previous evaluation of the measure.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-12; No Pass-0**; 4b. Usability: **H-1; M-12; L-0; I-0**

Rationale:

- This measure is currently in use for public reporting in the Texas Medicaid/CHIP Pay for Quality Program (P4Q).

- In 2016, the Dental Quality Alliance (DQA) expanded its scope of review of its measures by convening conference calls for two user groups – one comprised of representatives from six state Medicaid programs (Alabama, Florida, Kentucky, Oregon, Nevada, and Pennsylvania) and the other comprised of representatives from eight dental plans. Participants shared their experiences implementing DQA measures in their respective programs, including any challenges related to the DQA measures specifications and use of these measures in their quality improvement programs. The developer reports that participants did not have any significant issues related to the clarity or feasibility of implementing the measure specifications.
- The developer also noted that data retrieved suggest a trend in improvement over time, although these are initial performance data for one program. It stated that most measure users are just now getting their quality measurement programs underway.
- Ultimately, the Committee voted unanimously that the measure meets the NQF Use criterion. Additionally, the Committee raised no concerns regarding the measure's usability and voted that the measure meets NQF's Usability criterion.

---

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-13**; **N-0**

6. Public and Member Comment

NQF received one comment about this measure. The commenter requested clarification on whether the measure is specified for Medicaid plans only or if it also is specified for commercial plans with a dental benefit. The commenter also asked the Committee to look for opportunities for harmonization with existing Medicaid Early and Periodic Screening, Diagnostic, and Treatment (EPSTD) reporting requirements. The commenter further noted that, conceptually, any visit that satisfies measure 2511 would also satisfy measure 2517, and vice versa; and recommend that the Committee consider the value of having both measures in the Prevention and Population Health portfolio.

Developer Response:

> Thank you for your comment. Measure 2517 is calculated using administrative enrollment and claims data, and it is specified for reporting at the program (e.g., Medicaid or CHIP) or plan (e.g., MCO or DBA) level. Entities that do not have dental administrative data will not be able to report this measure. Reporting on the measure for a unit (e.g., provider level) or using a data source other than that for which the measure was developed may not be reliable.
>
> There is no counterpart on the CMS-416 EPSDT data for the DQA's Oral Evaluation measure. The DQA was formed at the request of CMS and maintains regular communication with CMS about its measure development activities in order to promote alignment and harmonization in dental quality measurement. Measure 2511, Utilization of Services, is an access measure – whether children are able to access the dental care system. Measure 2517 is a process

measure – whether children are receiving regular oral evaluations, including diagnostic services that are critical to evaluating oral disease and dentition development and to developing an appropriate oral health prevention regimen and treatment plan.

---

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals

2528 Prevention: Topical Fluoride for Children at Elevated Caries Risk, Dental Services

Submission

Description: Percentage of enrolled children aged 1-21 years who are at "elevated" risk (i.e., "moderate" or "high") who received at least 2 topical fluoride applications within the reporting year.

Numerator Statement: Unduplicated number of enrolled children aged 1-21 years who are at "elevated" risk (i.e., "moderate" or "high") who received at least 2 topical fluoride applications as a dental service

Denominator Statement: Unduplicated number of enrolled children aged 1-21 years who are at "elevated" risk (i.e., "moderate" or "high")

Exclusions: Medicaid/CHIP programs should exclude those individuals who do not qualify for dental benefits. The exclusion criteria should be reported along with the number and percentage of members excluded.

Level of Analysis: Health Plan, Integrated Delivery System

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Claims

Measure Steward: American Dental Association on behalf of the Dental Quality Alliance

---

STANDING COMMITTEE MEETING 02/09/2018

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Previous Evidence Evaluation Accepted**; 1b. Performance Gap: **Previous Performance Gap Evaluation Accepted**

Rationale:

- This maintenance measure assesses the percentage of enrolled children ages 1-21 years who are at elevated risk for caries who received at least two topical fluoride applications within the reporting year.
- The developer attested there was no new evidence since the last evaluation. Previously, the submission cited a systematic review of 71 controlled clinical trial studies. The

evidence received a grade of moderate by an expert panel, which is second on a three-point scale and denotes that evidence statements "are based on preliminary determination from the current best available evidence, but confidence in the estimate is constrained by one or more factors."

- Ultimately, given that the evidence had not changed since the measure's previous endorsement, the Committee opted not to re-vote on the Evidence criterion. The votes from the Committee's previous review will carry over to this maintenance evaluation.
- The developer provided plan and program level performance information from four sources: Texas Medicaid, Florida CHIP, and Florida Medicaid programs, as well as national commercial data from Dental Service of Massachusetts, Inc. The performance range of 18% to 35% in CY 2010 (year in which data were available for all five programs) indicates variation in topical fluoride application across programs.
- The developer reported improvement of approximately 2% in Texas Medicaid and approximately 4% in Texas CHIP. The developer attributes improvement on the measure's performance to the stimulus that measurement itself provides.
- The Committee noted the lack of new information on performance gap; the developer responded that the lack of recent data is due to the start-up phase for integration of the measures into contracts and for programs and plans to prepare for reporting, in combination with a lag period for reporting measures calculated using administrative claims data.
- The Committee accepted the rationale for the lack of updated performance gap information. Because it had not been updated since its 2014 evaluation of the measure, the Committee chose to accept the vote on Performance Gap from its previous evaluation of the measure.

---

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Previous Reliability Evaluation Accepted**; 2b. Validity: **Previous Validity Evaluation Accepted**

<u>Rationale</u>:

- Elevated risk is captured and defined in three new CPT caries risk assessment codes (high, moderate, low) that can be submitted as an indication of risk. Given that the codes are new and use is still growing, the measure also uses a "look back" methodology that looks back three years to determine if a child has undergone any restorative procedure that would indicate tooth decay to include in the measure.
- The developer submitted critical data element testing to satisfy requirements for reliability and empiric validity testing. The developer reviewed dental records for the Texas Medicaid program to evaluate the agreement between the claims data and dental charts by calculating the sensitivity, specificity, positive predictive value, and negative predictive value, as well as the kappa statistic.
- Agreement (concordance) for topical fluoride application was 89.9%. Sensitivity was 90.7% and specificity was 88.4%. The positive predictive value was 93.5% and negative predictive value was 83.9%. The kappa statistic value was 0.782, which indicates "substantial agreement" between the claims data and the dental charts.

- The developer also submitted results from a systematic assessment of the measure's face validity. The assessment used a modified Delphi method, among other activities, to assess the measure's validity and found unanimous agreement that the calculated measure scores can be used to evaluate quality of care.
- The measure excludes those children enrolled in Medicaid or CHIP who do not have dental coverage or whose coverage does not include the two fluoride applications called for in the measure. However, there is no similar exclusion for children on commercial insurance who may have the same lack of or gap in coverage. Because this measure is assessed at the plan level, the developer stated it is unlikely that its use will unduly penalize providers who do not meet the measure's requirements as a result of a patient's lack of coverage. Additionally, the developer noted that more and more dental insurers are expanding coverage to include the evidence-based recommendation for two or more fluoride treatments per year. Committee members, nevertheless urged the developer to make the same exclusion for commercial plans as is provided for Medicaid/CHIP.
- The Committee noted that while pediatric dentists often perform a risk assessment, most children receive dental services from general dentists who show variation in their understanding and documentation of risk assessments. Additionally, young children will not have had a restorative procedure in order to be captured via the look-back method. The developer responded that these children would most likely still receive the preventative fluoride treatment but would not be documented as at-risk, and therefore not included in the measure, unless the dentist used the new CDT codes. Further, the developer noted that plans have begun to incentivize the use of the new risk assessment CDT codes and that the inclusion of the CDT codes in the measure can promote further adoption.
- One Committee member noted that children insured under Medicaid or CHIP tend to change plans frequently. Risk assessment information does not follow a patient across plans and can therefore be lost when a child switches to a new dental plan.
- Ultimately, given that no new measure testing was provided, as permitted by NQF for maintenance consideration, the Committee opted to carry over the votes for the Reliability and Validity criterion from the measure's previous endorsement evaluation.

---

3. Feasibility: Previous Feasibility Evaluation Accepted

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c.Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- This measure relies on standard data elements in administrative claims data (e.g., patient ID, patient birthdate, enrollment information, CDT codes, date of service, and provider taxonomy). These data are readily available and can be easily retrieved because they are routinely used for billing and reporting purposes.
- Given that no new feasibility information was provided, the Committee opted to carry over the votes from the measure's previous endorsement evaluation for the Feasibility criterion.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-14; No Pass-0;** 4b. Usability: **H-1; M-12; L-0; I-1**

Rationale:

- This measure is currently used in Texas Medicaid and CHIP, as well as the Florida Statewide Medicaid Prepaid Dental Health Program.
- The Committee voted unanimously that the measure meets the NQF Use criterion. Additionally, the Committee raised no concerns regarding the measure's usability and voted in support on the Usability criterion.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: Y-14**;** N-0

6. Public and Member Comment

NQF received one comment on this measure. The commenter requested clarification on whether the measure is specified for Medicaid plans only or if it is also specified for commercial plans with a dental benefit. The commenter supported the Standing Committee's request for clarification of how to identify individuals who are at "high" or "moderate" risk.

Developer Response:

1. Thank you for your comment. Measure 2528 is calculated using administrative enrollment and claims data, and it is specified for reporting at the program (e.g., Medicaid or CHIP) or plan (e.g., MCO or DBA) level. Entities that do not have dental administrative data will not be able to report this measure. Reporting on the measure for a unit (e.g., provider level) or using a data source other than that for which the measure was developed may not be reliable.

   The focus of this measure is on children inferred at being elevated caries risk as a priority population to focus quality measurement. Testing data found that significant performance gaps existed within the elevated caries risk populations. During initial measure development, it was recognized that the ability to make reliable distinctions between at-risk levels (e.g., between "moderate" and "high" risk) was not well established. Consequently, the measure adopted a clearer cut dichotomous distinction of "low" risk and "elevated" risk. (The measure does not require distinguishing "moderate" risk from "high" risk.) The recent findings of an American Dental Association – American Academy of Pediatric Dentistry Caries Risk Assessment Expert Panel (available upon request), which reviewed the current state of science on caries risk assessment and developed guidance on risk categorization, found that current caries risk assessment tools share many common elements to assess risk and affirmed that

they have dichotomous predictive ability to quantify “low risk” and “elevated risk”: “Current tools have derived various methods to categorize risk based on expert consensus. The categorization of risk differs between the tools. However, all tools appear to qualify “low risk” in a similar manner: lack of disease and presence of protective factors. Current CRA tools could be effectively used in identifying “low risk” patients.” Consequently, the MDMC continues to support the focus of the measure on the priority population of children at elevated risk for developing dental caries.

---

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals

Measures Not Recommended

2508 Prevention: Sealants for 6-9 Year-Old Children at Elevated Caries Risk

Submission

Description: Percentage of enrolled children in the age category of 6-9 years at "elevated" risk (i.e., "moderate" or "high") who received a sealant on a permanent first molar tooth within the reporting year.

Numerator Statement: Unduplicated number of enrolled children age 6-9 years at "elevated" risk (i.e., "moderate" or "high") who received a sealant on a permanent first molar tooth as a dental service.

Denominator Statement: Unduplicated number of enrolled children age 6-9 years who are at "elevated" risk (i.e., "moderate" or "high")

Exclusions: Medicaid/ CHIP programs should exclude those individuals who do not qualify for dental benefits. The exclusion criteria should be reported along with the number and percentage of members excluded.

There are no other exclusions.

Level of Analysis: Health Plan, Integrated Delivery System

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Claims

Measure Steward: American Dental Association on behalf of the Dental Quality Alliance

STANDING COMMITTEE MEETING 02/09/2018

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Previous Evidence Evaluation Accepted**; 1b. Performance Gap: **Previous Performance Gap Evaluation Accepted**

Rationale:

- The developer provided an update to the evidence in the form of a recent Cochrane Review on the effectiveness of sealants. The results of the review continue to support the recommendations of the American Dental Associations Sealant Guideline, which is currently being updated.
- The developer provided program- and plan-level information from five different sources: Texas Medicaid, Texas Children's Health Insurance Plans (CHIP), Florida CHIP, and Florida Medicaid programs, as well as national commercial data from Dental Service of Massachusetts, Inc.
- Committee members agreed the data source and sample sizes are sufficient to assess the gaps in performance, which ranged from 20% to 30% in CY 2010. The developer noted this is indicative of variation in sealant replacement across the programs, but

Committee members expressed concern that the current performance gap is based on broad low performance and not a variation in sealant placements.

- The developer did not provide more recent performance data, stating that due to the start-up phase for integration of the measures, most of the entities that have adopted the measure are just beginning implementation and so there is limited data reporting.
- The developer provided disparities data and showed statistically significant differences for both race and ethnicity were detected for the two programs for which such data were available. Additionally, the developer evaluated whether the measure could detect disparities by income (within a given program), children's health status (based on their medical diagnoses), Medicaid program type, CHIP dental plan, commercial product line, and preferred language for program communications.
- Although the developer detected disparities based on all the differing factors, data on all the characteristics were not consistently available, so it presented disparities data just on those characteristics most consistently available and had the greatest standardization (i.e. race/ethnicity and geographic location).
- Committee members acknowledged the Medicaid data are captured differently and will have different disparities, but wonder if this conclusion will have an impact on the performance gap.
- The Committee agreed to accept its votes on Evidence and Performance Gap from its prior review.

---

2. Scientific Acceptability of Measure Properties: <u>The measure does not meet the Scientific Acceptability criteria</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Previous Reliability Evaluation Accepted**; 2b. Validity: **H-0**; **M-5**; **L-8**; **I-0**

<u>Rationale</u>:

- The developer previously provided empirical reliability testing using statistical tests, but did not update with new testing, as permitted by NQF's maintenance of endorsement policy. . NQF permits data-element level validity testing to suffice for reliability testing. This previous testing focused on assessing the accuracy of dental procedure codes reported in claims data as the data elements that contribute most to computing the measure score. To evaluate data element validity, the developer conducted reviews of dental records and compared them to extracted administrative claim data, which matched more than 90% of the time.
- For Validity, Committee members questioned whether children that received a sealant on a permanent first molar, within the target year, also means that they specifically have received the recommended sealants.
- One Committee member also expressed concerned that the exclusion of children without dental benefits is not taken into account when the measure is being computed. Developers noted that the number of children enrolled in Medicaid without dental benefits is minimal, but that the exclusion does apply to the Medicaid plans
- Committee members also questioned the ambiguity of elevated risk and suggested that improved coding is a technique that could improve performance without any unintended consequences arising.

- While the Committee generally supported the concept and direction of the measure, concerns remained regarding the measure's exclusions and questions about sealants generally vs. specific sealants recommended. Ultimately, the measure failed the Validity criterion.
- The developer submitted a request for reconsideration for this measure. The Committee voted not to reconsider the measure following deliberations on similar measure 2509, which was consensus not reached and was ultimately not recommended. The Committee agreed that the issues with 2509 would also exist in this measure.

---

5. Related and Competing Measures

No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **N/A**

6. Public and Member Comment

NQF received five post-evaluation comments on this measure. All five comments raised concern over the measure's specifications, specifically the lack of exclusions for individuals with zero sealable molars. One comment disagreed with the measure's inclusion of individuals with "elevated" risk in the denominator, noting there is evidence that current tools to assess caries risk are not reliable. One comment also requested the creation of implementation guidelines.

Developer Response:

- Thank you for your comment.
  The intent of Measure #2508 is to compare program (e.g., Medicaid) and plan performance over time related to the application of sealants for a population at inferred risk for caries by measuring the percentage of enrolled children 6-9 years of age, at elevated caries risk, who received a sealant on a permanent first molar tooth during the reporting year.
- The DQA thanks the DentaQuest for their comments.
  The intent of Measure #2508 is to compare program (e.g., Medicaid) and plan performance over time related to the application of sealants for a population at inferred risk for caries by measuring the percentage of enrolled children 6-9 years of age, at elevated caries risk, who received a sealant on a permanent first molar tooth during the reporting year.
- Thank you for your comment.
  This measure is calculated using administrative enrollment and claims data, and it is specified for reporting at the program (e.g., Medicaid or CHIP) or plan (e.g., MCO or DBA) level. Entities that do not have dental administrative data will not be able to report this measure. Reporting on the measure for a unit (e.g., provider level) or using a data source other than that for which the measure was developed may not be reliable.
  The focus of this measure is on children inferred at being elevated caries risk as a priority population to focus quality measurement. Testing data found that significant performance gaps existed within the elevated caries risk populations. During initial

measure development, it was recognized that the ability to make reliable distinctions between at-risk levels (e.g., between "moderate" and "high" risk) was not well established. Consequently, the measure adopted a clearer cut dichotomous distinction of "low" risk and "elevated" risk. (The measure does not require distinguishing "moderate" risk from "high" risk.) The recent findings of an American Dental Association – American Academy of Pediatric Dentistry Caries Risk Assessment Expert Panel (available upon request), which reviewed the current state of science on caries risk assessment and developed guidance on risk categorization, found that current caries risk assessment tools share many common elements to assess risk and affirmed that they have dichotomous predictive ability to quantify "low risk" and "elevated risk": "Current tools have derived various methods to categorize risk based on expert consensus. The categorization of risk differs between the tools. However, all tools appear to qualify "low risk" in a similar manner: lack of disease and presence of protective factors. Current CRA tools could be effectively used in identifying "low risk" patients." Consequently, the MDMC continues to support the focus of the measure on the priority population of children at elevated risk for developing dental caries.

- The DQA appreciates the support of the National Association of ACOs and its recognition of the importance of this measure for improving quality of care and oral health outcomes. The DQA would like to address the comment regarding accounting for the previously sealed molars as the same comment was also expressed by the NQF Standing Committee.

  Measure Intent: The intent of Measure #2508 is to compare program (e.g., Medicaid) and plan performance over time related to the application of sealants for a population at inferred risk for caries by measuring the percentage of enrolled children 6-9 years of age, at elevated caries risk, who received a sealant on a permanent first molar tooth during the reporting year.

  Intended to evaluate relative performance. The measure specifications note that:

  - This measure will not delineate those whose teeth have not erupted, those who have already received sealants in prior years, and those with decayed/filled teeth not candidates for sealants.
  - This measure is not designed to provide the absolute percentage of children who have ever had a sealant on a permanent second molar.
  - The measure is intended to be used for monitoring variations in sealant placement between reporting entities and disparities in sealant placement.

  Many of these limitations stem from lack of critical data within administrative claims including lack of ability to identify unerupted teeth, lack of diagnostic codes to identify decayed teeth, and lack of surface level data in many program level databases.

  Valid process measure. Testing indicated that the measure enables program and plan level process-of-care comparisons. Performance gaps and disparities in performance at a point in time can be identified. Face validity assessments by the MDMC as well as the stakeholder community at large affirmed that the measure is a valid process measure with a higher score signifying higher quality.

Consequently, the measure provides the information it claims to, and measure guidance explicitly clarifies what it is not designed to do in order to avoid mistaken interpretations of the measure score. The measure enables sound conclusions about the quality of care provided.

In addition, among all of the DQA measures, this measure has enjoyed the greatest adoption, which speaks to the measure's ability to serve as a valid quality indicator, including adoption by the Centers for Medicare & Medicaid Services in the CHIPRA Core Set of Child Quality Measures, with reporting by 34 states in 2016, and inclusion in the Covered California State Marketplace quality reporting.

Rationale for not accounting for prior sealant placement. Feasibility, reliability and validity concerns were identified. To accurately capture prior sealant data, a child's complete dental treatment history during the tooth eruption years would need to be captured in administrative claims data. Due to enrollment churn, these historical data are frequently not available. The lack of historical data could be addressed by requiring continuous enrollment in prior years during the tooth eruption period; however, the consequent substantial decrease in the denominator-eligible population raised significant face validity concerns about the representativeness of the resulting sample. Additionally, excluding children with prior sealants could create potentially biased measurement when there are variable observation windows across reporting entities for capturing prior sealant placement. A plan with more historical data will be able to identify more exclusions with a consequent increase in its measure score that is not reflective of improved quality but merely of having more historical data available.

Re-examination of lack of denominator exclusions. The MDMC reviewed data offered by two Dental Benefit Administrators (DBAs) that participate in the one of the same Medicaid and CHIP programs included in original testing to lend insight into the impact on measure performance when exclusions for prior sealant placement are incorporated. Data without incorporating enrollment criteria in years prior to the reporting year were provided.

Table 1 compares the measure scores provided by the two DBAs for the current measure and for the measure with prior sealants excluded, using a 3-year look-back period. (Note: Plan 1 did not have data available prior to 2013, so the 3-year look back could only be used for 2016.) As expected, the measure scores increased, ranging from an increase of 1.8 percentage points to 5.1 percentage points. The differences in the measure scores were statistically significant (based on non-overlapping 95% confidence intervals) in both programs.

Table 1: #2508 (Sealants, 6-9 years) Measure Score Comparisons with and without Exclusions for Previously Sealed Teeth

Denominator Measure Score 95% CI, Lower Bound 95% CI, Upper Bound

Plan 1

Medicaid 2016-DQA Current Measure 321038 24.49% 0.2434 0.2464

2016-Exclude Previously Sealed 263981 28.77% 0.2860 0.2895

CHIP 2016-DQA Current Measure 46767 22.97% 0.2259 0.2335

2016-Exclude Previously Sealed 38947 26.74% 0.2630 0.2718

Plan 2

Medicaid 2014-DQA Current Measure 161553 27.00% 0.2678 0.2722

2014-Exclude Previously Sealed 141771 30.00% 0.2976 0.3024

2015-DQA Current Measure 220022 25.70% 0.2552 0.2588

2015-Exclude Previously Sealed 184174 29.80% 0.2959 0.3001

2016-DQA Current Measure 243165 25.10% 0.2493 0.2527

2016-Exclude Previously Sealed 198213 29.70% 0.2950 0.2990

2017-DQA Current Measure 215350 24.30% 0.2412 0.2448

2017-Exclude Previously Sealed 171338 29.40% 0.2918 0.2962

CHIP 2014-DQA Current Measure 21092 25.10% 0.2451 0.2569

2014-Exclude Previously Sealed 18870 27.60% 0.2696 0.2824

2015-DQA Current Measure 17376 24.70% 0.2406 0.2534

2015-Exclude Previously Sealed 15617 27.10% 0.2640 0.2780

2016-DQA Current Measure 25147 23.10% 0.2258 0.2362

2016-Exclude Previously Sealed 23085 24.90% 0.2434 0.2546

2017-DQA Current Measure 23931 22.80% 0.2227 0.2333

2017-Exclude Previously Sealed 21894 24.60% 0.2403 0.2517

Impact of exclusions on relative performance: electronic patient record-level validation. We had detailed patient record-level data available from 77 dental practice these practices locations representing more than 60,000 children <21 years of age (>14,000 6-9 years). We used these data to compare the relative rankings of the 77 practices based on their measure scores calculated without any exclusions with the relative rankings based on their measure scores calculated excluding children with no sealable molars for any reason (prior sealants, restorations, extractions, unerupted teeth, missing teeth, and active caries). We used Kendall's tau correlation coefficient, a statistical test of associations based on ranks of data, to compare the relative rankings. This correlation coefficient is a more conservative measure of correlation than Spearman's rho. Values >0.70 indicate high correlation. The Kendall's tau correlation coefficient between the two approaches was 0.83 ($p<0.001$), signifying high positive correlation in the relative rankings for the two approaches (with and without exclusions) in calculating the measure scores. [Spearman's rho=0.94, $p<0.001$] These results further supported the conclusion that the measure scores calculated without exclusions enable comparable distinctions in performance.

Based on these evaluations, it appears that not accounting for exclusions does not compromise the measure's ability to distinguish performance between reporting entities.

---

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals

2509 Prevention: Sealants for 10-14 Year-Old Children at Elevated Caries Risk

Submission

Description: Percentage of enrolled children in the age category of 10-14 years at "elevated" risk (i.e., "moderate" or "high") who received a sealant on a permanent second molar tooth within the reporting year.

Numerator Statement: Unduplicated number of enrolled children age 10-14 years at "elevated" risk (i.e., "moderate" or "high") who received a sealant on a permanent second molar tooth as a dental service.

Denominator Statement: Unduplicated number of enrolled children age 10-14 years who are at "elevated" risk (i.e., "moderate" or "high")

Exclusions: Medicaid/CHIP programs should exclude those individuals who do not qualify for dental benefits. The exclusion criteria should be reported along with the number and percentage of members excluded.

Level of Analysis: Health Plan, Integrated Delivery System

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Claims

Measure Steward: American Dental Association on behalf of the Dental Quality Alliance

STANDING COMMITTEE MEETING 02/09/2018

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Previous Evidence Evaluation Accepted**; 1b. Performance Gap: **Previous Performance Gap Evaluation Accepted**

Rationale:

- This maintenance measure captures the percentage of children ages 10-14 years who are at elevated risk of dental caries and who have received a sealant on a permanent second molar within the reporting year.
- The measure mirrors evidence-based guidelines regarding effective caries prevention, as well as the specific tooth (second permanent molar) for which the evidence is the strongest and the timing (shortly after eruption) of sealant placement to maximize effectiveness.
- The measure received a Grade B, which is defined as directly based on category II evidence or extrapolated recommendation for category I evidence.
- A recent Cochrane Review on the effectiveness of sealants encapsulated all the evidence, which support the recommendations of the American Dental Association (ADA) Sealant Guideline. This guideline is currently being updated.

- Since the developer did not provide new evidence outside of the Cochrane Review, which was directionally the same as previous evidence, the Committee chose not to vote again and accepted its previous evaluation for the Evidence criterion.
- The performance range of 8% to 13% in CY 2010 is indicative of low prevalence in sealant placement and across programs. The developer also provided 2011 CMS performance data from state Medicaid programs, which ranged from 6% to 22% of children ages 10-14 years who received a sealant on a permanent second molar tooth.
- The developer did not provide more recent performance data, stating that due to the start-up phase for integration of the measures, most of the entities that have adopted the measure are just beginning implementation and there are limited data.
- The developer provided disparities data and showed statistically significant differences for both race and ethnicity were detected for the two programs for which such data were available. Additionally, the developer evaluated whether the measure could detect disparities by income (within a given program), children's health status (based on their medical diagnoses), Medicaid program type, CHIP dental plan, commercial product line, and preferred language for program communications.
- The developer also evaluated whether the measure could detect disparities (by income, children's health status, Medicaid program type, commercial product line, and preferred language for program communications), and disparities were detected for each of these factors, but data on all characteristics were not consistently available for all programs.
- Since the developer did not provide new performance gap information, the Committee chose not to vote again and accepted its previous evaluation for the Performance Gap criterion.

---

2. Scientific Acceptability of Measure Properties: <u>The measure did not meet the Scientific Acceptability criteria</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Previous Reliability Evaluation Accepted**; 2b. Validity: **H-0**; **M-8**; **L-6**; **I-1**

Re-votes from Standing Committee Post-Comment Meeting 04/30/2018

2b. Validity: **H-0**; **M-8**; **L-6**; **I-1**

<u>Rationale</u>:

- The developer previously provided empirical reliability testing using statistical tests, but did not update with new testing, as permitted by NQF's maintenance of endorsement policy. NQF permits data-element level validity testing to suffice for reliability testing. This previous testing focused on assessing the accuracy of dental procedure codes reported in claims data as the data elements that contribute most to computing the measure score. To evaluate data element validity, the developer conducted reviews of dental records and compared them to extracted administrative claim data, which was >88% for the second molar, if sealant applied and >95% for application of sealant.
- For validity, Committee members expressed concern that children who received sealants on their second molars might not have met the recommended clinical guidelines. Specifically, Committee members questioned whether a child has a sealant on a permanent second molar in the target year actually meant that they had received the recommended sealant that year, since it could have happened in a prior year.

- The Committee did not reach consensus on the Validity criterion during the measure evaluation meeting. The Committee re-voted on the measure's Validity during the post-comment web meeting.
- During the post-comment call, Committee members agreed the updated information provided by the developer explained the Dental Quality Alliance's intended use of the measure. Some Committee members, however, maintained their previous concerns regarding sealants being placed at the appropriate age, and not whether the sealants were placed during the target year. Ultimately, the measure did not receive sufficient votes to meet the Validity criterion.

---

3. Feasibility: Previous Feasibility Evaluation Accepted

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c.Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- This measure relies on standard data elements in administrative claims data (e.g., patient ID, birthdate, enrollment info, CDT codes, date of service, and provider taxonomy), which are readily available and can be easily retrieved because they are routinely used for billing and reporting purposes. There have been no reports of feasibility issues with implementing this measure, as provided by the developer.
- The Committee opted to accept its previous vote on Feasibility.

---

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-13; No Pass-0**; 4b. Usability: **H-0; M-11; L-0; I-1**

Rationale:

- The measure is used for public reporting and accountability by the Texas HHS Commission: Texas Medicaid/CHIP Pay-for-Quality (P4Q) program.
- In 2016, the Dental Quality Alliance (DQA) expanded its scope of review for its measures by convening participants who shared experiences implementing DQA measures, including any challenges related to the measures specifications and use of the measures in their quality improvement programs. The developer reports participants did not have any significant issues related to the clarity or feasibility of implementing the measure.
- The developer noted that the data retrieved suggest a trend in improvement over time, but that these are initial performance data for one program. It stated that most measure users are just now getting their quality measurement programs underway, and therefore do not have substantial data to provide.
- The Committee voted to approve both the Use and Usability criteria.

---

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-X**; **N-X**

6. Public and Member Comment

NQF received two comments on this measure. One comment agreed with the Committee's concern about the need for an exclusion of patients with previously sealed molars. Additionally, however, the commenter noted support for the measure's endorsement. The second comment called for clarification as to whether the measure is specified exclusively for Medicaid plans or if it also applied to commercial plans with a dental benefit. The commenter also agreed with the Committee's request for clarification on how individuals are classified as "high" or "moderate" risk.

Developer Response:

- Thank you for your comment. This measure is calculated using administrative enrollment and claims data, and it is specified for reporting at the program (e.g., Medicaid or CHIP) or plan (e.g., MCO or DBA) level. Entities that do not have dental administrative data will not be able to report this measure. Reporting on the measure for a unit (e.g., provider level) or using a data source other than that for which the measure was developed may not be reliable.
  The focus of this measure is on children inferred at being elevated caries risk as a priority population to focus quality measurement. Testing data found that significant performance gaps existed within the elevated caries risk populations. During initial measure development, it was recognized that the ability to make reliable distinctions between at-risk levels (e.g., between "moderate" and "high" risk) was not well established. Consequently, the measure adopted a clearer cut dichotomous distinction of "low" risk and "elevated" risk. (The measure does not require distinguishing "moderate" risk from "high" risk.) The recent findings of an American Dental Association – American Academy of Pediatric Dentistry Caries Risk Assessment Expert Panel (available upon request), which reviewed the current state of science on caries risk assessment and developed guidance on risk categorization, found that current caries risk assessment tools share many common elements to assess risk and affirmed that they have dichotomous predictive ability to quantify "low risk" and "elevated risk": "Current tools have derived various methods to categorize risk based on expert consensus. The categorization of risk differs between the tools. However, all tools appear to qualify "low risk" in a similar manner: lack of disease and presence of protective factors. Current CRA tools could be effectively used in identifying "low risk" patients." Consequently, the MDMC continues to support the focus of the measure on the priority population of children at elevated risk for developing dental caries.
- The DQA appreciates the support of the National Association of ACOs and its recognition of the importance of this measure for improving quality of care and oral health outcomes. The DQA would like to address the comment regarding accounting for the previously sealed molars as the same comment was also expressed by the NQF Standing Committee.

Measure Intent: The intent of Measure #2509 is to compare program (e.g., Medicaid) and plan performance related to the application of sealants for a population at inferred risk for caries by measuring the percentage of enrolled children 10-14 years of age, at elevated caries risk, who received a sealant on a permanent second molar tooth during the reporting year.
Intended to evaluate relative performance. The measure specifications note that:

- This measure will not delineate those whose teeth have not erupted, those who have already received sealants in prior years, and those with decayed/filled teeth not candidates for sealants.
- This measure is not designed to provide the absolute percentage of children who have ever had a sealant on a permanent second molar.
- The measure is intended to be used for monitoring variations in sealant placement between reporting entities and disparities in sealant placement.

Many of these limitations stem from lack of critical data within administrative claims including lack of ability to identify unerupted teeth, lack of diagnostic codes to identify decayed teeth, and lack of tooth surface level data in many program level databases.

Valid process measure. Testing indicated that the measure enables program and plan level process-of-care comparisons. Performance gaps and disparities in performance at a point in time can be identified. Face validity assessments by the MDMC as well as the stakeholder community at large affirmed that the measure is a valid process measure with a higher score signifying higher quality.

Consequently, the measure provides the information it claims to, and measure guidance explicitly clarifies what it is not designed to do in order to avoid mistaken interpretations of the measure score. The measure enables sound conclusions about the quality of care provided.

Rationale for not accounting for prior sealant placement. Feasibility, reliability and validity concerns were identified. To accurately capture prior sealant data, a child's complete dental treatment history during the tooth eruption years would need to be captured in administrative claims data. Due to enrollment churn, these historical data frequently are not available. The lack of historical data could be addressed by requiring continuous enrollment in prior years during the tooth eruption period; however, the consequent substantial decrease in the denominator-eligible population raised significant face validity concerns about the representativeness of the resulting sample. Additionally, excluding children with prior sealants could create potentially biased measurement when there are variable observation windows across reporting entities for capturing prior sealant placement. A plan with more historical data will be able to identify more exclusions with a consequent increase in its measure score that is not reflective of improved quality but merely of having more historical data available.

The MDMC reviewed data offered by two Dental Benefit Administrators (DBAs) that participate in the one of the same Medicaid and CHIP programs included in original testing to lend insight into the impact on measure performance when exclusions for

prior sealant placement are incorporated. Data without incorporating enrollment criteria in years prior to the reporting year were provided.

Table 1 compares the measure scores provided by the two DBAs for the current measure and for the measure with exclusions for children who already had all four molars previously sealed, using a 3-year look-back period. (Note: Plan 1 did not have data available prior to 2013, so the 3-year look back could only be used for 2016.) As expected, the measure scores increased, but the increases were less than 1 percentage point, ranging from 0.30% percentage point to 0.76% percentage point. The differences in the measure scores were statistically significant (based on non-overlapping 95% confidence intervals) in the Medicaid program, but not in the CHIP program.

Table 1: #2509 (Sealants, 10-14 years) Measure Score Comparisons with and without Exclusions for Previously Sealed Teeth

Denominator Measure Score 95% CI, Lower Bound 95% CI, Upper Bound

Plan 1

Medicaid 2016-DQA Current Measure 347260 17.04% 0.1691 0.1716

2016-Exclude Previously Sealed 327778 17.79% 0.1766 0.1792

CHIP

2016-DQA Current Measure 46110 17.24% 0.1690 0.1759

2016-Exclude Previously Sealed 43823 17.94% 0.1758 0.1829

Plan 2

Medicaid 2014-DQA Current Measure 157734 17.70% 0.1751 0.1789

2014-Exclude Previously Sealed 151899 18.20% 0.1801 0.1839

2015-DQA Current Measure 215113 17.60% 0.1744 0.1776

2015-Exclude Previously Sealed 204487 18.30% 0.1813 0.1847

2016-DQA Current Measure 260807 17.00% 0.1686 0.1714

2016-Exclude Previously Sealed 248681 17.70% 0.1755 0.1785

2017-DQA Current Measure 264111 16.70% 0.1656 0.1684

2017-Exclude Previously Sealed 248829 17.40% 0.1725 0.1755

CHIP 2014-DQA Current Measure 29510 14.60% 0.1420 0.1500

2014-Exclude Previously Sealed 28687 15.00% 0.1459 0.1541

2015-DQA Current Measure 22175 15.00% 0.1453 0.1547

2015-Exclude Previously Sealed 21493 15.30% 0.1482 0.1578

2016-DQA Current Measure 31012 15.00% 0.1460 0.1540

2016-Exclude Previously Sealed 30308 15.30% 0.1489 0.1571

2017-DQA Current Measure 30835 15.00% 0.1460 0.1540

2017-Exclude Previously Sealed 29990 15.30% 0.1489 0.1571

Impact of exclusions on relative performance: electronic patient record-level validation. We had detailed patient record-level data available from 77 dental

practice locations representing more than 60,000 children <21 years of age (>19,000 10-14 years). We used these data to compare the relative rankings of the 77 practices based on their measure scores calculated without any exclusions with the relative rankings based on their measure scores calculated excluding children with no sealable molars for any reason (prior sealants, restorations, extractions, unerupted teeth, missing teeth, and active caries). We used Kendall's tau correlation coefficient, a statistical test of associations based on ranks of data, to compare the relative rankings. This correlation coefficient is a more conservative measure of correlation than Spearman's rho. Values >0.70 indicate high correlation. The correlation coefficient between the two approaches was 0.96 ($p<0.001$), signifying nearly perfect positive correlation in the relative rankings for the two approaches (with and without exclusions) in calculating the measure scores. [Spearman's rho=0.996, $p<0.001$] These results further supported the conclusion that the measure scores calculated without exclusions enable comparable distinctions in performance.

Based on these evaluations, it appears that not accounting for exclusions does not compromise the measure's ability to distinguish performance between reporting entities.

---

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals