NATIONAL QUALITY FORUM
Driving measurable health improvements together

MEASURE WORKSHEET

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

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

Brief Measure Information

NQF #: 2528

Measure Title: Prevention: Topical Fluoride for Children, Dental Services

Measure Steward: American Dental Association

Brief Description of Measure: Percentage of children aged 1 through 20 years who received at least 2 topical fluoride applications as dental services within the reporting year.

The measure is specified for reporting at the program (e.g., Medicaid, CHIP, Health Insurance Marketplaces) and plan (e.g., dental and health plans) levels for both public and private/commercial reporting.

Developer Rationale: SPRING 2022 MAINTENANCE SUBMISSION

Dental caries remains one of the most common, yet preventable, diseases of childhood. As noted in our original submission, dental decay children has significant short- and long-term adverse consequences on children's health and overall well-being. Updated national surveillance data for the period 2011-2016 indicates that 23% of children aged 2-5 years experience dental caries related lesions, increasing to 52% among children aged 6- 8 years. Untreated decay was 10% among children aged 2-5 years and 16% among children aged 6-8 years. On permanent teeth, the prevalence of caries related tooth lesions was 17% for children 6-11 years and 57% among adolescents aged 12-19 years. Low-income children are twice as likely to have dental decay as higher-income children. Untreated decay was 5% among children aged 6-11 years and 17% among adolescents aged 12-19 years. (Centers for Disease Control and Prevention, 2019)

Oral health disparities are well documented and persist. Poor and near-poor children (children living in households with <=200% of the federal poverty level) are approximately twice as likely to have dental caries and untreated decay compared to higher income children (>200% FPL). National surveillance data indicate that Mexican American and non-Hispanic black children are more likely to have dental caries and untreated decay than non-Hispanic white children. For example, the prevalence of dental caries related lesions in primary teeth among Mexican American children aged 6-8 years was 73% compared to 44% for non-Hispanic white children. (Centers for Disease Control and Prevention, 2019)

Although dental caries can managed and caries related lesions can be treated and restored, it is important to prevent the disease process from developing in the first place. As noted in the evidence section, multiple

systematic reviews with meta-analyses find evidence supporting professionally applied topical fluoride, starting as early as six months of age and applied at least twice per year, as beneficial in preventing dental caries and associated decay(USPSTF 2021, Marinho 2013, Weyant 2013).

References

Centers for Disease Control and Prevention. Oral Health Surveillance Report: Trends in Dental Caries and Sealants, Tooth Retention, and Edentulism, United States, 1999–2004 to 2011–2016. Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services; 2019. Available at:

https://www.cdc.gov/oralhealth/publications/OHSR-2019-index.html

Marinho VCC, Worthington HV, Walsh T, Clarkson JE. 2013. Fluoride Varnishes for Preventing Dental Caries in Children and Adolescents (Review). Cochrane Database of Systematic Reviews 2013, Issue 7. Art. No.: CD002279. DOI: 10.1002/14651858.CD002279.pub2. PMID: 23846772. Available at:

https://www.cochrane.org/CD002279/ORAL_fluoride-varnishes-for-preventing-dental-caries-in-children-and-adolescents .

US Preventive Services Task Force. Screening and Interventions to Prevent Dental Caries in Children Younger Than 5 Years, Final Recommendation Statement. December 7, 2021. *Available at:*

https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/prevention-of-dental-caries-in-children-younger-than-age-5-years-screening-and-interventions1#bootstrap-panel--12

Weyant RJ, Tracy SL, Anselmo TT, Beltrán-Aguilar ED, et al; American Dental Association Council on Scientific Affairs Expert Panel on Topical Fluoride Caries Preventive Agents. J Am Dent Assoc. 2013 Nov;144(11):1279-91. Topical fluoride for caries prevention: executive summary of the updated clinical recommendations and supporting systematic review.

PRIOR SUBMISSION

Inequalities in oral health status and inadequate use of oral health care services are well documented. Dental caries is the most common chronic disease in children in the United States (NCHS 2012). In 2009–2010, 14% of children aged 3 –5 years had untreated dental caries. Among children aged 6–9 years, 17% had untreated dental caries, and among adolescents aged 13–15, 11% had untreated dental caries (Dye, Li, and Thorton-Evans 2012). Dental decay among children has significant short- and long-term adverse consequences (Tinanoff and Reisine 2009). Childhood caries is associated with increased risk of future caries (Gray, Marchment, and Anderson 1991; O´Sullivan and Tinanoff 1996; Reisine, Litt, and Tinanoff 1994), missed school days (Gift, Reisine, and Larach 1992; Hollister and Weintraub 1993), hospitalization and emergency room visits (Griffin et al. 2000; Sheller, Williams, and Lombardi 1997) and, in rare cases, death (Casamassimo et al. 2009).

Identifying caries early is important to reverse the disease process, prevent progression of caries, and reduce incidence of future lesions. Evidence suggests that topical fluoride applied to children starting as early as six months of age is beneficial in preventing dental caries (Weyant et al. 2013). However, approximately three quarters of children younger than age 6 years did not have at least one visit to a dentist in the previous year (Edelstein & Chinn 2009). Evidence-based clinical recommendations suggest that topical fluoride should be applied at least every three to six months in children at elevated risk for caries (Weyant et al. 2013).

The proposed measure, Topical Fluoride for Children at Elevated Caries Risk – Dental Services, captures whether children at moderate or high caries risk received at least two topical fluoride applications as dental services. Because topical fluoride is indicated at 3-6 month intervals (2-4 times per year) for children at elevated caries risk, at least two applications are indicated during the reporting year. This measure directly reflects evidence-based guidelines regarding an effective caries prevention measure (professionally applied topical fluoride), including the frequency required for clinical effectiveness (at least every three-six months). Topical Fluoride allows plans and programs to assess whether children at risk for caries are receiving evidence-based preventive services and target performance improvement initiatives accordingly.

Note: Procedure codes contained within claims data are the most feasible and reliable data elements for quality metrics in dentistry, particularly for developing programmatic process measures to assess the quality

of care provided by programs (e.g., Medicaid, CHIP) and health/dental plans. In dentistry, diagnostic codes are not commonly reported and collected, precluding direct outcomes assessments. Although some programs are starting to implement policies to capture diagnostic information, evidence-based process measures are the most feasible and reliable quality measures at programmatic and plan levels at this point in time.

[Complete citations provided in 1c4 and in Evidence Submission Form.]

Numerator Statement: Unduplicated number of children who received at least 2 topical fluoride applications as dental services

Denominator Statement: Unduplicated number of children aged 1 through 20 years

Denominator Exclusions: There are no measure-specific exclusions. There is a standard exclusion as part of determining denominator eligibility: Medicaid/CHIP programs should exclude those individuals who do not qualify for dental benefits.

Measure Type: Process

Data Source: Claims, Enrollment Data

Level of Analysis: Health Plan, Program

IF Endorsement Maintenance – Original Endorsement Date: 9/18/2014

Most Recent Endorsement Date: 6/6/2018

IF this measure is paired/grouped, NQF#/title: Prevention: Topical Fluoride for Children

#3701 - Prevention: Topical Fluoride for Children, Oral Health Services

#3700 - Prevention: Topical Fluoride for Children, Dental or Oral Health Services

#2528 - Prevention: Topical Fluoride for Children, Dental Services

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

Although this measure can be reported as a stand-alone measure, it is being grouped with two complementary measures to enable more robust quality improvement efforts. The DQA considered submitting a single measure with three numerators (denominator population is the same). But NQF evaluation criteria state: “Measures with multiple measure components that are assessed for each patient, but that result in multiple scores for an accountable entity rather than a single score. These generally should be submitted as separate measures and indicated as paired/grouped measures.” (Measure Evaluation Criteria and Guidance, September 2021, p. 52) Based on this and discussions with NQF staff, we are submitting as three distinct measures.

This measure – NQF 2528: Topical Fluoride for Children, Dental Services – focuses on topical fluoride delivered as a “dental” service (by or under the supervision of a dentist). Because many children, especially very young children, do not have a dental home, many state Medicaid programs and MCOs pay for the application of topical fluoride as an “oral health” service (by a physician or health care provider other than a dentist nor under supervision of a dentist). Consequently, state Medicaid programs, as well as commercial integrated medical-dental benefit MCOs or integrated medical-dental healthcare delivery sites, have a strong interest in tracking whether children receive **any** topical fluoride regardless of provider type ("dental” or “oral health" services). They also have a strong interest in understanding whether, and for whom, topical fluoride is being delivered through "dental" providers and "oral health" providers.

Measures of topical fluoride provision by provider type, in addition to a measure of overall provision, are important because multi-pronged quality improvement strategies may be used to improve rates of topical fluoride application among a population of children. Dental providers and/or medical providers may be the focus of these efforts. Without measures that track the effectiveness by provider type, it is more difficult for programs and plans to assess which efforts are most effective. In addition, the accountability and delivery systems are typically distinct. Improving fluoride application by dental providers is accomplished through the dental delivery system and related financing/reimbursement structures, whereas topical fluoride application

by medical providers is accomplished through medical delivery system and related financing/reimbursement structures. Some measure users will benefit by implementing and using all three measures (such as Medicaid programs and private payers/delivery systems that include both medical and dental). Other users, focused specifically on either dental or medical care delivery, respectively, will be able to report using one of the measures: either the measure related to "dental" services or the measure related to "oral health" services.

The need for three grouped measures comes directly from user community requests. In considering a topical fluoride measure for inclusion in the Centers for Medicare and Medicaid Services' Core Set of Children's Health Care Quality Measures, a specific request was made for a single measure that included three numerators (dental or oral health services, dental services, and oral health services) because of the recognized need by Medicaid programs to track not only overall receipt of topical fluoride but also topical fluoride provided through the dental and medical delivery systems specifically. Review by the DQA's Measures Development and Maintenance Committee, which includes representation of providers, community health centers and payers, affirmed the value of reporting three numerators across public and private/commercial measure applications for the reasons described above.

Consequently, the DQA is submitting two complementary measures for endorsement to be "grouped" with the existing endorsed measure: (1) Topical Fluoride for Children, Dental or Oral Health Services and (2) Topical Fluoride for Children, Oral Health Services. This grouping provides users with measurement options that appropriately support population-based assessments of quality. It enables measure users, including Medicaid programs and their contracted MCOs, integrated medical-dental MCOs and integrated medical-dental delivery systems, to examine the overall provision of topical fluoride by provider type, which has been identified by stakeholders as integral for quality improvement and accountability purposes.

Preliminary Analysis: Maintenance of Endorsement

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

Criteria 1: Importance to Measure and Report

1a. Evidence

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

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

The developer provides the following description for this measure:

- This is a maintenance process measure at the program and health plan level that evaluates the percentage of children aged 1 through 20 years who received at least 2 topical fluoride applications as dental services within the reporting year. This measure is paired with two measures that also focus on topical fluoride but are delivered through oral health services and through either dental or oral health services.
- The developer provides a logic model that depicts topical fluoride applied to children starting as early as six months of age is beneficial in preventing dental caries.

The developer provides the following evidence for this measure:

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

Summary of prior review in 2017

- The Standing Committee accepted the previous evidence evaluation from the 2014 submission as there was no new evidence presented in 2017. Therefore, the following information is from the 2014 review.
- The developer presented a systematic review done by the American Dental Association (ADA).
    - The evidence-based clinical guidelines recommend the specific topical fluoride agents for people who are at elevated risk of developing dental caries.
    - Additionally, the guidelines in the report recommend applying 2.26 percent fluoride varnish at least every three to six months for children younger than six years old and for children six-18, they recommend 2.26 percent fluoride varnish at least every three to six months or 1.23 percent acidulated phosphate fluoride (APF) gel for four minutes at least every three to six months.

- Seventy-one studies were included in evidence reviews. All studies included were controlled clinical trials. Seventeen randomized and five non-randomized control trials evaluated 2.26 percent fluoride varnish while 11 randomized and four non-randomized control trials evaluated 1.23 percent APF gel.
- The only potential harms identified in the study were nausea and vomiting when topical fluoride is ingested and dental fluorosis while tooth enamel is developing.
- The evidence received a moderate grade by an expert panel which was the second highest rating it could receive.
- The clinical recommendations for fluoride among children and adolescents received an evidence grade of “”in favor” which was the second highest recommendation out of a six-point scale.

Changes to evidence from last review

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

☒ The developer provided updated evidence for this measure:

- The developer changed the denominator to reflect all children ages 1-20 rather than those classified as at elevated risk. The developer acknowledged that while the methodology to identify children at elevated risk was reliable, there were concerns that a large portion of at-risk children were being excluded from the measure.
- The developers identified a Cochrane systematic review as evidence.
  - There were 22 studies included in this review. All were randomized or quasi-randomized controlled trials. The evidence found that fluoride varnish on permanent dentition is associated on average with a 43 percent reduction in decayed, missing, and filled tooth surfaces. For fluoride varnish on primary dentition, there was a 37 percent reduction. In general, they stated there was little information on possible adverse effects.
  - The review received a quality rating of moderate. The authors define a moderate rating as one where more research will likely impact the confidence in the estimate of the effect. This was the second highest option in the scale used, the higher option being high quality and the lower options being low quality and very low quality.
  - The study concluded that fluoride varnish in both permanent and primary teeth have a substantial caries-inhibiting effect.
  - Because this is not a clinical guideline, no formal recommendations were made. Rather, the author provides an implication for practice. The implication is that applying a fluoride varnish two to four times a year is associated with a reduction in caries.
- The developer also cited the United States Preventive Services Task Force (USPSTF) systematic review and recommendation on the *Prevention of Dental Caries in Children Younger than 5 Years*.
  - This study addressed two key questions. The first is asking how effective preventive treatment is in preventing dental caries in children younger than 5 years old. The second is asking what harms are associated with preventive oral health interventions in children younger than 5 years old.
  - There were 32 studies and 1 systematic review that included 19 studies included in the USPSTF review. Twenty-two studies addressed the “how effective…” key question stated above. Fifteen of which specifically addressed topical fluoride application. These studies were all randomized controlled trials. Four studies reported on adverse events associated with topical fluoride. These studies were all randomized controlled trials. The trials found that there were decreased caries increments and decreased likelihood of incident caries. Further, the studies found that there were no reported differences between fluoride varnish and

placebo in risk of fluorosis or adverse events. The only "adverse events" reported were that children did not like the smell of the fluoride varnish. Additionally, one study found that a few children vomited due to the smell, texture, or taste.

- The studies within this review all supported that the recommendation that primary care clinicians should apply fluoride varnish to the primary teeth of all infants and children starting from primary tooth eruption. This recommendation received a B rating which suggests that the service should be offered and provided. This rating is the second highest option in a five-letter grade scale.
- The authors gave the evidence a moderate rating which is the second highest option available in the scale. The authors stated that the evidence concluded there was a moderate certainty that a moderate benefit of preventing dental caries with fluoride varnish applications exists in all children younger than age 5.

Questions for the Committee:

- *The evidence provided by the developer is updated, directionally the same, and stronger compared to that for the previous NQF review. Does the Committee agree there is no need for repeat discussion and vote on Evidence?*
- *What is the relationship of this measure to patient outcomes?*
- *How strong is the evidence for this relationship?*
- *Is the evidence directly applicable to the process of care being measured?*

Guidance from the Evidence Algorithm

Process measure based on systematic review and grading (Box 3) → QQC provided in the submission (Box 4) → Quantity: High; Quality: Moderate; Consistency: Moderate/High (Box 5b) → Moderate

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

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

Maintenance measures – increased emphasis on gap and variation

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

- Data on performance gap was derived from 14 state Medicaid programs including Alaska, Arizona, Delaware, Idaho, Michigan, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, Oregon, South Carolina, Washington, and Wyoming.
    - These states were specifically selected because of their quality of data and diverse geographic locations, population size, demographic characteristics, and Medicaid dental delivery system. The data came from Medicaid enrollment and claims data contained within the Transformed Medicaid Statistical Information System (T-MIS) Analytic Files (TAFs).
    - There were over seven million enrollees aged zero to 20 years across the 14 programs for each of the calendar years (2016, 2017, and 2018) included in the analysis.
    - In the most recent year of data, 2018, measure scores ranged from 14 percent to 28 percent. A similar amount of variation exists in 2017 and 2016 as well. The developer states this variation indicates a gap in care.
    - The developer also added that 72 percent of children in the highest performing states did not receive at least two fluoride treatments in 2018, which the developer states indicates that there is an opportunity for improvement.

Disparities

- The developer stratified measures scores from 2018 into four categories: 1) age; 2) rural/urban geographic location (based on zip code); 3) race and ethnicity; and 4) sex assigned at birth. They did not report stratification when more than 10 percent of the data for the variable was missing.
- The developer noted disparities in each stratification category, but that the greatest disparities existed around age and race/ethnicity.
    - Children in the youngest and oldest age cohorts had the lowest performance scores
    - Additionally, non-Hispanic Black and American Indian/Alaskan Native children had lower measure scores than non-Hispanic White children. The developer also stated that measure scores were higher for non-Hispanic Asian children and Hispanic children compared with non-Hispanic White children.

Questions for the Committee:

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

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

Committee Pre-evaluation Comments:

1a. Evidence

- I have concerns about the approach of this metric. The USPSTF guidelines and other data are geared toward HIGH RISK youth (generally age less than 5)- instead, this metric is being expanded to age 1 through 20, and I worry that there is a conflict of interest in the broadening of this recommended measure.
- There is some updated evidence from the previous review. Evidence supports the measure.
- Rating moderate: Supported by systematic review w QQC
- Evidence is relevant. Additional evidence strengthens the measure.
- Strong evidence reflected in systematic reviews and practice guidelines. If anything, changing the focus from "children at elevated caries risk" to all children and adding two new measures for the same treatment delivered by oral health service providers (3700 and 3701) makes this evidence more relevant to the measures, and I believe that the evidence does not need to be discussed at this meeting.
- Strong evidence provided
- This is a process measure at the program and health plan level. Evidence was upgraded from 2014, updated systematic literature review that presents quality, quantify and consistency of the evidence. Also evidence is graded. Fluoride varnish in both permanent and primary teeth have a substantial caries-inhibiting effect.
- The evidence for this measure is moderate and new studies have not changed the quality of evidence.

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

- Systematic review data offered is conducted by developer themselves (ADA)
- There is data regarding performance and disparity gaps.
- Rating moderate
- Gap in care persists, is significant. Evidence of disparities in care is also present.
- Strong evidence of overall performance gaps and disparities, especially by age and race/ethnicity.

- Strong evidence of both gaps and inequities
- Gap in care evident from 2018 data 4 groups: age, rural/urban, geographic location by zip code; race and ethnicity, and sex assigned at birth.
- Disparities by age, geographic location, race/ethnicity and sex were reported, underscoring the need for improvement.

Criteria 2: Scientific Acceptability of Measure Properties

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

Evaluators: Staff

2a. Reliability: Specifications and Testing

For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

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

For maintenance measures – less emphasis if no new testing data provided.

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

Specifications:

- Since the last review, the denominator specifications have changed. Originally, it was restricted to children at elevated risk for dental caries. However, due to concerns that a significant percentage of at-risk children being excluded from measurement, they changed it to include all children. This change was made because of feedback from measure users.
- Additionally, the developer has created two new measures that are under consideration by NQF (NQF #3700 and NQF #3701). This measure – NQF 2528 Topical Fluoride for Children, Dental Services – focuses on topical fluoride delivered as a "dental" service (by or under the supervision of a dentist).
    - The developer states that since many children, especially very young children, do not have a dental home, many state Medicaid programs and managed care organizations (MCOs) pay for the application of topical fluoride as an "oral health" service (by a physician or health care provider other than a dentist nor under supervision of a dentist). Consequently, state Medicaid programs, as well as commercial integrated medical-dental benefit MCOs or integrated medical-dental delivery sites, have a strong interest in tracking whether children receive any topical fluoride regardless of provider type ("dental or oral health" services). They also have a strong interest in understanding whether, and for whom, topical fluoride is being delivered through "dental" providers and "oral health" providers. Consequently, the developer is submitting three complementary measures for endorsement.

Reliability Testing:

- The level of analysis states health plan and program; however, the new testing is only conducted at the program level. The developer justified this by stating that the program data is transferrable to the plan level. They noted further that the only potential concern would be denominator size. They stated that the denominator requirement from the measure captures a broad population and to-date they have not encountered issues with small denominators.

- Reliability testing conducted at the accountable entity level:
    - The developers used a random-split sample methodology and intraclass correlation coefficient (ICC). For each state Medicaid program, they randomly split the population and the denominator, numerator, and measure score were calculated for each sample. The ICC is used to calculate agreement between the samples.
        - The developer states the variation between split samples for each of the Medicaid programs is relatively small, signifying that the samples were similar. It is also noted that the measure scores have overlapping 95 percent confidence intervals (CI). Additionally, the ICC scores were reported by year.
            - 2018 (n=14): ICC = 0.997, with a 95 percent confidence interval of (0.9923,0.9992) and a p-value of less than 0.0001
            - 2017 (n=14): ICC = 0.999, with a 95 percent confidence interval of (0.9962,0.9996) and a p-value of less than 0.0001
            - 2016 (n=10): ICC = 0.999, with a 95 percent confidence interval of (0.9983,0.9999) and a p-value of less than 0.0001
    - Additionally, an evaluation of relative rankings between split samples and year were done. For the relative ranking between split samples, they were evaluating whether measure scores remained stable. For the relative ranking between years, they were evaluating if any dramatic changes happened that would threaten reliability. To report the findings, they calculated Kendall's Tau-b which is a rank correlation coefficient that measures associations based on concordant and discordant pairs. They also reported the Spearman's rank correlation coefficient. They noted that while they did both, they felt Kendall's Tau-b was the more appropriate statistical test given the relatively small sample size (n=14).
        - The developer noted that the relative rankings based on measure score are stable across split samples.
        - The relative rankings between years were as follows:
            - 2017 & 2018 (n=14): Kendall's Tau-b = 0.8462 (p-value = less than 0.0001) and Spearman's rank = 0.9429 (p-value = less than 0.0001)
            - 2016 & 2017 (n=10): Kendall's Tau-b = 0.6889 (p-value = 0.0073) and Spearman's rank = 0.8303 (p-value = 0.0029)
- Previous submission testing (patient/encounter level):
    - In the previous submission, the developer used patient/encounter validity testing as reliability testing. The developer did not update validity patient/encounter level testing because the data elements within the measure did not change. Therefore, the original patient/encounter level testing that used chart reviews holds for this measure.
        - In the previous submission, the developer validated the encounter data by comparing claims data against dental charts. The developer noted that record and procedure codes on the claims had a 94.04 percent agreement rate.
        - Additionally, the developer assessed whether the preventive service of topical fluoride application was accurately captured by claims data. The developer reported concordance (89.9 percent), sensitivity (90.7 percent), specificity (88.4 percent), positive predictive value (93.5 percent), and negative predictive value (83.9 percent). The developer also reported a kappa statistic which was 0.782.

Questions for the Committee regarding reliability:

- *Do you have any concerns that the measure cannot be consistently implemented (i.e., are measure specifications adequate)?*
- *Are there concerns that the measure was not tested at the plan level and will program level testing translate to the plan level?*

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

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

For maintenance measures – less emphasis if no new testing data provided.

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

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

Validity Testing

- The developer provided accountable entity level validity testing and noted that the patient/encounter level validity testing from the previous submission still remains valid and additional testing was not needed.
- The level of analysis states health plan and program; however, the new testing is only conducted at the program level. The developer justified this by stating that the program data is transferrable to the plan level. They noted further that the only potential concern would be denominator size. They stated that the denominator requirement from the measure captures a broad population and to-date they have not encountered issues with small denominators.
- Validity testing conducted at the accountable entity level:
    - The developers tested a hypothesized positive relationship between two other dental NQF-endorsed measures (stated below) using Kendall's Tau-b and Spearman's rank correlation coefficient to measure association. Once again, the developer noted that they reported both Kendall's Tau-b and Spearman's rank, but that Kendall's Tau-b is the more statistically appropriate test considering the smaller sample size. The results are as follows:
        - Utilization of Services, Dental Services compared to Topical Fluoride
            - 2018 (n=14): Kendall's Tau-b = 0.6703 (p-value = 0.0010) and Spearman's rank = 0.8286 (p-value = 0.0003)
            - 2017 (n=14): Kendall's Tau-b = 0.6264 (p-value = 0.0022) and Spearman's rank = 0.8022 (p-value = 0.0006)
            - 2016 (n=10): Kendall's Tau-b = 0.8222 (0.0013) and Spearman's rank = 0.9394 (p-value = 0.0001)
        - Oral Evaluation, Dental Services compared to Topical Fluoride
            - 2018 (n=14): Kendall's Tau-b = 0.7802 (p-value = 0.0001) and Spearman's rank = 0.9253 (p-value = less than 0.0001)
            - 2017 (n=14): Kendall's Tau-b = 0.8022 (p-value = 0.0001) and Spearman's rank = 0.9253 (p-value = less than 0.0001)
            - 2016 (n=10): Kendall's Taub = 0.8667 (p-value = 0.0007) and Spearman's rank = 0.9515 (p-value = less than 0.0001)
- Previous submission testing (patient/encounter level and face validity):

- In the previous submission, the developer presented face validity which was scored on the Delphi rating system. The measure received a mean importance score of seven, mean feasibility score of eight, and a mean validity score of seven for specific evidence-based preventive services, all out of a nine-point scale. The developers notes that a rating of seven through nine is considered scientifically sound and valid.
- While the developer did not provide updated data element testing, they did clarify that even though there was a denominator change, the critical data elements did not change.
    - In the previous submission, the developer validated the encounter data by comparing claims data against dental charts. They noted that record and procedure codes on the claims had a 94.04 percent correlation rate.
    - Additionally, the developer assessed whether the preventive service of topical fluoride application was accurately captured by claims data. The developer reported concordance (89.91 percent), sensitivity (90.7 percent), specificity (88.4 percent), positive predictive value (93.5 percent), and negative predictive value (83.9 percent). The developer also reported a kappa statistic which was 0.782.

Exclusions

- The measure does not use exclusions.

Risk-Adjustment

- The measure is not risk adjusted or stratified.

Meaningful Differences

- The developer provided performance score data for each program across three calendar years with 95 percent confidence intervals which were used to calculate the mean, median, standard deviation, and percentile distributions used for a chi-square test. The developers noted that the 95 percent confidence intervals did not overlap across any of the programs.
- The developer calculated the interquartile range for the measure scores and then conducted a chi-square test to evaluate the statistical significance of the differences in the measure scores between the lowest and highest quartiles. The following results were reported:
    - 2018 – Mean = 0.2079, interquartile range = 0.0595, p-value = less than 0.0001
    - 2017 – Mean = 0.1979, interquartile range = 0.0700, p-value = less than 0.0001
    - 2016 – Mean = 0.1809, interquartile range = 0.0724, p-value = less than 0.0001
- The developer states that these results suggest the measure can identify statistically and clinically meaningful differences in performance.

Missing Data

- The developers assessed missing data in two ways.
    - The first is the Medicaid and CHIP Business Information Solutions (MACBIS) conducted data quality assessments of the Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAFs). Then a value (low concern, medium concern, high concern, unusable, and unclassified) was assigned to each of the states based on if the data is usable, reliable, and accurate. Then the developers reviewed the results for age, Medicaid enrollment, claims file completeness (claims volume), claims file completeness (service users), and service use

(procedure codes). The second is the developer did an analysis of date of birth, beneficiary ID, CDT codes, and rendering provider taxonomy for dental procedure codes. The developer reported the same values and cut points that MACBIS uses.

- There were two states in calendar year 2016 for which the level of concern for missing data was medium and/or high.
    - In the state that had medium concern for missing data, the developer chose to include their data for measure reporting purposes as examination of measure scores found that performance was within expected ranges and similar between years.
    - In the state that had medium and high concerns, they excluded 2016 reporting from the testing as examination of measure scores found that the missing data significantly impacted performance.
- For the fields assessed by the developer, the rate of missing and invalid data was generally less than one percent.
- In 2016, there were three states for which rendering provider taxonomy for dental procedure codes were greater than 50 percent and were deemed unusable. These were therefore excluded from the measure testing.

○ The developer noted that because there was generally low concern and low rates of missing data, no rules were developed for handling it. They also stated that because the data are standard fields contained within administrative claims, the measure users are encouraged to improve data collection and quality as part of the quality improvement efforts rather than using statistical methods to address missing data.

Comparability

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

Questions for the Committee regarding validity:

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

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

Committee Pre-evaluation Comments:

2a1. Reliability - Specifications

- Denominator has been over-broadened
- No concerns about reliability testing.
- Rating: moderate
- No concerns.
- Changing the denominator from “children at elevated caries risk” to all children probably makes specifications clearer, as I do not know how “elevated caries risk” was defined.
- No concerns
- The reliability testing is done at the program level, and is now testing with for all children rather than only children at high risk for caries. The developer confirms the methods are appropriate still.

- Specifications have changed since the last review to include all children based upon feedback from measure users.

2a. Reliability - Testing

- No concerns
- No
- No.
- Testing at the program vs. plan level is a minor concern. Split sample ICCs and relative rankings compared – appropriate and strong results. Previous submission included testing at the patient/encounter level with strong results.
- No concerns
- No
- Reliability rating is moderate. No concerns identified.

2b1. Validity - Testing

- No
- No
- Rating: moderate
- No concerns.
- No concerns. Appropriate convergent validity tests with strong results. Face validity from previous submission still relevant.
- No concerns
- No concerns
- The validity ratings are moderate. No concerns identified.

2b2-2b3. Threats to validity (Exclusions, Risk Adjustment)

- What is rate of development of carries - is this tracked, and do we have disparities data on this? I understand that the measure looks at fluoride treatment, but its unclear to me what the impact of the disparities are in terms of outcome
- Risk adjustment is not used.
- No risk adjustment included for a large age range
- No exclusions. No risk adjustment.
- No concerns.
- I do wonder about whether HPSAs/dental provider access should be taken into account (as a risk adjustor, to stratify)? Do "oral health" providers fill the gaps?
- No measure for risk adjustment.
- This measure does not have exclusions nor is it risk adjusted or stratified. No concerns identified.

2b4-2b7. Threats to validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)

- The missing state data is from 2016- are there updates since then?
- Missing data is addressed
- No
- No concerns.
- No concerns.
- Missing data do pose a small threat
- No threat to validity
- No concerns identified.

Criterion 3. Feasibility

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

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

- The developer noted that since the last review, the have removed the elevated risk determination from the denominator, which should reduce data requirements.
- The data elements are coded by someone other than the person obtaining the original information and all data elements are defined fields in electronic claims. The measure uses data elements that are found in standard fields in administrative claims data which are routinely collected.
- Additionally, the measure was designed to avoid using software or other materials that require licensing fees. The measure specifications are accessible through a website and are free.

Questions for the Committee:

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

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

Committee Pre-evaluation Comments:

3. Feasibility

- –
- The required elements are routinely generated. Because of the reliance on claims data, uninsured children may not be captured, and they may be at highest risk.
- Rating: high
- No problems noted. Removing elevated risk seems to improve feasibility.
- Removing elevated risk designation makes the measure more feasible.
- No concerns
- All data elements are in defined fields in electronic claims.
- There is high feasibility for the measure because it uses electronic claims data.

Criterion 4: Use and Usability

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

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

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

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial

endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

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

Accountability program details

- Currently, the measure is used in:
    - Texas Health and Human Services Commission for payment program purposes
    - Florida Medicaid for payment program purposes
    - Other State Medicaid Agencies for quality improvement purposes
    - Covered California for quality improvement purposes
- The developer also notes that the measure has been:
    - Adopted by CMS for Child Core Health Care Quality Measurement for fiscal year 2022 reporting by state Medicaid and CHIP.
    - Included in the Center for Oral Health Systems Integration and Improvement (COHSII) Oral Health Quality Indicators for the Maternal and Child Health Population, which is funded by the Health Services and Resources Administration (HRSA) Maternal and Child Health Bureau for 2022 reporting.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- During the development phase results and data were shared with stakeholders including the CMS Oral Health Technical Advisory Group, NCQA, and the Center for Oral Health Systems Integration and Improvement. The Dental Quality Alliance (DQA) also provides technical assistance for measures through webinars, resource documents, and staff support. The DQA has also released a State Oral Health Quality Dashboard that has information on key oral healthcare measures.
- DQA has a process in place for reviewing and updating all measures. The process is overseen by the DQA's Measures Development and Maintenance Committee (MDMC). The review process includes public commenting, evaluation of comments, user group feedback, and code set reviews. Based on feedback from the community/stakeholders/measure users DQA changed the denominator to not include elevated risk in the denominator. This change was initiated to support a more population-based approach for this measure.

***Questions for the Committee**:*

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

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

4b. Usability (4b1. Improvement; 4b2. Benefits of measure)

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

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

Improvement results

- The developer noted a 10-point increase in performance from 2014 to 2018 in Texas Medicaid and Texas CHIP.

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

Unexpected findings (positive or negative) during implementation

- No negative impacts have been identified by the developer. However, developer does make references towards making a measure that is more population-based which is they state is why elevated risk for dental caries was removed from the denominator.

Potential harms

- The developer states the potential for harm is minimal.
- The developer states the benefits are as expected.

***Questions for the Committee**:*

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

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

Committee Pre-evaluation Comments:

4a. Use

- –
- This measure is a bit confusing with the aligned measures that incorporate who offered the fluoride treatment. Those being measured have been given performance results.
- Rating: moderate
- Change (eliminating elevated risk) is responsive to those being measured. Measure is in use or in consideration for use widely.
- Already in widespread use
- No concerns
- Used for public reporting, accountability - current and planned.

- This measure meets the expectations for feedback on performance, public reporting, use in an accountability program and is being adopted by the CMS for Child Core Health Care Quality Measurement in 2022. No concerns identified.

4b. Usability

- No direct harms
- No identified unintended consequences.
- Unable to determine benefit/harm trends reliably with info supplied
- No concerns.
- No comments.
- No concerns
- More harm than benefit. Developer states harm is minimal.
- Usability benefits outweigh potential risks.

Criterion 5: Related and Competing Measures

Related measures

- NQF #2511 Utilization of Services, Dental Services
- NQF #2517 Oral Evaluation, Dental Services
- NQF #2689 Ambulatory Care Sensitive Emergency Department Visits for Dental Caries in Children
- NQF #2695 Follow-Up after Emergency Department Visits for Dental Caries in Children

Harmonization

- The developer states that the measures have harmonized to the extent possible. Additionally, the developer noted that the above stated measures all address the same population, however, the denominators are different. Therefore, the measures are complementary to each other but, in fact, distinct.

Committee Pre-evaluation Comments:

5: Related and Competing Measures

- 2 other fluoride/children measures
- There are related measures based on the type of provider offering fluoride treatment. They are harmonized but the value of each may not be great.
- Group measures with distinction agreed
- No concerns.
- Addition of 3700 and 3701 is an important improvement.
- No concerns
- There are several related measures. The developer advises that the measure specifications are harmonized to the extent possible.
- Related measures address the same population, but are distinct and complimentary. No concerns.

Public and NQF Member Comments (Submitted as of June 15, 2022)

Member Expression of Support

- No member submitted an expression of support.

Comments

- No NQF member and public comments were received in advance of the Standing Committee evaluation.

Scientific Acceptability Evaluation

RELIABILITY: SPECIFICATIONS

1. **Have measure specifications changed since the last review? ☒ Yes ☐ No**
2. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented? ☒ Yes ☐ No**
3. **Briefly summarize any changes to the measure specifications and/or concerns about the measure specifications.**
   - In the previous measure review, the Standing Committee noted that while pediatric dentists often perform a risk assessment, most children receive dental services from general dentists who vary on their understanding and level of documentation of risk assessments. Additionally, the Standing Committee stated that children insured under Medicaid or CHIP often change plans and typically, their risk assessment information does not follow across plans and therefore lost.
   - Since the last review, the developer changed the denominator from being exclusively children at elevated risk for dental caries to all children.
   - The level of analysis states health plan and program; however, the new testing is only conducted at the program level. The developer justified this by stating that the program data is transferrable to the plan level. They noted further that the only potential concern would be denominator size. They stated that the denominator requirement from the measure captures a broad population and to-date they have not encountered issues with small denominators.

RELIABILITY: TESTING

4. **Did the developer conduct new reliability testing? ☒ Yes ☐ No**

   4a. If no, summarize the Standing Committee's previous feedback:

   - N/A

   4b. If yes, describe any differences between the new and old testing and summarize any relevant Standing Committee's feedback from the previous review:

   - In the previous submission, the developer's patient/encounter level validity testing stood for reliability and validity. The developer did not do updated testing, they noted that though the specifications have changed slightly, the data elements have not.
   - In this submission, the developer performed accountable entity level testing using a random split-sample methodology and reported an intraclass correlation coefficient (ICC). Additionally, they evaluated the relative ranking between split-samples and years reporting Kendall's Tau-b and Spearman's rank correlation coefficients.
5. **Reliability testing level: ☒ Accountable-Entity Level ☐ Patient/Encounter Level ☐ Neither**
6. **Reliability testing was conducted with the data source and level of analysis indicated for this measure:**

   ☒ Yes ☐ No

7. If accountable-entity level and/or patient/encounter level reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ **Yes** ☐ **No**

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

   - The developers presented reliability testing at the accountable-entity level using a random split-sample methodology and reporting an intraclass correlation coefficient (ICC). Additionally, they evaluated relative rankings between split-samples and years reporting Kendall's Tau-b and Spearman's rank correlation coefficients.

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

   - For reliability testing at the accountable entity level the developer reported the following results
     - ICC results per calendar year
       - The variation between split samples for each of the Medicaid programs is relatively small, signifying that the samples were similar. Additionally, the ICC scores were reported by year.
         - 2018 (n=14): ICC = 0.997, with a 95% confidence interval of (0.9923,0.9992) and a p-value of less than 0.0001
         - 2017 (n=14): ICC = 0.999, with a 95% confidence interval of (0.9962,0.9996) and a p-value of less than 0.0001
         - 2016 (n=10): ICC = 0.999, with a 95% confidence interval of (0.9983,0.9999) and a p-value of less than 0.0001
     - Relative rankings between split-samples
       - The developer reported that the relative rankings between split-samples remained stable over time.
     - Relative rankings between years
       - For 2017 & 2018 (n=14): Kendall's Tau-b = 0.8462 (p-value = <0.0001) and Spearman's rank = 0.9429 (p-value = <0.0001)
       - For 2016 & 2017 (n=10): Kendall's Tau-b = 0.6889 (p-value = 0.0073) and Spearman's rank = 0.8303 (p-value = 0.0029)
   - Previous submission testing (patient/encounter level)
     - In the previous submission, the developer used patient/encounter validity testing as reliability testing. The developer did not update validity patient/encounter level testing because the data elements within the measure did not change. Therefore, the original patient/encounter level testing that used chart reviews holds for this measure.
       - In the previous submission, the developer validated the encounter data by comparing claims data against dental charts. The developer noted that record and procedure codes on the claims had a 94.04 percent agreement rate.
       - Additionally, the developer assessed whether the preventive service of topical fluoride application was accurately captured by claims data. The developer reported concordance (89.9 percent), sensitivity (90.7 percent), specificity (88.4 percent), positive predictive value (93.5 percent), and negative predictive value (83.9 percent). The developer also reported a kappa statistic which was 0.782.

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

    ☒ **Yes** ☐ **No** ☐ **Not applicable**

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

☐ **Yes** ☐ **No** ☒ **Not applicable** (patient/encounter level testing was not performed)

12. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

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

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

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

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

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

- The ICC values were near perfect, and all the relative ranking coefficients were above 0.6. Additionally, though the developers provided a rationale for why program level data transfers to plan level data, that is ultimately up to the committee to determine appropriateness.

VALIDITY: TESTING

14. **Did the developer conduct new validity testing? ☒ Yes ☐ No**

14a. If no, summarize the Standing Committee's previous feedback:

- N/A

14b. If yes, describe any differences between the new and old testing and summarize any relevant Standing Committee's feedback from the previous review:

- The Standing Committee did not offer any specific feedback on the validity of the measure during its last review.
- In the previous submission, the developer's reported patient/encounter level validity testing as well as face validity testing.
- In this submission, they provided empirical construct validity by hypothesizing a positive relationship between this measure (NQF #2528) with NQF #2511 Utilization of Services, Dental Services and NQF #2517 Oral Evaluation, Dental Services.

15. **Validity testing level (check all that apply):**

☒ **Accountable-Entity Level** ☐ **Patient or Encounter-Level** ☐ **Both**

NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

16. **If patient/encounter level validity testing was provided, was the method described and appropriate for assessing the accuracy of ALL critical data elements? NOTE:** Data element validation from the literature is acceptable.

☐ **Yes**

☐ **No**

☒ **Not applicable** (patient/encounter level testing was not performed)

17. **Method of establishing validity at the accountable-entity level:**

☐ **Face validity**

☒ **Empirical validity testing at the accountable-entity level**

☐ **N/A (accountable-entity level testing not conducted)**

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

    ☒ **Yes**

    ☐ **No**

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

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

- The developers conducted accountable entity level empirical validity testing using two comparator measures and analyzing their correlation/hypothesized relationships with this measure. To do this, the developers calculated Kendall's Tau-b, which measures association based on number of discordant and concordant pairs. In this instance, they assessed concordance. They also reported Spearman's rank correlation but felt Kendall's tau was the more appropriate testing given the relatively small sample size
    - The developers did note that they were limited in available dental quality measures, because of this the testing could only be done against other NQF-endorsed DQA measures that are process measures.
    - NQF #2511 Utilization of Services, Dental Services
    - NQF #2517 Oral Evaluation, Dental Services

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

- For accountable entity level validity testing the developer reported the following results
    - Utilization of Services, Dental Services compared to Topical Fluoride
        - 2018 (n=14): Kendall's Tau-b = 0.6703 (p-value = 0.0010) and Spearman's rank = 0.8286 (p-value = 0.0003)
        - 2017 (n=14): Kendall's Tau-b = 0.6264 (p-value = 0.0022) and Spearman's rank = 0.8022 (p-value = 0.0006)
        - 2016 (n=10): Kendall's Tau-b = 0.8222 (0.0013) and Spearman's rank = 0.9394 (p-value = 0.0001)
    - Oral Evaluation, Dental Services
        - 2018 (n=14): Kendall's Tau-b = 0.7802 (p-value = 0.0001) and Spearman's rank = 0.9253 (p-value = <0.0001)
        - 2017 (n=14): Kendall's Tau-b = 0.8022 (p-value = 0.0001) and Spearman's rank = 0.9253 (p-value = <0.0001)
        - 2016 (n=10): Kendall's Taub = 0.8667 (p-value = 0.0007) and Spearman's rank = 0.9515 (p-value = <0.0001)
- Previous submission testing (patient/encounter level and face validity):
    - In the previous submission, the developer presented face validity which was scored on the Delphi rating system. The measure received a mean importance score of seven, mean feasibility score of eight, and a mean validity score of seven for specific evidence-based preventive services, all out of a nine-point scale. The developers notes that a rating of seven through nine is considered scientifically sound and valid.
    - While the developer did not provide updated data element testing, they did clarify that even though there was a denominator change, the critical data elements did not change.
        - In the previous submission, the developer validated the encounter data by comparing claims data against dental charts. They noted that record and procedure codes on the claims had a 94.04 percent correlation rate.

- Additionally, the developer assessed whether the preventive service of topical fluoride application was accurately captured by claims data. The developer reported concordance (89.91 percent), sensitivity (90.7 percent), specificity (88.4 percent), positive predictive value (93.5 percent), and negative predictive value (83.9 percent). The developer also reported a kappa statistic which was 0.782.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

21. **Please describe any concerns you have with measure exclusions.**
    - This measure has no exclusions.
22. **Risk Adjustment**

    22a. **Risk-adjustment method**

    ☒ None (only answer Question 20b and 20e) ☐ Statistical model ☐ Stratification

    ☐ Other method assessing risk factors (please specify)

    22b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

    ☐ Yes ☐ No ☐ Not applicable

    22c. **Social risk adjustment:**

    22c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☐ Not applicable

    22c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

    22c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

    22d.**Risk adjustment summary:**

    22d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

    22d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

    22d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

    22d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

    22d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

    22e**. Assess the risk-adjustment approach**
    - N/A
23. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    For cost/resource use measures, does this measure identify meaningful differences about cost and resource use between the measured entities?

    - The developer calculated the interquartile range for the measure rates and then conducted a chi-square test to evaluate the statistical significance of the differences in the measure scores between the lowest and highest quartiles. The following results were reported:
        - 2018 – Mean = 0.2079, interquartile range = 0.0595, p-value = less than 0.0001
        - 2017 – Mean = 0.1979, interquartile range = 0.0700, p-value = less than 0.0001
        - 2016 – Mean = 0.1809, interquartile range = 0.0724, p-value = less than 0.0001
    - The developer states that these results suggest that the measure can identify statistically and clinically meaningful differences in performance.
24. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**

- No concerns because only one data source is used.

25. **Please describe any concerns you have regarding missing data.**
    - No concerns with missing data.

For cost/resource use measures ONLY:

If not cost/resource use measure, please skip to question 25.

26. **Are the specifications in alignment with the stated measure intent?**

    ☐ **Yes** ☐ **Somewhat** ☐ **No (If "Somewhat" or "No", please explain)**

27. **Describe any concerns of threats to validity related to attribution, the costing approach, carve outs, or truncation (approach to outliers):**
28. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

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

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if accountable-entity level testing has NOT been conducted)

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

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

29. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**
    - The results suggest moderate validity.

FOR COMPOSITE MEASURES ONLY: Empirical analyses to support composite construction

30. **What is the level of certainty or confidence that the empirical analysis demonstrates that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct?**

    ☐ **High**

    ☐ **Moderate**

    ☐ **Low**

    ☐ **Insufficient**

31. **Briefly explain rationale for rating of EMPIRICAL ANALYSES TO SUPPORT COMPOSITE CONSTRUCTION**
    - N/A

ADDITIONAL RECOMMENDATIONS

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

Criteria 1: Importance to Measure and Report

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

---

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

[Response Begins]

Yes

[Yes Please Explain]

Since the last maintenance cycle, we have one significant update. The United States Preventive Services Task Force (USPSTF) updated its systematic review and recommendations on the *Prevention of Dental Caries in Children Younger than 5 Years*. Although these guidelines focus on fluoride varnish application by medical primary care providers, the systematic review examined the effectiveness of topical fluoride application in preventing dental caries among children younger than five years irrespective of the type of provider applying topical fluoride. The additional studies that the USPSTF identified in its review supported its prior and ongoing recommendation that "primary care clinicians apply fluoride varnish to the primary teeth of all infants and children starting at the age of primary tooth eruption" (USPSTF 2021).

Our original measure development and evidence submission relied on the American Dental Association (ADA) Council on Scientific Affairs evidence-based clinical recommendations, which were based on a systematic review and meta-analyses. These guidelines have not been updated since 2013 (Weyant et al. 2013).

The Centers for Medicare and Medicaid Services (CMS) that has adopted this measure for its Core Set reporting, and oral health measurement user community requested that the DQA re-consider the decision to restrict the denominator population to children at elevated caries risk. Consequently, we examined these systematic reviews and recommendations with a particular focus on the findings related to the benefits and harms of topical fluoride application relative to a child's caries risk. In the process of this re-evaluation, we also identified a Cochrane systematic review that had not been identified in our prior evidence review (Marinho 2013). This review was conducted at approximately the same time that the ADA conducted its review. A comparison of the included studies reveals significant overlap between these two systematic reviews. However, the Marinho meta-analyses include subgroup analyses that specifically addressed the question of whether the effect of fluoride varnishes differs according to the baseline level of caries severity. Consequently, we have added this review to our evidence submission.

Collectively, all three systematic reviews found evidence in support of topical fluoride as effective for caries prevention, and the **evidence statements** were not qualified by a patient's caries risk status. All of the reviews also noted very low risk of harm from professionally applied topical fluoride. There is variation in how this evidence has been translated into clinical practice recommendations.

(1) The ADA's clinical recommendations, upon which the measure was originally based, prioritize populations at elevated caries risk for professionally applied topical fluoride. These recommendations make the following observations:

- "Some of the studies were conducted before the 1970s, when dental caries rates were higher, the percentage of the population receiving fluoridated water was substantially lower, and the percentage of people use fluoridated dentrifice was much lower."
- "The study populations often could not be categorized in terms of caries risk, and the panel could not assign risk categories to the populations as they are defined today."
- "Patients at low risk of developing caries may not need additional fluoride interventions."
- "Professional judgment is required to interpret the clinical relevance of preventive measures for individual patients. The combination of evidence from clinical studies, the patient's caries risk status, the practitioner's professional judgment and the patient's needs and preferences should guide decision making."
- "Studies regarding caries prevention and the economic benefit of topical fluoride in different caries risk populations are needed."

(2) The Cochrane systematic review conducted by Marinho et al. (2013) concluded that "fluoride varnish, compared to placebo or no treatment, is effective in the prevention of caries in children and adolescents." Marihno et al specifically conducted subgroup analyses based on initial caries severity:

> "We explored this heterogeneity in addressing the second, third and fourth objectives of this review which were to examine whether there was any relationship between the caries-preventive effectiveness of fluoride varnish and the initial level of caries severity, background exposure to fluoride (water supply, dentifrice, other fluoride sources), concentration of fluoride, frequency of application and whether prophylaxis was undertaken prior to the application of the varnish. The univariate meta-regressions found no significant associations between any of these pre-specified factors and the estimates of D(M)FS or d(m)fs prevented fractions, despite substantial variations between trials in these factors"

Marinho et al.'s conclusions and practice implications regarding the benefits of topical fluoride application are not qualified by a patient's caries risk status.

(3) The USPSTF recommendations considered the role of caries risk status in its recommendation and concluded:

"All children with erupted teeth can potentially benefit from the periodic application of fluoride varnish, regardless of the levels of fluoride in their water. Although the evidence to support fluoride varnish is drawn from higher-risk populations, the provision of fluoride varnish to all children is reasonable since the prevalence of risk factors is high in the U.S. population."

Thus, the evidence supports the application of professionally applied topical fluoride as an effective caries prevention intervention. But practice guidance varies with respect to whether this intervention should be focused on higher risk populations.

Measure Modifications Based on Evidence Review and User Feedback

The findings of all three reviews support topical fluoride as effective for caries prevention. Only Marinho et al. specifically examined whether there was evidence of variations in effectiveness based on disease severity and did not find statistically significant associations.

With limited evidence regarding variations in the effectiveness of topical fluoride in preventing dental caries based on a child's risk factors for developing caries, the original measure focused on children for whom the evidence of effectiveness was expected to be greatest – those at elevated risk for dental caries.

The ADA clinical recommendations note that preventive interventions should be informed not only by the evidence, but also the provider's clinical judgment and an evaluation of the patient's individual caries risk status. Caries risk assessment takes into account multiple factors, including both biological risk factors and social determinants of health. The ideal methodology for assessing topical fluoride application for children at elevated caries risk is to base the identification of elevated risk on the findings of a provider's caries risk assessment as to whether the child is at low or elevated risk of dental caries.

When the measure was originally developed, caries risk assessment finding codes were in the process of being adopted into the Current Dental Terminology (CDT) dental procedure codes. Recognizing that documentation of these codes would take time, the original measure included two approaches for identifying elevated risk using administrative claims data. In addition to using the newly (at the time) developed caries risk assessment codes, an alternate (and previously established) methodology of looking for caries-related treatment codes was also used. This alternate methodology looked for procedure codes, such as restorations (i.e., fillings), that are indicative of caries-related treatments to identify children with prior caries experience, which is one of the primary predictors of future caries experience. Although this approach is an accepted, validated methodology for identifying patients with prior caries experience, it was recognized that it served as a sampling methodology and would not capture all children at elevated risk for dental caries as it captures only one important indication of risk, prior caries experience, which is identifiable only after caries related lesions have already occurred. This sampling approach may miss children who have multiple risk factors that are not reflected in claims data. It also fails to identify children who are at risk for developing caries but who have not yet experienced caries related lesions Key measure users, including the Centers for Medicare and Medicaid Services, have requested removal of the elevated risk methodology to support a more prospective and population-based approach to caries prevention.

The DQA's Measure Development and Maintenance Committee reviewed the user community's feedback and the evidence and re-affirmed that topical fluoride application is effective in caries prevention. It further determined that removing elevated risk would be aligned with the intent of the measure to support quality improvement efforts around primary caries prevention.

[Response Ends]

1a. Evidence

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

2021 Submission:

Updated evidence information here.

2018 Submission:

Evidence from the previous submission here.

1a.01. Provide a logic model.

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

[Response Begins]

2022 SUBMISSION

Topical Fluoride for Children, Dental Services indicates the percentage of children who received at least two topical fluoride applications as dental services during the reporting year. Evidence suggests that topical fluoride applied to children starting as early as six months of age is beneficial in preventing dental caries (USPTF 2021, Weyant et al. 2013; Marinho et al. 2013). This measure directly reflects the findings of systematic reviews regarding an effective caries prevention measure (professionally applied topical fluoride), including the frequency required for clinical effectiveness (at least every three-six months). As described in 1b.01 (Importance), dental caries is the most common chronic disease in children in the U.S., and a significant

percentage of children have untreated dental caries. Dental decay causes significant short- and long-term adverse consequences for children's health and functioning, including pain, tooth loss, impaired quality of life, and negative effects on school performance. As detailed below, professionally applied topical fluoride has demonstrated effectiveness in reducing caries among children, thereby improving oral health, overall health, and overall well-being. The main change to the measure from the prior submission is the removal of elevated risk criteria from the measure's denominator.

2014 SUBMISSION

Topical Fluoride for Children at Elevated Caries Risk indicates the percentage of children at moderate to high risk for caries who received at least two topical fluoride applications as dental services during the reporting year. Evidence suggests that topical fluoride applied to children starting as early as six months of age is beneficial in preventing dental caries (Weyant et al. 2013). Evidence-based clinical recommendations also suggest that topical fluoride should be applied at least every three to six months in children at elevated risk for caries (Weyant et al. 2013). This measure directly reflects evidence-based guidelines regarding an effective caries prevention measure (professionally applied topical fluoride), including the frequency required for clinical effectiveness (at least every three-six months). As described in 1b1 (Importance), dental caries is the most common chronic disease in children in the U.S. and a significant percentage of children have untreated dental caries. Dental decay causes significant short- and long-term adverse consequences for children's health and functioning. As detailed below, professionally applied topical fluoride has demonstrated effectiveness in reducing caries among children at elevated caries risk, thereby improving oral health, overall health, and overall well-being.

[Response Ends]

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

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

[Response Begins]

Clinical Practice Guideline recommendation (with evidence review)

US Preventive Services Task Force Recommendation

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

[Response Ends]

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

Evidence - Systematic Reviews Table (Repeatable)

Group 1 - Evidence - Systematic Reviews Table

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

[Response Begins]

2022 SUBMISSION:

Fluoride Varnishes for Preventing Dental Caries in Children and Adolescents (Review). Marinho VCC, Worthington HV, Walsh T, Clarkson JE. 2013. Cochrane Database of Systematic Reviews 2013, Issue 7. Art. No.: CD002279. DOI: 10.1002/14651858.CD002279.pub2. PMID: 23846772. Available at:

https://www.cochrane.org/CD002279/ORAL_fluoride-varnishes-for-preventing-dental-caries-in-children-and-adolescents

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Conclusions (abstract):

"The conclusions of this updated review remain the same as those when it was first published. The review suggests a substantial caries- inhibiting effect of fluoride varnish in both permanent and primary teeth, however the quality of the evidence was assessed as moderate, as it included mainly high risk of bias studies, with considerable heterogeneity."

Implications for Practice (specific section in conclusion section of review):

"This review has found that the application of fluoride varnishes two to four times a year, either in the permanent or primary dentition, is associated with a substantial reduction in caries increment. We found that this relative effect applies in populations with different levels of caries risk and exposure to other sources of fluoride. We also found no evidence that this relative effect was dependent on frequency of varnish application, length of follow-up, whether prophylaxis was undertaken prior to application of the varnish, concentration of fluoride in the varnish and use of a placebo rather than a no treatment control, although these results should be interpreted with caution. The review does not provide any information on the likelihood of side effects with this treatment and inconclusive information on acceptability."

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

The authors rate the quality of the evidence as **moderate**, which is defined as: "Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate."

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

High quality	Further research is very unlikely to change our confidence in the estimate of effect
Moderate quality	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low quality	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low quality	We are very uncertain about the estimate

Table with definitions of the Cochrane Review's evidence grading system

[Response Ends]

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

[Response Begins]

This is not a clinical guideline so there are no formal recommendations with assigned grades. As noted above, the authors offer their perspective on "implications for practice."

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Not applicable.

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Number of studies:

22 trials published between 1975-2012 in which a total of 12,455 children were randomized to treatment with either fluoride varnish or placebo/no treatment.

Type of studies:

- Included: Randomised or quasi-randomised controlled trials using or indicating blind outcome assessment, in which fluoride varnish is compared concurrently to a placebo or no treatment group during at least one year.
- Excluded: Randomised or quasi-randomised controlled trials using within-group paired comparison designs (e.g. split-mouth trials), or with open outcome assessment or no indication of blind outcome assessment, or lasting less than one year, or controlled trials where random or quasi-random allocation was not used or indicated.

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Measure of treatment effect: Prevented fraction (PF) was the measure of treatment effect presented for caries increment. The prevented fraction is calculated as the mean increment in the control group minus the mean increment in the intervention group divided by the mean increment in the control group. For an outcome such as caries increment (where discrete counts are considered to approximate to a continuous scale and are treated as continuous outcome), this measure was considered more appropriate than the mean difference or standardized mean difference since it allowed combination of different ways of measuring caries increment and a meaningful investigation of heterogeneity between trials. It is also simple to interpret.

Overall findings:

- The evidence from meta-analysis of the 13 trials assessing the effect of fluoride varnish on the permanent dentition is that the use of fluoride varnish is associated on average with a 43% (95% CI 30% to 57%) reduction in decayed, missing and filled tooth surfaces.
- The meta-analysis of the 10 trials assessing the effect of fluoride varnish on the primary dentition suggests a 37% (95% CI 24% to 51%) reduction in decayed, missing and filled tooth surfaces.

There was considerable statistical heterogeneity in both these estimates.

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

The authors collected adverse effects information from the included trials and noted the following:

- There was little information concerning possible adverse effects or acceptability of treatment.

Three studies provided data, reporting no adverse effects.

[Response Ends]

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

[Response Begins]

The recently released updated USPSTF guidelines included in Group 2 Evidence include more recent studies.

[Response Ends]

Group 2 - Evidence - Systematic Reviews Table

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

[Response Begins]

2022 SUBMISSION:

Full Report: Screening and Interventions to Prevent Dental Caries in Children Younger Than 5 Years, Final Recommendation Statement. US Preventive Services Task Force. December 7, 2021. *Available at:*

https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/prevention-of-dental-caries-in-children-younger-than-age-5-years-screening-and-interventions1#bootstrap-panel--12

Recommendation Statement: Screening and Interventions to Prevent Dental Caries in Children Younger Than 5 Years: US Preventive Services Task Force Recommendation Statement. U.S. Preventive Services Task Force. 2021. *JAMA.* 2021;326(21):2172–2178. doi:10.1001/jama.2021.20007. PMID: 34874412. Available at:

https://jamanetwork.com/journals/jama/fullarticle/2786823

Screening and Interventions to Prevent Dental Caries in Children Younger Than 5 Years: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. Chou R, Pappas M, Dana T, Selph S, Hart E, Fu RF, Schwarz E. 2021. JAMA. 2021 Dec 7;326(21):2179-2192. doi: 10.1001/jama.2021.15658. PMID: 34874413. Available at:

https://jamanetwork.com/journals/jama/fullarticle/2786824

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Recommendation:

"The USPSTF recommends that primary care clinicians apply fluoride varnish to the primary teeth of all infants and children starting at the age of primary tooth eruption."

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

The USPTF Recommendations specifically addressed the following questions:

- How Effective Are Preventive Interventions (Dietary Fluoride Supplementation, Topical Fluoride Application, Silver Diamine Fluoride, or Xylitol) in Preventing Dental Caries in Children Younger Than Age 5 Years?
- What Are the Harms of Specific Oral Health Interventions to Prevent Dental Caries in Children Younger Than Age 5 Years (Parental or Caregiver/Guardian Oral Health Education, Referral to a Dental Health Care Professional, and Preventive Interventions)?

We focus our reporting from this review on the studies and findings related to topical fluoride application specifically. These findings were reported separately within the review.

Evidence statement related to topical fluoride application: "The USPSTF concludes **with moderate certainty** that there is a **moderate net benefit** of preventing future dental caries with fluoride varnish application in all children younger than 5 years."

Moderate	The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: • The number, size, or quality of individual studies. • Inconsistency of findings across individual studies. • Limited generalizability of findings to routine primary care practice. • Lack of coherence in the chain of evidence. As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion.

Table providing the USPSTF definition of moderate

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Level of Certainty	Description
High	The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.
Moderate	The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: • The number, size, or quality of individual studies. • Inconsistency of findings across individual studies. • Limited generalizability of findings to routine primary care practice. • Lack of coherence in the chain of evidence. As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion.
Low	The available evidence is insufficient to assess effects on health outcomes. Evidence is insufficient because of: • The limited number or size of studies. • Important flaws in study design or methods. • Inconsistency of findings across individual studies. • Gaps in the chain of evidence. • Findings not generalizable to routine primary care practice. • Lack of information on important health outcomes. More information may allow estimation of effects on health outcomes.

Table presents the USPSTF grading system, level of certainty and description

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Population	Recommendation	Grade
Children younger than 5 years	The USPSTF recommends that primary care clinicians apply fluoride varnish to the primary teeth of all infants and children starting at the age of primary tooth eruption.	B

Table describes the USPSTF recommendation and grade

B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer or provide this service.

Table provides detailed explanation of grade B in the USPSTF recommendation

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Grade	Definition	Suggestions for Practice
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer or provide this service.
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer or provide this service.
C	The USPSTF recommends selectively offering or providing this service to individual patients based on professional judgment and patient preferences. There is at least moderate certainty that the net benefit is small.	Offer or provide this service for selected patients depending on individual circumstances.
D	The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.	Discourage the use of this service.

Grade	Definition	Suggestions for Practice
I Statement	The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.	Read the clinical considerations section of USPSTF Recommendation Statement. If the service is offered, patients should understand the uncertainty about the balance of benefits and harms.

Table outlines the USPSTF recommendation grading system, including definitions and suggestions for practice

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Number of studies:

32 studies and 1 systematic review (19 studies) were included in the entire review (covering all key questions). Comparisons were against placebo or no intervention. Outcomes were dental caries (incidence or caries burden, measured based on the number of decayed, missing, or filled teeth [dmft] or decayed, missing, or filled surfaces), morbidity, quality of life, and harms (including fluorosis).

1. 22 studies (representing 25 articles) addressed key question "How effective are preventive interventions (dietary fluoride supplementation, topical fluoride application, silver diamine fluoride, or xylitol) in preventing dental caries in children younger than 5 years?" **15 of these studies addressed topical fluoride application specifically.**
2. 4 (representing 6 articles) studies reported on adverse events associated with topical fluoride.

Type of studies:

- 15 studies on topical fluoride: RCTs (n=9,541); inconsistent (high statistical heterogeneity), precise, moderate strength of evidence

4 studies reporting on adverse events: RCTs (n=4,141); consistency cannot be determined (single trials reported different adverse events), precise; low-moderate strength of evidence

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

Based on 15 trials (5 trials in the prior USPSTF review and 10 new trials), topical fluoride (administered as fluoride varnish in all trials except for one) versus placebo or no varnish was associated with:

- decreased caries increment (13 trials, N=5733, mean difference -0.94, 95% CI, -1.74 to -0.34, I2=86%)

- decreased likelihood of incident caries (12 trials, N=8177, RR 0.80, 95% CI, 0.66 to 0.95, I2=79%; absolute risk difference [ARD] -7%, 95% CI, -12% to - 2%)

The authors identified high statistical heterogeneity across studies, but precise estimates.

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

- Four trials (n=4,141) reported no differences between fluoride varnish versus placebo or no varnish in risk of fluorosis or the likelihood of any adverse event.
- Two studies (n=2,864) reported children did not like the smell of the fluoride varnish, and one study reported that a few children vomited due to the smell, texture, or taste.
- These findings led the USPSTF to conclude: "There is adequate evidence to bound the harms for dietary fluoride supplementation and topical fluoride application as no greater than small, based on limited evidence of harms."

[Response Ends]

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

[Response Begins]

2022 SUBMISSION:

This recommendation from the USPSTF was released in December 2021.

[Response Ends]

Group 3 - Evidence - Systematic Reviews Table

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

[Response Begins]

2022 SUBMISSION:

Our original measure development and evidence submission relied on the American Dental Association (ADA) Council on Scientific Affairs evidence-based clinical recommendations, which were based on a systematic review and meta-analyses. These guidelines have not been updated since 2013 (Weyant et al. 2013) and remain part of the evidence base supporting the measure.

2014 SUBMISSION:

Full Report: Weyant RJ, Tracy SL, Anselmo TT, Beltrán-Aguilar ED, et al; American Dental Association Council on Scientific Affairs Expert Panel on Topical Fluoride Caries Preventive Agents. Topical fluoride for caries prevention: full report of the updated clinical recommendations and supporting systematic review. Available at:

http://ebd.ada.org/contentdocs/Topical_fluoride_for_caries_prevention_2013_update_-_full_manuscript.pdf

Condensed version: Weyant RJ, Tracy SL, Anselmo TT, Beltrán-Aguilar ED, et al; American Dental Association Council on Scientific Affairs Expert Panel on Topical Fluoride Caries Preventive Agents. J Am Dent Assoc. 2013 Nov;144(11):1279-91. Topical fluoride for caries prevention: executive summary of the updated clinical recommendations and supporting systematic review. Available at:

http://ebd.ada.org/contentdocs/JADA_updated_executive_summary_Nov_2013.pdf

[Response Ends]

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

[Response Begins]

2014 SUBMISSION

"Clinical recommendations for use of professionally applied or prescription-strength, home-use topical fluorides for caries prevention in patients at elevated risk of developing caries." The recommendations are "in favor" for:

- "2.26 percent fluoride varnish at least every three to six months" for children younger than 6 years
- "2.26 percent fluoride varnish at least every three to six months OR 1.23 percent fluoride (APF) gel for four minutes at least every three to six months" for children 6-18 years

(Weyant et al., 2013, Table 4)

[Response Ends]

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

[Response Begins]

2014 SUBMISSION

The recommendations are based on the following evidence statements that were graded as **moderate** level of certainty:

- < 6 years: There is a benefit of 2.26 percent fluoride varnish application at least twice per year for caries prevention.
- 6-18 years: There is a benefit of 2.26 percent fluoride varnish application at least twice per year for caries prevention.
- 6-18 years: There is a benefit of APF gel (1.23 percent fluoride) application up to every three months for four minutes for caries.

    † No studies were found regarding professionally applied fluoride APF gels with an application time of less than three minutes.

Moderate: "This statement is based on preliminary determination from the current best available evidence; as more information becomes available, the magnitude or direction of the observed effect could change, and this change could be large enough to alter the conclusion.

[Response Ends]

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

[Response Begins]

2014 SUBMISSION

High: This statement is strongly established by the best available evidence; the conclusion is unlikely to be affected strongly by the results of future studies. The body of evidence usually includes consistent results from well-designed, well-conducted studies in representative populations. This conclusion is unlikely to be strongly affected by the results of future studies.

Moderate: This statement is based on preliminary determination from the current best available evidence, but confidence in the estimate is constrained by one or more factors, such as: the number, size, or risk of bias of individual studies; inconsistency of findings across individual studies; limited applicability due to the populations of interest; or lack of coherence in the chain of evidence. As more information becomes available, the magnitude or direction of the observed effect could change, and this change could be large enough to alter the conclusion."

Low: The available evidence is insufficient to support the statement, or the statement is based on extrapolation from the best available evidence. Evidence is insufficient or the reliability of estimated effects is limited by factors such as: the limited number or size of studies; important flaws in study design or methods leading to high risk of bias; inconsistency of findings across individual studies; gaps in the chain of evidence; findings not applicable to the populations of interest; or a lack of information on important health outcomes. More information could allow a reliable estimation of effects on health outcomes. (Weyant 2013, full report, pp. 18-19)

The grading system was adapted from that used by the U.S. Preventive Services Task Force

[Response Ends]

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

[Response Begins]

2014 SUBMISSION

For children at elevated risk of developing caries, the recommendations are "in favor" for:

- "2.26 percent fluoride varnish at least every three to six months" for children younger than 6 years
- "2.26 percent fluoride varnish at least every three to six months OR 1.23 percent fluoride (APF) gel for four minutes at least every three to six months" for children 6-18 years

(Weyant et al., 2013, p. 1282, Table 4)

Grade: The grade for the recommendations is **"in favor"** which is **defined** as: "Evidence favors providing this intervention." This is the **second highest recommendation out of a six-point scale**. The grading system was adapted from that used by the U.S. Preventive Services Task Force. (Weyant et al. 2013, p. 1281, Table 3)

[Response Ends]

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

[Response Begins]

2014 SUBMISSION

Strong: Evidence strongly supports providing this intervention.

In Favor: Evidence favors providing this intervention.

Weak: Evidence suggests implementing this intervention after alternatives have been considered.

Expert Opinion For:† Evidence is lacking; the level of certainty is low. Expert opinion guides this recommendation

Expert Opinion Against:† Evidence is lacking; the level of certainty is low. Expert opinion suggests not implementing this intervention.

Against: Evidence suggests not implementing this intervention or discontinuing ineffective procedure

† The USPSTF system defines this category of evidence as "insufficient"; "grade I indicates that the evidence is insufficient to determine the relationship between benefits and harms (i.e., net benefit)." The corresponding recommendation grade "I" is defined as follows: "The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined."

[Response Ends]

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

[Response Begins]

2014 SUBMISSIONs

71 studies included in evidence reviews, representing 82 citations. All studies included were controlled clinical trials.

- 17 randomized and five non-randomized clinical trials evaluated 2.26% fluoride varnish with control groups of no treatment, oral health counseling or placebo varnish. (6 randomized, 2 non-randomized concerned primary dentition; 11 randomized, 2 non-randomized concerned permanent; 1 combined results) The varnish was professionally applied every 3 to 12 months, with the majority of studies applying varnish every 6 months.
- 11 randomized and 4 non-randomized clinical trials evaluated 1.23% APF gel with comparison groups of no treatment, placebo, prophylaxis or non-fluoride placebo gel. All studies except one were on permanent teeth. All studies applied fluoride gel for four minutes.

[Response Ends]

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

[Response Begins]

2014 SUBMISSION

Recommendations:

- "2.26 percent fluoride varnish at least every three to six months" for children younger than 6 years
- "2.26 percent fluoride varnish at least every three to six months OR 1.23 percent fluoride (APF) gel for four minutes at least every three to six months" for children 6-18 years

Estimates of Benefit in Support of Recommendations:

(1) 2.26% Fluoride Varnish

"The results of meta-analyses for primary teeth indicate that the application of 2.26% fluoride varnish has a statistically significant effect (SMD -0.19 [95% CI: -0.31, -0.08) on caries prevention as measured by increment or incidence using surface-level data." Weyant et al., full report, 2013, p. 25

"The results of meta-analyses for permanent teeth indicate that 2.26% fluoride varnish has a statistically significant effect (SMD= -0.38 [95% CI: -0.53, -0.24]) on caries prevention as measured by increment or incidence using surface-level data." Weyant et al., full report, 2013, p. 25

Evidence Profile (Weyant et al., full report, 2013, pp. 26-27):

(a) Primary teeth (children under age 6):

- Level of certainty: Moderate
- Benefit: Yes (smaller caries increment or incidence with topical fluoride use).
    - Standardized mean difference=-0.19 [-0.31, -0.08]
    - Prevented fraction=0.27
    - Number needed to treat for control rate of 1 DMFS per year = 4
- Adverse events or harms: Little potential for harms if swallowed
- Benefit-harm assessment (Net benefit rating): Benefits outweigh potential harms
- Strength of clinical recommendation: In favor

(b) Permanent teeth (children):

- Level of certainty: Moderate
- Benefit: Yes (smaller caries increment or incidence with topical fluoride use).
    - Standardized mean difference=-0.38 [-0.53, -0.24]
    - Prevented fraction=0.36
    - Number needed to treat for control rate of 1 DMFS per year = 3
- Adverse events or harms: None if used as manufacturers recommend
- Benefit-harm assessment (Net benefit rating): Benefits outweigh potential harms
- Strength of clinical recommendation: In favor

The table below (Table 8 from the report) summarizes the findings.

Outcome Measures	Number and type† of studies	Number of participants††	Standardized Mean Difference [95% Confidence Interval] (negative favors intervention, positive favors control)
Meta-analysis results: Primary teeth	*	*	*
d(e/m)fs, increment or incidence+	6 RCT and 2 CCT	3, 409**	**-0.19 [-0.31, -0.08]**
Meta-analysis results: Permanent teeth	*	*	*
D(M)FS, increment or incidence+	8 RCT and 1 CCT	2, 574	**-0.38 [-0.53, -0.24]**
Root caries, meta-analysis results	*	*	*
Root caries increment	2 RCT	132	**-0.67 [-1.14, -0.20]**

Outcome Measures	Number and type† of studies	Number of participants††	Standardized Mean Difference [95% Confidence Interval] (negative favors intervention, positive favors control)
Individual study results	*	*	*
Combined dentition	1 CCT	390	**DMFS + dmfs: -1.47 [-1.70, -1.25]** **DMFT + dmft: -1.15 [-1.37, -0.94]**
DMFT	1 CCT	77	**-0.13 [-0.58, 0.32]**
DS occlusal surfaces	1 RCT	79	**-0.54 [-1.06, -0.03]**

Table summarizes the systematic review's findings, including both meta-analysis results and individual study results.

*Cell left intentionally blank.

Notes: †RCT = randomized controlled trial; CCT = controlled clinical trial (non-randomized); ††including all participants (not using cluster-adjusted number of participants or numbers of clusters); +all stages used if cavitated data not available, parentheses indicate the component was included in some of the combined results and not others.

(2) 1.23% fluoride (APF) gel

"The panel concluded with moderate certainty that there is a benefit of APF gel (1.23% fluoride) application up to every three months for 4 minutes for caries prevention in the permanent teeth of 6-14 year olds. This statement is based on meta-analysis of 12 studies with moderate to high bias scores and including over 4,000 participants; although there was some inconsistency, there was low statistical heterogeneity (I2=43) between the studies." (Weyant, full report, 2013, p. 33)

Evidence Profile (Weyant et al., full report, 2013, p. 34):

Permanent teeth (children):

- Level of certainty: Moderate
- Benefit: Yes (smaller caries increment or incidence with topical fluoride use).
    - Standardized mean difference=-0.25 [-0.33, -0.16]
    - Prevented fraction=0.27
    - Number needed to treat for control rate of 1 DMFS per year = 4
- Adverse events or harms: None if used as manufacturers recommend
- Benefit-harm assessment (Net benefit rating): Benefits outweigh potential harms
- Strength of clinical recommendation: In favor

[Response Ends]

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

[Response Begins]

2014 SUBMISSION:

Potential harms evaluated included: (1) nausea and vomiting associated with the ingestion of topical fluorides and (2) dental fluorosis while tooth enamel is developing until approximately age 6, due to daily ingestion of topical fluoride, such as from toothpaste or from prescription home gels.

"There is less of a concern with professionally-applied topical fluorides that have much longer intervals between applications [citing Wong et al. 2010]. Additionally, fluoride varnish has less potential for harms than other forms of high concentration topical fluoride because the amount of fluoride that is placed in the mouth with fluoride varnish is approximately one-tenth that of other professionally-applied products [citing Beltran-Aguilar et al. 2000]. The panel judged that the benefits outweighed the potential for harms for all professionally-applied or prescription-strength topical fluorides and age groups except for children under age 6, where the risk of swallowing and associated events (particularly nausea and vomiting) outweighed the potential benefits for all professionally-applied or prescription-strength topical fluorides except 2.26% fluoride varnish." (Weyant et al., 2013, p. 10)

Citations

Beltran-Aguilar ED, Goldstein JW, Lockwood SA. Fluoride varnishes - A review of their clinical use, cariostatic mechanism, efficacy and safety. JADA 2000;131(May):589-96.

Weyant RJ, Tracy SL, Anselmo TT, Beltrán-Aguilar ED, et al; American Dental Association Council on Scientific Affairs Expert Panel on Topical Fluoride Caries Preventive Agents. Topical fluoride for caries prevention: full report of the updated clinical recommendations and supporting systematic review. Available at:

http://ebd.ada.org/contentdocs/Topical_fluoride_for_caries_prevention_2013_update_-_full_manuscript.pdf

Wong MC, Glenny AM, Tsang BW, et al. Topical fluoride as a cause of dental fluorosis in children. Cochrane Database of Systematic Reviews 2010;Jan 20(1).

[Response Ends]

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

[Response Begins]

2022 SUBMISSION

The recently released updated USPSTF guidelines included in Group 2 Evidence include more recent studies.

2014 SUBMISSION

This review was published on November 2013 and reflects the latest evidence.

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

Dental caries remains one of the most common, yet preventable, diseases of childhood. As noted in our original submission, dental decay children has significant short- and long-term adverse consequences on children's health and overall well-being. Updated national surveillance data for the period 2011-2016 indicates that 23% of children aged 2-5 years experience dental caries related lesions, increasing to 52% among children aged 6- 8 years. Untreated decay was 10% among children aged 2-5 years and 16% among children aged 6-8 years. On permanent teeth, the prevalence of caries related tooth lesions was 17% for children 6-11 years and 57% among adolescents aged 12-19 years. Low-income children are twice as likely to have dental decay as higher-income children. Untreated decay was 5% among children aged 6-11 years and 17% among adolescents aged 12-19 years. (Centers for Disease Control and Prevention, 2019)

Oral health disparities are well documented and persist. Poor and near-poor children (children living in households with <=200% of the federal poverty level) are approximately twice as likely to have dental caries and untreated decay compared to higher income children (>200% FPL). National surveillance data indicate that Mexican American and non-Hispanic black children are more likely to have dental caries and untreated decay than non-Hispanic white children. For example, the prevalence of dental caries related lesions in primary teeth among Mexican American children aged 6-8 years was 73% compared to 44% for non-Hispanic white children. (Centers for Disease Control and Prevention, 2019)

Although dental caries can managed and caries related lesions can be treated and restored, it is important to prevent the disease process from developing in the first place. As noted in the evidence section, multiple systematic reviews with meta-analyses find evidence supporting professionally applied topical fluoride, starting as early as six months of age and applied at least twice per year, as beneficial in preventing dental caries and associated decay(USPSTF 2021, Marinho 2013, Weyant 2013).

References

Centers for Disease Control and Prevention. Oral Health Surveillance Report: Trends in Dental Caries and Sealants, Tooth Retention, and Edentulism, United States, 1999–2004 to 2011–2016. Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services; 2019. Available at:

https://www.cdc.gov/oralhealth/publications/OHSR-2019-index.html

Marinho VCC, Worthington HV, Walsh T, Clarkson JE. 2013. Fluoride Varnishes for Preventing Dental Caries in Children and Adolescents (Review). Cochrane Database of Systematic Reviews 2013, Issue 7. Art. No.: CD002279. DOI: 10.1002/14651858.CD002279.pub2. PMID: 23846772. Available at:

https://www.cochrane.org/CD002279/ORAL_fluoride-varnishes-for-preventing-dental-caries-in-children-and-adolescents .

US Preventive Services Task Force. Screening and Interventions to Prevent Dental Caries in Children Younger Than 5 Years, Final Recommendation Statement. December 7, 2021. *Available at:*

https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/prevention-of-dental-caries-in-children-younger-than-age-5-years-screening-and-interventions1#bootstrap-panel--12

Weyant RJ, Tracy SL, Anselmo TT, Beltrán-Aguilar ED, et al; American Dental Association Council on Scientific Affairs Expert Panel on Topical Fluoride Caries Preventive Agents. J Am Dent Assoc. 2013 Nov;144(11):1279-91. Topical fluoride for caries prevention: executive summary of the updated clinical recommendations and supporting systematic review.

PRIOR SUBMISSION

Inequalities in oral health status and inadequate use of oral health care services are well documented. Dental caries is the most common chronic disease in children in the United States (NCHS 2012). In 2009–2010, 14% of children aged 3 –5 years had untreated dental caries. Among children aged 6–9 years, 17% had untreated dental caries, and among adolescents aged 13–15, 11% had untreated dental caries (Dye, Li, and Thorton-Evans 2012). Dental decay among children has significant short- and long-term adverse consequences (Tinanoff and Reisine 2009). Childhood caries is associated with increased risk of future caries (Gray, Marchment, and Anderson 1991; O´Sullivan and Tinanoff 1996; Reisine, Litt, and Tinanoff 1994), missed school days (Gift, Reisine, and Larach 1992; Hollister and Weintraub 1993), hospitalization and emergency room visits (Griffin et al. 2000; Sheller, Williams, and Lombardi 1997) and, in rare cases, death (Casamassimo et al. 2009).

Identifying caries early is important to reverse the disease process, prevent progression of caries, and reduce incidence of future lesions. Evidence suggests that topical fluoride applied to children starting as early as six months of age is beneficial in preventing dental caries (Weyant et al. 2013). However, approximately three quarters of children younger than age 6 years did not have at least one visit to a dentist in the previous year (Edelstein & Chinn 2009). Evidence-based clinical recommendations suggest that topical fluoride should be applied at least every three to six months in children at elevated risk for caries (Weyant et al. 2013).

The proposed measure, Topical Fluoride for Children at Elevated Caries Risk – Dental Services, captures whether children at moderate or high caries risk received at least two topical fluoride applications as dental services. Because topical fluoride is indicated at 3-6 month intervals (2-4 times per year) for children at elevated caries risk, at least two applications are indicated during the reporting year. This measure directly reflects evidence-based guidelines regarding an effective caries prevention measure (professionally applied topical fluoride), including the frequency required for clinical effectiveness (at least every three-six months). Topical Fluoride allows plans and programs to assess whether children at risk for caries are receiving evidence-based preventive services and target performance improvement initiatives accordingly.

Note: Procedure codes contained within claims data are the most feasible and reliable data elements for quality metrics in dentistry, particularly for developing programmatic process measures to assess the quality of care provided by programs (e.g., Medicaid, CHIP) and health/dental plans. In dentistry, diagnostic codes are not commonly reported and collected, precluding direct outcomes assessments. Although some programs are starting to implement policies to capture diagnostic information, evidence-based process measures are the most feasible and reliable quality measures at programmatic and plan levels at this point in time.

[Complete citations provided in 1c4 and in Evidence Submission Form.]

[Response Ends]

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

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

Data Source Description

We used Medicaid enrollment and claims data contained within the Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAFs) available from the Centers for Medicare & Medicaid Services (https://www.medicaid.gov/medicaid/data-systems/macbis/transformed-medicaid-statistical-information-system-t-msis/index.html).

Dates: Calendar years 2016, 2017, 2018

Number of Measured Entities: Data from 14 state Medicaid programs were included for this submission: Alaska, Arizona, Delaware, Idaho, Michigan, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, Oregon, South Carolina, Washington, and Wyoming. These states were selected based both on the quality of their data submissions to CMS and because they represent diversity in geographic location, population size, population demographic characteristics, and Medicaid dental delivery system.

Number of Patients, all measured entities included in testing, by year:

Total enrollees 0-20 years across all 14 programs:

2018: 7,720,412

2017: 7,854,440

2016: 7,850,885

In all cases, statewide program data are used (i.e., location is statewide)

Number of Patients, 0-20 Years Enrolled at Least 1 Month, by State Medicaid Program and Year:

Medicaid Program	Dates	# Mem	Dental Delivery
Alaska	CY 2018	101,273	FFS
*	CY 2017	99,296	FFS
*	CY 2016	94,550	FFS
Arizona	CY 2018	974,161	Managed care carve in
*	CY 2017	994,391	Managed care carve in

Medicaid Program	Dates	# Mem	Dental Delivery
*	CY 2016	981,695	Managed care carve in
Delaware	CY 2018	118,646	FFS
*	CY 2017	118,295	FFS
*	CY 2016	120,348	FFS
Idaho	CY 2018	214,879	Dental only PAHP
*	CY 2017	220,084	Dental only PAHP
*	CY 2016	201,253	Dental only PAHP
Michigan	CY 2018	1,163,658	Dental only PAHP
*	CY 2017	1,182,388	Dental only PAHP
*	CY 2016	1,182,388	Dental only PAHP
Mississippi	CY 2018	444,432	Managed care carve in
*	CY 2017	456,123	Managed care carve in
*	CY 2016	492,813	Managed care carve in
Nevada	CY 2018	379,289	Dental only PAHP & FFS
*	CY 2017	378,460	FFS
*	CY 2016	370,394	Managed care carve in & FFS
New Mexico	CY 2018	376,379	Managed care carve in
*	CY 2017	387,255	Managed care carve in
*	CY 2016	383,056	Managed care carve in
North Carolina	CY 2018	1,231,829	FFS
*	CY 2017	1,259,699	FFS

Medicaid Program	Dates	# Mem	Dental Delivery
*	CY 2016	1,241,882	FFS
Oklahoma	CY 2018	531,222	PCCM/FFS
*	CY 2017	553,905	PCCM/FFS
*	CY 2016	557,138	PCCM/FFS
Oregon	CY 2018	435,074	Dental only PAHP
*	CY 2017	463,301	Dental only PAHP
*	CY 2016	479,469	Dental only PAHP
South Carolina	CY 2018	767,719	FFS
*	CY 2017	762,747	FFS
*	CY 2016	752,206	FFS
Washington	CY 2018	932,270	FFS
*	CY 2017	945,583	FFS
*	CY 2016	939,142	FFS
Wyoming	CY 2018	49,581	FFS
*	CY 2017	52,127	FFS
*	CY 2016	54,551	FFS

Table showing program enrollment and dental delivery system type for 14 state Medicaid programs in each year 2016 through 2018

*Cell left intentionally blank

Performance Scores

(1) Performance scores, overall summary for all included state Medicaid programs

*	Mean	SD	Minimum	10th Percentile	25th Percentile	Median	75th Percentile	90th Percentile	Maximum
2018 (n=14)	0.21	0.04	0.14	0.16	0.18	0.21	0.24	0.26	0.28

*	Mean	SD	Minimum	10th Percentile	25th Percentile	Median	75th Percentile	90th Percentile	Maximum
2017 (n=14)	0.20	0.04	0.14	0.14	0.16	0.19	0.23	0.25	0.27
2016 (n=10)	0.18	0.06	0.09	0.09	0.15	0.17	0.23	0.26	0.27

Table summarizing descriptive statistics for performance scores of 14 state Medicaid programs.

*Cell left intentionally blank

(2) Performance Scores with 95% CIs by State Medicaid Program and Year:

CY 2018 [lowest to highest performing]

Program	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
AK, 2018	14.46%	(	0.0013	,	0.1421	,	0.1472	)
NV, 2018	15.95%	(	0.0008	,	0.1579	,	0.1610	)
MI, 2018	16.20%	(	0.0004	,	0.1612	,	0.1628	)
WY, 2018	17.77%	(	0.0024	,	0.1731	,	0.1824	)
OR, 2018	17.82%	(	0.0007	,	0.1767	,	0.1796	)
OK, 2018	19.23%	(	0.0008	,	0.1908	,	0.1938	)
AZ, 2018	19.33%	(	0.0005	,	0.1923	,	0.1942	)
SC, 2018	22.17%	(	0.0006	,	0.2206	,	0.2228	)
MS, 2018	22.88%	(	0.0008	,	0.2272	,	0.2303	)
NC, 2018	23.24%	(	0.0005	,	0.2315	,	0.2333	)
DE, 2018	23.72%	(	0.0015	,	0.2342	,	0.2402	)
ID, 2018	24.87%	(	0.0011	,	0.2465	,	0.2509	)
NM, 2018	25.62%	(	0.0009	,	0.2545	,	0.2579	)
WA, 2018	27.79%	(	0.0005	,	0.2769	,	0.2790	)

Table showing performance scores for each program in calendar year 2018, with standard deviation and 95% confidence intervals

*Cell left intentionally blank

CY 2017 [lowest to highest performing]

Program	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
NV, 2017	13.96%	(	0.0007	,	0.1381	,	0.1411	)
AK, 2017	14.45%	(	0.0013	,	0.1419	,	0.1471	)
MI, 2017	16.06%	(	0.0004	,	0.1598	,	0.1614	)

Program	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
OR, 2017	17.64%	(	0.0007	,	0.1749	,	0.1779	)
WY, 2017	17.75%	(	0.0024	,	0.1729	,	0.1821	)
OK, 2017	18.71%	(	0.0007	,	0.1857	,	0.1886	)
AZ, 2017	18.89%	(	0.0005	,	0.1879	,	0.1898	)
ID, 2017	19.79%	(	0.0010	,	0.1960	,	0.1999	)
MS, 2017	20.34%	(	0.0007	,	0.2020	,	0.2048	)
SC, 2017	22.40%	(	0.0006	,	0.2229	,	0.2251	)
NC, 2017	23.06%	(	0.0005	,	0.2297	,	0.2315	)
DE, 2017	23.90%	(	0.0015	,	0.2360	,	0.2420	)
NM, 2017	25.11%	(	0.0008	,	0.2494	,	0.2527	)
WA, 2017	27.45%	(	0.0005	,	0.2734	,	0.2755	)

Table showing performance scores for each program in calendar year 2017, with standard deviation and 95% confidence intervals.

*Cell left intentionally blank

CY 2016 [lowest to highest performing]

Program	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
NV, 2016	8.94%	(	0.0006	,	0.0882	,	0.0906	)
AZ, 2016	9.16%	(	0.0003	,	0.0909	,	0.0922	)
OR, 2016	15.22%	(	0.0007	,	0.1509	,	0.1535	)
MI, 2016	15.35%	(	0.0004	,	0.1527	,	0.1542	)
AK, 2016	16.13%	(	0.0015	,	0.1584	,	0.1642	)
WY, 2016	17.44%	(	0.0022	,	0.1701	,	0.1787	)
NC, 2016	22.30%	(	0.0004	,	0.2222	,	0.2239	)
DE, 2016	22.58%	(	0.0015	,	0.2229	,	0.2288	)
NM, 2016	23.98%	(	0.0008	,	0.2383	,	0.2414	)
WA, 2016	27.40%	(	0.0005	,	0.2729	,	0.2750	)

Table showing performance scores for each program in calendar year 2016, with standard deviation and 95% confidence interval

*Cell left intentionally blank

PRIOR SUBMISSION

Data Sources:

We used data from four sources and refer to “program” level information and “plan” level information. We included data for publicly insured children in the Texas Medicaid, Florida CHIP, and Florida Medicaid programs

as well as national commercial data from Dental Service of Massachusetts, Inc. Florida and Texas represent two of the largest and most diverse states. The two states also represent the upper and lower bounds of dental utilization based on dental utilization data available from the Centers for Medicare and Medicaid Services. The four programs collectively represent different delivery system models. The Texas Medicaid data represented dental fee-for-service. The Florida CHIP data included data from two dental MCOs. The Florida Medicaid data include dental fee-for-service and prepaid dental data. The commercial data included members in indemnity and preferred provider organization (PPO) product lines. Data from calendar years 2010 and 2011 were used for all programs except Florida Medicaid. Full-year data for CY 2011 were not available for Florida Medicaid. Therefore, we report only CY 2010 data for Florida Medicaid.

In the data summaries, "Programs" refer to population data from (1) Texas Medicaid, (2) Florida CHIP, (3) Commercial Data, and (4) Florida Medicaid. "Plans" refer to data from the two dental plans that served Florida CHIP members in both 2010 and 2011.

Below we provide summary data for each of the four programs and two plans individually.

Programs

Our source data for the testing included children 0-20 years in each program. The numbers of children ages 0-20 years enrolled at least one month in each program were as follows:

Texas Medicaid, 2011: 3,544,247

Texas Medicaid, 2010: 3,393,963

Florida CHIP, 2011: 317,146

Florida CHIP, 2010: 315,975

Commercial, 2011: 184,152

Commercial, 2010: 189,968

Florida Medicaid, 2010: 2,068,670

Plans

Within these programs, we had claims data available in both years for two dental managed care plans in Florida CHIP. We also report rates for those two plans separately.

Plan 1, 2010: 77,255

Plan 2, 2010: 116,388

Plan 1, 2011: 140,986

Plan 2, 2011: 168,191

Data 1b.2. Performance Scores for Topical Fluoride, Dental Services

Program, Year, Measure Score as % (Measure Score, SD, Lower 95% CI, Upper 95% CI)

Program 1, CY 2011: 37.13% (0.3713 , 0.0004 , 0.3704 , 0.3722)

Program 2, CY 2011: 27.15% (0.2715 , 0.0020 , 0.2676 , 0.2754)

Program 3, CY 2011: 22.04% (0.2204 , 0.0020 , 0.2165 , 0.2243)

Program 1, CY 2010: 34.96% (0.3496 , 0.0005 , 0.3487 , 0.3505)

Program 2, CY 2010: 22.63% (0.2263 , 0.0019 , 0.2225 , 0.2301)

Program 3, CY 2010: 35.04% (0.3504 , 0.0023 , 0.3458 , 0.3550)

Program 4, CY 2010: 18.16% (0.1816 , 0.0009 , 0.1799 , 0.1833)

Plan 1, CY 2011: 25.50% (0.2550 , 0.0030 , 0.2491 , 0.2609)

Plan 2, CY 2011: 28.69% (0.2869 , 0.0027 , 0.2815 , 0.2923)

Plan 1, CY 2010: 23.24% (0.2324 , 0.0048 , 0.2230 , 0.2418)

Plan 2, CY 2010 : 23.76% (0.2376 , 0.0034 , 0.2309 , 0.2443)

The measure score range of 18% to 35% in CY 2010 (year in which data were available for all four programs) indicates a significant performance gap overall. Two-thirds or more of children identified as being at elevated risk for caries do not receive the evidence-based recommendations of at least two topical fluoride applications during the reporting year. In addition, these results demonstrate the ability of the measure to identify variations in performance between programs.

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

In 2018, the performance on this measure ranged from 14% of children receiving at least two topical fluoride applications during the calendar year in the lowest performing state to 28% in the highest performing state. Consequently, there was a two-fold variation between the highest and lowest performing states, indicating significant variation across state Medicaid programs. Moreover, 72% of children in the highest performing state did not receive at least two fluoride applications in 2018, demonstrating a significant opportunity for overall performance improvement in all state Medicaid programs.

These results are consistent with our original testing data on a more limited set of programs, which also found two-fold variation between the highest and lowest performing programs as well as an overall performance gap.

PRIOR SUBMISSION

The measure testing findings are consistent with other data indicating that children have sub-optimal utilization of dental services in general and preventive dental services in particular. Although comprehensive dental benefits are covered under Medicaid and the Children's Health Insurance Program (CHIP), there are significant variations in use of dental services overall across states, ranging from approximately 25% to 69% (CMS EPSDT Data, FY 2011). Similar variation between states is observed among children 0-20 years of age enrolled in commercial dental plans (ADA 2013). With respect to preventive dental services more specifically, 14% to 58% of children enrolled in Medicaid/CHIP for at least 90 continuous days receive any preventive dental services (CMS EPSDT Data, FY 2011). Even among the highest performing states, 42% of publicly-insured children do not receive any type of preventive dental service during the year.

[Complete citations provided in 1c4 and in Evidence Submission Form Template.]

[Response Ends]

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

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

The same data source, measured entities, and patients were used as described in 1b.03) above.

We report stratified measure scores for CY 2018.

We have reported the measure scores by the following four demographic variables:

- age (required stratification),
- rural/urban geographic location based on the patient's zip code,
- race and ethnicity, and
- sex assigned at birth.

We do not report stratifications where data are missing > 10% for the stratification variable. Table 1b.04-A summarizes the extent of missing data across all 50 state Medicaid programs plus the District of Columbia (n=51) as well as among our sample of 14 state Medicaid programs specifically. We note the ongoing deficiencies in race and ethnicity in state Medicaid enrollment data.

Table 1b.04-A Percentage of States Plus District of Columbia Missing Data on Age, Geographic Location, Race/Ethnicity and Sex, CY 2018

*	missing<=10%	10%<missing<=20%	20%<missing<=50%	missing>50%
All 50 States + DC (n=51)	*	*	*	*
Age	100%	0%	0%	0%
Biological Sex	100%	0%	0%	0%
Geographic Location	96%	0%	2%	2%
Race/Ethnicity	31%	20%	37%	12%
Testing Sample (n=14)	*	*	*	*
Age	100%	0%	0%	0%
Biological Sex	100%	0%	0%	0%
Geographic Location	100%	0%	0%	0%
Race/Ethnicity	50%	21%	21%	7%

Table showing percentage of state Medicaid programs missing data on age, geographic location, race/ethnicity and sex for calendar year 2018

*Cell left intentionally blank

Table 1b.04-B below provides descriptive statistics that summarize performance scores for all 14 state Medicaid programs stratified by age, biological sex, geographic location, and race/ethnicity. On average:

- By age, performance scores were lowest for the youngest and oldest age groups (e.g., 8% for 1-2 years and 5% for 19-20 years compared with 28% for 6-7 years and 8-9 years).
- By geographic location, performance was somewhat lower for children living in rural areas (20%) compared with those living in urban areas (22%).
- By race and ethnicity, performance was lower for non-Hispanic black and non-Hispanic AIAN children (16% in both cases) and highest for non-Hispanic Asian and Hispanic children (26% and 29%, respectively).
- By biological sex, performance was somewhat lower for male children (20%) compared with female children (21%).

Table 1b.04-B Performance Scores Stratified by Age, Geographic Location, Race/Ethnicity and Sex: Descriptive Statistics, CY 2018

*	Mean	SD	Minimum	10th Percentile	25th Percentile	Median	75th Percentile	90th Percentile	Maximum
Overall Performance Score	0.21	0.04	0.14	0.16	0.18	0.21	0.24	0.26	0.28
Age Group (n=14)	*	*	*	*	*	*	*	*	*
1-2 yrs	0.08	0.04	0.04	0.05	0.06	0.08	0.09	0.11	0.20
3-5 yrs	0.24	0.06	0.18	0.18	0.18	0.24	0.27	0.30	0.37
6-7 yrs	0.28	0.05	0.20	0.22	0.23	0.29	0.32	0.33	0.37
8-9 yrs	0.28	0.04	0.19	0.22	0.26	0.28	0.32	0.33	0.35
10-11 yrs	0.27	0.05	0.18	0.21	0.24	0.27	0.31	0.32	0.33
12-14 yrs	0.23	0.04	0.16	0.16	0.20	0.25	0.27	0.28	0.28
15-18 yrs	0.15	0.05	0.05	0.07	0.11	0.17	0.18	0.21	0.22
19-20 yrs	0.05	0.03	0.00	0.02	0.04	0.06	0.06	0.09	0.09
Geographic Location (n=14)	*	*	*	*	*	*	*	*	*
Rural	0.20	0.05	0.11	0.15	0.16	0.19	0.22	0.27	0.29
Urban	0.22	0.05	0.12	0.15	0.18	0.23	0.25	0.27	0.28
Race/ Ethnicity	*	*	*	*	*	*	*	*	*
White, non-Hispanic (n=7)	0.20	0.04	0.14	0.14	0.16	0.21	0.23	0.25	0.25
Black, non-Hispanic (n=6)	0.16	0.05	0.09	0.09	0.11	0.17	0.18	0.21	0.21
Asian, non-Hispanic (n=6)	0.26	0.07	0.14	0.14	0.19	0.29	0.30	0.33	0.33
AIAN, non-Hispanic (n=7)	0.16	0.03	0.11	0.11	0.13	0.15	0.17	0.22	0.22
Hawaiian/Pacific Islander (n=5)	0.18	0.09	0.10	0.10	0.13	0.15	0.19	0.34	0.34
Multiracial, non-Hispanic (n=3)	0.18	0.04	0.14	0.14	0.14	0.20	0.21	0.21	0.21
Hispanic, all races (n=7)	0.29	0.10	0.16	0.16	0.20	0.28	0.38	0.42	0.42
Sex (n=14)	*	*	*	*	*	*	*	*	*
Female	0.21	0.04	0.15	0.16	0.18	0.21	0.24	0.26	0.28
Male	0.20	0.04	0.14	0.16	0.17	0.20	0.23	0.25	0.28

Table showing mean, median, minimum, maximum and percentiles for scores stratified by geographic location, race/ethnicity and sex for calendar year 2018

*Cell left intentionally blank

We believe it is meaningful to examine disparities for each program rather than just examining an overall summary because the disparities patterns may vary across states, and each state Medicaid program is urged to examine disparities specific to the populations served.

By age, children in the youngest and oldest age cohorts were less likely to receive at least two topical fluoride applications compared with those in the middle cohorts. In most states, measure performance was lower for children living in rural areas compared with those living in urban areas although there were some state Medicaid programs for which performance was higher among children living in rural areas or for which performance was similar between rural and urban areas. On average, measure scores were lower for non-Hispanic black children and American Indian/Alaskan Native (AIAN) children compared with non-Hispanic white children. However, measure scores were higher for AIAN children in Alaska. Measure scores were typically higher for non-Hispanic Asian children and Hispanic children compared with non-Hispanic white children. Measure scores were consistently higher for female children compared with male children, but the differences were typically within 1-2 percentage points.

Table 1b.04-C Performance Scores Stratified by Age, Geographic Location, Race/Ethnicity and Sex, CY 2018

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
Overall Performance Score	14.46%	19.33%	23.72%	24.87%	16.20%	22.88%	15.95%
Age Group	*	*	*	*	*	*	*
1-2 yrs	5.94%	8.86%	7.44%	10.21%	4.90%	6.43%	5.89%
3-5 yrs	17.97%	21.51%	24.61%	28.09%	17.78%	25.93%	17.72%
6-7 yrs	19.97%	25.30%	31.13%	33.43%	26.64%	31.72%	22.31%
8-9 yrs	19.40%	26.06%	31.66%	32.88%	26.82%	30.86%	22.28%
10-11 yrs	18.28%	25.31%	31.47%	31.47%	24.90%	29.20%	21.02%
12-14 yrs	16.04%	21.74%	27.78%	26.17%	20.73%	25.20%	18.06%
15-18 yrs	10.84%	15.21%	20.82%	17.55%	5.32%	17.41%	11.36%
19-20 yrs	4.00%	5.70%	8.81%	6.08%	0.01%	8.12%	3.76%
Missing	0.0%	0.0%	0.00%	0.00%	0.00%	0.00%	0.00%
Geographic Location	*	*	*	*	*	*	*
Rural	17.51%	11.42%	28.75%	24.07%	19.21%	20.92%	15.50%
Urban	11.95%	20.83%	22.56%	25.28%	15.49%	25.67%	15.97%
Missing	0.2%	2.7%	0.0%	0.00%	0.0%	0.01%	0.05%
Race/ Ethnicity	*	*	*	*	*	*	*
White, non-Hispanic	13.71%	NR	21.16%	25.03%	NR	NR	15.56%
Black, non-Hispanic	10.61%	NR	18.19%	NR	NR	NR	9.06%
Asian, non-Hispanic	14.28%	NR	29.56%	NR	NR	NR	18.62%
AIAN, non-Hispanic	16.84%	NR	13.04%	17.38%	NR	NR	11.31%

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
Hawaiian/Pacific Islander	9.55%	NR	33.67%	NR	NR	NR	14.88%
Multiracial, non-Hispanic	13.77%	NR	NR	NR	NR	NR	20.45%
Hispanic, all races	15.72%	NR	37.60%	41.82%	NR	NR	19.59%
Missing	5.9%	34.2%	0.0%	0.00%	91.32%	12.53%	3.19%
Sex	*	*	*	*	*	*	*
Female	15.07%	19.59%	24.23%	25.44%	16.48%	23.72%	16.20%
Male	13.89%	19.08%	23.23%	24.33%	15.94%	22.06%	15.70%
Missing	0.0%	0.0%	0.0%	0.00%	0.00%	0.00%	0.00%
*	**New Mexico**	**North Carolina**	**Oklahoma**	**Oregon**	**South Carolina**	**Washington**	**Wyoming**
Overall Performance Score	25.62%	23.24%	19.23%	17.82%	22.17%	27.79%	17.77%
Age Group	*	*	*	*	*	*	*
1-2 yrs	11.30%	8.13%	3.84%	7.21%	8.21%	19.91%	8.19%
3-5 yrs	30.00%	25.67%	17.63%	27.17%	22.83%	36.79%	22.56%
6-7 yrs	32.98%	31.20%	22.59%	28.19%	29.38%	36.68%	23.05%
8-9 yrs	32.50%	32.09%	26.29%	25.70%	29.98%	35.39%	25.94%
10-11 yrs	31.88%	31.13%	26.23%	23.51%	28.73%	32.75%	23.04%
12-14 yrs	28.23%	27.57%	24.47%	16.32%	25.04%	26.84%	19.82%
15-18 yrs	21.63%	18.82%	17.10%	9.67%	17.36%	17.30%	6.82%
19-20 yrs	9.30%	5.66%	6.10%	3.49%	5.43%	3.97%	2.37%
Missing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Geographic Location	*	*	*	*	*	*	*
Rural	22.37%	20.71%	17.08%	16.40%	19.37%	27.38%	15.10%
Urban	27.28%	24.21%	20.56%	18.41%	23.05%	27.87%	22.81%
Missing	0.05%	0.03%	0.02%	2.35%	0.53%	0.00%	0.00%
Race/ Ethnicity	*	*	*	*	*	*	*
White, non-Hispanic	22.59%	22.02%	17.67%	NR	NR	NR	NR
Black, non-Hispanic	20.60%	17.93%	16.90%	NR	NR	NR	NR
Asian, non-Hispanic	32.97%	27.84%	29.79%	NR	NR	NR	NR
AIAN, non-Hispanic	21.57%	13.68%	15.07%	NR	NR	NR	NR
Hawaiian/Pacific Islander	NR	19.06%	13.49%	NR	NR	NR	NR

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
Multiracial, non-Hispanic	NR	20.59%	NR	NR	NR	NR	NR
Hispanic, all races	27.79%	37.21%	26.12%	NR	NR	NR	NR
Missing	1.13%	0.35%	5.27%	21.56%	37.04%	10.58%	17.8%
Sex	*	*	*	*	*	*	*
Female	26.34%	23.78%	19.92%	18.05%	22.73%	28.04%	18.23%
Male	24.91%	22.73%	18.57%	17.59%	21.62%	27.55%	17.35%
Missing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Table describes the performance scores stratified by age, geographic location, race/ethnicity and sex for calendar year 2018

NR: Not reportable due to missing data>10% or specific category has 0 denominator or is not reported by the state.

*Cell left intentionally blank

We also tested for significantly significant differences in stratified measure scores within each state using bivariate logistic regression. Table 1b.04-D summarizes the number of states with odds ratios that are greater than the reference category ($p<0.05$), not statistically significant different from the reference category, less than the reference category ($p<0.05$), or not reportable due to missing data for that specific category. Most of the disparities in performance observed above were statistically significant. The odds ratio ranges also provide important perspective on the extent of variation both between stratification categories as well as between states. For example, the odds ratios demonstrate the extent to which performance drops off when moving from the reference group of 8-9 year olds to the youngest children of 1-2 years (OR range from 0.11 to 0.45 across states) and to the oldest children of 19-20 years (OR range from 0.00 to 0.21 across states). In the six states for which performance for non-Hispanic black children could be compared to non-Hispanic white children (reference), all six states had OR<1 for non-Hispanic black children, ranging from 0.54 to 0.95.

Table 1b.04-D Bivariate Logistic Regression, Reporting Number of States with Statistically Significant Differences from the Reference Category for Each Stratification Variable, CY 2018

Topical Fluoride, Dental Services	*	*	*	*
*	**OR< Ref (p<0.05)**	**Not significant**	**OR> Ref (p<0.05)**	**Not reportable**
AGE (n=14 states)	*	*	*	*
1-2 yrs	**14**	**0**	**0**	**0**
Odds ratio range	(0.11 to 0.45)	*	*	*
3-5 yrs	**12**	**0**	**2**	**0**
Odds ratio range	(0.39 to 0.91)	*	(1.06 to 1.09)	*
6-7 yrs	**5**	**7**	**2**	**0**
Odds ratio range	(0.82 to 0.97)	(0.98 to 1.04)	(1.06 to 1.14)	*
8-9 yrs	**Ref**	**Ref**	**Ref**	**Ref**

Topical Fluoride, Dental Services	*	*	*	*
10-11 yrs	**10**	**4**	**0**	**0**
Odds ratio range	(0.86 to 0.96)	(0.93 to 1.00)	*	*
12-14 yrs	**14**	**0**	**0**	**0**
Odds ratio range	(0.56 to 0.91)	*	*	*
15-18 yrs	**14**	**0**	**0**	**0**
Odds ratio range	(0.15 to 0.58)	*	*	*
19-20 yrs	**14**	**0**	**0**	**0**
Odds ratio range	(0.00 to 0.21)	*	*	*
GEOGRAPHIC LOCATION	**n=14 states**	*	*	*
Rural	**Ref**	**Ref**	**Ref**	**Ref**
Urban	**3**	**1**	**10**	**0**
Odds ratio range	(0.64 to 0.78)	1.03	(1.03 to 1.7)	*
BIOLOGICAL SEX	**n=14 states**	*	*	*
Female	**Ref**	**Ref**	**Ref**	**Ref**
Male	**13**	**1**	**0**	**0**
Odds ratio range	(0.91 to p.98)	0.94	*	*
RACE/ETHNICITY	**n=7 states**	*	*	*
White, non-Hispanic	**Ref**	**Ref**	**Ref**	**Ref**
Black, non-Hispanic	**6**	**0**	**0**	**1**
Odds ratio range	(0.54 to 0.95)	*	*	*
Asian, non-Hispanic	**0**	**1**	**5**	**1**
Odds ratio range	*	1.05	(1.23 to 1.98)	*
AIAN, non-Hispanic	**6**	**0**	**1**	**0**
Odds ratio range	(0.56 to 0.94)	*	1.27	*
Hawaiian/Pacific Islander	**2**	**2**	**1**	**2**
Odds ratio range	(0.66 to 0.73)	(0.84 to 0.93)	1.93	*
Multiracial, non-Hispanic	**1**	**2**	**0**	**4**
Odds ratio range	0.92	(1.00 to 1.23)	*	*
Hispanic, all races	**0**	**0**	**7**	**0**

Topical Fluoride, Dental Services	*	*	*	*
Odds ratio range	*	*	(1.18 to 2.25)	*
Ref=reference category; NS=not significant; NR=not reportable due to missing data; NA - not applicable - lower bound of age range for Topical Fluoride is 1 yr	*	*	*	

Table shows the bivariate logistic regression, reporting on the number of states with statistically significant differences from the reference category for each stratification variable in calendar year 2018

*Cell left intentionally blank

In summary, our results reflect an overall performance gap, variations in performance between states, disparities within each state, and variations in the nature and extent of disparities between states.

PRIOR SUBMISSION

The same data sources were used as described in 1b.2. The data below summarizes performance data by age, geographic location, and race/ethnicity for CY 2011 (CY 2010 for one program) with the p-values from chi-square tests used to detect whether there were statistically significant differences in performance between groups. The results demonstrate that there are disparities by age, geographic location and race/ethnicity. In addition, we also evaluated whether the measure could detect disparities by income (within 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. We additionally detected disparities by income, health status, CHIP plan, and Medicaid program type, but data on all of these characteristics were not consistently available for all programs so we are presenting disparities data on those characteristics that were most consistently available and had the greatest standardization

Data1b.4. Disparities in Performance by Child Age, Geographic Location and Race/Ethnicity

PROGRAM 1

Overall performance score: 37.13%

Scores by Age

Age 1-2 years: 6.21%

Age 3-5 years: 43.07%

Age 6-7 years: 43.64%

Age 8-9 years: 42.03%

Age 10-11 years: 40.50%

Age 12-14 years: 34.83%

Age 15-18 years: 24.93%

Age 19-20 years: 11.75%

p-value from Chi-square test: <0.0001

Scores by Geographic Location

Urban: 37.87%

Rural: 32.50%

p-value from Chi-square test: <0.0001

Scores by Race

Non-Hispanic White: 30.37%
Non-Hispanic Black: 29.68%
Hispanic: 40.84%
p-value from Chi-square test: <0.0001
PROGRAM 2
Overall performance score: 27.15%
Scores by Age
Age 1-2 years: n/a
Age 3-5 years: 30.00%
Age 6-7 years: 37.81%
Age 8-9 years: 34.88%
Age 10-11 years: 31.60%
Age 12-14 years: 27.14%
Age 15-18 years: 18.60%
Age 19-20 years: n/a
p-value from Chi-square test: <0.0001
Scores by Geographic Location
Urban: 26.96%
Rural: 30.64%
p-value from Chi-square test: <0.0001
Scores by Race
Non-Hispanic White: n/a
Non-Hispanic Black: n/a
Hispanic: n/a
p-value from Chi-square test n/a
PROGRAM 3
Overall performance score: 22.04%
Scores by Age
Age 1-2 years: 25.93%
Age 3-5 years: 34.24%
Age 6-7 years: 34.11%
Age 8-9 years: 33.97%
Age 10-11 years: 32.26%
Age 12-14 years: 28.78%
Age 15-18 years: 15.08%
Age 19-20 years: 2.22%
p-value from Chi-square test: <0.0001
Scores by Geographic Location
Urban: 22.13%
Rural: 19.71%
p-value from Chi-square test: 0.025

Scores by Race
Non-Hispanic White: n/a
Non-Hispanic Black: n/a
Hispanic: n/a
p-value from Chi-square test n/a
PROGRAM 4
Overall performance score: 18.16%
Scores by Age
Age 1-2 years: 17.17%
Age 3-5 years: 21.43%
Age 6-7 years: 21.19%
Age 8-9 years: 21.44%
Age 10-11 years: 19.47%
Age 12-14 years: 16.86%
Age 15-18 years: 12.53%
Age 19-20 years: 7.45%
p-value from Chi-square test: <0.0001
Scores by Geographic Location
Urban: 18.16%
Rural: 17.32%
p-value from Chi-square test: 0.025
Scores by Race
Non-Hispanic White: 21.64%
Non-Hispanic Black: 15.02%
Hispanic: 17.74%
p-value from Chi-square test: <0.0001
Note: N/A for age indicates that those ages are not within the program's age eligibility. N/A for race/ethnicity indicates that those programs did not collect race/ethnicity data or had high rates of missing data

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

In this maintenance submission, we have reported on a greater number of state Medicaid programs than previously so we did not undertake a literature search for additional information related to disparities in receipt of topical fluoride application. Our findings of disparities are consistent with those identified by Bouchery (2013) in Medicaid-enrolled children's use of dental services described below in our prior submission: the oldest and youngest children are less likely to receive services; non-Hispanic black children are

less likely to receive services and Hispanic children are more likely to receive services compared with non-Hispanic white children.

PRIOR SUBMISSION

There is extensive literature documenting disparities in dental service use among children by age, race/ethnicity, and geographic region, including within vulnerable populations, much of which is summarized in three major national reports on oral health: the Surgeon General's report on Oral Health in America in 2000, the IOM report, Improving Access to Oral Health Care for Vulnerable and Underserved Populations, and the IOM report, Advancing Oral Health in America.

With respect to preventive dental services, there are documented disparities. Using data from the National Survey of Children's Health, Edelstein and Chinn (2009) noted disparities in access to preventive dental services by race and income: "Stepwise disparities in access to preventive dental services are evident by race and income in ways that parallel Medical Expenditure Panel Survey findings. White parents report higher use of preventive dental services than do black or Hispanic parents (77%, 66%, and 61%, respectively). Poor parents report less use of services than do low income, middle class, and higher-income parents (58%, 66%, 77%, and 82%, respectively)" (Edelstein & Chinn, 2009, p.418). A recent analysis by Bouchery (2013) of the Medicaid Analytic eXtract files for nine states found variations in the percentage of children receiving a preventive dental visit by age, race and ethnicity, and geographic area. Specifically, relative to the reference group of 9 year olds, the percentage point change in the probability of having a dental preventive services was -27.6 for 3 years old; -8.6 for 6 years, -2.2 for 12 years and -15.4 for 15 years (all significant at $p<0.0001$); relative to the reference group of white, non-Hispanic, the percentage point change was -1.8 for black non-Hispanic and 7.8 for Hispanic ($p<0.0001$ for both); relative to the reference group of small metro area, the percentage point change was 5.9 for large metro area ($p<0.0001$).

Sources

Bouchery, E. 2013. "Utilization of Dental Services among Medicaid-Enrolled Children." Medicare & Medicaid Research Review. 3(3) E1-16. Available at: https://www.cms.gov/mmrr/Downloads/MMRR2013_003_03_b04.pdf.

Dietrich, T., C. Culler, R. Garcia, and M. M. Henshaw. 2008. Racial and ethnic disparities in children's oral health: The National Survey of Children's Health. Journal of the American Dental Association 139(11):1507-1517.

Dye BA, Li X, Thorton-Evans G. Oral health disparities as determined by selected healthy people 2020 oral health objectives for the United States, 2009-2010. NCHS Data Brief 2012(104):1-8.U.S. Dept. of Health and Human Services, National Institute of Dental and Craniofacial Research.

Edelstein, B. L. and C. H. Chinn. 2009. "Update on Disparities in Oral Health and Access to Dental Care for America´s Children." Acad Pediatr 9(6): 415-9.

Institute of Medicine (U.S.). Committee on an Oral Health Initiative. Advancing oral health in America. Washington, D.C.: National Academies Press; 2011.

Institute of Medicine and National Research Council. Improving access to oral health care for vulnerable and underserved populations. Washington, D.C.: National Academies Press; 2011.

Kenney, G. M., J. R. McFeeters, and J. Y. Yee. 2005. Preventive dental care and unmet dental needs among low-income children. American Journal of Public Health 95(8):1360-1366.

Lewis, C., W. Mouradian, R. Slayton, and A. Williams. 2007. Dental insurance and its impact on preventative dental care visits for U.S. children. Journal of the American Dental Association 138(3):369-380.

U.S. Dept. of Health and Human Services, National Institute of Dental and Craniofacial Research. Oral health in America : a report of the Surgeon General. Rockville, Md.: U.S. Public Health Service, Dept. of Health and Human Services; 2000.

[Response Ends]

Criteria 2: Scientific Acceptability of Measure Properties

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

---

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

[Response Begins]

Yes

[Yes Please Explain]

There was one main change to the measure specifications based on feedback by measure users requesting a more population-based measure: removal of elevated risk criteria from the denominator. We have conducted updated testing of the revised measure.

[Response Ends]

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

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

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

[Response Begins]

Denominator Change. Based on feedback from the stakeholder community, the denominator has been updated. The denominator was originally restricted to children at elevated risk for dental caries to focus on the priority populations that would benefit most from application of topical fluoride. However, there were acknowledged limitations with the approach used to identify elevated risk populations. The validated methodology for identifying elevated risk relied on procedure codes within claims data that signified either (a) the child received a risk assessment and was identified as being at elevated risk or (b) the child had a history of caries-related treatment. While this methodology identified a sample of children who could be reliably inferred to be at elevated caries risk, there were concerns that a significant percentage of at-risk children were being excluded from measurement, including those without a treatment history and particularly young children who are most likely to not have a treatment history.

Process. The determination to make this change was initiated by feedback from measure users, including the Centers for Medicare and Medicaid Services (CMS). During the DQA's 2021 Annual Measures Review Call for Public Comment (one month comment period), a specific request was issued to the stakeholder community on the denominator definition in a document titled "DQA Requests Stakeholder Feedback on the Denominator Definition of the DQA Topical Fluoride Measure." The document described the measure as originally specified, limitations, and options for modification (including no change as an option, removing elevated risk for only the youngest children, and removing elevated risk altogether). The DQA Measures Development and Maintenance Committee (MDMC) evaluated the public comments as well as data for 11 state Medicaid

programs to evaluate the impact of the change. Based on its evaluation of user community feedback and data, the MDMC recommended removal of elevated risk from the denominator with guidance in the User Guide for how to optionally stratify by elevated risk for users who wish to continue to examine fluoride receipt by caries risk status. This recommendation was approved by the DQA membership at its June 2021 meeting.

Grouped Measures. In addition, also in response to the user community's request, this measure is being grouped with two complementary measures. This measure – NQF 2528: Topical Fluoride for Children, Dental Services – focuses on topical fluoride delivered as a "dental" service (by or under the supervision of a dentist). Because many children, especially very young children, do not have a dental home, many state Medicaid programs and managed care organizations (MCOs) pay for the application of topical fluoride as an "oral health" service (by a physician or health care provider other than a dentist nor under supervision of a dentist). Consequently, state Medicaid programs, as well as commercial integrated medical-dental benefit MCOs or integrated medical-dental delivery sites, have a strong interest in tracking whether children receive **any** topical fluoride regardless of provider type ("dental or oral health" services). They also have a strong interest in understanding whether, and for whom, topical fluoride is being delivered through "dental" providers and "oral health" providers. Consequently, we are submitting two complementary measures for endorsement to be "grouped" with the existing endorsed measure: (1) Topical Fluoride for Children, Dental or Oral Health Services (NQF 3700) and (2) Topical Fluoride for Children, Oral Health Services (NQF 3701).

These updates (denominator update and grouped measure approach) enable more population-based assessments of quality by expanding the measure to a broader population of children and enabling measure users, including Medicaid programs and their contracted MCOs, to examine the overall provision of topical fluoride by provider type, which has been identified by stakeholders as integral for quality improvement and accountability purposes. The suite of 3 measures has been incorporated into the CMS Child Core Set for reporting beginning in FY 2022.

[Response Ends]

sp.01. Provide the measure title.

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

[Response Begins]

Prevention: Topical Fluoride for Children, Dental Services

[Response Ends]

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

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

[Response Begins]

Percentage of children aged 1 through 20 years who received at least 2 topical fluoride applications as dental services within the reporting year.

The measure is specified for reporting at the program (e.g., Medicaid, CHIP, Health Insurance Marketplaces) and plan (e.g., dental and health plans) levels for both public and private/commercial reporting.

[Response Ends]

sp.03. Provide a rationale for why this measure must be reported with other measures to appropriately interpret results.

[Response Begins]

Although this measure can be reported as a stand-alone measure, it is being grouped with two complementary measures to enable more robust quality improvement efforts. The DQA considered submitting a single measure with three numerators (denominator population is the same). But NQF evaluation criteria state: "Measures with multiple measure components that are assessed for each patient, but that result in multiple scores for an accountable entity rather than a single score. These generally should be submitted as separate measures and indicated as paired/grouped measures." (Measure Evaluation Criteria and Guidance, September 2021, p. 52) Based on this and discussions with NQF staff, we are submitting as three distinct measures.

This measure – NQF 2528: Topical Fluoride for Children, Dental Services – focuses on topical fluoride delivered as a "dental" service (by or under the supervision of a dentist). Because many children, especially very young children, do not have a dental home, many state Medicaid programs and MCOs pay for the application of topical fluoride as an "oral health" service (by a physician or health care provider other than a dentist nor under supervision of a dentist). Consequently, state Medicaid programs, as well as commercial integrated medical-dental benefit MCOs or integrated medical-dental healthcare delivery sites, have a strong interest in tracking whether children receive **any** topical fluoride regardless of provider type ("dental" or "oral health" services). They also have a strong interest in understanding whether, and for whom, topical fluoride is being delivered through "dental" providers and "oral health" providers.

Measures of topical fluoride provision by provider type, in addition to a measure of overall provision, are important because multi-pronged quality improvement strategies may be used to improve rates of topical fluoride application among a population of children. Dental providers and/or medical providers may be the focus of these efforts. Without measures that track the effectiveness by provider type, it is more difficult for programs and plans to assess which efforts are most effective. In addition, the accountability and delivery systems are typically distinct. Improving fluoride application by dental providers is accomplished through the dental delivery system and related financing/reimbursement structures, whereas topical fluoride application by medical providers is accomplished through medical delivery system and related financing/reimbursement structures. Some measure users will benefit by implementing and using all three measures (such as Medicaid programs and private payers/delivery systems that include both medical and dental). Other users, focused specifically on either dental or medical care delivery, respectively, will be able to report using one of the measures: either the measure related to "dental" services or the measure related to "oral health" services.

The need for three grouped measures comes directly from user community requests. In considering a topical fluoride measure for inclusion in the Centers for Medicare and Medicaid Services' Core Set of Children's Health Care Quality Measures, a specific request was made for a single measure that included three numerators (dental or oral health services, dental services, and oral health services) because of the recognized need by Medicaid programs to track not only overall receipt of topical fluoride but also topical fluoride provided through the dental and medical delivery systems specifically. Review by the DQA's Measures Development and Maintenance Committee, which includes representation of providers, community health centers and payers, affirmed the value of reporting three numerators across public and private/commercial measure applications for the reasons described above.

Consequently, the DQA is submitting two complementary measures for endorsement to be "grouped" with the existing endorsed measure: (1) Topical Fluoride for Children, Dental or Oral Health Services and (2) Topical Fluoride for Children, Oral Health Services. This grouping provides users with measurement options that

appropriately support population-based assessments of quality. It enables measure users, including Medicaid programs and their contracted MCOs, integrated medical-dental MCOs and integrated medical-dental delivery systems, to examine the overall provision of topical fluoride by provider type, which has been identified by stakeholders as integral for quality improvement and accountability purposes.

[Response Ends]

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

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

Please do not select:

- *Surgery: General*

[Response Begins]

Dental

Dental: Caries

[Response Ends]

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

[Response Begins]

Disparities Sensitive

Primary Prevention

[Response Ends]

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

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

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

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]

Children (Age < 18)

[Response Ends]

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

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

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

Please do not select:

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

[Response Begins]

Health Plan

Other

[Response Ends]

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

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

[Response Begins]

Outpatient Services

[Response Ends]

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

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

[Response Begins]

None available.

[Response Ends]

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

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

[Response Begins]

Available in attached Excel or csv file

[Response Ends]

Attachment: 2528_NQF2528_sp11_NUCC provider taxonomy codes_2022SpringMaintenance.xlsx

sp.12. State the numerator.

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

DO NOT include the rationale for the measure.

[Response Begins]

Unduplicated number of children who received at least 2 topical fluoride applications as dental services

[Response Ends]

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

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

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

[Response Begins]

Please see section sp 22.

[Response Ends]

sp.14. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

Unduplicated number of children aged 1 through 20 years

[Response Ends]

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

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

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

[Response Begins]

Please see section sp 22.

[Response Ends]

sp.16. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

There are no measure-specific exclusions. There is a standard exclusion as part of determining denominator eligibility: Medicaid/CHIP programs should exclude those individuals who do not qualify for dental benefits.

[Response Ends]

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

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

[Response Begins]

There are no measure-specific exclusions.

[Response Ends]

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

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

[Response Begins]

This measure is stratified by age (in years) using the following categories:

1-2; 3-5; 6-7; 8-9; 10-11; 12-14; 15-18; 19-20

No new data are needed for this stratification. Please see sp. 22 and attached specifications for complete measure details.

[Response Ends]

sp.19. Select the risk adjustment type.

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

[Response Begins]

No risk adjustment or risk stratification

[Response Ends]

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

Attachment: If available, please provide a sample report.

[Response Begins]

Rate/proportion

[Response Ends]

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

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

[Response Begins]

Better quality = Higher score

[Response Ends]

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

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

[Response Begins]

Topical Fluoride for Children, Dental Services, Measure Score Calculation

(1) Check if the subject meets age criteria at the last day of the reporting year:

[1]

(a) If child is >=1 and <21,

[2] then proceed to next step.

(b) If age criteria are not met or there are missing or invalid field codes (e.g., date of birth), then STOP processing. This subject does not get counted.

(2) Check if subject is continuously enrolled for the reporting year (12 months) with a gap of no more than 31 days (one-month gap for programs that determine eligibility on a monthly basis):

[3]

(a) If subject meets continuous enrollment criterion, then proceed to next step.

(b) If subject does not meet enrollment criterion, then STOP processing. This subject does not get counted.

YOU NOW HAVE THE DENOMINATOR (DEN): SUBJECTS WHO MEET THE AGE AND ENROLLMENT CRITERIA

(3) Check if subject received at least two fluoride applications as **dental services** during the reporting year – at least two unique dates of service when topical fluoride was provided. Service provided on each date of service should satisfy the following criteria:

(a) If [SERVICE CODE] = CDT D1206 or D1208,

[4] AND

(b) If [RENDERING PROVIDER TAXONOMY] code = any of the NUCC maintained Provider Taxonomy Codes in Table 1 below, then include in the numerator;

[5] proceed to next step.

(c) If both a AND b are not met, then the service was not a “dental” service; STOP processing. This subject is already included in the denominator but will not be included in the numerator.

> **Note 1:** No more than one fluoride application can be counted for the same member on the same date of service.
>
> **Note 2**: In this step, all **claims** with missing or invalid SERVICE CODE or with missing or invalid NUCC maintained Provider Taxonomy Codes should be excluded.

YOU NOW HAVE NUMERATOR (NUM) COUNT: Subjects who received at least two fluoride applications as *dental* services

(4) Report

(a) Unduplicated number of subjects in denominator (DEN)

(b) Unduplicated number of subjects in numerator (NUM)

(c) Measure rate (NUM/DEN)

(d) Rate stratified by age

Table 1: NUCC maintained Provider Taxonomy Codes classified as "Dental Service"++

Note: See Excel file attached in sp.11) for code descriptions.

122300000X	**1223P0106X**	**1223X0008X**	**125Q00000X**	**126800000X**
1223D0001X	1223P0221X	1223X0400X	261QF0400X	261QD0000X
1223D0004X	1223P0300X	124Q00000X+	261QR1300X	204E00000X
1223E0200X	1223P0700X	125J00000X	1223X2210X	261QS0112X
1223G0001X	1223S0112X	125K00000X	122400000X	*

Table showing NUCC-maintained Provider Taxonomy Codes classified as "Dental Service"

Alt text: Table showing NUCC-maintained Provider Taxonomy Codes classified as "Dental Service"

*Cell left intentionally blank

++Services provided by County Health Department dental clinics may also be included as "dental" services.

+Only dental hygienists who provide services under the supervision of a dentist should be classified as "dental" services. Services provided by independently practicing dental hygienists should be classified as "oral health" services and are not applicable for this measure.

[1] **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.

[2] **Age**: Medicaid/CHIP programs use under age 21(<21) as upper bound of age range; Exchange quality reporting use under age 19 (<19) as the upper bound of the age range; other programs check with program officials. The age criteria should be reported with the measure score.

[3] **Enrollment in "same" plan vs. "any" plan**: At the state program level (e.g., Medicaid/CHIP) a criterion of "any" plan applies versus at the health plan (e.g., MCO) level a criterion of "same" plan applies. The criterion used should be reported with the measure score. While this prevents direct aggregation of results from plan to program, each entity is given due credit for the population it serves. Thus, states with multiple MCOs should not merely "add up" the plan level scores but should calculate the state score from their database to allow inclusion of individuals who may be continuously enrolled but might have switched plans in the interim.

[4] **Topical Fluoride codes**: For reporting years prior to 2013, use CDT codes D1203 or D1204 or D1206.

[5] **Identifying "dental" services:** Programs and plans that do not use standard NUCC maintained provider taxonomy codes should use a valid mapping to identify providers whose services would be categorized as "dental" or "oral health" services.

[Response Ends]

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

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

Claims

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

sp.30. Provide the data collection instrument.

[Response Begins]

No data collection instrument provided

[Response Ends]

2a. Reliability

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

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

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

Yes

[Response Ends]

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

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

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

Yes

[Response Ends]

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

[Response Begins]

No

[Response Ends]

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

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

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

[Response Begins]

No additional risk adjustment analysis included

[Response Ends]

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

• Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.

• All required sections must be completed.

• For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.

• If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.

• An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.

• Contact NQF staff with any questions. Check for resources at the Submitting Standards webpage.

• For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the 2021 Measure Evaluation Criteria and Guidance.

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

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

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

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

AND

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

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

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

OR

• rationale/data support no risk adjustment/ stratification.

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

OR

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

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

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

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

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

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

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

Definitions

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

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

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

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

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

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

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

2021 Submission:

Updated testing information here.

2018 Submission:

Testing from the previous submission here.

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

[Response Begins]

Claims

Other (specify)

[Other (specify) Please Explain]

Enrollment Data

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

Since the original testing was conducted, there have been no changes in the data source, level of analysis, or setting. Additional testing focused on one main area: assessing reliability and validity at the measure score/accountable entity level (versus data element level).

Data Sources Used to Assess Measure Score Reliability and Validity

We used Medicaid enrollment and claims data contained within the Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAFs) available from the Centers for Medicare & Medicaid Services (

https://www.medicaid.gov/medicaid/data-systems/macbis/transformed-medicaid-statistical-information-system-t-msis/index.html).

We selected a sample of 14 states: Alaska, Arizona, Delaware, Idaho, Michigan, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, Oregon, South Carolina, Washington, and Wyoming. These states were selected based both on the quality of their data submissions to CMS and because they represent diversity in geographic location, population size, population demographic characteristics, and Medicaid dental delivery system.

PRIOR TESTING

The testing datasets were consistent with the measure specifications for the target populations and reporting entities. This measure was specified for administrative enrollment and claims data for children with private or public insurance coverage. We used data from four sources and refer to "program" level information and "plan" level information. We included data for publicly insured children in the Texas Medicaid, Florida CHIP, and Florida Medicaid programs as well as national commercial data from Dental Service of Massachusetts, Inc. Florida and Texas represent two of the largest and most diverse states. The two states also represent the upper and lower bounds of dental utilization based on dental utilization data available from the Centers for Medicare and Medicaid Services. The four programs collectively represent different delivery system models. The Texas Medicaid data represented dental fee-for-service. The Florida CHIP data included data from two dental MCOs. The Florida Medicaid data include dental fee-for-service and prepaid dental data. The commercial data included members in indemnity and preferred provider organization (PPO) product lines.

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

01-01-2016 through 12-31-2018

PRIOR TESTING

We used data from calendar years 2010 and 2011 for all programs except Florida Medicaid. Full-year data for 2011 were not available for Florida Medicaid

[Response Ends]

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

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

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

Please do not select:

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

[Response Begins]

Health Plan

Other (specify)

[Other (specify) Please Explain]

Program (e.g., Medicaid, CHIP)

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

Measured Entities used for Testing Measure Score Reliability

14 State Medicaid Programs: Alaska, Arizona, Delaware, Idaho, Michigan, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, Oregon, South Carolina, Washington, and Wyoming

Total enrollees 0-20 years across all 14 programs:

2018: 7,720,412

2017: 7,854,440

2016: 7,850,885

In all cases, statewide program data are used (i.e., location is statewide).

Number of Beneficiaries, 0-20 Years Enrolled at Least 1 Month, by State Medicaid Program and Year:

Medicaid Program	Dates	# Mem	Dental Delivery
Alaska	CY 2018	101,273	FFS
*	CY 2017	99,296	FFS
*	CY 2016	94,550	FFS
Arizona	CY 2018	974,161	Managed care carve in
*	CY 2017	994,391	Managed care carve in
*	CY 2016	981,695	Managed care carve in
Delaware	CY 2018	118,646	FFS
*	CY 2017	118,295	FFS
*	CY 2016	120,348	FFS
Idaho	CY 2018	214,879	Dental only PAHP
*	CY 2017	220,084	Dental only PAHP
*	CY 2016	201,253	Dental only PAHP
Michigan	CY 2018	1,163,658	Dental only PAHP
*	CY 2017	1,182,388	Dental only PAHP
*	CY 2016	1,182,388	Dental only PAHP
Mississippi	CY 2018	444,432	Managed care carve in
*	CY 2017	456,123	Managed care carve in
*	CY 2016	492,813	Managed care carve in
Nevada	CY 2018	379,289	Dental only PAHP & FFS
*	CY 2017	378,460	FFS
*	CY 2016	370,394	Managed care carve in & FFS

Medicaid Program	Dates	# Mem	Dental Delivery
New Mexico	CY 2018	376,379	Managed care carve in
*	CY 2017	387,255	Managed care carve in
*	CY 2016	383,056	Managed care carve in
North Carolina	CY 2018	1,231,829	FFS
*	CY 2017	1,259,699	FFS
*	CY 2016	1,241,882	FFS
Oklahoma	CY 2018	531,222	PCCM/FFS
*	CY 2017	553,905	PCCM/FFS
*	CY 2016	557,138	PCCM/FFS
Oregon	CY 2018	435,074	Dental only PAHP
*	CY 2017	463,301	Dental only PAHP
*	CY 2016	479,469	Dental only PAHP
South Carolina	CY 2018	767,719	FFS
*	CY 2017	762,747	FFS
*	CY 2016	752,206	FFS
Washington	CY 2018	932,270	FFS
*	CY 2017	945,583	FFS
*	CY 2016	939,142	FFS
Wyoming	CY 2018	49,581	FFS
*	CY 2017	52,127	FFS
*	CY 2016	54,551	FFS

Table showing program enrollment and dental delivery system type for 14 state Medicaid programs in each year 2016 through 2018.

*Cell left intentionally blank

Plan-Level Data. This measure is specified for reporting at both the program (e.g., Medicaid and CHIP) and plan (e.g., dental and health plans) levels for both public and private/commercial reporting. Our original testing, the basis for NQF endorsement and continued maintenance, included plan-specific data. DQA measures have been implemented by dental plans operating in commercial, Health Insurance Marketplace, and Medicaid/CHIP markets. However, our updated testing for this maintenance cycle does not include plan-specific data. The T-MSIS data used for our updated testing currently do not enable reliable identification of which dental services are provided by which dental plans. Based on prior testing of this and other administrative claims-based dental measures, we would not expect to see marked differences in the reliability or validity of plan-level reporting compared with program-level reporting given that the data sources (administrative claims) and measure specifications are the same. The only potential concern would be if the plan level denominators were too small to yield reliable results. However, the denominator requirements for this measure and other dental measures capture a broad population, and we have not encountered issues with small denominator sizes in our testing or in feedback from the user community. The DQA membership includes dental plan representatives that operate in state Medicaid/CHIP programs throughout the United States, and insufficient denominator sizes have not arisen as a concern. The DQA is also in frequent communication with the Centers for Medicare and Medicaid Services, as well as with Mathematica which serves as CMS's technical resource to state Medicaid programs for quality measure implementation; there have been no issues raised related to challenges with plan-level implementation.

The T-MSIS claims data are missing the managed care plan identifier for more than 90% of dental services in the states that we examined. This does **not** represent a feasibility issue for Medicaid programs and their participating plans to calculate the measures. We know from working with state Medicaid programs and state Health Insurance Marketplaces that it is highly feasible to have plan level reporting of dental quality measures. Rather, this reflects a limitation of the database that we used for testing. This is a relatively new database (released for public use in September 2020) for which data completeness and quality are continually being improved. In addition, because the measure specifications were updated during the DQA's 2021 Annual Measure Review, there is no public reporting yet of the revised measure.

PRIOR TESTING

Level of Analysis: Program, 4 Measured Entities

(1) Texas Medicaid

(a) Size: # Members 0-20 years, CY 2011: 3,554,247; # Members 0-20 years, CY 2010: 3,393,963

(b) Location: Texas – Statewide

(c) Delivery Type – FFS

(2) Florida CHIP

(a) Size: # Members 0-20 years, CY 2011: 317,146; # Members 0-20 years, CY 2010: 315,975

(b) Location: Florida – Statewide

(c) Delivery Type – Dental MCO (2 plans)

(3) Commercial

(a) Size: # Members 0-20 years, CY 2011: 184,152; # Members 0-20 years, CY 2010: 189,968

(b) Location: National

(c) Delivery Type – Indemnity/FFS & PPO product lines

(4) Florida Medicaid

(a) Size: # Members 0-20 years, CY 2010: 2,068,670

(b) Location: Florida – Statewide

(c) Delivery Type – FFS and Prepaid Dental

Note: At the time of testing, complete data were not available for Florida Medicaid for CY 2011.

Level of Analysis: Plan, 2 Measured Entities

The FL CHIP program had two separate dental plans that participate in the program in 2010 and 2011.

(1) FL CHIP – Plan 1

(a) Size: # Members 0-20 years, CY 2011: 140,986; # Members 0-20 years, CY 2010: 77,255

(b) Location: Florida – Statewide

(c) Delivery Type – Dental MCO

(2) FL CHIP – Plan 2

(a) Size: # Members 0-20 years, CY 2011: 168,191; # Members 0-20 years, CY 2010: 116,388

(b) Location: Florida – Statewide

(c) Delivery Type – Dental MCO

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

Patient Characteristics: Measure Score Reliability Testing

Tables 2a.06)A-C below provides the patient characteristics for children enrolled in Medicaid (with comprehensive benefits) for at least one month included in the T-MSIS analytic files for the 14 Medicaid programs for each year CY 2016 through CY 2018. In CY 2018, program enrollment ranged from 49,581 in Wyoming Medicaid to 1,231,829 in North Carolina Medicaid. Age and biological sex distributions were similar across programs. There was substantial variation in the geographic location and race/ethnicity distributions between states. Four states were excluded from testing analysis in CY2016 due to data quality issues (see 2b.08 regarding analysis of missing data).

Table 2a.06-A, Part 1, State Medicaid Program Patient Characteristics, <21 Years Old, CY2018 (T-MSIS Data)

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
Total # Patients	101,273	974,161	118,646	214,879	1,163,658	444,432	379,289
Age Group	*	*	*	*	*	*	*
<1 yr	5.35%	5.08%	5.29%	5.12%	5.48%	6.02%	5.69%
1-2 yrs	12.07%	10.75%	11.09%	11.49%	11.12%	11.80%	11.95%

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
3-5 yrs	16.76%	15.36%	15.49%	16.87%	15.51%	16.42%	16.63%
6-7 yrs	10.23%	9.70%	10.03%	10.33%	9.88%	9.98%	10.21%
8-9 yrs	9.90%	9.78%	10.20%	10.18%	9.81%	9.93%	9.97%
10-11 yrs	9.41%	10.32%	10.28%	10.48%	9.72%	10.42%	10.09%
12-14 yrs	12.94%	14.47%	13.87%	14.51%	13.72%	14.17%	13.62%
15-18 yrs	15.79%	17.06%	16.23%	16.63%	16.52%	16.55%	15.36%
19-20 yrs	7.53%	7.49%	7.51%	4.40%	8.23%	4.71%	6.48%
Missing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Geographic Location	*	*	*	*	*	*	*
Rural	44.49%	13.36%	18.44%	33.46%	19.27%	57.16%	5.86%
Urban	55.34%	84.19%	81.56%	66.20%	80.72%	42.84%	94.07%
Missing	<1%	2.45%	<1%	<1%	<1%	<1%	<1%
Race/ Ethnicity	*	*	*	*	*	*	*
White, non-Hispanic	32.50%	44.05%	34.56%	97.48%	0.00%	29.61%	24.94%
Black, non-Hispanic	3.62%	9.15%	41.42%	<1%	0.00%	53.53%	21.42%
Asian, non-Hispanic	5.37%	1.59%	1.97%	<1%	0.00%	<1%	2.86%
AIAN, non-Hispanic	36.17%	9.22%	<1%	2.15%	0.00%	<1%	1.31%
Hawaiian/Pacific Islander	4.86%	<1%	<1%	0.00%	0.00%	<1%	1.46%
Multiracial, non-Hispanic	8.11%	0.00%	0.00%	0.00%	0.00%	0.00%	<1%
Hispanic, all races	2.94%	<1%	21.73%	<1%	8.99%	<1%	43.90%
non-Hispanic, race unspecified	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Missing	6.43%	34.93%	<1%	<1%	91.01%	14.75%	4.09%
Sex	*	*	*	*	*	*	*
Female	48.77%	49.19%	49.41%	48.98%	49.20%	49.68%	49.27%
Male	51.23%	50.81%	50.59%	51.02%	50.80%	50.32%	50.73%
Missing	0.00%	0.00%	<1%	0.00%	0.00%	<1%	0.00%

Table showing demographic characteristics for individuals <21 years old enrolled in each of 14 state Medicaid programs in CY 2018

*Cell left intentionally blank

Table 2a.06-A, Part 2, State Medicaid Program Patient Characteristics, <21 Years Old, CY2018 (T-MSIS Data)

*	New Mexico	North Carolina	Oklahoma	Oregon	South Carolina	Washington	Wyoming
Total # Patients	376,379	1,231,829	531,222	435,074	767,719	932,270	49,581
Age Group	*	*	*	*	*	*	*
<1 yr	4.68%	5.80%	6.42%	5.42%	4.90%	4.79%	5.80%
1-2 yrs	10.02%	11.68%	13.00%	11.19%	10.86%	10.49%	12.37%
3-5 yrs	15.55%	16.29%	17.85%	15.13%	16.42%	15.76%	16.54%
6-7 yrs	9.91%	10.01%	10.69%	9.64%	10.30%	10.21%	10.44%
8-9 yrs	9.75%	9.73%	10.31%	9.67%	10.40%	10.27%	10.27%
10-11 yrs	10.27%	9.96%	10.43%	9.91%	10.65%	10.36%	10.50%
12-14 yrs	14.33%	13.52%	13.99%	13.60%	14.36%	14.11%	13.92%
15-18 yrs	17.24%	16.11%	13.15%	16.81%	16.62%	16.83%	15.80%
19-20 yrs	8.24%	6.90%	4.17%	8.63%	5.48%	7.19%	4.34%
Missing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Geographic Location	*	*	*	*	*	*	*
Rural	33.89%	26.30%	38.40%	25.39%	22.78%	15.50%	65.44%
Urban	66.06%	72.91%	61.58%	72.78%	76.71%	84.49%	34.56%
Missing	<1%	<1%	<1%	1.84%	<1%	<1%	<1%
Race/ Ethnicity	*	*	*	*	*	*	*
White, non-Hispanic	20.83%	37.88%	39.65%	32.07%	24.71%	40.01%	54.43%
Black, non-Hispanic	1.93%	35.24%	13.79%	2.50%	27.71%	6.91%	2.20%
Asian, non-Hispanic	<1%	1.66%	1.70%	1.46%	<1%	3.43%	<1%
AIAN, non-Hispanic	15.25%	1.29%	17.07%	1.79%	<1%	3.06%	8.92%
Hawaiian/Pacific Islander	0.00%	<1%	<1%	<1%	<1%	3.01%	<1%
Multiracial, non-Hispanic	0.00%	3.98%	0.00%	<1%	0.00%	2.18%	0.00%
Hispanic, all races	60.05%	19.10%	21.79%	39.72%	5.07%	29.04%	11.46%
non-Hispanic, race unspecified	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Missing	1.32%	<1%	5.54%	21.79%	41.60%	12.36%	22.42%
Sex	*	*	*	*	*	*	*

*	New Mexico	North Carolina	Oklahoma	Oregon	South Carolina	Washington	Wyoming
Female	49.42%	49.39%	49.59%	49.23%	49.25%	49.21%	48.91%
Male	50.58%	50.61%	50.41%	50.77%	50.74%	50.79%	51.09%
Missing	0.00%	0.00%	0.00%	<1%	<1%	0.00%	0.00%

Table showing demographic characteristics for individuals<21 years old enrolled in each of 14 state Medicaid programs in CY 2018

*Cell left intentionally blank

Table 2a06-B, Part 1, State Medicaid Program Patient Characteristics, <21 Years Old, CY2017 (T-MSIS Data)

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
Total # Patients	99,296	994,391	118,295	220,084	1,163,174	456,123	378,460
Age Group	*	*	*	*	*	*	*
<1 yr	5.72%	5.16%	5.49%	5.46%	5.34%	5.99%	5.88%
1-2 yrs	12.66%	11.04%	11.33%	11.92%	11.39%	11.77%	12.14%
3-5 yrs	16.46%	15.43%	15.29%	17.02%	15.45%	16.33%	16.57%
6-7 yrs	10.29%	9.68%	10.39%	10.34%	9.95%	10.22%	10.32%
8-9 yrs	9.86%	10.27%	10.53%	10.45%	9.95%	10.43%	10.30%
10-11 yrs	9.30%	10.32%	10.02%	10.34%	9.72%	10.62%	10.16%
12-14 yrs	12.54%	14.00%	13.56%	14.02%	13.44%	13.46%	13.17%
15-18 yrs	15.74%	16.63%	15.83%	16.07%	16.54%	16.42%	14.94%
19-20 yrs	7.45%	7.46%	7.56%	4.37%	8.22%	4.77%	6.51%
Missing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Geographic Location	*	*	*	*	*	*	*
Rural	44.74%	12.99%	18.45%	33.61%	19.31%	57.00%	6.02%
Urban	55.06%	81.41%	81.55%	66.00%	80.67%	42.99%	93.91%
Missing	<1%	5.61%	<1%	<1%	<1%	<1%	<1%
Race/ Ethnicity	*	*	*	*	*	*	*
White, non-Hispanic	32.60%	44.14%	35.10%	97.51%	0.00%	30.83%	25.33%
Black, non-Hispanic	3.66%	8.92%	41.31%	<1%	0.00%	55.30%	20.85%
Asian, non-Hispanic	5.47%	1.61%	1.93%	<1%	0.00%	<1%	2.82%
AIAN, non-Hispanic	36.14%	9.24%	<1%	2.06%	0.00%	<1%	1.35%
Hawaiian/Pacific Islander	4.83%	<1%	<1%	0.00%	0.00%	<1%	1.23%
Multiracial, non-Hispanic	8.21%	0.00%	0.00%	0.00%	0.00%	0.00%	<1%
Hispanic, all races	2.94%	2.03%	21.35%	<1%	8.95%	<1%	44.03%

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
non-Hispanic, race unspecified	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Missing	6.14%	33.72%	<1%	<1%	91.05%	11.78%	4.37%
Sex	*	*	*	*	*	*	*
Female	48.92%	49.17%	49.47%	49.01%	49.19%	49.75%	49.28%
Male	51.08%	50.83%	50.53%	50.99%	50.81%	50.25%	50.72%
Missing	0.00%	0.00%	<1%	0.00%	0.00%	<1%	0.00%

Table showing demographic characteristics for individuals<21 years old enrolled in each of 14 state Medicaid programs in CY 2017

*Cell left intentionally blank

Table 2a06-B, Part 2, State Medicaid Program Patient Characteristics, <21 Years Old, CY2017 (T-MSIS Data)

*	New Mexico	North Carolina	Oklahoma	Oregon	South Carolina	Washington	Wyoming
Total # Patients	387,255	1,259,699	553,905	463,301	762,747	945,583	52,127
Age Group	*	*	*	*	*	*	*
<1 yr	4.74%	5.82%	6.27%	5.41%	5.07%	4.94%	5.83%
1-2 yrs	10.37%	11.67%	13.08%	11.50%	11.32%	10.74%	12.68%
3-5 yrs	15.72%	16.30%	17.68%	15.08%	16.28%	15.78%	16.68%
6-7 yrs	9.98%	10.20%	10.76%	9.75%	10.41%	10.27%	10.66%
8-9 yrs	10.26%	10.15%	10.70%	10.07%	10.82%	10.58%	10.54%
10-11 yrs	10.22%	9.90%	10.39%	9.78%	10.64%	10.23%	10.35%
12-14 yrs	13.83%	13.07%	13.62%	13.24%	13.65%	13.66%	13.46%
15-18 yrs	16.87%	16.00%	13.25%	16.68%	16.60%	16.68%	15.30%
19-20 yrs	8.00%	6.88%	4.26%	8.49%	5.21%	7.12%	4.49%
Missing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Geographic Location	*	*	*	*	*	*	*
Rural	33.71%	26.47%	38.71%	24.60%	22.87%	15.45%	65.42%
Urban	66.25%	72.71%	61.25%	71.97%	76.62%	84.55%	34.58%
Missing	<1%	<1%	<1%	3.43%	<1%	<1%	0.00%
Race/ Ethnicity	*	*	*	*	*	*	*
White, non-Hispanic	20.82%	38.41%	40.18%	33.66%	26.23%	40.37%	56.33%
Black, non-Hispanic	1.94%	35.12%	13.51%	2.46%	29.24%	6.66%	2.12%

*	New Mexico	North Carolina	Oklahoma	Oregon	South Carolina	Washington	Wyoming
Asian, non-Hispanic	<1%	1.68%	1.67%	1.46%	<1%	3.46%	<1%
AIAN, non-Hispanic	15.04%	1.28%	17.07%	1.76%	<1%	2.88%	8.63%
Hawaiian/Pacific Islander	0.00%	<1%	<1%	<1%	<1%	2.89%	<1%
Multiracial, non-Hispanic	0.00%	3.91%	0.00%	<1%	0.00%	1.78%	0.00%
Hispanic, all races	60.22%	18.69%	21.45%	36.70%	5.42%	28.38%	11.87%
non-Hispanic, race unspecified	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Missing	1.34%	<1%	5.66%	23.21%	38.17%	13.59%	20.37%
Sex	*	*	*	*	*	*	*
Female	49.37%	49.43%	49.54%	49.19%	49.33%	49.18%	48.93%
Male	50.63%	50.57%	50.46%	50.81%	50.67%	50.82%	51.07%
Missing	0.00%	0.00%	0.00%	<1%	<1%	0.00%	0.00%

Table showing demographic characteristics for individuals<21 years old enrolled in each of 14 state Medicaid programs in CY 2017

*Cell left intentionally blank

Table 2a06-C, Part 1, State Medicaid Program Patient Characteristics, <21 Years Old, CY2016 (T-MSIS Data)

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
Total # Patients	94,550	981,695	120,348	201,253	1,182,388	492,813	370,394
Age Group	*	*	*	*	*	*	*
<1 yr	6.43%	5.47%	5.32%	5.27%	5.38%	5.60%	5.97%
1-2 yrs	12.42%	11.59%	11.12%	12.65%	11.33%	11.20%	12.18%
3-5 yrs	16.26%	15.02%	15.68%	17.15%	15.33%	15.61%	16.60%
6-7 yrs	10.56%	10.06%	10.63%	10.98%	10.01%	10.64%	10.70%
8-9 yrs	9.87%	10.57%	10.60%	11.23%	9.94%	11.12%	10.63%
10-11 yrs	9.11%	10.03%	9.92%	10.42%	9.50%	10.35%	9.76%
12-14 yrs	12.50%	13.64%	13.34%	13.93%	13.28%	13.71%	12.90%
15-18 yrs	15.85%	16.34%	15.84%	14.44%	17.07%	17.08%	14.76%
19-20 yrs	7.00%	7.28%	7.56%	3.93%	8.16%	4.69%	6.51%
Missing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Geographic Location	*	*	*	*	*	*	*
Rural	45.02%	12.88%	18.31%	33.d63%	18.03%	56.69%	<1%

*	Alaska	Arizona	Delaware	Idaho	Michigan	Mississippi	Nevada
Urban	54.77%	80.34%	81.69%	66.36%	75.62%	43.30%	7.87%
Missing	<1%	6.78%	<1%	<1%	6.34%	<1%	91.70%
Race/ Ethnicity	*	*	*	*	*	*	*
White, non-Hispanic	32.49%	43.12%	35.80%	98.02%	53.04%	5.63%	25.99%
Black, non-Hispanic	3.80%	8.69%	41.11%	<1%	27.90%	7.05%	20.23%
Asian, non-Hispanic	5.52%	1.61%	1.96%	<1%	<1%	<1%	2.78%
AIAN, non-Hispanic	36.49%	9.39%	<1%	1.95%	<1%	<1%	1.45%
Hawaiian/Pacific Islander	4.84%	<1%	<1%	0.00%	<1%	<1%	1.01%
Multiracial, non-Hispanic	8.16%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Hispanic, all races	3.07%	3.88%	20.83%	<1%	8.19%	<1%	43.93%
non-Hispanic, race unspecified	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Missing	5.64%	32.99%	<1%	<1%	9.43%	86.63%	4.61%
Sex	*	*	*	*	*	*	*
Female	48.95%	49.25%	49.54%	48.50%	49.22%	49.67%	49.34%
Male	51.05%	50.75%	50.45%	51.50%	50.78%	50.32%	50.66%
Missing	0.00%	0.00%	<1%	0.00%	0.00%	<1%	0.00%

Table showing demographic characteristics for individuals<21 years old enrolled in each of 14 state Medicaid programs in CY 2016

*Cell left intentionally blank

Table 2a06-C, Part 2, State Medicaid Program Patient Characteristics, <21 Years Old, CY2016 (T-MSIS Data)

*	New Mexico	North Carolina	Oklahoma	Oregon	South Carolina	Washington	Wyoming
Total # Patients	383,056	1,241,882	557,138	479,469	752,206	939,142	54,551
Age Group	*	*	*	*	*	*	*
<1 yr	4.94%	5.58%	6.52%	5.53%	5.12%	5.20%	6.11%
1-2 yrs	10.76%	11.91%	13.06%	11.70%	11.52%	10.84%	12.52%
3-5 yrs	15.78%	16.42%	17.58%	15.17%	16.15%	15.77%	16.93%
6-7 yrs	10.17%	10.65%	11.14%	10.02%	10.69%	10.55%	10.97%
8-9 yrs	10.55%	10.41%	10.85%	10.28%	10.98%	10.70%	11.04%
10-11 yrs	9.91%	9.62%	10.04%	9.59%	10.13%	9.90%	9.86%
12-14 yrs	13.65%	12.94%	13.41%	13.02%	13.35%	13.35%	13.03%
15-18 yrs	16.58%	15.91%	13.26%	16.61%	16.59%	16.61%	14.96%

*	New Mexico	North Carolina	Oklahoma	Oregon	South Carolina	Washington	Wyoming
19-20 yrs	7.67%	6.57%	4.14%	8.07%	5.45%	7.09%	4.58%
Missing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Geographic Location	*	*	*	*	*	*	*
Rural	33.89%	26.58%	38.85%	22.21%	23.14%	15.40%	0.00%
Urban	66.06%	72.95%	61.12%	67.14%	76.37%	84.60%	0.00%
Missing	<1%	<1%	<1%	10.65%	<1%	<1%	100.00%
Race/ Ethnicity	*	*	*	*	*	*	*
White, non-Hispanic	20.97%	38.76%	40.55%	36.46%	28.36%	41.07%	59.76%
Black, non-Hispanic	1.93%	35.36%	13.44%	2.55%	31.64%	6.54%	2.23%
Asian, non-Hispanic	<1%	1.68%	1.68%	1.35%	<1%	3.50%	<1%
AIAN, non-Hispanic	15.13%	1.28%	16.73%	1.76%	<1%	2.56%	8.23%
Hawaiian/Pacific Islander	0.00%	<1%	<1%	<1%	<1%	2.81%	<1%
Multiracial, non-Hispanic	0.00%	3.76%	0.00%	<1%	0.00%	1.49%	0.00%
Hispanic, all races	59.94%	18.34%	21.26%	33.92%	5.82%	27.92%	0.00%
non-Hispanic, race unspecified	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Missing	1.38%	<1%	5.91%	23.08%	33.22%	14.11%	29.08%
Sex	*	*	*	*	*	*	*
Female	49.40%	49.46%	49.55%	49.20%	49.39%	49.24%	48.79%
Male	50.60%	50.54%	50.45%	50.79%	50.61%	50.76%	51.21%
Missing	0.00%	0.00%	0.00%	<1%	<1%	0.00%	0.00%

Table showing demographic characteristics for individuals<21 years old enrolled in each of 14 state Medicaid programs in CY 2016

*Cell left intentionally blank

PRIOR TESTING

Note that there were only three programs in CY 2011 because Florida Medicaid did not have complete claims data available for CY 2011 at the time testing was conducted.

Table 2a.06)-D, Patient Characteristics, 0-20 Years Old, 2011

Descriptive Characteristics of Individuals 0-20 Years Enrolled at Least One Month, CY 2011	*	*	*	*	*
*	Program 1	Program 2	Program 3	Plan 1	Plan 2
Total Number Patients	3,544,247	317,146	184,152	140,986	168,191
Age Group Distribution	*	*	*	*	*
Age <1 years	7.05%	N/A	1.54%	N/A	N/A
Age 1-2 years	14.32%	N/A	5.75%	N/A	N/A
Age 3-5 years	19.46%	3.81%	12.68%	4.12%	3.60%
Age 6-7 years	11.21%	13.05%	9.57%	13.71%	12.55%
Age 8-9 years	9.85%	15.00%	10.18%	15.76%	14.41%
Age 10-11 years	9.03%	15.71%	10.55%	16.27%	15.25%
Age 12-14 years	11.63%	23.73%	16.09%	23.06%	24.31%
Age 15-18 years	13.19%	28.70%	22.13%	27.08%	29.88%
Age 19-20 years	4.27%	N/A	11.50%	N/A	N/A
Geographic Location	*	*	*	*	*
Urban	83.63%	92.94%	95.95%	93.01%	92.91%
Rural	15.15%	5.02%	3.86%	4.83%	5.15%
Missing	1.22%	2.04%	0.19%	2.16%	1.94%
Race and Ethnicity	*	*	*	*	*
Non-Hispanic White	17.36%	N/A	N/A	N/A	N/A
Non-Hispanic Black	15.08%	N/A	N/A	N/A	N/A
Hispanic	58.07%	N/A	N/A	N/A	N/A
Other & Unknown	9.49%	N/A	N/A	N/A	N/A

Table showing demographic characteristics for individuals<21 years old enrolled in each of 3 Programs and 2 Plans in CY 2011

*Cell left intentionally blank

Table 2a.06)-E, Patient Characteristics, 0-20 Years Old, 2010

Descriptive Characteristics of Individuals 0-20 Years Enrolled at Least One Month, CY 2010	*	*	*	*	*	*
*	Program 1	Program 2	Program 3	Program 4	Plan 1	Plan 2

Descriptive Characteristics of Individuals 0-20 Years Enrolled at Least One Month, CY 2010	*	*	*	*	*	*
Total Number Patients	3,393,963	315,975	189,968	2,068,670	77,255	116,388
Age Group Distribution	*	*	*	*	*	*
Age <1 years	7.35%	N/A	1.45%	6.05%	N/A	N/A
Age 1-2 years	15.16%	N/A	5.67%	14.23%	N/A	N/A
Age 3-5 years	19.48%	3.64%	12.73%	19.26%	5.72%	4.22%
Age 6-7 years	11.12%	13.32%	9.69%	10.47%	15.68%	12.54%
Age 8-9 years	9.70%	15.14%	10.24%	9.19%	16.99%	14.21%
Age 10-11 years	8.75%	15.84%	10.60%	8.74%	16.41%	15.18%
Age 12-14 years	11.23%	23.70%	16.20%	11.87%	21.40%	24.05%
Age 15-18 years	12.99%	28.37%	22.12%	14.73%	23.79%	29.81%
Age 19-20 years	4.22%	N/A	11.31%	5.47%	N/A	N/A
Geographic Location	*	*	*	*	*	*
Urban	83.20%	92.08%	96.70%	91.47%	92.10%	92.11%
Rural	15.56%	5.07%	3.17%	7.30%	5.00%	5.19%
Missing	1.24%	2.85%	0.13%	1.23%	2.89%	2.70%
Race and Ethnicity	*	*	*	*	*	*
Non-Hispanic White	18.21%	N/A	N/A	29.89%	N/A	N/A
Non-Hispanic Black	15.45%	N/A	N/A	29.39%	N/A	N/A
Hispanic	59.42%	N/A	N/A	29.65%	N/A	N/A
Other & Unknown	6.92%	N/A	N/A	11.06%	N/A	N/A

Table showing demographic characteristics for individuals<21 years old enrolled in each of 3 Programs and 2 Plans in CY 2010

*Cell left intentionally blank

[Response Ends]

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

Measure score reliability and validity tests were conducted using Medicaid claims and enrollment data contained within the T-MSIS analytic files from 14 state programs.

PRIOR TESTING

These data were used for all testing aspects except two:

A. Part of the face validity assessments involved expert consensus processes, including conducting an environmental scan of measure concepts and using the RAND-UCLA modified Delphi process to rate the importance, feasibility and validity. Please see section 2b2.2 for a complete description.

B. Data element validation using medical chart reviews did not include all programs. Due to the cost of these activities, chart reviews were conducted only for the Texas Medicaid program. Texas has the third largest Medicaid program in the U.S. with significant diversity represented. In addition, the research team conducting the testing is the External Quality Review Organization for Texas and has years of experience conducting medical chart audits for the Texas Medicaid program for ongoing quality assurance purposes. Thus, an established infrastructure and expertise was in place to conduct chart reviews for these programs.

[Response Ends]

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

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

[Response Begins]

Testing for Spring 2022 Maintenance Evaluation

The measure scores were stratified by the following patient characteristics (when sufficient data were available): age, geographic location (rural or urban), race and ethnicity, and biological sex. These will be reported in section 1b: Importance to Measure and Report: Gap in Care/Disparities.

[Response Ends]

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

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

Choose one or both levels.

[Response Begins]

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

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

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

(1) Accountable Entity Measure Score Reliability Testing using a random split-sample methodology:

Reliability indicates the extent to which repeated measurements yield consistent results. We conducted accountable entity measure score level reliability testing using a **random split-sample methodology and the intraclass correlation coefficient** (ICC) of the measure scores. For each of the 14 state Medicaid programs, we randomly split the population of children aged <21 years present in the T-MSIS demographic and eligibility file. The denominator, numerator, and measure score were calculated for each sample. Thus, the measure score is calculated twice for each state Medicaid program among two distinct and randomly selected sets of children contained within the analytic files. We used the ICC to calculate the agreement between the randomly selected samples (Koo & Li 2016; McGraw & Wong 1996; Shrout & Fleiss 1979). A higher ICC value indicates greater agreement and, therefore, greater reliability. We follow the guidance in Koo and Li (2016) regarding the interpretation of reliability using the 95% confidence interval of the ICC: <0.5 = poor; 0.5–0.75 = moderate; 0.75–0.9 = good; and > 0.9 = excellent.

(2) Evaluation of Relative Rankings: Between Split Samples

We compared the **relative rankings** for the split samples to evaluate whether the relative measure scores for the state Medicaid programs remained stable between the split samples.

(3) Evaluation of Relative Rankings: Between Years

We compared the **relative rankings** of the overall measure scores between 2017 and 2018 and between 2016 and 2017 to evaluate whether there were any dramatic changes that could suggest a threat to reliability. Using the measure scores for the 14 state Medicaid programs in 2017 and 2018 and 10 state Medicaid programs in 2016 and 2017, we calculated Kendall's Tau-b, which is a rank correlation coefficient that measures association based on the number of concordant and discordant pairs. For reference, we also report the more commonly reported Spearman's rank correlation coefficients. Although the strength of these associations is stronger, we felt Kendall's tau was the more appropriate test to report given the relatively small sample size.

(4) Data Element Reliability/Validity

The original data element validity/reliability testing using chart reviews holds for the updated measure specifications. The critical data elements have not changed. Consequently, we did not repeat data element reliability and validity testing since the reliability and validity of the data elements used are well-established from prior testing and in the peer-reviewed literature as reported in our prior submission.

Plan-Level Data. This measure is specified for reporting at both the program (e.g., Medicaid and CHIP) and plan (e.g., dental and health plans) levels for both public and private/commercial reporting. Our original testing, the basis for NQF endorsement and continued maintenance, included plan-specific data. DQA measures have been implemented by dental plans operating in commercial, Health Insurance Marketplace, and Medicaid/CHIP markets. However, our updated testing for this maintenance cycle does not include plan-specific data. The T-MSIS data used for our updated testing currently do not enable reliable identification of which dental services are provided by which dental plans. Based on prior testing of this and other administrative claims-based dental measures, we would not expect to see marked differences in the reliability or validity of plan-level reporting compared with program-level reporting given that the data sources (administrative claims) and measure specifications are the same. The only potential concern would be if the plan level denominators were too small to yield reliable results. However, the denominator requirements for

this measure and other dental measures capture a broad population, and we have not encountered issues with small denominator sizes in our testing or in feedback from the user community. The DQA membership includes dental plan representatives that operate in state Medicaid/CHIP programs throughout the United States, and insufficient denominator sizes have not arisen as a concern. The DQA is also in frequent communication with the Centers for Medicare and Medicaid Services, as well as with Mathematica which serves as CMS's technical resource to state Medicaid programs for quality measure implementation; there have been no issues raised related to challenges with plan-level implementation.

The T-MSIS claims data are missing the managed care plan identifier for more than 90% of dental services in the states that we examined. This does **not** represent a feasibility issue for Medicaid programs and their participating plans to calculate the measures. We know from working with state Medicaid programs and state Health Insurance Marketplaces that it is highly feasible to have plan level reporting of dental quality measures. Rather, this reflects a limitation of the database that we used for testing. This is a relatively new database (released for public use in September 2020) for which data completeness and quality are continually being improved. In addition, because the measure specifications were updated during the DQA's 2021 Annual Measure Review, there is no public reporting yet of the revised measure.

References

Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016 Jun;15(2):155-63. doi: 10.1016/j.jcm.2016.02.012. Epub 2016 Mar 31.

McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996;1:30–46.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86:420–*428.*

PRIOR TESTING

Clarification note: Following NQF guidance at the time of original testing, reliability for original endorsement was established through:

1. Data element reliability was demonstrated through data element validation using chart audits as described in Section 2b1 below.
2. Evaluation and refinement of measure specifications to ensure that they were detailed, clear, and complete.
3. Voting by the DQA membership on whether reliability criterion was met.

Data Elements:

- See section 2b2 for validity testing of data elements.
- Note: Unlike measures that rely on medical record data for which issues such as inter-rater reliability are likely to introduce measurement concerns or measures that rely on survey data for which issues such as internal consistency may be a concern, this measure relies on standard data fields commonly used in administrative data for a wide range of billing and reporting purposes.

Measure Score – Threats to Measure Reliability

An important component of assessing reliability is assessing, testing, and addressing threats to measure reliability.

1. Evaluation of Clarity and Completeness of Measure Specifications

For a measure to be reliable – to allow for meaningful comparisons across entities – the measure specifications must be unambiguous: the denominator criteria, numerator criteria, exclusions, and scoring need to be clearly specified. The initial measure specifications were developed by the Dental Quality Alliance (DQA). The Dental Quality Alliance includes 30 members, representing a broad range of stakeholders, including federal agencies involved with oral health services, dental professional associations, medical professional associations, dental and medical health insurance commercial plans, state Medicaid and CHIP programs, quality accrediting bodies, and the general public. The initial specifications were developed based on (1) the evidence regarding the effectiveness of professionally applied topical fluoride in caries prevention, (2) an

environmental scan, and (3) face validity assessments of the measure concept. These specifications were contained in the competitive Request for Proposals to conduct measure testing; a research team from the University of Florida was selected to conduct testing. The research team independently carefully evaluated whether the measure specifications identified all necessary data elements to calculate the numerators and denominators for each measure. In addition, the research team carefully reviewed the logic flow and made revision recommendations to improve the reliability of the resulting calculations. The DQA also solicited public comment on an Interim Report and posted the measurement specifications online for public comment. The research team worked with the DQA to evaluate and address all comments provided. Throughout the eight-month testing period, there were numerous reviews and revisions of the specifications conducted jointly by the research team and the DQA to ensure clear and detailed measure specifications.

2. Other Threats to Reliability - Sample Size

Our measured entities include very large numbers of patients; small sample size is not a concern.

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

(1) Measure Score Reliability Testing: Split-Half

(A) Performance Scores with 95% CI

Tables 2a.11)A-C below provides the performance scores for the split samples with their 95% confidence intervals.

Table 2a.11)A, Split Sample Topical Fluoride for Children, Dental Services, Measure Scores and 95% CI, CY2018 (T-MSIS Data)

Program	Sample	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
AK, 2018	1	14.32%	(	0.0018	,	0.1396	,	0.1468	)
*	2	14.57%	(	0.0019	,	0.1421	,	0.1494	)
AZ, 2018	1	19.38%	(	0.0007	,	0.1924	,	0.1951	)
*	2	19.28%	(	0.0007	,	0.1914	,	0.1941	)
DE, 2018	1	23.76%	(	0.0022	,	0.2333	,	0.2419	)

Program	Sample	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
*	*	*	*	*	*	*	*	*	*
*	2	23.65%	(	0.0022	,	0.2323	,	0.2408	)
ID, 2018	1	24.83%	(	0.0016	,	0.2452	,	0.2515	)
*	2	24.90%	(	0.0016	,	0.2459	,	0.2521	)
MI, 2018	1	16.22%	(	0.0006	,	0.1611	,	0.1634	)
*	2	16.18%	(	0.0006	,	0.1607	,	0.1629	)
MS, 2018	1	23.05%	(	0.0011	,	0.2282	,	0.2327	)
*	2	22.72%	(	0.0011	,	0.2249	,	0.2294	)
NV, 2018	1	16.02%	(	0.0011	,	0.1580	,	0.1623	)
*	2	15.86%	(	0.0011	,	0.1564	,	0.1608	)
NM, 2018	1	25.72%	(	0.0012	,	0.2549	,	0.2596	)
*	2	25.50%	(	0.0012	,	0.2527	,	0.2574	)
NC, 2018	1	23.23%	(	0.0007	,	0.2310	,	0.2335	)
*	2	23.26%	(	0.0007	,	0.2313	,	0.2339	)
OK, 2018	1	19.24%	(	0.0011	,	0.1903	,	0.1945	)
*	2	19.22%	(	0.0011	,	0.1901	,	0.1943	)
OR, 2018	1	17.96%	(	0.0010	,	0.1775	,	0.1816	)
*	2	17.69%	(	0.0010	,	0.1749	,	0.1790	)
SC, 2018	1	22.07%	(	0.0008	,	0.2191	,	0.2222	)
*	2	22.27%	(	0.0008	,	0.2211	,	0.2242	)
WA, 2018	1	27.75%	(	0.0008	,	0.2760	,	0.2790	)
*	2	27.83%	(	0.0008	,	0.2768	,	0.2798	)
WY, 2018	1	17.33%	(	0.0033	,	0.1668	,	0.1799	)

Program	Sample	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
*	2	18.23%	(	0.0034	,	0.1756	,	0.1890	)

Table showing split sample performance scores with standard deviation and 95% confidence intervals for 14 state Medicaid programs in 2018

*Cell intentionally left blank

Table 2a.11)B, Split Sample Topical Fluoride for Children, Dental Services, Measure Scores and 95% CI, CY2017 (T-MSIS Data)

Program	Sample	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
AK, 2017	1	14.47%	(	0.0019	,	0.1410	,	0.1484	)
*	2	14.41%	(	0.0019	,	0.1404	,	0.1478	)
AZ, 2017	1	18.91%	(	0.0007	,	0.1878	,	0.1904	)
*	2	18.87%	(	0.0007	,	0.1874	,	0.1900	)
DE, 2017	1	24.01%	(	0.0022	,	0.2358	,	0.2444	)
*	2	23.73%	(	0.0022	,	0.2330	,	0.2416	)
ID, 2017	1	19.84%	(	0.0014	,	0.1957	,	0.2011	)
*	2	19.75%	(	0.0014	,	0.1948	,	0.2003	)
MI, 2017	1	16.03%	(	0.0006	,	0.1592	,	0.1614	)
*	2	16.08%	(	0.0006	,	0.1597	,	0.1620	)
MS, 2017	1	20.38%	(	0.0010	,	0.2018	,	0.2058	)
*	2	20.32%	(	0.0010	,	0.2012	,	0.2052	)
NV, 2017	1	13.77%	(	0.0011	,	0.1357	,	0.1398	)
*	2	14.15%	(	0.0011	,	0.1394	,	0.1435	)
NM, 2017	1	25.15%	(	0.0012	,	0.2492	,	0.2538	)
*	2	25.07%	(	0.0012	,	0.2484	,	0.2530	)
NC, 2017	1	23.11%	(	0.0006	,	0.2299	,	0.2324	)

Program	Sample	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
*	2	23.00%	(	0.0006	,	0.2287	,	0.2312	)
OK,2017	1	18.65%	(	0.0010	,	0.1845	,	0.1885	)
*	2	18.78%	(	0.0010	,	0.1857	,	0.1898	)
OR, 2017	1	17.73%	(	0.0010	,	0.1752	,	0.1794	)
*	2	17.56%	(	0.0010	,	0.1736	,	0.1777	)
SC, 2017	1	22.37%	(	0.0008	,	0.2221	,	0.2252	)
*	2	22.43%	(	0.0008	,	0.2228	,	0.2259	)
WA, 2017	1	27.53%	(	0.0008	,	0.2738	,	0.2767	)
*	2	27.38%	(	0.0008	,	0.2723	,	0.2753	)
WY, 2017	1	17.99%	(	0.0034	,	0.1733	,	0.1865	)
*	2	17.50%	(	0.0033	,	0.1685	,	0.1815	)

Table showing split sample performance scores with standard deviation and 95% confidence intervals for 14 state Medicaid programs in 2017

*Cell left intentionally blank

Table 2a.11)C, Split Sample Topical Fluoride for Children, Dental Services, Measure Scores and 95% CI, CY2016 (T-MSIS Data)

Program	Sample	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
AK, 2016	1	16.25%	(	0.0021	,	0.1584	,	0.1666	)
*	2	16.02%	(	0.0021	,	0.1561	,	0.1642	)
AZ, 2016	1	9.16%	(	0.0005	,	0.0907	,	0.0926	)
*	2	9.15%	(	0.0005	,	0.0905	,	0.0924	)
DE, 2016	1	22.63%	(	0.0021	,	0.2222	,	0.2305	)
*	2	22.52%	(	0.0021	,	0.2210	,	0.2293	)
MI, 2016	1	15.37%	(	0.0006	,	0.1526	,	0.1549	)

Program	Sample	Score	*	SD	*	95% CI, lower bound	*	95% CI, upper bound	*
*	*	*	*	*	*	*	*	*	*
*	2	15.32%	(	0.0006	,	0.1521	,	0.1544	)
NV, 2016	1	8.90%	(	0.0009	,	0.0873	,	0.0906	)
*	2	8.98%	(	0.0009	,	0.0882	,	0.0915	)
NM, 2016	1	24.07%	(	0.0011	,	0.2385	,	0.2429	)
*	2	23.89%	(	0.0011	,	0.2367	,	0.2411	)
NC, 2016	1	22.28%	(	0.0006	,	0.2215	,	0.2240	)
*	2	22.33%	(	0.0006	,	0.2220	,	0.2245	)
OR, 2016	1	15.24%	(	0.0009	,	0.1506	,	0.1542	)
*	2	15.19%	(	0.0009	,	0.1501	,	0.1537	)
WA, 2016	1	27.47%	(	0.0007	,	0.2732	,	0.2762	)
*	2	27.33%	(	0.0007	,	0.2718	,	0.2747	)
WY, 2016	1	17.21%	(	0.0031	,	0.1661	,	0.1782	)
*	2	17.68%	(	0.0031	,	0.1707	,	0.1829	)

Table showing split sample performance scores with standard deviation and 95% confidence intervals for 14 state Medicaid programs in 2016

*Cell left intentionally blank

(B) ICC

The ICCs for each year 95% CI are shown in the table below:

Intraclass Correlation Coefficients	*	*	*	*	*	*	*
Year	**ICC**	**95% CI**	*	*	*	*	**p-value**
2018 (n-14 states)	0.997	(	0.9923	,	0.9992	)	<0.0001
2017 (n=14 states)	0.999	(	0.9962	,	0.9996	)	<0.0001
2016 (n=10 states)	0.999	(	0.9983	,	0.9999	)	<0.0001

Table showing intraclass correlation coefficient results for split sample measure scores in each year, 2016 through 2018

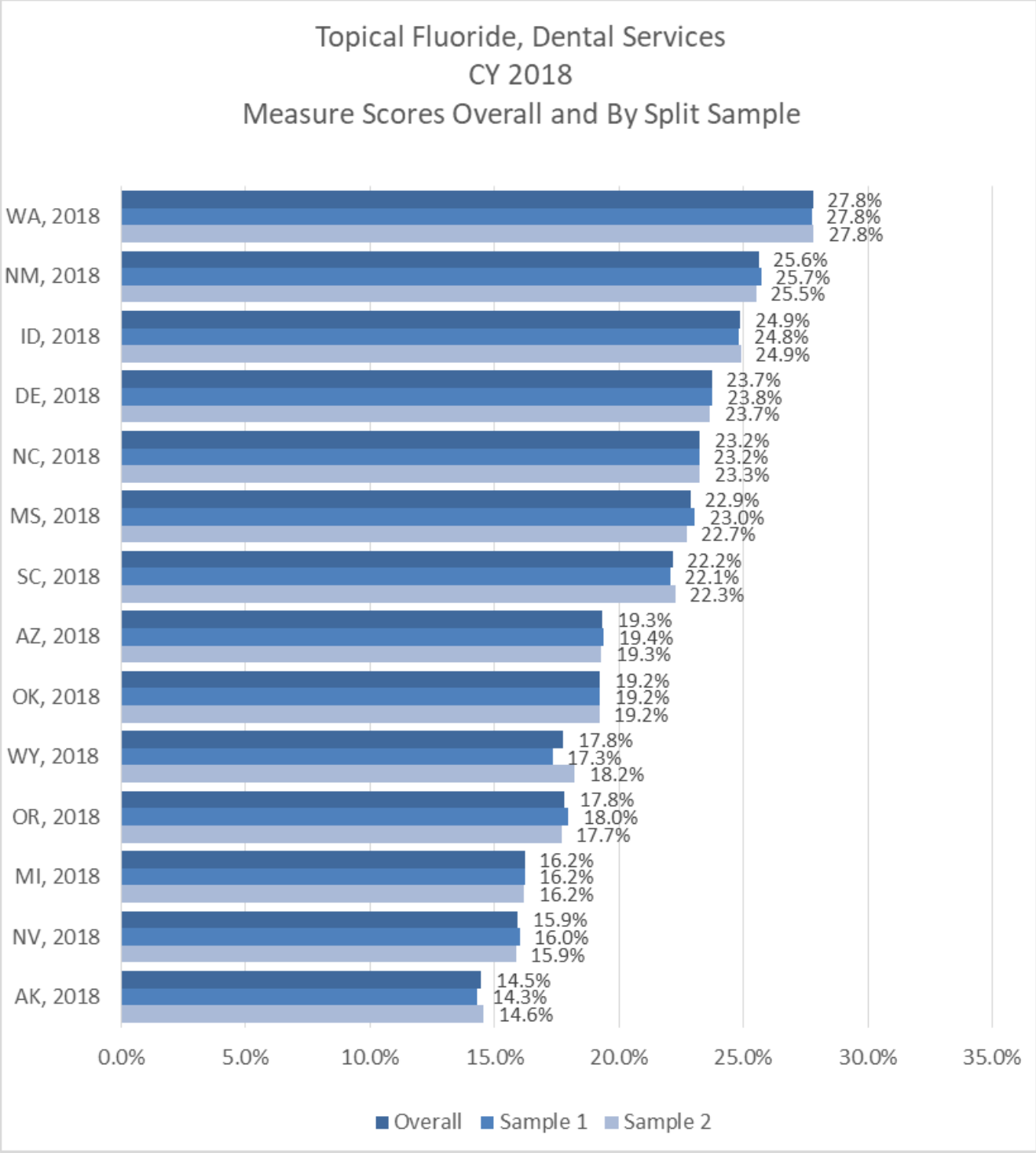
*Cell left intentionally blank

Following the guidance in Koo and Li (2016), this indicates "excellent" reliability given that even the lower bound of the 95% CIs are greater than 0.90.

(2) Relative Rankings: Split Samples

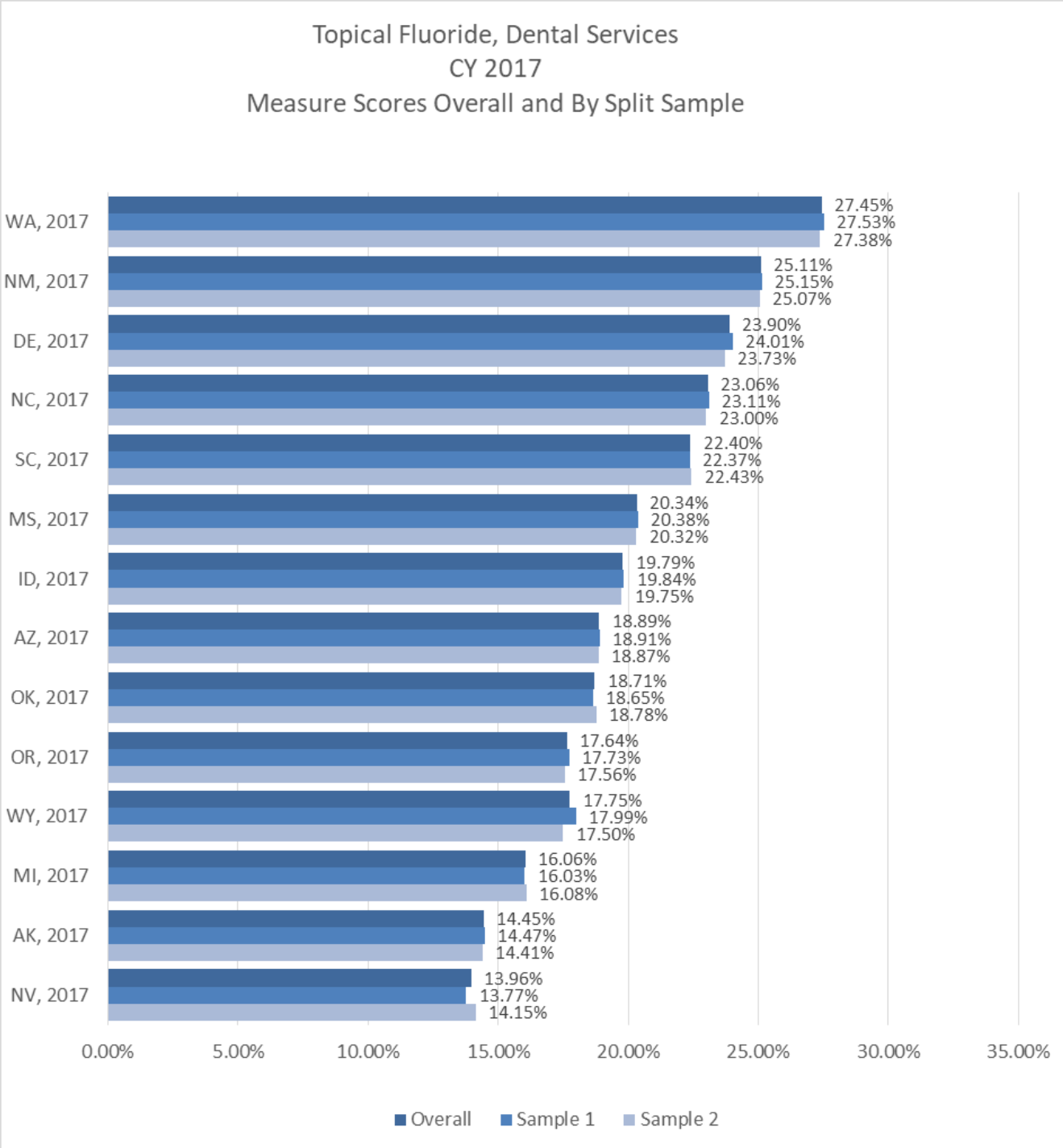
The figures below show the relative rankings of the state Medicaid programs, overall and for the split samples, for each year 2016-2018. The only changes in relative rankings were not significant. In 2018 and 2017, comparing split samples 1 and 2, Oregon and Wyoming switched places (between ranks 10 and 11). The measure scores for these two states have overlapping 95% CIs. The rankings were identical across split samples in 2016. Consequently, the relative rankings demonstrate high stability.

Figure 2a.11)-1, Relative Rankings, Overall Score and Split Samples for Topical Fluoride for Children, Dental Services, CY 2018



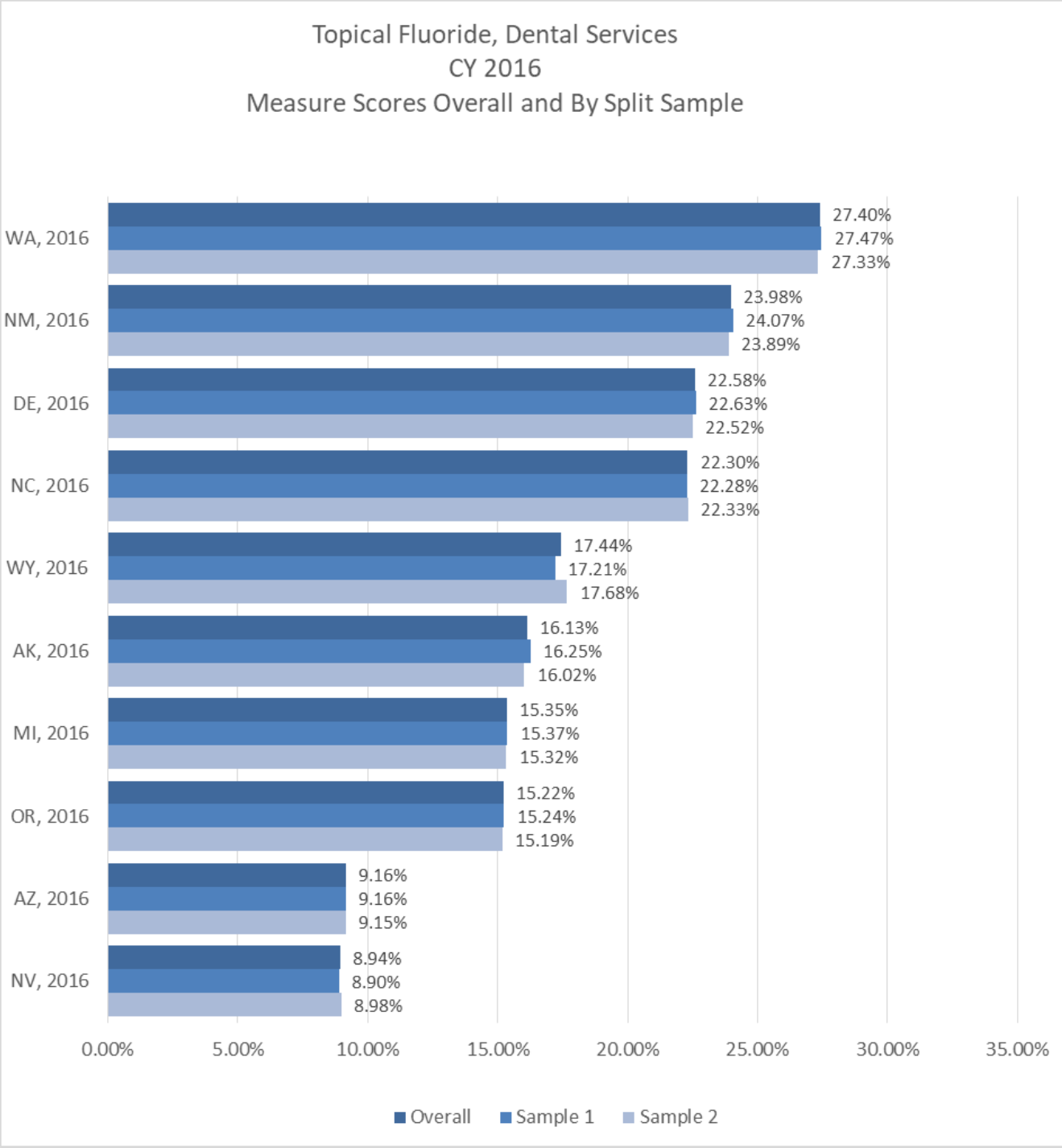
Relative Rankings, Overall Score and Split Samples for Topical Fluoride for Children, Dental Services, CY 2018

Figure 2a.11)-2, Relative Rankings, Overall Score and Split Samples for Topical Fluoride for Children, Dental Services, CY 2017



Relative Rankings, Overall Score and Split Samples for Topical Fluoride for Children, Dental Services, CY 2017

Figure 2a.11)-3, Relative Rankings, Overall Score and Split Samples for Topical Fluoride for Children, Dental Services, CY 2016



Relative Rankings, Overall Score and Split Samples for Topical Fluoride for Children, Dental Services, CY 2017

(3) Relative Rankings: Between Years

The correlation results are in the table below:

Year-to-Year Comparisons	Kendall's tau	p-value	Spearman's rank correlation coefficient	p-value
2017 & 2018 (n=14)	0.8462	<0.0001	0.9429	<0.0001
2016 & 2017 (n=10)	0.6889	0.0073	0.8303	0.0029

Table showing the relative rankings between years using both Kendall's tau and Spearman's rank correlation coefficient

The Kendall tau-b results indicate "moderate" to "strong" association in the state measure scores between years.

[Response Ends]

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

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

[Response Begins]

Testing for Spring 2022 Maintenance Evaluation

The testing results indicate that the accountable entity measure scores are reliable:

1. The measure scores between the split samples for each state are similar and have overlapping 95% CIs.
2. The ICC of the split sample measure scores is >0.9, including the lower bound of the 95% CI for the ICC. Following the guidance of Koo and Li (2016), this indicates "excellent" reliability.
3. The relative rankings based on measure scores are stable across the split samples.
4. The relative rankings based on measure scores are stable between years. The Kendall's tau values of 0.6889 indicates a "moderate" (2016-2017 comparisons) and 0.8462 "strong" (2017-2018 comparisons) degree of association. We have not located a definitive source regarding absolute cut points for what constitutes "weak", "moderate", or "strong" association. But based on what we have found in the literature collectively (e.g., Akoglu, 2018), we consider it is a fair characterization to classify the association as moderate to strong

Our original testing demonstrated the reliability of the data elements used to calculate the measures (by establishing data element validity). This updated testing demonstrates the reliability of the performance measure scores at the accountable entity (i.e., Medicaid program) level.

Reference: Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med. 2018;18(3):91-93. Published 2018 Aug 7. doi:10.1016/j.tjem.2018.08.001

PRIOR TESTING

Data element reliability: See section 2b for validity testing of data elements, which also establishes data element reliability.

[Response Ends]

2b. Validity

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

[Response Begins]

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

Accountable Entity Level (e.g. hospitals, clinicians)

Empirical validity testing

Systematic assessment of face validity of performance measure score as an indicator of quality or resource use (i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance)

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

Previous testing focused on data element validity and accountable entity measure score face validity. Therefore, we focused additional testing on accountable entity level validity testing. Ideally, to do so, we would identify a "gold standard" outcomes indicator of quality that is hypothesized to be associated with this process measure. However, we are very limited in available dental quality measures. The DQA was formed at the request of CMS precisely because there was a lack of validated dental quality measures. Consequently, the only rigorously tested quality measures against which this measure can be validated are other NQF-endorsed DQA measures. Moreover, these are mainly process measures. Dental claims data do not contain diagnosis codes, which significantly limits the ability to develop outcome measures, and the field of dentistry is in nascent stages of developing PRO-PMs.

Comparators for Correlation Analysis and Rationale for Hypothesized Relationships

We selected two other NQF-endorsed DQA measures that we hypothesized to be positively correlated with Topical Fluoride for Children, Dental Services:

1. Utilization of Services, Dental Services. This is a broad process measure that evaluates whether a child received any dental service during the reporting period, regardless of the type of service. It serves as an indirect indicator for access, indicating whether the child has had any contact with the dental care delivery system. State Medicaid programs in which a greater percentage of children are able to access the dental care delivery system (i.e., are able to get some type of dental service) were expected to have higher measure scores for Topical Fluoride than those in which fewer children are accessing the dental care system.
2. Oral Evaluation, Dental Services. This measures whether a child received a comprehensive or periodic oral evaluation during the reporting period. This is similar to a medical well-child visit as it represents a visit during which diagnostic care is provided to evaluate current oral disease and the risk for future disease, evaluate dentition development, and develop appropriate prevention and treatment plans. Children who receive a comprehensive or periodic oral evaluation represent a subset of those contained within the Utilization of Services measure. The subset of children who have a periodic or comprehensive oral evaluation are more likely to be established into routine care (versus seeking only problem- or emergency-based dental care). As a result, like Utilization of Services, we hypothesized a positive correlation between the Oral Evaluation and Topical Fluoride measures. Moreover, we anticipated a potentially stronger association since Oral Evaluation is more likely to reflect children who are established into routine care and, therefore, are more likely to receive recommended prevention.

Methodology

Using the measure scores for the 14 state Medicaid programs, we calculated Kendall's Tau-b, which is a rank correlation coefficient that measures association based on the number of concordant and discordant pairs. We assessed concordance between: (1) Topical Fluoride and Utilization of Services and (2) Topical Fluoride and Oral Evaluation. For reference, we also report the more commonly reported Spearman's rank

correlation coefficients. Although the strength of these associations is stronger, we felt Kendall's tau was the more appropriate test to report given the relatively small sample size of 14 states.

PRIOR TESTING

We assessed (1) critical data element validity, (2) measure score validity, and (3) potential threats to validity.

1. CRITICAL DATA ELEMENT VALIDITY

Topical Fluoride measures the percentage of children aged 1-21 years at moderate to high risk for dental caries who had at least 2 topical fluoride applications during the reporting year. The critical data elements for this measure include: (1) member ID (to link between claims and enrollment data), (2) date of birth, (3) monthly enrollment indicator, (4) date of service, and (5) CDT codes. The first four items are core fields used in virtually all measures relying on administrative data and essential for any reporting or billing purposes. As such, it was determined that these fields have established reliability and validity. Thus, **critical data element validity testing focused on assessing the accuracy of the dental procedure codes reported in the claims data as the data elements that contribute most to the measure score.** To evaluate data element validity, we conducted reviews of dental records for the Texas Medicaid program. Validation of clinical codes in administrative claims data are most often conducted using manual abstraction from the patient's full chart as the authoritative source. As described in detail below, we evaluated 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.

Data Sources & Methodology

A. Data Sources

A random sample of encounters for members ages 3-18 years with at least one outpatient dental visit was selected for dental record reviews. The targeted number of records was 400. The expected response rate for returning records was 65%. Therefore, 600 records were requested. All outpatient dental records for members during an eight-month period were requested. Table 2b2.2-1 below summarizes the number of records requested and received. The number of eligible records received (414) exceeded the total targeted number of 400 records.

Table 2b2.2-1 Dental Records Requested and Received

# Requested	# Received	% Received
600	414	69%

Dental Records Requested and Received

B. Record Review Methodology

There were two components to the record reviews used to evaluate data element validity:

1. Encounter data validation (EDV) that provided an **overall assessment** of the accuracy of dental procedure codes found in the administrative claims data compared to dental records for the same dates of service.
2. Validation of **topical fluoride application procedure codes specifically.**

The record reviews were conducted by two coders certified as registered health information technicians (RHITs). At weekly intervals during the record review process, the two RHITs randomly selected a sample of records to evaluate inter-rater reliability. A total of 100 records and 1,830 fields were reviewed by both individuals with 100% agreement.

C. Encounter Data Validation – Overall Assessment

For the first component of validation, encounter data validation, the research team followed standard Encounter Data Validation processes following External Quality Review protocols from CMS that it has used in ongoing quality assurance activities for the Texas Health and Human Services Commission. [Centers for Medicare and Medicaid Services, External Quality Review Encounter Data Validation Protocol

(http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Quality-of-Care/Quality-of-Care-External-Quality-Review.html)]. The first three procedure codes were reviewed for each claim. A total of 1,135 procedure codes were reviewed. The RHITs were provided with a pre-populated data entry form with the codes from the claims data for the patient with the specified provider on a particular date of service. They evaluated whether the code in the claims data was supported by the dental record.

D. Critical Data Element Validation – Topical Fluoride Application Procedure Codes

Data Extraction. For the second component of validation, assessing whether the specific preventive service of topical fluoride application is accurately captured by claims data, chart abstraction forms were developed by the research team. The chart abstraction forms and process were reviewed and approved by the DQA R&D Committee. Claims data were validated against dental records by comparing the dental records to the codes in the claims data for a randomly selected date of service. Prior to conducting the reviews, a sample of 30 records from prior encounter data validation activities was used to test the data abstraction tool and refinements were made accordingly. During the chart abstraction testing process, the RHITs met with the research team, which included two dentists (including a pediatric dentist), to review questions about interpreting the records. They then evaluated the 414 dental records using the data abstraction form. The results were recorded in an Access database. Specifically, the chart abstracting process involved identifying and recording whether there was any evidence of fluoride application during the visit. The programming team extracted data from the administrative claims data for the same members and dates of service, recording the presence or absence of topical fluoride procedure codes. The data files from the record review team and the programming team were merged into a single data file.

Statistical Analysis. To assess validity, we calculated sensitivity (accuracy of administrative data indicating a service was received when it is present in the chart), specificity (accuracy of administrative data indicating a service was not received when it is absent in the chart), positive predictive value (extent to which a procedure that is present in the administrative data is also present in the charts), and negative predictive value (extent to which a procedure that is absent from the administrative data is also absent in the chart). Positive and negative predictive values are influenced by sensitivity and specificity **as well as the prevalence of the procedure**. Thus, interpretation of "high" and "low" values is not straightforward. In addition, although charts are typically used as the authoritative source for validating claims data, some question whether charts always represent an "authoritative" source versus being better characterized as a "reference" standard. The kappa statistic has been recommended as "a more 'neutral' description of agreement between the 2 data sources . . . ." (Quan H, Parsons GA, Ghali WA, Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data, Med Care, 2004;42(8):801-809.) Thus, the kappa statistic also was used to compare the degree of agreement between the two data sources. A kappa statistic value of 0 reflects the amount of agreement that would be expected to be observed by chance. A kappa statistic value of 1 indicates perfect agreement. Guidance on interpreting the kappa statistic is: <0 (poor/less chance of agreement; 0.00-0.20 (slight agreement); 0.21-0.40 (fair agreement); 0.41-0.60 (moderate agreement); 0.61-0.80 (substantial agreement); 0.81-0.99 (almost perfect agreement). (Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics. Jun 1977;33(2):363-374.)

2. MEASURE SCORE - FACE VALIDITY

Face validity of this measure was assessed at several stages during the measure development and testing processes.

A. Face Validity Assessment – Measure Development

Face validity was **systematically assessed by recognized experts**. The Dental Quality Alliance (DQA) was formed at the request of the Centers of Medicare and Medicaid Services (CMS) specifically for the purpose of bringing together recognized expertise in oral health to develop quality measures through consensus

processes. As noted in the letter from Cindy Mann, JD, Director of the Center for Medicaid & CHIP Services within CMS: "The dearth of tested quality measures in oral health has been a concern to CMS and other payers of oral health services for quite some time." (See Appendix)

During the measurement development process, the DQA Research and Development Committee, purposely comprised of individuals with recognized and appropriate expertise in oral health to lead quality measure development, undertook an environmental scan of existing pediatric oral health performance measures, which involved the following: (1) Literature Search, (2) Measure Solicitation, (3) Review of Measure Concepts, (4)Delphi Ratings of Measure Concepts, (5) Scan Results Analysis, (6) Gap Analysis, (7) Identification of Measures. A more detailed description of this process, the findings and the resulting measure concepts that were pursued is provided in reports published by the DQA. (Dental Quality Alliance. Pediatric Oral Health Quality and Performance Measures: Environmental Scan. 2012; Dental Quality Alliance. Pediatric Oral Health Quality & Performance Measure Concept Set: Achieving Standardization & Alignment. 2012. Both reports available at: http://ada.org/7503.aspx.)

(1) Literature Search. The Committee began its work by identifying existing performance and quality measure concepts (description, numerator, and denominator) on pediatric populations defined as children younger than 21 years. Staff conducted a comprehensive online search for publicly available measure concepts. This search was conducted initially in August – September 2011 and then updated on February 8, 2012. The following searches were conducted: (1) PubMed Search. Staff used two specific search strategies to search Medline. Search 1: (performance OR process OR outcome OR quality) AND measure AND (oral or dental) AND (children OR child OR pediatric OR paediatric) – 1121 citations. Search 2 - "Quality Indicators, Health Care"[Mesh] AND (dental OR oral) - 150 citations. Staff included five articles based on title and abstract review of these citations. Measure concepts presented within these articles were included in the list of concepts for R&D Committee review. (2) Web Search. Staff then performed an internet search with keywords similar to the ones used for the PubMed search. (3) Search of relevant organization websites. Staff began this search through the links provided within the National Library of Medicine database of relevant organizations (f). Example of organizations involved in quality measurement include the National Quality Measures Clearinghouse (NQMC), National Quality Forum (NQF), and Maternal and Child Health Bureau (MCHB).

(2) Solicitation of Measures. In addition, the R&D Committee contacted staff at the Agency for Healthcare Research and Quality (AHRQ) in August 2011 to obtain the measures collected by the Subcommittee on Children's Healthcare Quality for Medicaid and CHIP programs (SNAC). The Committee solicited measures from other entities, such as the DentaQuest Institute, involved in measure development activities.

(3) Review of Measure Concepts. Using inclusion/exclusion criteria, the R&D Committee reviewed the measure concepts and identified the measures that would be reviewed and rated in greater depth.

(4) Delphi Ratings. The RAND-UCLA modified Delphi approach was used to rate the remaining measure concepts, applying the criteria and scoring system for importance, validity, and feasibility consistent with the process that was used by the SNAC. There were two rounds of Delphi ratings to identify a starter set of pediatric oral health performance measures. [Brook RH. The RAND/UCLA appropriateness method. In: McCormick KA, Moore SR, Siegel R, United States. Agency for Health Care Policy and Research. Office of the Forum for Quality and Effectiveness in Health Care., editors. Clinical practice guideline development : methodology perspectives.]

(5) Scan Results. There were a total of 112 measure concepts identified through the environmental scan: 59 met the inclusion criteria for being processed through the Delphi rating process and 53 did not. Among the 59 measures that were evaluated through the Delphi rating process, 38 were deemed "low-scoring measure concepts" and 21 were deemed "high-scoring measure concepts."

(6) Gap Analysis. The R&D Committee then identified the gaps in existing measures, including both gaps in terms of the care domains addressed (e.g., use of services, prevention, care continuity) as well as gaps based

on good measurement practices (e.g., standardized measurement methodology, evidence-based, etc.). Although the Committee did identify content areas that were not addressed, **a key finding was the lack of standardized, clearly-specified, validated measures.**

(7) Identification of Measures. The findings were used to identify a starter set of measures that would achieve the following objectives: (a) uniformly assess the quality of care for comparison of results across private/public sectors and across state/community and national levels; (b) inform performance improvement projects longitudinally and monitor improvements in care; (c) identify variations in care, and (d) develop benchmarks for comparison.

B. Face Validity Assessment – Measure Testing

The research team and the DQA R&D Committee continued to assess face validity throughout the testing process. Face validity also was gauged through feedback solicited through public comment periods. In March 2013, an Interim Report describing the measures, testing process, and preliminary results was sent to a broad range of stakeholders, including representatives of federal agencies, dental professionals/professional associations, state Medicaid and CHIP programs, community health centers, and pediatric medical professionals/professional associations. Each comment received was carefully reviewed and addressed by the research team and DQA, which entailed additional sensitivity testing and refinement of the measure specifications. Draft measure specifications were subsequently posted on the DQA's website in a public area and public comment was invited. National presentations, including presentations at the National Oral Health Conference, were made by the research team and DQA in the spring and summer of 2013, which included reference to the website containing the measure specifications and invitations to provide feedback. All comments received were reviewed and addressed by the research team and DQA, including additional sensitivity testing and refinement of the measure specifications.

The final face validity assessment was conducted at the July 2013 Dental Alliance Quality meeting at which the full membership, representing a broad range of stakeholders. A detailed presentation of the testing results was provided. The membership then participated in an open consensus process with observed unanimous agreement that the calculated measure scores can be used to evaluate quality of care.

Sample Presentations

Aravamudhan K. Dental Quality Alliance Measures. Presentation at 2013 National Oral Health Conference Pre-Conference Workshop on Objectives, Indicators, Measures and Metrics. 2013.

Herndon JB. DQA Pediatric Oral Health Performance Measure Set: Overview of Measures and Validation Process. Presentation at 2013 National Oral Health Conference Pre-Conference Workshop on Objectives, Indicators, Measures and Metrics. 2013.

Herndon JB. DQA Pediatric Oral Health Performance Measure Set: Overview of Measures and Validation Process. Presentation at 2013 Texas Medicaid and CHIP Managed Care Quality Forum. 2013.

3. ADDITIONAL VALIDITY TESTING - IDENTIFYING ELEVATED RISK WITH CLAIMS DATA

Evidence based guideline indicate that fluoride is most effective for children at higher risk for caries. Thus, inclusion in the denominator is limited to children identified as being at moderate to high risk for caries. Administrative claims data for dental claims typically do not include diagnostic codes. Procedure codes for risk assessment that identify moderate and high risk were included in the measure logic. However, because these are newer codes, additional logic was included to identify children with recent history of restorations, which are indicative of caries. **A systematic review found that prior caries experience to be an important predictor of future risk** (Zero D, Fontana M, Lennon AM. 2001. Clinical applications and outcomes of using indicators of risk in caries management. J Dent Educ. 2001 Oct;65(10):1126-32.) Expert consensus and

validation through chart reviews was done to finalize the procedure codes (indicated in the measure specifications) used to identify elevated risk. The test data results reported in this application demonstrate that it is feasible to use these validated codes to identify children at elevated risk who should receive preventive services.

4. ADDITIONAL VALIDITY EVALUATION – ASSESSMENT OF THREATS TO VALIDITY

A. Exclusions

As described in 2b3. of this form, there are no exclusions for this measure.

B. Risk Adjustment

Risk adjustment is not applicable for this process measure.

C. Missing Data

As described in measure evaluation criteria 3c1, this measure relies on standard data elements in claims data that are already collected and widely used for a range of reporting and billing purposes with very low rates of missing or invalid data (which we empirically assessed and reported in 3c1).

D. Multiple Sets of Specifications

This does not apply to the proposed measure.

E. Ability to Identify Statistically Significant and Meaningful Differences in Performance

As described in 2b5 of this form, this measure is able to identify statistically significant and meaningful differences in performance. We also demonstrate with empirical data and statistical testing the ability of this measure to detect disparities in 1b4 (Importance).

[Response Ends]

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

Examples may include correlations or t-test results.

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

The rank correlation results, using Kendall's tau were:

1. Topical Fluoride and Utilization of Services

*	Kendall's tau	p-value	Spearman's rank correlation coefficient	p-value
Topical Fluoride & Utilization of Services	*	*	*	*
2018 (n=14)	0.6703	0.0010	0.8286	0.0003
2017 (n=14)	0.6264	0.0022	0.8022	0.0006
2016 (n=10)	0.8222	0.0013	0.9394	0.0001

Table showing the rank correlation results using Kendall's tau and Spearman's rank correlation efficient for topical fluoride and utilization of services measures

*Cell left intentionally blank

2. Topical Fluoride and Oral Evaluation

Topical Fluoride & Oral Evaluation	Kendall's tau	p-value	Spearman's rank correlation coefficient	p-value
2018 (n=14)	0.7802	0.0001	0.9253	<0.0001
2017 (n=14)	0.8022	0.0001	0.9253	<0.0001
2016 (n=10)	0.8667	0.0007	0.9515	<0.0001

Table showing the rank correlation results using Kendall's tau and Spearman's rank correlation efficient for topical fluoride and oral evaluation measures

PRIOR TESTING

1. CRITICAL DATA ELEMENT VALIDITY

A. Encounter Data Validation – Overall Assessment

Encounter data validation of 1,135 procedure codes in the claims data against dental charts found agreement for 94% of the procedure codes (Table 2b2.3-1). Only 4.2% of procedure codes reported in the administrative data were not supported by evidence in the dental record. For 1.8% of the records reviewed, the documentation was insufficient to determine whether the service indicated by the procedure code had been rendered or not.

Table 2b2.3-1 Agreement between Records and Administrative Data for Procedures

Number of Procedure Codes	Record and Procedure Code on Claim Correlate	Record Did Not Correlate with Procedure Code on Claim	Unable to Determine Correlation
1,135	94.04%	4.22%	1.75%

Table showing the agreement between records and administrative data for procedures

95% confidence intervals indicated in parentheses

B. Critical Data Element Validation – Topical Fluoride Application Procedure Codes

To assess whether the specific preventive service of topical fluoride application is accurately captured by claims data, the 414 records, representing 631 dates of service, were reviewed. Table 2b2.3-2 below summarizes the agreement between the dental records and administrative data for topical fluoride applications. 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%. As noted above, the kappa statistic provides a more neutral description of agreement and extends a comparison of simple agreement by taking into account agreement occurring by chance, thereby providing a more rigorous and conservative measure of agreement between the two data sources. The kappa statistic value was 0.782, which is at the high end of the "substantial agreement" category.

Table 2b2.3-2 Agreement between Record and Administrative Data for Specific Services

*	Concordance	Prevalence	Sensitivity	Specificity	PPV	NPV	Kappa
Fluoride	89.91%	0.647	0.907	0.884	0.935	0.839	0.782
95% confidence intervals indicated in parentheses	*	*	(0.857-0.942)	(0.806-0.934)	(0.888-0.963)	(0.757-0.898)	(0.710-0.853)

*	Concordance	Prevalence	Sensitivity	Specificity	PPV	NPV	Kappa
Dates of service: 317	*	*	*	*	*	*	*
# indeterminate: 0	*	*	*	*	*	*	*

Table showing the agreement between records and administrative data for specific services

*Cell left intentionally blank

Our findings are similar to those in the peer-reviewed literature. A study was conducted in 2004 that used data from 3,751 patient visits in 120 dental practices participating in the Ohio Practice-Based Research Network to examine the concordance of chart and billing data with direct observation of dental procedures. For fluoride, they found lower sensitivity (80%), higher specificity (98%) and similar kappa value (0.81) of billing data compared to direct observation. (Demko CA, Victoroff KZ, Wotman S. 2008. "Concordance of chart and billing data with direct observation in dental practice" Community Dent Oral Epidemiol. 36(5):466-74.)

2. FACE VALIDITY

The measures concept of preventive dental services identified using CDT codes (within which topical fluoride falls) was identified through the Delphi rating process as a high-scoring measure concept with a mean importance score of 7, mean feasibility score of 8, and mean validity score of 7 for specific evidence-based preventive services, all out of a 9-point scale. [Rating of 1-3: not scientifically sound and invalid; 4-6 – uncertain scientific soundness and uncertain validity; 7-9 – scientifically sound and valid.] Thus, the measure has face validity. However, gaps were identified with existing preventive services measures, including defining "preventive services" too broadly (encompassing services without sound evidence of their effectiveness in caries prevention), lack of clear specifications and lack of standardization. Although the scan included two measure concepts that were specific to fluoride, they were deemed to be low scoring because they pertained to "fluoride supplements" or "fluoride exposure assessment." Scientific soundness was limited due to lack of clarity in measure description.

Content Validity. In addition, the measure also demonstrates **content validity** – the extent to which the measure specifications reflect the intended domain of care. This measure directly reflects evidence-based guidelines regarding an effective caries prevention measure (professionally applied topical fluoride), including the frequency required for clinical effectiveness (at least every three-six months). Please see the Measure Evidence Form for more details.

[Response Ends]

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

The findings were as hypothesized.

1. Correlations were in the expected direction. Both correlations demonstrated a positive relationship. Higher measure scores for Utilization of Services were associated with higher measure scores for Topical Fluoride. Higher measure scores for Oral Evaluation were associated with higher measures scores for Topical Fluoride.
2. Correlations were statistically significant.
3. Correlations demonstrated moderate to strong strength of association with Kendall's tau ranging from 0.67 to 0.82 for the association between Topical Fluoride and Utilization of Services and 0.78-0.87 for

the association between Topical Fluoride and Oral Evaluation. We have not located a definitive source regarding absolute cut points for what constitutes "weak", "moderate", or "strong" association. But based on what we have found in the literature collectively (e.g., Akoglu 2018), we consider it is a fair characterization to classify the association as moderate to strong.

4. As hypothesized, the strength of association was greater when examining the relationship between the Topical Fluoride and Oral Evaluation measure scores compared with the association between the Topical Fluoride and Utilization of Services measures scores.

Collectively, these results complement our prior testing, which established data element validity and measure score face validity, by lending further support of the validity of the measure scores at the accountable entity level through empirical testing.

Reference: Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med. 2018;18(3):91-93. Published 2018 Aug 7. doi:10.1016/j.tjem.2018.08.001

PRIOR TESTING

As noted above, the overall agreement between the administrative claims data and dental record data was high based on both simple agreement and using the more conservative Kappa statistic. Overall, we interpret these findings as evidence that validates the accuracy of administrative claims data for performance measurement purposes. These empirical findings, combined with our face validity and content validity assessments of the measure score, lead us to conclude that both the data elements and the measure score represent valid measures of the evidence-based preventive service topical fluoride application.

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

(1) We evaluated performance score data with 95% confidence intervals to assess whether there were statistically significant differences in the measure scores between programs.

(2) We calculated the mean, median, standard deviation, and percentile distribution (10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th}).

(3) To illustrate meaningful differences in performance, we calculated the interquartile range for the measure rates. We used the chi-square test to evaluate statistically significant differences in measure scores between programs in the lowest and highest quartiles.

PRIOR TESTING

This is a new measure. As noted in 1b, there were variations in the measure scores across the four programs included in the testing. For convenience we have included the performance score data from 1b below. In addition to providing the 95% confidence intervals for each score, we used chi-square tests to analyze whether there were statistically significant differences between (1) the 3 programs with performance data for 2011, (2) the 4 programs with performance data for 2010, (3) the two dental MCOs in FL CHIP in CY 2010 and (4) the two dental MCOs in FL CHIP in CY 2011. Because the measure score is the proportion of children who received two topical fluoride applications, the dichotomous outcome of had/did not have two topical fluoride applications can be used to conduct chi-square significance testing in order to evaluate whether there are statistically significant differences in the measure scores between programs and between plans.

Table 1b.2. Performance Scores

Program/Plan Year	Measure Score as %	*	Measure Score	SD	*	*	Lower 95% CI	*	Upper 95% CI	*
Program 1, CY 2011:	37.13%	(	0.3713	,	0.0004	,	0.3704	,	0.3722	)
Program 2, CY 2011:	27.15%	(	0.2715	,	0.0020	,	0.2676	,	0.2754	)
Program 3, CY 2011:	22.04%	(	0.2204	,	0.0020	,	0.2165	,	0.2243	)
Program 1, CY 2010:	34.96%	(	0.3496	,	0.0005	,	0.3487	,	0.3505	)
Program 2, CY 2010:	22.63%	(	0.2263	,	0.0019	,	0.2225	,	0.2301	)
Program 3, CY 2010:	35.04%	(	0.3504	,	0.0023	,	0.3458	,	0.3550	)
Program 4, CY 2010:	18.16%	(	0.1816	,	0.0009	,	0.1799	,	0.1833	)
Plan 1, CY 2011:	25.50%	(	0.2550	,	0.0030	,	0.2491	,	0.2609	)
Plan 2, CY 2011:	28.69%	(	0.2869	,	0.0027	,	0.2815	,	0.2923	)
Plan 1, CY 2010:	23.24%	(	0.2324	,	0.0048	,	0.2230	,	0.2418	)
Plan 2, CY 2010 :	23.76%	(	0.2376	,	0.0034	,	0.2309	,	0.2443	)

Table showing performance scores for each program and plan each year, with standard deviation and 95% confidence intervals

*Cell left intentionally blank

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

(1) Performance Scores with 95% CIs: CY 2018

Program	Score	*	SD	*	95% CI , lower bound	*	95% CI , upper bound	*
AK, 2018	14.46%	(	0.0013	,	0.1421	,	0.1472	)
NV, 2018	15.95%	(	0.0008	,	0.1579	,	0.1610	)
MI, 2018	16.20%	(	0.0004	,	0.1612	,	0.1628	)
WY, 2018	17.77%	(	0.0024	,	0.1731	,	0.1824	)
OR, 2018	17.82%	(	0.0007	,	0.1767	,	0.1796	)
OK, 2018	19.23%	(	0.0008	,	0.1908	,	0.1938	)
AZ, 2018	19.33%	(	0.0005	,	0.1923	,	0.1942	)
SC, 2018	22.17%	(	0.0006	,	0.2206	,	0.2228	)
MS, 2018	22.88%	(	0.0008	,	0.2272	,	0.2303	)
NC, 2018	23.24%	(	0.0005	,	0.2315	,	0.2333	)
DE, 2018	23.72%	(	0.0015	,	0.2342	,	0.2402	)
ID, 2018	24.87%	(	0.0011	,	0.2465	,	0.2509	)
NM, 2018	25.62%	(	0.0009	,	0.2545	,	0.2579	)
WA, 2018	27.79%	(	0.0005	,	0.2769	,	0.2790	)

Table showing performance scores for each program in calendar year 2018, with standard deviation and 95% confidence intervals

*Cell left intentionally blank

CY 2017

Program	Score	*	SD	*	95% CI , lower bound	*	95% CI , upper bound	*
NV, 2017	13.96%	(	0.0007	,	0.1381	,	0.1411	)
AK, 2017	14.45%	(	0.0013	,	0.1419	,	0.1471	)
MI, 2017	16.06%	(	0.0004	,	0.1598	,	0.1614	)
OR, 2017	17.64%	(	0.0007	,	0.1749	,	0.1779	)
WY, 2017	17.75%	(	0.0024	,	0.1729	,	0.1821	)
OK, 2017	18.71%	(	0.0007	,	0.1857	,	0.1886	)
AZ, 2017	18.89%	(	0.0005	,	0.1879	,	0.1898	)
ID, 2017	19.79%	(	0.0010	,	0.1960	,	0.1999	)
MS, 2017	20.34%	(	0.0007	,	0.2020	,	0.2048	)
SC, 2017	22.40%	(	0.0006	,	0.2229	,	0.2251	)
NC, 2017	23.06%	(	0.0005	,	0.2297	,	0.2315	)
DE, 2017	23.90%	(	0.0015	,	0.2360	,	0.2420	)

Program	Score	*	SD	*	95% CI , lower bound	*	95% CI , upper bound	*
NM, 2017	25.11%	(	0.0008	,	0.2494	,	0.2527	)
WA, 2017	27.45%	(	0.0005	,	0.2734	,	0.2755	)

Table showing performance scores for each program in calendar year 2017, with standard deviation and 95% confidence intervals

*Cell left intentionally blank

CY 2016

Program	Score	*	SD	*	95% CI , lower bound	*	95% CI , upper bound	*
NV, 2016	8.94%	(	0.0006	,	0.0882	,	0.0906	)
AZ, 2016	9.16%	(	0.0003	,	0.0909	,	0.0922	)
OR, 2016	15.22%	(	0.0007	,	0.1509	,	0.1535	)
MI, 2016	15.35%	(	0.0004	,	0.1527	,	0.1542	)
AK, 2016	16.13%	(	0.0015	,	0.1584	,	0.1642	)
WY, 2016	17.44%	(	0.0022	,	0.1701	,	0.1787	)
NC, 2016	22.30%	(	0.0004	,	0.2222	,	0.2239	)
DE, 2016	22.58%	(	0.0015	,	0.2229	,	0.2288	)
NM, 2016	23.98%	(	0.0008	,	0.2383	,	0.2414	)
WA, 2016	27.40%	(	0.0005	,	0.2729	,	0.2750	)

Table showing performance scores for each program in calendar year 2016, with standard deviation and 95% confidence intervals

*Cell left intentionally blank

(2) Performance Distribution and Chi-Square Results, CY2016—CY2018

*	2018 (n=14)	2017 (n=14)	2016 (n=10)
Mean	0.2079	0.1979	0.1809
SD	0.0408	0.0416	0.0610
Minimum	0.1446	0.1396	0.0894
10th Percentile	0.1595	0.1445	0.0905
25th Percentile	0.1777	0.1606	0.1535
Median	0.2075	0.1934	0.1679
75th Percentile	0.2372	0.2306	0.2258
90th Percentile	0.2562	0.2511	0.2569
Maximum	0.2779	0.2745	0.2740

*	2018 (n=14)	2017 (n=14)	2016 (n=10)
Interquartile Range	0.0595	0.0700	0.0724
Chi-Square p-value	<0.0001	<0.0001	<0.0001

Table showing the descriptive statistics for each calendar year

*Cell left intentionally blank

PRIOR TESTING

For both years, statistically significant differences were detected in the measure scores between programs in both years and between plans in one of the two years (Table 2b5.2).

Table 2b5.2. Chi-Square Test of Differences in Measure Scores

*	Chi-Square Value	*p*- value
Program Results, 2011	5887.1	<0.0001
Program Results, 2010	23554.5	<0.0001
Plan Results, 2011	61.2	<0.0001
Plan Results, 2010	0.8	0.3711

Table showing the chi-square test of differences in measure scores

*Cell left intentionally blank

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

The performance scores with 95% confidence intervals demonstrate the ability of the measure to detect statistically significant differences in performance between measured entities. The mean and median measure scores in CY 2018 were both 21% with scores ranging from 14% to 28%. The maximum value of 28% indicates that even in the highest performing state, most Medicaid beneficiaries are not receiving at least two topical fluoride applications, which offer primary prevention against dental caries. This indicates a significant performance gap and opportunity for improvement overall among all measured entities. In addition, there is practically significant variation in performance between states with a two-fold difference between the highest and lowest performing states. The interquartile range varied from 5.95% in 2018 to 7.24% in 2016. There were statistically significant differences in measure rates between programs in the lowest and highest quartiles in each of the three years ($p<0.0001$) Consequently, the measure enables identification of both statistically significant and clinically meaningful differences in performance.

PRIOR TESTING

Statistically significant differences between measured entities were detected at both the program and plan reporting levels, with program-level performance scores ranging by approximately 17 percentage points. At the plan level, statistically significant differences were detected in 2011, but not in 2010. This is consistent

with a greater difference in performance between the two plans in 2011 (25.50% and 28.69%) than in 2010 when the rates were almost equal (23.24% and 23.76%). This is precisely the purpose of performance measurement - to detect when there are differences in performance. In 2010, there was no appreciable difference in performance between the two plans. Collectively, however, it is clear that this measure detects differences in performance on the measure scores when they do exist. Our findings are consistent with evidence reported elsewhere in this application documenting a performance gap and disparities in performance. Thus, Topical Fluoride informs performance improvement efforts by allowing plans and programs to identify and monitor performance gaps and disparities both at any given point in time and over time.

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

Updated testing used data from 14 Medicaid programs using data submitted by the states to the Centers for Medicaid and Medicare Services contained within the Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAFs). We assessed data quality and missing/invalid data through two methods.

1. CMS T-MSIS Data Quality Atlas. The Medicaid and CHIP Business Information Solutions (MACBIS) conducted data quality assessments of T-MSIS enrollment, claims, expenditures and service use for each state and for each year/release. (Centers for Medicare & Medicaid Services. *DQ Atlas*. Available at https://www.medicaid.gov/dq-atlas/ as of November 2021.) There is a background and methodology report for each topic assessed.

For each state, the Atlas assigns one of the values listed below to indicate the extent to which a state's TAF data are usable, reliable, and accurate for analyzing a particular topic.

- **Low concern**: No major problems were identified that would affect the usability of the TAF data for analyzing a given topic.
- **Medium concern**: Some problems were identified that may affect the usability of the TAF data for analyzing a topic.
- **High concern**: Major problems in the completeness or reliability of the TAF data are likely to impede an analysis of a topic.
- **Unusable**: Extreme problems in the completeness or reliability of the TAF data will prevent a topic from being analyzed.
- **Unclassified**: The topic is either not applicable to a state, or there were not enough TAF or benchmark data for a reliable analysis, or a methodological issue prevented a state's data from being classified into one of the four categories above.

We reviewed the results of these assessments for the following topics (with their descriptions contained within the Quality Atlas) that are relevant to the calculation of Topical Fluoride for Children:

- **Age.** This analysis examines the completeness and distribution of beneficiary age information in the TAF.

- **Medicaid enrollment.** This analysis examines how well the TAF data on the number of total Medicaid beneficiaries align with an external benchmark, the Performance Indicators data set.

- **Claims file completeness: Claims Volume - other services (includes outpatient).** Examining the volume of service use records adjusted for program size can identify outlier states that may have incomplete claims, encounter records, or eligibility data in the TAF. This analysis examines the volume of OT header records, the volume of OT line records, and the average number of lines per header.

- **Claims file completeness: Service Users - other services (includes outpatient).** Examining the overall percentage of beneficiaries with any service use can identify outlier states that may have incomplete claims, encounter, or eligibility data in the TAF. Low rates of service use may also indicate problems in linking service use and eligibility records. This analysis examines the percentage of beneficiaries in each state with an OT record indicating the receipt of ambulatory, physician, or other medical services during the year.

- **Service use - Procedure Codes - other services (includes outpatient).** This analysis examines how often the procedure code is missing on professional claims in the OT file and how often the non-missing values on these claims represent valid national or state-specific codes.

2. **Additional Evaluations.** We conducted our own assessments of the following data fields:
   - **Date of Birth**. We evaluated how frequently date of birth was missing.

   - **Beneficiary ID.** We evaluated how frequently beneficiary ID was missing among children <21 years.

   - **Dental Procedure Codes (CDT codes).** For each year, we used the list of active and valid procedure codes for each year available from the American Dental Association to evaluate how often non-missing values represent non-valid or non-active codes.

   - **Rendering Provider Taxonomy for Dental Procedure Codes**. We evaluated how often valid, active dental procedure codes had missing data on the rendering provider type (i.e., did not have an accompanying provider taxonomy code).

   For consistency with the cut-points used by MACBIS for the Data Quality Atlas, we defined the following categories based on the percentage of missing data:

   -**Low concern:** Missing $\leq$ 10%

   -**Medium concern:** 10% < Missing $\leq$ 20%

   -**High concern:** 20% < Missing $\leq$ 50%

   -**Unusable** Missing > 50%

PRIOR TESTING

As described in measure evaluation criteria 3c1, this measure relies on standard data elements in claims data that are already collected and widely used for a range of reporting and billing purposes with very low rates of missing or invalid data (which we empirically assessed and reported in 3c1).

From 3c.1. in Prior Submission

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. A key advantage of using administrative claims data is that the time and cost of data collection for performance measurement purposes are relatively low because these data are already collected for other purposes.

Data 3c.1 Percentage of Missing and Invalid Values for Critical Data Elements

PROGRAM 1

Member ID: 0.00%

Date of Birth: 0.00%

Monthly enrollment indicator: 0.00%

Dental Procedure Codes - CDT: 0.00%

Date of Service: 0.01%

Rendering Provider ID: 0.28%

PROGRAM 2

Member ID: 0.27%

Date of Birth: 0.00%

Monthly enrollment indicator: 0.00%

Dental Procedure Codes - CDT: 0.28%

Date of Service: 0.00%

Rendering Provider ID: 0.18%

PROGRAM 3

Member ID: 0.00%

Date of Birth: 0.00%

Monthly enrollment indicator: 0.00%

Dental Procedure Codes - CDT: 0.01%

Date of Service: 0.00%

Rendering Provider ID: 0.61%

PROGRAM 4

Member ID: 0.43%

Date of Birth: 0.02%

Monthly enrollment indicator: 0.00%

Dental Procedure Codes - CDT: 0.00%

Date of Service: 0.00%

Rendering Provider ID: 0.67%

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

Data Quality Atlas Assessments

The tables below provide the results of the assessments of data completeness, missing data, and invalid data. For the data content areas addressed by MACBIS, as reported in the T-MSIS Quality Atlas, most states were assessed as having data of "low concern" for all content areas in all three years. For CY 2016, there were two states for which all fields were not low concern.

(1) New Mexico was assessed as having outpatient claims volume of "medium concern" in 2017. However, Service Users was "low concern" which is another indicator of claims data quality completeness and specifically assesses the percentage of beneficiaries with any service use. Examination of the measure scores found that performance was within expected ranges and similar between years (2018: 25.6%; 2017: 25.1%; 2016: 24.0%). Consequently, we are comfortable with the data quality for measure reporting purposes.

(2) In 2016, Mississippi was assessed as having outpatient claims volume of “high concern” and Medicaid-only enrollment of “medium concern.” Examination of the measure scores indicates that these data deficiencies significantly impact the measure scores (2018: 22.9%; 2017: 20.3%; 2016: 0.97%). Consequently, we excluded 2016 reporting for Mississippi from our testing.

Additional Evaluations

For the additional data fields assessed (beneficiary ID, age, dental procedure codes, and dental provider taxonomy codes), the rates of missing and invalid data were commonly <1%. In 2016, there were three states for which Rendering Provider Taxonomy for Dental Procedure Codes was >50% and deemed unusable: Idaho, Oklahoma, and South Carolina. Consequently, we excluded 2016 reporting for these three states from our measure testing.

Because there was generally low concern (LC)/low rates of missing/invalid data for the critical data elements used to calculate the measure scores, no rules were developed for handling missing data. In addition, because these are standard data fields contained within administrative claims data, when state Medicaid or CHIP programs are identified as having incomplete or poor quality data, they are encouraged to improve data collection and quality as part of their quality improvement efforts rather than using statistical methods to address missing data. The data field that had the most significant issues with missing data – Rendering Provider Taxonomy – only had issues in 2016 with dramatically improved data quality in 2017 and 2018; all three states had no missing data in 2017 and 2018, supporting the feasibility both of this data element and of improving data quality within the T-MSIS.

Percentage of Missing and Invalid Values for Critical Data Elements from State Medicaid Programs for CY2018

*	AK	AZ	DE	ID	MI	MS	NV
From T-MSIS Quality Atlas	*	*	*	*	*	*	*
Age	LC	LC	LC	LC	LC	LC	LC
Medicaid-Only Enrollment	LC	LC	LC	LC	LC	LC	LC
Claims Volume - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Service Users - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Procedure Codes - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Additional Checks	*	*	*	*	*	*	*

*	AK	AZ	DE	ID	MI	MS	NV
Beneficiary ID	0.01%	0.01%	0.01%	0.00%	0.02%	0.01%	0.01%
Dental Procedure Codes - CDT (% invalid)	0.01%	0.00%	0.01%	0.00%	0.00%	0.02%	0.00%
Rendering Provider Taxonomy for Dental Procedure Codes	0.01%	0.91%	0.10%	0.00%	0.00%	3.02%	4.28%
*	**NM**	**NC**	**OK**	**OR**	**SC**	**WA**	**WY**
From T-MSIS Quality Atlas	*	*	*	*	*	*	*
Age	LC	LC	LC	LC	LC	LC	LC
Medicaid-Only Enrollment	LC	LC	LC	LC	LC	LC	LC
Claims Volume - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Service Users - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Procedure Codes - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Additional Checks	*	*	*	*	*	*	*
Beneficiary ID	0.00%	0.01%	0.01%	0.01%	0.02%	0.01%	0.04%
Date of Birth	0.00%	0.00%	0.00%	0.00%	0.54%	0.00%	1.00%
Dental Procedure Codes - CDT (% invalid)	0.00%	0.01%	0.00%	0.01%	0.00%	0.00%	0.05%
Rendering Provider Taxonomy for Dental Procedure Codes	0.00%	0.18%	0.00%	0.02%	0.00%	0.00%	0.00%

Table showing percentage of missing and invalid values for critical data elements for state Medicaid programs for patients under age twenty-one in calendar year 2018

LC=Low concern, MC=Medium concern, HC=High concern

*Cell left intentionally blank

Percentage of Missing and Invalid Values for Critical Data Elements from State Medicaid Programs for CY2017

*	AK	AZ	DE	ID	MI	MS	NV
From T-MSIS Quality Atlas	*	*	*	*	*	*	*
Age	LC	LC	LC	LC	LC	LC	LC
Medicaid-Only Enrollment	LC	LC	LC	LC	LC	LC	LC
Claims Volume - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Service Users - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC

*	AK	AZ	DE	ID	MI	MS	NV
Procedure Codes - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Additional Checks	*	*	*	*	*	*	*
Beneficiary ID	0.01%	0.01%	0.01%	0.00%	0.02%	0.01%	0.01%
Date of Birth	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Dental Procedure Codes - CDT (% invalid)	0.02%	0.00%	0.01%	0.00%	0.00%	0.03%	0.11%
Rendering Provider ID for Dental Procedure Codes	0.01%	1.64%	0.13%	0.00%	0.00%	2.23%	4.85%
*	**NM**	**NC**	**OK**	**OR**	**SC**	**WA**	**WY**
From T-MSIS Quality Atlas	*	*	*	*	*	*	*
Age	LC	LC	LC	LC	LC	LC	LC
Medicaid-Only Enrollment	LC	LC	LC	LC	LC	LC	LC
Claims Volume - Other Services (Outpatient)	MC	LC	LC	LC	LC	LC	LC
Service Users - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Procedure Codes - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Additional Checks	*	*	*	*	*	*	*
Beneficiary ID	0.01%	0.01%	0.01%	0.01%	0.02%	0.01%	0.04%
Date of Birth	0.00%	0.00%	0.00%	0.00%	0.61%	0.00%	1.35%
Dental Procedure Codes - CDT (% invalid)	0.00%	0.01%	0.00%	0.01%	0.00%	0.00%	0.06%
Rendering Provider Taxonomy for Dental Procedure Codes	0.00%	0.15%	0.00%	0.02%	0.00%	0.00%	0.00%

Table showing percentage of missing and invalid values for critical data elements for state Medicaid programs for patients under age twenty-one in calendar year 2017

LC=Low concern, MC=Medium concern, HC=High concern

*Cell left intentionally blank

Percentage of Missing and Invalid Values for Critical Data Elements from State Medicaid Programs for CY2016

*	AK	AZ	DE	ID	MI	MS	NV
From T-MSIS Quality Atlas	*	*	*	*	*	*	*
Age	LC	LC	LC	LC	LC	LC	LC
Medicaid-Only Enrollment	LC	LC	LC	LC	LC	MC	LC

*	AK	AZ	DE	ID	MI	MS	NV
Claims Volume - Other Services (Outpatient)	LC	LC	LC	LC	LC	HC	LC
Service Users - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Procedure Codes - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Additional Checks	*	*	*	*	*	*	*
Beneficiary ID	0.00%	0.01%	0.02%	0.00%	0.05%	0.01%	0.01%
Date of Birth	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%
Dental Procedure Codes - CDT (% invalid)	0.02%	0.00%	0.11%	0.00%	0.00%	0.07%	0.25%
Rendering Provider ID for Dental Procedure Codes	0.05%	2.23%	5.11%	99.14%	0.02%	1.77%	8.12%
*	**NM**	**NC**	**OK**	**OR**	**SC**	**WA**	**WY**
From T-MSIS Quality Atlas	*	*	*	*	*	*	*
Age	LC	LC	LC	LC	LC	LC	LC
Medicaid-Only Enrollment	LC	LC	LC	LC	LC	LC	LC
Claims Volume - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Service Users - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Procedure Codes - Other Services (Outpatient)	LC	LC	LC	LC	LC	LC	LC
Additional Checks	*	*	*	*	*	*	*
Beneficiary ID	0.01%	0.01%	0.01%	0.02%	0.02%	0.01%	0.54%
Date of Birth	0.00%	0.00%	0.00%	0.01%	0.53%	0.00%	0.00%
Dental Procedure Codes - CDT (% invalid)	0.02%	0.01%	0.00%	0.01%	0.00%	0.01%	0.05%
Rendering Provider ID for Dental Procedure Codes	0.00%	0.11%	95.39%	0.01%	73.54%	0.09%	9.68%

Table showing percentage of missing and invalid values for critical data elements for state Medicaid programs for patients under age twenty-one in calendar year 2016

LC=Low concern, MC=Medium concern, HC=High concern

*Cell left intentionally blank

[Response Ends]

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

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

[Response Begins]

TESTING FOR SPRING 2022 MAINTENANCE EVALUATION

As noted above, this measure relies on standard data elements in administrative claims. These data are readily available and can be easily retrieved because they are routinely used for billing and reporting purposes. Despite the fact that these are routine data elements, we undertook due diligence to ensure the data completeness within the specific databases that we used for measure testing. Overall, there was low concern/low rates of missing/invalid data for the critical data elements used. In 2016, there were four states with sufficiently poor data quality on one critical data element that we excluded those states performance scores for 2016 from our testing. Because these are routine data elements that are already collected for other important purposes, particularly claims processing, no rules were developed for handling missing data other than not reporting on the performance measure when data quality is poor. As noted above, programs and plans are instead encouraged to improve their data quality rather than developing statistical techniques to overcome poor data quality.

[Response Ends]

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

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

[Response Begins]

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

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

Examples may include correlation, and/or rank order.

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

N/A or no exclusions

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

No risk adjustment or stratification

[Response Ends]

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

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

Other (specify)

[Other (specify) Please Explain]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

2b.27. Provide risk model discrimination statistics.

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

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

Criterion 3. Feasibility

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

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

[Response Begins]

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

[Response Ends]

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

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

[Response Begins]

ALL data elements are in defined fields in electronic claims

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

No updates.

PRIOR SUBMISSION

This measure is specified for reporting at the program and plan level and there are currently no efforts to develop an eMeasure (eCQM) of this measure at these levels.

Our understanding is that the Feasibility Score Card is only for eMeasures; consequently, we have not submitted this. Feasibility criteria were met during the initial endorsement review.

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

This measure relies on standard fields in administrative claims data that are routinely collected for other purposes. We have not received feedback indicating significant barriers to measure implementation. The removal of elevated risk from the denominator will reduce the data requirements because there is no need to review the three prior years of claims data. It will also simplify the programming from not having to implement the logic to identify elevated risk.

[Response Ends]

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

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

Attach the fee schedule here, if applicable.

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

No updates.

PRIOR SUBMISSION

This measure is intended to be transparent and available for widespread adoption. As such, it was purposefully designed to avoid using software or other proprietary materials that would require licensing fees. The measure specifications, including a companion User Guide, is accessible through a website and available free of charge for non-commercial purposes. The main requirements of users is to ensure the quality of their source data and expertise to program the measures within their information systems, following the clear and detailed specifications. Technical assistance is available to users.

[Response Ends]

Criterion 4: Use and Usability

4a. Use

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

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

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

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

Name of program and sponsor

URL

Purpose

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

Level of measurement and setting

[Response Begins]

Public Reporting

[Public Reporting Please Explain]

1. Topical Fluoride for Children, Dental Services (as currently specified without elevated risk required for denominator inclusion), has been adopted by the Centers for Medicare and Medicaid Services (CMS) for Child Core Health Care Quality Measurement for federal fiscal year 2022 reporting by state Medicaid and the Children's Health Insurance Program (CHIP) programs (https://www.medicaid.gov/medicaid/quality-of-care/downloads/2022-child-core-set.pdf). The CMS Core Set measures are foundational to assessing access to and quality of health care being provided to Medicaid and CHIP beneficiaries. The two related measures proposed to be "grouped" with this measure (Topical Fluoride for Children, Dental or Oral Health Services and Topical Fluoride for Children, Oral Health Services) also have been adopted for inclusion in the Child Core Set. Reporting on the Child Core Set of measures will change from voluntary to mandatory reporting in federal fiscal year 2024.
2. This measure has been included in the Center for Oral Health Systems Integration and Improvement (COHSII) Oral Health Quality Indicators for the Maternal and Child Health Population, which is funded by Health Services and Resources Administration (HRSA) Maternal and Child Health Bureau, for 2022 reporting. The measure will be reported with the two related measures proposed to be "grouped" with this measure (Topical Fluoride for Children, Dental or Oral Health Services and Topical Fluoride for Children, Oral Health Services).

Payment Program

[Payment Program Please Explain]

1. Program and Sponsor: Texas Health and Human Services Commission

https://www.hhs.texas.gov/sites/default/files/documents/about-hhs/process-improvement/quality-efficiency-improvement/p4q/p4q-dental-measures-2018-2022.pdf

Purpose: Payment Program

The currently endorsed measure has been adopted by the Texas Health and Human Services Commission as part of the Texas CHIP and Medicaid Dental Services Pay-for-Quality (P4Q) program. This measure was also present in earlier iterations of the Texas Medicaid and CHIP quality programs since initial endorsement. We are referencing current use for this update.

Geographic Area and Number/Percentage of Accountable Entities and Patients:

This applies to the state of Texas CHIP and Medicaid programs (statewide application). There are two dental plans (i.e., the accountable entities) that serve Texas CHIP and Medicaid. In December 2021, there were 3,884,768 children enrolled in Texas Medicaid and CHIP (https://hhs.texas.gov/about-hhs/records-statistics/data-statistics/healthcare-statistics). Level of Measurement and Setting: The measure is implemented at the plan and program level within the Texas Medicaid and CHIP programs.

2. Florida Medicaid

https://ahca.myflorida.com/Medicaid/statewide_mc/pdf/Contracts/2020-10-01_Dental/Dental_Attachment_II_Oct-2020.pdf

Purpose: Payment Program

Florida Medicaid has been reporting on the currently endorsed version of this measure since 2018.

As of January 31, 2022, there were a total of 4,595,576 enrollees, of which, 2,523,322 were 0-18 year olds.

https://ahca.myflorida.com/medicaid/finance/data_analytics/enrollment_report/index.shtml

Level of Measurement and Setting. The measure is implemented at the program level at the Florida Medicaid level.

Quality Improvement (Internal to the specific organization)

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

1. Covered California, the California Health Benefit Exchange

https://hbex.coveredca.com/insurance-companies/PDFs/QDP-Model-Contract-for-2017-2022-Amended-for-2022.pdf

https://hbex.coveredca.com/insurance-companies/PDFs/QDP-Attachment-14-for-2017-2022-Amended-for-2022.pdf

Purpose: Quality Improvement

The currently endorsed version of this measure is included in the Covered California Qualified Health Plan Issuer Contract for 2017-2022 For the Individual Market and the Covered California Qualified Dental Plan Issuer Contract for 2017-2022. The measure is to be reported annually.

Geographic Area and Number/Percentage of Accountable Entities and Patients: This applies statewide. In February 2022 there were 255,580 enrollees, of which 24,300 were 0-18 years old in CC health plans (which may offer dental benefits and would therefore report on the dental quality measures). There were 6,290 children enrolled specifically in Qualified Dental Plans. (http://hbex.coveredca.com/data-research/)

Level of Measurement and Setting. The measure is implemented at the plan level with the Covered California program.

2. State Medicaid Agencies

https://www.msdanationalprofile.com/profiles/2021/national/all/general

(Note: To access the data, a public user account must be created. We can help facilitate access to the data if needed.)

Purpose: Quality Improvement

The Medicaid|Medicare|CHIP Services Dental Association conducts an annual survey of state Medicaid programs and collects data specifically on which programs report Dental Quality Alliance measures.

In its 2021 profile (the most recent available), 8 states reported that they currently use this measure in the Medicaid and/or CHIP programs.

Geographic Area and Number/Percentage of Accountable Entities and Patients:

The 8 states are: California, Connecticut, District of Columbia, North Carolina, Oregon, Texas, Virginia and Wyoming. Data are not provided on the number of accountable entities included.

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Public reporting

Regulatory and Accreditation Program

Other (specify)

[Other (specify) Please Explain]

1. Topical Fluoride for Children, Dental Services (as currently specified without elevated risk required for denominator inclusion), has been adopted by the Centers for Medicare and Medicaid Services (CMS) for Child Core Health Care Quality Measurement for federal fiscal year 2022 reporting by state Medicaid and the Children's Health Insurance Program (CHIP) programs (https://www.medicaid.gov/medicaid/quality-of-care/downloads/2022-child-core-set.pdf). The CMS Core Set measures are foundational to assessing access to and quality of health care being provided to Medicaid and CHIP beneficiaries. The two related measures proposed to be "grouped" with this measure (Topical Fluoride for Children, Dental or Oral Health Services and Topical Fluoride for Children, Oral Health Services) also have been adopted for inclusion in the Child Core Set. Reporting on the Child Core Set of measures will change from voluntary to mandatory reporting in federal fiscal year 2024.

2. This measure is currently under consideration for use by the National Committee for Quality Assurance for Healthcare Effectiveness Data and Information Set (HEDIS) plan-level reporting.

3. This measure has been included in the Center for Oral Health Systems Integration and Improvement (COHSII) Oral Health Quality Indicators for the Maternal and Child Health Population, which is funded by Health Services and Resources Administration (HRSA) Maternal and Child Health Bureau, for 2022 reporting. The measure will be reported with the two related measures proposed to be "grouped" with this measure (Topical Fluoride for Children, Dental or Oral Health Services and Topical Fluoride for Children, Oral Health Services).

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

Not applicable.

[Response Ends]

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

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

The DQA provides technical assistance to these and other users of DQA measures through webinars, resource document development, and one-on-one staff support. In addition to these activities, DQA has now released a State Oral Healthcare Quality Dashboard on key oral healthcare measures, including Topical Fluoride for Children, Dental Services, that provides a snapshot of oral healthcare quality.

During the 2021 measure review cycle, performance data were shared with key stakeholders, including the CMS Oral Health Technical Advisory Group, NCQA, and the Center for Oral Health Systems Integration and Improvement.

PRIOR SUBMISSION

Per the annual survey conducted by the Medicaid | Medicare | CHIP Services Dental Association (MSDA), 9 Medicaid/CHIP agencies are implementing this measure. The measure is part of measure set included in the Request for Proposal (RFP) released by the Michigan Healthy Kids Dental Program. This measure is included in the Pay-For-Quality program and publicly reported in the Texas Medicaid and CHIP programs. Additionally, this measure is a requirement for the Qualified Dental Plans to report to the Covered California, the state-based marketplace in California.

The DQA provides technical assistance to these and other users of DQA measures through webinars, resource document development, and one-on-one staff support. The DQA has an Implementation Committee dedicated to developing implementation and improvement resources.

In order to ensure transparency, incorporate learnings from implementation, establish proper protocols for timely assessment of the evidence and measure properties, and to comply with the NQF's endorsement agreement, the DQA has established an annual measure review and maintenance process. This measure review process is overseen by the DQA's Measures Development and Maintenance Committee (MDMC) which is comprised of subject matter experts. This annual review process includes: (1) call for public comments, (2) evaluation of the comments, (3) user group feedback, and (4) code set reviews.

In 2016, the DQA expanded its scope of review of its measures by convening conference calls for two user groups – one comprised of representatives from 6 state Medicaid programs (Alabama, Florida, Kentucky, Oregon, Nevada, and Pennsylvania) and the other comprised of representatives from 8 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. Participants did not have any significant issues related to the clarity or feasibility of implementing the measure specifications.

This is the first 3-year maintenance endorsement review for this measure. As indicated above, the measure is being implemented in multiple programs. Because measure implementation requires a 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, most of the entities that have adopted the measures are just getting underway and there is limited data reporting. Implementation has mostly focused on addressing questions related to how to use the measures in the context of broader quality improvement and clarifying questions related to the specifications.

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

The DQA provides technical assistance to these and other users of DQA measures through webinars, resource document development, and one-on-one staff support.

In addition to these activities, DQA has now released a State Oral Healthcare Quality Dashboard on key oral healthcare measures, including Topical Fluoride for Children, Dental Services, that provides a snapshot of the oral healthcare quality.

During the 2021 measure review cycle, performance data were shared with key stakeholders, including the CMS Oral Health Technical Advisory Group, NCQA, and the Center for Oral Health Systems Integration and Improvement.

PRIOR SUBMISSION

In an effort to facilitate implementation of the DQA measures, the DQA provides technical assistance on an ongoing basis to users of DQA measures through webinars, resource document development and one-on-one staff support.

In 2016, the DQA expanded its scope of review of its measures by convening conference calls for two user groups – one comprised of representatives from 6 state Medicaid programs (Alabama, Florida, Kentucky, Oregon, Nevada, and Pennsylvania) and the other comprised of representatives from 8 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. Participants did not have any significant issues related to the clarity or feasibility of implementing the measure specifications.

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

To ensure transparency, establish proper protocols for timely assessment of the evidence and measure properties, and to comply with the NQF's endorsement agreement, the DQA has established an annual measure review and maintenance process. This measure review process is overseen by the DQA's Measures Development and Maintenance Committee (MDMC) which is comprised of subject matter experts. This annual review process includes: (1) month-long call for public comments, (2) evaluation of the comments, (3) user group feedback, and (4) code set reviews.

The DQA provides technical assistance on an ongoing basis to users of DQA measures through webinars, resource document development and one-on-one staff support.

The DQA's measure Topical Fluoride for Children, Dental Services was recently updated from its currently endorsed version based on feedback from the measure implementers, such as the CMS, an examination of the evidence supporting the measure, and additional testing data. This change was initiated by feedback from measure users seeking a more population-based approach. The 2021 Annual Measures Review Call for Public Comment included a specific request to the stakeholder community on the denominator definition with a document describing the measure as originally specified, limitations, and options for modification (including no change as an option, removing elevated risk for only the youngest children, and removing elevated risk altogether). The document was titled "DQA Requests Stakeholder Feedback on the Denominator Definition of the DQA Topical Fluoride Measure." Comments were provided by individuals representing dental plans, Medicaid programs, clinicians, and federal agencies. The DQA Measures Development and Maintenance Committee (MDMC) evaluated the comments as well as data for 11 state Medicaid programs to evaluate the impact of the change. DQA staff also gave webinar presentations to key stakeholder groups about the proposed measure changes and solicited feedback. Based on its evaluation of user community feedback, evidence, and testing data, the MDMC recommended removal of elevated risk from the criteria used to establish the denominator, with guidance in the User Guide on how to stratify by elevated risk, for users who wish to continue to examine fluoride receipt by caries risk status. This recommendation was approved by the DQA membership at its June 2021 meeting. The review and update process is detailed in DQA's

2021 Annual Measure Review Final Report .

PRIOR SUBMISSION

In order to ensure transparency, establish proper protocols for timely assessment of the evidence and measure properties, and to comply with the NQF's endorsement agreement, the DQA has established an annual measure review and maintenance process. This measure review process is overseen by the DQA's Measures Development and Maintenance Committee (MDMC) which is comprised of subject matter experts. This annual review process includes: (1) call for public comments, (2) evaluation of the comments, (3) user group feedback, and (4) code set reviews.

DQA provides technical assistance on an ongoing basis to users of DQA measures through webinars, resource document development and one-on-one staff support.

In 2016, the DQA expanded its scope of review of its measures by convening conference calls for two user groups – one comprised of representatives from 6 state Medicaid programs (Alabama, Florida, Kentucky, Oregon, Nevada, and Pennsylvania) and the other comprised of representatives from 8 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. Participants did not have any significant issues related to the clarity or feasibility of implementing the measure specifications.

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

Please see response in 4a.07)

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

Please see response in 4a.07)

PRIOR SUBMISSION

There have been no significant issues related to the clarity or feasibility of implementing the measure specifications.

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

Please see response in 4a.07)

PRIOR SUBMISSION

There have been no significant issues related to the clarity or feasibility of implementing the measure specifications.

[Response Ends]

4b. Usability

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

Although we used three years of data for reliability and testing purposes, we do not have information about what improvement efforts were or were not in place for those states. Although the currently endorsed measure has been in use by several entities, the data frequently are not publicly reported. The main program

that we are aware of that has incorporated the measure into quality improvement and payment programs - and publicly reports the measures – is the Texas Medicaid/CHIP programs.

We provide updated data on the performance in these programs at the program and plan levels below. Performance improved by 10 percentage points at both the program and plan levels between 2014, when the measure was first endorsed, to 2018 in both the Medicaid and CHIP programs.

Texas Medicaid

Year, Program Denominator, Program Overall Score, DentaQuest(Plan) Score, MCNA(Plan) Score

2018, 1730126, 50.04, 51.10, 48.32

2017, 1760138, 49.77, 50.94, 47.92

2016, 1610549, 49.11, 49.79, 48.13

2015, 1334887, 41.75, 44.70, 38.15

2014, 1090952, 39.97, 41.57, 37.62

Texas CHIP

Year, Program Denominator, Program Overall Score, DentaQuest(Plan) Score, MCNA(Plan) Score

2018, 133116, 44.85, 48.32, 44.87

2017, 139986, 43.14, 46.89, 43.99

2016, 108297, 42.97, 46.28, 43.85

2015, 79693, 37.50, 41.44, 37.71

2014, 108704, 33.01, 35.45, 32.99

Topical Fluoride for Children, Dental Services along with two related measures proposed to be "grouped" with this measure (Topical Fluoride for Children, Dental or Oral Health Services and Topical Fluoride for Children, Oral Health Services) has been

adopted by the Centers for Medicare and Medicaid Services (CMS) for Child Core Health Care Quality Measurement for federal fiscal year 2022 reporting by state Medicaid and the Children's Health Insurance Program (CHIP). The CMS Core Set measures are foundational to assessing access to and quality of health care being provided to Medicaid and CHIP beneficiaries. Reporting on the child core set of measures will change from voluntary to mandatory reporting in federal fiscal year 2024. Mandatory reporting as part of the CMS Child Core Set will greatly facilitate the ability to evaluate changes in performance over time and how improvement varies across state Medicaid programs.

PRIOR SUBMISSION

This is the first 3-year maintenance endorsement review for this measure. As indicated above, the measure is being implemented in multiple programs. Because measure implementation requires a 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, most of the entities that have adopted the measures either have only limited baseline scores or will start reporting measures within the next year.

We are only aware of repeat measurements within the Texas Medicaid/CHIP programs (https://thlcportal.com/qoc/dental), which started implementing this measure after it was approved by the Dental Quality Alliance and before NQF endorsement, as follows:

Texas Medicaid

Year, Program Denominator, Program Overall Score, DentaQuest(Plan) Score, MCNA(Plan) Score

2014, 1090952, 39.97, 41.57, 37.62

2015, 1334887, 41.75, 44.70, 38.15

Texas CHIP

Year, Program Overall, DentaQuest(Plan), MCNA(Plan)

2014, 108704, 33.01, 35.45, 32.99

2015, 79693, 37.50, 41.44, 37.71

These data suggest a trend in improvement over time. However, as noted above, these are initial performance data for one program. Most measure users are just now getting their quality measurement programs underway.

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

The measure seeks to improve receipt of an evidence-based process of care among children, that currently has significant performance gaps (i.e., most children do not receive at least two professionally applied topical fluoride applications during the year). As noted in the evidence section, the potential for harm is minimal. No negative unintended impacts have been identified. The main concern expressed by measure users is the desire for apopulation-based approach, given the inherent limitations in administrative claims data to identify children at elevated risk for dental caries – basically to ensure that this evidence-based prevention is reaching a broader population of children. Based on this feedback as described above and elsewhere in this application, the measure was modified to remove this criterion from the denominator, while providing guidance in the User Guide on how to stratify the measure by elevated caries risk status.

PRIOR SUBMISSION

No unintended or negative consequences have been identified.

[Response Ends]

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

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

The benefits are as expected.

[Response Ends]

Criterion 5: Related and Competing Measures

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

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

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

(Can search and select measures.)

[Response Begins]

2689: Ambulatory Care Sensitive Emergency Department Visits for Dental Caries in Children

2511: Utilization of Services, Dental Services

2517: Oral Evaluation, Dental Services

2695: Follow-Up after Emergency Department Visits for Dental Caries in Children

[Response Ends]

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

(Can search and select measures.)

[Response Begins]

[Response Ends]

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

[Response Begins]

Not applicable.

[Response Ends]

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

[Response Begins]

Yes. The NQF-endorsed DQA measures (#2511, #2517, #2689 and #2695) all address broadly the same population - children enrolled in Medicaid and CHIP. But the denominators are specified differently. The measures are complementary to one another but distinct.

[Response Ends]

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

[Response Begins]

Not applicable.

[Response Ends]

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

Provide analyses when possible.

[Response Begins]

Not applicable.

[Response Ends]

Appendix

Supplemental materials may be provided in an appendix.:

No appendix

Contact Information

Measure Steward (Intellectual Property Owner): American Dental Association

Measure Steward Point of Contact: Ojha, Diptee, ojhad@ada.org

Alliance, Dental, dqa@ada.org

Colangelo, Erica, colangeloe@ada.org

Measure Developer if different from Measure Steward: American Dental Association

Measure Developer Point(s) of Contact: Ojha, Diptee, ojhad@ada.org

Herndon, Jill, jill.herndon@keyanalyticsconsulting.com

Alliance, Dental, dqa@ada.org

Colangelo, Erica, colangeloe@ada.org

Additional Information

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

[Response Begins]

No appendix

[Response Ends]

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

Describe the members' role in measure development.

[Response Begins]

SPRING 2022 MAINTENANCE SUBMISSION

This project is headed by the DQA through its Measure Development and Maintenance Committee (formerly Research and Development Committee). The following individuals are currently responsible for executing and overseeing all scientific aspects of this project, including maintenance reviews.

2022 Measures Development and Maintenance Committee (MDMC)

1. Craig W. Amundson, DDS, Senior Dental Advisor, HealthPartners. Dr. Amundson serves as chair for the Committee.
2. Frederick Eichmiller, DDS, Vice President & Science Officer Emeritus, Delta Dental of Wisconsin
3. Chris Farrell, RDH, BSDH, MPA, Oral Health Program Director, Michigan Department of Health and Human Services
4. An Nyugen, Chief Dental Officer, Clinica Family Health
5. Chris Okunseri, B.D.S., M.Sc., Director, Predoctoral Program, Dental Public Health, Marquette University
6. Bob Russell, DDS, MPH, MPA, CPM, FACD, FICD, Previous State Public Health Dental Director, Chief, Bureau of Oral and Health Delivery Systems, Iowa
7. Tim Wright, DDS, MS, Distinguished Professor, University of North Carolina School of Dentistry, Editor-in-Chief, Journal of American Dental Association
8. Rob Margolin, DDS, Executive Committee Liaison to the MDMC
9. Paul Casamassimo, DDS, MS, Dr. Casamassimo is the current chair of the DQA
10. Ralph Cooley, DDS, Dr. Cooley is the current Chair-Elect of the DQA.

PRIOR SUBMISSION

This project is headed by the DQA through its Measure Development and Maintenance Committee (formerly Research and Development Committee). The following individuals were responsible for executing and overseeing all scientific aspects of this project.

• Craig W. Amundson, DDS, General Dentist, HealthPartners, National Association of Dental Plans. Dr. Amundson serves as chair for the Committee.

• Mark Casey, DDS, MPH, Dental Director, North Carolina Department of Health and Human Services Division of Medical Assistance

• Natalia Chalmers, DDS, PhD, Diplomate, American Board of Pediatric Dentistry, Director, Analytics and Publication, DentaQuest Institute

• Frederick Eichmiller, DDS, Vice President & Science Officer, Delta Dental of Wisconsin

• Chris Farrell, RDH, BSDH, MPA, Oral Health Program Director, Michigan Department of Health and Human Services

This group oversees the maintenance. All work of this Committee was distributed for review and formal vote and approval by the entire Dental Quality Alliance. (http://ada.org/dqa) The DQA is made up of representatives from 38 stakeholder organizations.

[Response Ends]

3. Indicate the year the measure was first released.

[Response Begins]

2013

[Response Ends]

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

[Response Begins]

January 2022

[Response Ends]

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

[Response Begins]

Annual

[Response Ends]

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

[Response Begins]

2023

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

Dental Quality Alliance measures and related data specifications, developed by the Dental Quality Alliance (DQA), are intended to facilitate quality improvement activities. These Measures are intended to assist stakeholders in enhancing quality of care. These performance Measures are not clinical guidelines and do not establish a standard of care. The DQA has not tested its Measures for all potential applications.

Measures are subject to review and may be revised or rescinded at any time by the DQA. The Measures may not be altered without the prior written approval of the DQA. The DQA shall be acknowledged as the measure steward in any and all references to the measure.

Measures developed by the DQA, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measures require a license agreement between the user and DQA. Neither the DQA nor its members shall be responsible for any use of these Measures.

THE MEASURES ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND

Limited proprietary coding is contained in the Measure specifications for convenience.

For Proprietary Codes:

The code on Dental Procedures and Nomenclature is published in Current Dental Terminology (CDT), Copyright © 2021 American Dental

Association (ADA). All rights reserved.

This material contains National Uniform Claim Committee (NUCC) Health Care Provider Taxonomy codes

(http://www.nucc.org/index.php?option=com_content&view=article&id=14&Itemid=125). Copyright © 2021 American Medical

Association. All rights reserved.

Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. The DQA, American Dental Association (ADA), and its members disclaim all liability for use or accuracy of any terminologies or other coding contained in the specifications.

THE SPECIFICATIONS ARE PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND.

[Response Ends]

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

[Response Begins]

In 2008, the Centers for Medicare and Medicaid Services (CMS) asked the ADA to lead the development of a broad coalition of organizations that would lead dentistry to improve the oral health of Americans through quality measurement and quality improvement. The ADA subsequently established the Dental Quality Alliance (DQA). The DQA is a multi-stakeholder alliance comprised of approximately 38 stakeholders (with organizations as members) from across the oral health community, including federal agencies, third-party payers, professional associations, and an individual member from the general public. The DQA’s mission is to advance the field of performance measurement to improve oral health, patient care, and safety through a consensus building process.

[Response Ends]