

Pediatric Oral Health-Related Well-Being and Distress

Section 1. Basic Measure Information

1.A. Measure Name

Pediatric Oral Health-Related Well-Being (POHR-WB) and Distress (POHR-D) Measures

1.B. Measure Number

0211

1.C. Measure Description

Please provide a non-technical description of the measure that conveys what it measures to a broad audience.

The Pediatric Oral Health-Related Well-Being and Distress measures (prospectively referred to as the POHR-WB and POHR-D measures) are evidence-based, patient-reported outcome measures of pediatric oral health. The POHR-WB and POHR-D item banks comprise 14 and 11 questions, respectively. They are being presented together as they reflect the complementary and inter-related concepts of one's satisfaction with appearance and the social-emotional impact of these perceptions. Child report versions are appropriate for children aged 8-17 years, and parent-proxy versions are appropriate for parents of children aged 5-17 years. The measures use a 7-day report period.

The general domain of oral health-related well-being was selected based on an extensive literature review of oral health measures and prioritization by experts and children/parents. The item pools were developed using a rigorous mixed-methods approach in accordance with Patient Reported Outcome Measurement Information System® (PROMIS) standards. PROMIS is a federally funded initiative to develop patient-reported outcomes measures; it includes over 60 measures of patient-reported health. Methods included a systematic literature review of existing oral health PROMs; semi-structured interviews with pediatric dentists, parents, and youth; cognitive interviews with parents and youth; readability assessments; and classical and modern psychometric testing.

Development of the POHR-WB was conducted with 998 children and youth (ages 8-17 years) and 626 parents (of children ages 5-17 years). Development of the POHR-D was conducted with 997 children and youth (ages 8-17 years) and 626 parents (of children ages 5-17 years). Participants resided in all 50 States. The measures have excellent reliability, which indicates that they yield very precise measurements of a child's oral health well-being and distress. Likewise, the validity is excellent, which suggests that indeed it evaluates a child's oral health-related well-being and distress. The POHR-WB and POHR-D have also been validated with other measures of children's health, including the Healthy Pathways Body Image Child and Parent Report scale.

The POHR-WB and POHR-D measures focus on the individual's satisfaction with oral health-related appearance and the social-emotional impact of these perceptions. The POHR-WB and POHR-D are outcome measures that will be useful in a wide variety of applications and across pediatric populations. Patient-reported outcome measures, such as the POHR-WB and POHR-D, provide information about health and the impact of healthcare that is unique and complementary to information derived from measures of clinical outcomes and healthcare processes. Such measures have the potential to improve pediatric healthcare quality assessment at the point of care and through system-level performance evaluations.

1.D. Measure Owner

Children's Hospital of Philadelphia (CHOP).

1.E. National Quality Forum (NQF) ID (if applicable)

Not applicable.

1.F. Measure Hierarchy

Please note here if the measure is part of a measure hierarchy or is part of a measure group or composite measure. The following definitions are used by AHRQ:

- 1. Please identify the name of the collection of measures to which the measure belongs (if applicable). A collection is the highest possible level of the measure hierarchy. A collection may contain one or more sets, subsets, composites, and/or individual measures.**

The POHR-WB and POHR-D will be incorporated into the PROMIS measure collection. The domain framework, which is hierarchically structured is shown in Figure 1 (see Supporting Documents). Note that the overall concept that PROMIS measures is self-reported health, and the highest level dimensions are physical, mental, and social health. This structure is reflective of the WHO tripartite conceptualization of health. The POHR-WB and POHR-D measures integrate one's view of physical, mental, and social health.

- 2. Please identify the name of the measure set to which the measure belongs (if applicable). A set is the second level of the hierarchy. A set may include one or more subsets, composites, and/or individual measures.**

Not applicable.

- 3. Please identify the name of the subset to which the measure belongs (if applicable). A subset is the third level of the hierarchy. A subset may include one or more composites, and/or individual measures.**

Not applicable.

- 4. Please identify the name of the composite measure to which the measure belongs (if applicable). A composite is a measure with a score that is an aggregate of scores**

from other measures. A composite may include one or more other composites and/or individual measures. Composites may comprise component measures that can or cannot be used on their own.

Not applicable.

1.G. Numerator Statement

We used an IRT-based approach for scoring. The raw sum med score is ordinal, which means that a difference of one unit does not have the same meaning across the full scale. With interval scaling, a difference of one unit has the same meaning across the latent trait.

The score is computed by finding the raw summed score and converting that to an IRT-based (two-parameter graded response model) score, called Theta, which has a mean of 0. The theta score is then converted to a T-score, which has mean of 50.

1.H. Numerator Exclusions

The child-report measure is applicable to children 8 to 17 years of age. We recommend the exclusion of children with cognitive limitations or developmental delays that affect a child's ability to self-report. The parent proxy-report measure is applicable to children 5-17 years old. Both these age ranges are consistent with PROMIS standards.

1.I. Denominator Statement

The denominator is the total number of individuals within the population.

1.J. Denominator Exclusions

The denominator has the same exclusions as the numerator.

The child-report measure is applicable to children 8 to 17 years of age. We recommend the exclusion of children with cognitive limitations or developmental delays that affect a child's ability to self-report. The parent proxy-report measure is applicable to children 5-17 years old. Both these age ranges are consistent with PROMIS standards.

1.K. Data Sources

Check all the data sources for which the measure is specified and tested.

Survey: health professional report; parent/caregiver report; child report. Electronic medical record.

If other, please list all other data sources in the field below.

Patient registry; personal health record.

Note: The POHR-WB and POHR-D are designed for questioning of children or their parents/guardians. The survey can be completed using paper-and-pencil, web-based forms in the electronic health record (e.g., EHRs via patient portal), independent web-based data collection systems, patient-reported registries, and personal health records.

Section 2: Detailed Measure Specifications

Provide sufficient detail to describe how a measure would be calculated from the recommended data sources, uploading a separate document (+ Upload attachment) or a link to a URL. Examples of detailed measure specifications can be found in the CHIPRA Initial Core Set Technical Specifications Manual 2011 published by the Centers for Medicare & Medicaid Services. Although submission of formal programming code or algorithms that demonstrate how a measure would be calculated from a query of an appropriate electronic data source are not requested at this time, the availability of these resources may be a factor in determining whether a measure can be recommended for use.

Please see the Supporting Documents for detailed measure specifications.

Section 3. Importance of the Measure

In the following sections, provide brief descriptions of how the measure meets one or more of the following criteria for measure importance (general importance, importance to Medicaid and/or CHIP, complements or enhances an existing measure). Include references related to specific points made in your narrative (not a free-form listing of citations).

3.A. Evidence for General Importance of the Measure

Provide evidence for all applicable aspects of general importance:

- **Addresses a known or suspected quality gap and/or disparity in quality (e.g., addresses a socioeconomic disparity, a racial/ethnic disparity, a disparity for Children with Special Health Care Needs (CSHCN), a disparity for limited English proficient (LEP) populations).**
- **Potential for quality improvement (i.e., there are effective approaches to reducing the quality gap or disparity in quality).**
- **Prevalence of condition among children under age 21 and/or among pregnant women.**
- **Severity of condition and burden of condition on children, family, and society (unrelated to cost).**
- **Fiscal burden of measure focus (e.g., clinical condition) on patients, families, public and private payers, or society more generally, currently and over the life span of the child.**

- **Association of measure topic with children’s future health – for example, a measure addressing childhood obesity may have implications for the subsequent development of cardiovascular diseases.**
- **The extent to which the measure is applicable to changes across developmental stages (e.g., infancy, early childhood, middle childhood, adolescence, young adulthood).**

Known Quality Gap or Disparity in Quality

The Pediatric Oral Health-Related Well-Being and Distress Item Banks are outcome measures designed to evaluate the complementary and inter-related concepts of one’s satisfaction with oral health-related appearance and the social-emotional impact of these perceptions.

Dental appearance is an important contributor to an individual’s facial and physical appearance, and it plays a role in self-perception as well as social interactions (Eli, Bar-Tal, Kostovetzki, 2001). Characteristics such as tooth color, shape, and position; disease status; and quality of restoration are all factors affecting overall dental appearance. Therefore, while the POHR-WB and POHR-D measures reflect aesthetic concepts, self-assessments of appearance may be indicators of oral health status (e.g., untreated dental caries, discolored or missing teeth, malocclusion etc.) and also serve as a driver for health-seeking behavior (Tin-Oo, Saddki, Hassan, 2011). In particular, the POHR-D measure includes items that are intended to capture the impact of satisfaction with appearance on one’s mental and social health. As reflected in the PROMIS conceptual framework, evaluation of self is critical to mental and social well-being and has important implications for positive health and behaviors, as well as other long-term consequences (Mann, Hosman, Schaalma, 2004).

The importance of measuring oral health was highlighted in the landmark, inaugural Surgeon General’s Report on Oral Health, which underscored the integral nature of oral health to general health and well-being. The report noted that poor oral health has significant implications for children’s overall health, growth and development, and learning (Jackson, Vann, Kotch, et al., 2011; U.S. Department of Health and Human Services [HHS], 2000). Dental caries is the most common chronic disease of children in the United States, far exceeding even asthma—one of the most common pediatric conditions (American Dental Association [ADA], 2012). According to the National Health and Nutrition Examination Survey (NHANES), the prevalence of caries in primary teeth (untreated or filled, excluding missing teeth) in 2- to 5-year-olds increased from approximately 24 to 28 percent between the periods 1988 to 1994 and 1999 to 2004, with the majority of these children having not received appropriate treatment. Among 6-15 year-olds, one in two children are affected by dental caries (Dye, Tan, Smith, et al, 2007).

The profound disparities that exist in dental caries disproportionately affect minority and economically disadvantaged children (HHS, 2010). NHANES found that over 50 percent of children ages 2 to 11 years in families below the Federal poverty threshold experienced primary tooth dental caries, compared with one-third of children in families with incomes above 200 percent of the poverty threshold (Dye, et al., 2007). Such disparities are likely to be paralleled by similar differences in levels of POHR-WB and Distress.

Potential for Quality Improvement

An essential component of patient-centric quality improvement (QI) initiatives is to measure the impact of the intervention on patient outcomes. Studies have shown that young children are capable of answering their own questions related to oral health, and that parent-proxy answered questions closely reflect their child's oral health (Filstrup, Briskie, da Fonseca, et al., 2003).

The goal of most QI activities is to improve the reliability of processes of care that have likely or established associations with patient outcomes. Typically QI initiatives do not measure the impact of the intervention on health. QI activity that reduces symptom burden or improves children's functional status will have a positive effect on their global health (see Figure 3 in the Supporting Documents for our adaptation of the Wilson and Cleary (1995) model of health related quality of life).

In the case of oral health, there are bio-physiological processes related to the development of dental caries or other oral health conditions that bear an impact on dental appearance. These conditions may then lead to physical symptoms of disease or decay, which in turn can lead to a variety of oral health impacts, whether related to physical, mental, or social functioning. The POHR-WB and POHR-D measures may support the evaluation of QI interventions that are needed to bridge the gap in quality among different populations.

Prevalence and Severity of Condition

The POHR-WB and POHR-D Item Banks are domain-specific rather than disease-specific. Oral health concepts are applicable to all children and pregnant women. However, it is well-documented that oral health conditions profoundly impact the well-being of children, families, and society, while disproportionately affecting poorer populations (HHS, 2000). The POHR-WB and POHR-D are designed to measure the oral health of children with or without a specific oral health issue.

Fiscal Burden on Patients, Families, and Public and Private Payers

In 2010, dental services accounted for approximately 5 percent of national healthcare spending (Centers for Medicare & Medicaid Services [CMS], 2014). Dental costs continue to rise: Medicaid spending on dental services in 2012 was \$7.3 billion – 60 percent higher than 5 years prior (CMS, 2014). When left untreated, oral health disorders in children commonly extend to adulthood, affecting an individual's well-being and incurring additional financial costs that could have been avoided with proper childhood dental care (Mouradian, Wehr, Crall, 2000).

Not only is there a shortage of dental providers to the Medicaid population, children are 3 times as likely as adults to not have dental insurance or medical insurance. Both these factors disproportionately affect low income and minority families (HHS, 2000). Significant fiscal barriers exist that inhibit children from receiving the appropriate preventive dental care and treatment, which has implications for oral health status and appearance.

The consequences of poor oral healthcare for children are numerous and have long-term impacts. These include: the development of progressive dental disease and a greater risk of costly dental treatment and infectious and systemic disease as adults (HHS, 2000; Mouradian, et al., 2000)

Importantly, in psychometric testing of the POHR-WB, we found no age-related differential item functioning. This means that for a given score, younger and older children are equally likely to endorse a response option for any given item. That is, younger and older children are using the measure similarly to describe their health.

3.B. Evidence for Importance of the Measure to Medicaid and/or CHIP

Comment on any specific features of this measure important to Medicaid and/or CHIP that are in addition to the evidence of importance described above, including the following:

- **The extent to which the measure is understood to be sensitive to changes in Medicaid or CHIP (e.g., policy changes, quality improvement strategies).**
- **Relevance to the Early and Periodic Screening, Diagnostic and Treatment benefit in Medicaid (EPSDT).**
- **Any other specific relevance to Medicaid/CHIP (please specify).**

CMS is the single largest payer of health services for children in the United States, and yet less than half of the nation's 32 million children enrolled in Medicaid receive any dental services in a given year, and even fewer receive a preventive dental service (CMS, 2014). These low rates persist despite the inclusion of oral healthcare in Medicaid's Early and Periodic Screening, Diagnostic and Treatment (EPSDT) benefit for children. While low oral healthcare provider participation in Medicaid is an ongoing barrier to access – due in part to issues such as low reimbursement rates and administrative requirements – an additional 3.2 million children were expected to gain dental benefits via Medicaid expansion by 2018, taxing Medicaid programs with an already inadequate supply of providers (Health Management Associates [HMA], 2009). At the same time, dental costs continue to rise: Medicaid spending on dental services in 2012 was \$7.3 billion – 60 percent higher than 5 years prior (CMS, 2014). Given this enormous investment and anticipated beneficiary expansion, State Medicaid agencies have a critical responsibility to improve access to lower-cost oral healthcare prevention and treatment services. While from a cost perspective Medicaid coverage is limited to medically required services, the relevance of the POHR-WB and POHR-D is still significant, particularly among children for whom disparities in oral health status exist.

3.C. Relationship to Other Measures (if any)

Describe, if known, how this measure complements or improves on an existing measure in this topic area for the child or adult population, or if it is intended to fill a specific gap in an existing measure category or topic. For example, the proposed measure may enhance an existing measure in the initial core set, it may lower the age range for an existing adult-

focused measure, or it may fill a gap in measurement (e.g., for asthma care quality, inpatient care measures).

Not applicable.

Section 4. Measure Categories

CHIPRA legislation requires that measures in the initial and improved core set, taken together, cover all settings, services, and topics of health care relevant to children. Moreover, the legislation requires the core set to address the needs of children across all ages, including services to promote healthy birth. Regardless of the eventual use of the measure, we are interested in knowing all settings, services, measure topics, and populations that this measure addresses. These categories are not exclusive of one another, so please indicate "Yes" to all that apply.

Does the measure address this category?

- a. **Care Setting – ambulatory:** Yes.
- b. **Care Setting – inpatient:** Yes.
- c. **Care Setting – other – please specify:** Yes; home, school, long-term care. The POH-WB and POH-D are applicable across all settings where oral health evaluations are useful.
- d. **Service – preventive health, including services to promote healthy birth:** No.
- e. **Service – care for acute conditions:** Yes.
- f. **Service – care for children with special health care needs/chronic conditions:** Yes.
- g. **Service – other (please specify):** Yes; preventive health, health promotion, behavioral health, oral health. Although the POHR-WB and POHR-D are primarily focused on oral health, this domain extends across service types and is restricted to children 5 years of age and older.
- h. **Measure Topic – duration of enrollment:** No.
- i. **Measure Topic – clinical quality:** Yes.
- j. **Measure Topic – patient safety:** Yes.
- k. **Measure Topic – family experience with care:** No.
- l. **Measure Topic – care in the most integrated setting:** No.
- m. **Measure Topic other (please specify):** Yes; outcomes, patient-centered care.
- n. **Population – pregnant women:** No.
- o. **Population – neonates (28 days after birth) (specify age range):** No.
- p. **Population – infants (29 days to 1 year) (specify age range):** No.
- q. **Population – pre-school age children (1 year through 5 years) (specify age range):** Yes for age 5; the parent-reported version extends to children age 5. Both parent report and child report cover school-age and adolescents.
- r. **Population – school-aged children (6 years through 10 years) (specify age range):** Yes.
- s. **Population – adolescents (11 years through 20 years) (specify age range):** Yes.
- t. **Population – other (specify age range):** No.
- u. **Other category (please specify):** Not applicable.

Section 5. Evidence or Other Justification for the Focus of the Measure

The evidence base for the focus of the measures will be made explicit and transparent as part of the public release of CHIPRA deliberations; thus, it is critical for submitters to specify the scientific evidence or other basis for the focus of the measure in the following sections.

5.A. Research Evidence

Research evidence should include a brief description of the evidence base for valid relationship(s) among the structure, process, and/or outcome of health care that is the focus of the measure. For example, evidence exists for the relationship between immunizing a child or adolescent (process of care) and improved outcomes for the child and the public. If sufficient evidence existed for the use of immunization registries in practice or at the State level and the provision of immunizations to children and adolescents, such evidence would support the focus of a measure on immunization registries (a structural measure).

Describe the nature of the evidence, including study design, and provide relevant citations for statements made. Evidence may include rigorous systematic reviews of research literature and high-quality research studies.

We conducted a systematic review of the published literature based on procedures outlined by Klem and colleagues to build a pediatric oral health conceptual framework (Klem, Saghafi, Abromitis, et al., 2009). Biomedical librarians at the University of Pennsylvania worked with our team to develop sets of terms related to dental/oral health, quality of life, measurement, and self/proxy report (Table 6; see Supporting Documents). The terms were applied to the Medline databases to identify articles that described or applied PRO instruments to assess children's oral health.

Studies were eligible for inclusion if they were (1) written in the English language, (2) published in peer-reviewed journals or government reports in 1985 or later, (3) included specific measure(s) of child oral health outcomes (as indicated by the presence of Dental/Oral Health, Quality of Life, Measurement, and Self/Proxy Report Terms), and (4) enrolled children and youth to 21 years of age. Two study investigators reviewed all article titles and abstracts to identify those that met the above criteria for potential eligibility. The full text of each potentially eligible paper was reviewed independently by two investigators. Differences between investigators were settled by consensus.

Characteristics of the PRO measures used in the eligible papers were abstracted using a standardized electronic abstraction form. The information abstracted included measure name, adult/child respondent, reporter, language, specific population, and reason. The top 25 most frequently found measures were identified and compiled (Table 8; see Supporting Documents). Individual items from the top 25 measures were binned to construct the domain framework.

Constructing the domain framework involved classifying the concept text of each measure item into domains (12 identified) and sub-domains (55 identified). There were 11 key concepts derived from this analysis, with social-emotional well-being as the most common (Table 9; see Supporting Documents). After this analysis was complete, the domain framework was reviewed and revised by six dental faculty members at the University of Pennsylvania. Additionally, framework development was supplemented by interviews with 13 content experts, 28 children (ages 8-17 years), and 25 parents of children ages 5-17 years. A qualitative analysis of these interviews was performed using a saturation mapping technique.

Once the domain framework was finalized, the items for the measure were created. Sets of items called “item banks” were developed and then calibrated using item response theory (IRT) methods. IRT modeling informs the selection of items that reliably measure the full range of a health experience (e.g., from very poor to excellent family relationships) with minimal gaps in coverage and item redundancies. The models also provide a framework for investigating measurement bias as indicated by differential item functioning (DIF) and for developing computerized adaptive test (CAT) algorithms (Teresi, Fleishman, 2007). In CAT, items are presented to respondents based on their prior responses, tailoring the questions so that only a few items (typically five) are needed to reliably assess a full latent trait (Bjorner, Chang, Thissen, et al., 2007).

5.B. Clinical or Other Rationale Supporting the Focus of the Measure (optional)

Provide documentation of the clinical or other rationale for the focus of this measure, including citations as appropriate and available.

1. Response to stakeholders: In stakeholder discussions that helped to identify the measure priorities for PQMP, several State leaders advocated for greater attention to child outcomes, such as well-being and distress.
2. Patient-centeredness: With the emergence of PCORI and renewed emphasis in AHRQ and across the public sector on patient-centered care, new measures of outcome from the person’s perspective are needed.

POHR-WB and POHR-D are responsive to stakeholders, deal with a high prevalence issue from a patient-centered perspective, and contribute to the existing gap in outcome measures for children.

Section 6. Scientific Soundness of the Measure

Explain the methods used to determine the scientific soundness of the measure itself. Include results of all tests of validity and reliability, including description(s) of the study sample(s) and methods used to arrive at the results. Note how characteristics of other data systems, data sources, or eligible populations may affect reliability and validity.

6.A. Reliability

Reliability of the measure is the extent to which the measure results are reproducible when conditions remain the same. The method for establishing the reliability of a measure will depend on the type of measure, data source, and other factors.

Explain your rationale for selecting the methods you have chosen, show how you used the methods chosen, and provide information on the results (e.g., the Kappa statistic). Provide appropriate citations to justify methods.

Internal Consistency Reliability

We calculated Cronbach's alpha, a measure of internal consistency that assesses how closely related a set of items are as a group. A "high" value of alpha is often used as evidence that the items measure a latent construct, in this case oral health-related well-being and distress (Tables 10 and 11; see Supporting Documents).

Precision Across the Full Range of the Latent Trait

Reliability estimates are summaries of the precision of a measure computed from a sample. It is possible, and indeed likely, that the precision of a measure will differ by level of a latent trait.

To address this attribute of reliability, we have examined the test information function from a two-parameter graded response IRT model. The test information function (TIF) represents the reciprocal of the plot of the standard error of the measure, by the level of the latent trait, called "theta." A theta of zero is the average level of oral health-related well-being and distress (Figures 4-7; see Supporting Documents).

6.B. Validity

Validity of the measure is the extent to which the measure meaningfully represents the concept being evaluated. The method for establishing the validity of a measure will depend on the type of measure, data source, and other factors.

Explain your rationale for selecting the methods you have chosen, show how you used the methods chosen, and provide information on the results (e.g., R2 for concurrent validity).

Structural Validity

This type of validity asks the question "how many different things, or dimensions, does this measure assess?" A unidimensional structure is a requirement for IRT modeling. It is desirable for the measure to be unidimensional for all ages, a phenomenon termed factor invariance.

Initially, the unidimensionality assumption was tested by fitting exploratory factor analytic models. These analyses were conducted to assess the extent to which the original aesthetics item pool measures a singular and dominant health trait. The ratio of the first two eigenvalues and the percent variance explained by the single factor indicate that although there was a strong first

factor, enough evidence existed for the extraction of a second factor, which became the Oral Distress item bank (Table 12; see Supporting Documents).

The ratio of eigenvalues supports the extraction of two factors for 8-17 year-old children and youth.

The Well-Being and Distress banks were shown to be unidimensional using confirmatory factor analysis (CFA) to assess model fit for each distinct bank. The comparative fit index (CFI) was used to assess model fit. A CFI value of 0.95 or above is considered excellent fit. Table 13 (see Supporting Documents) shows the CFA model fit for both banks.

Concurrent Validity

We conducted a set of analyses that correlated the well-being and distress measures with other measures that we expected to be strongly related.

To examine convergent and discriminant validity, we evaluated associations with two PROMIS pediatric measures of pain and global health. We also evaluated correlations with the Child Oral Impact Profile-Reduced and the Project Health Pathways Positive Body Image scale (Table 14; see Supporting Documents).

Differential Item Functioning (DIF)

In the context of IRT, DIF is observed when the probability of item response differs across comparison groups such as gender, age, and race, after conditioning on (controlling for) level of the underlying state or trait measured. Uniform DIF occurs if the probability of response is consistently higher (or lower) for one of the comparison groups across all levels of the state or trait. Non-uniform DIF is observed when the probability of response is in a different direction for the groups compared at different levels of the state or trait.

We used the R package, *lordif*, to fit ordinal logistic regression analyses (regressing the item responses on overall oral health well-being, distress or pain level (theta), and a putative DIF characteristic (e.g., gender). We tested the following characteristics for DIF across each of the questions in each bank: age, gender, and race/ethnicity. Results indicate no significant DIF by any of these characteristics.

These findings support the validity of the items, suggesting that there is no ‘built-in’ bias of the items that would add error to estimates of differences by groups defined by these characteristics. Lack of DIF is an important pre-condition for conducting disparity analyses.

Section 7. Identification of Disparities

CHIPRA requires that quality measures be able to identify disparities by race, ethnicity, socioeconomic status, and special health care needs. Thus, we strongly encourage nominators to have tested measures in diverse populations. Such testing provides evidence for assessing measure’s performance for disparities identification. In the sections below,

describe the results of efforts to demonstrate the capacity of this measure to produce results that can be stratified by the characteristics noted and retain the scientific soundness (reliability and validity) within and across the relevant subgroups.

7.A. Race/Ethnicity

Known group comparisons for race and ethnicity were pending at the time this report was prepared.

7.B. Special Health Care Needs

Known group comparisons for special healthcare need status were pending at the time this report was prepared.

7.C. Socioeconomic Status

Socioeconomic status information was not available for analysis.

7.D. Rurality/Urbanicity

Known group comparisons for geographic area were pending at the time this report was prepared.

7.E. Limited English Proficiency (LEP) Populations

LEP information was not available for analysis.

Section 8. Feasibility

Feasibility is the extent to which the data required for the measure are readily available, retrievable without undue burden, and can be implemented for performance measurement. Using the following sections, explain the methods used to determine the feasibility of implementing the measure.

8.A. Data Availability

1. What is the availability of data in existing data systems? How readily are the data available?

This is a new measure and has not been used. Data for the POHR-WB and POHR-D are not available in any existing data collection system. POHR-WB and POHR-D data are acquired through self-report and parent-proxy report. Person-reported outcome (PRO) measures can be conducted via self-administration or interviewer administration. Data can be captured through many of the typical administration methods, including: face-to-face, telephone (interview or interactive voice response), mail, and internet.

2. If data are not available in existing data systems or would be better collected from future data systems, what is the potential for modifying current data systems or creating new data systems to enhance the feasibility of the measure and facilitate implementation?

With the imminent development of a short form, the brevity of the POHR-WB and POHR-D measures will make a variety of administration methods feasible. However, within the context of performance measurement activities at the State, health plan, or other level of the healthcare delivery system, implementation of these methods must be considered in terms of their sustainability over time.

We previously engaged Medicaid and CHIP leaders from the States of New Jersey, Pennsylvania, and Massachusetts to gather feedback on alternative approaches for implementing PRO performance measures. State leaders acknowledged that scarcity in resources poses barriers to any data collection effort that is not already systematically in use.

The following represent potential PRO administration contexts suggested by State leaders:

1. Inclusion in an existing survey. Many Medicaid and CHIP programs and health plans have a long history of fielding periodic surveys as part of their QI activities. As an example, many States utilize surveys from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) to obtain performance feedback regarding Medicaid clients' satisfaction with healthcare and to provide feedback at the health plan, provider group, and individual provider levels. The family of CAHPS surveys comprises widely disseminated tools that are used throughout the healthcare industry. Person-reported measures could be added to an existing survey, like the CAHPS.

2. Independent population-based survey. An alternative to integrating questions directly into another survey is to administer the POHR-WB and POHR-D measures potentially even parallel and based on the same sampling procedures. As independent surveys, these tools could be internet-based or administered via the telephone or using a paper-and-pencil instrument. Health plans could integrate PRO measures as part of the new beneficiary on-boarding process, whereby new members may participate in health assessment surveys. States or plans could also focus on high-risk populations, administering the tools via case management systems that are already in place.

3. Patient portal integrated with an EHR. The use of patient portals as a mechanism for engaging patients and collecting important information continues to increase. The support of pilot studies and implementation efforts integrating PROs into EHRs have received major Federal support from the National Institutes of Health (NIH), Agency for Healthcare Research and Quality (AHRQ), Patient Centered Outcomes Research Institute (PCORI), Centers for Medicare & Medicaid Services (CMS), Food and Drug Administration (FDA), and the Office of the National Coordinator for Health Information Technology (ONC) (Wu, 2013). The integration of PROs into Epic Systems Corporation's suite of EHR products also facilitates movement toward automated integration of PRO data into the EHR environment. As States and health plans continue to evolve in terms of their health information technology (IT) capacity and interconnectedness among different systems, the potential for administering PROs via patient

portals will become increasingly viable. Input provided by health plans also raises the possibility of administering the survey at the point of care.

The second option was identified by States as the most opportunistic strategy, given the existence of and familiarity with administering surveys for a variety of QI and performance measurement activities.

CHOP's extensive experience and success in administering PRO measures to thousands of children through multiple platforms and contexts (clinics, schools, etc.) provides insight into developing a strategy for implementing PROs in Medicaid and CHIP. In our work, we have observed important features of PRO administration that include low rates of non-completion, short completion duration, and reasonable time and resources to collect, process, and analyze data. These reflect important characteristics that are extremely relevant to publicly financed insurance programs where resources, including personnel time, energy, and funding, are very limited.

As States move progressively toward greater use and integration of health information systems, we believe that the opportunity for collecting and extracting PRO data via EHRs will be great. Software applications aimed at harnessing both clinical information and person-reported outcomes have been developed and are in early stages of adoption. The key will be the identification of PRO measures that reflect the need and interests of relevant stakeholders.

8.B. Lessons from Use of the Measure

1. Describe the extent to which the measure has been used or is in use, including the types of settings in which it has been used, and purposes for which it has been used.

Not applicable.

2. If the measure has been used or is in use, what methods, if any, have already been used to collect data for this measure?

Not applicable.

3. What lessons are available from the current or prior use of the measure?

Not applicable.

Section 9. Levels of Aggregation

CHIPRA states that data used in quality measures must be collected and reported in a standard format that permits comparison (at minimum) at State, health plan, and provider levels. Use the following table to provide information about this measure's use for reporting at the levels of aggregation in the table.

For the purpose of this section, please refer to the definitions for provider, practice site, medical group, and network in the Glossary of Terms.

If there is no information about whether the measure could be meaningfully reported at a specific level of aggregation, please write "Not available" in the text field before progressing to the next section.

Level of aggregation (Unit) for reporting on the quality of care for children covered by Medicaid/ CHIP†:

State level* Can compare States

Intended use: Is measure intended to support meaningful comparisons at this level? (Yes/No)

Yes.

Data Sources: Are data sources available to support reporting at this level?

No.

Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Not determined.

In Use: Have measure results been reported at this level previously?

No.

Reliability & Validity: Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

Unintended consequences: What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

Other geographic level: Can compare other geographic regions (e.g., MSA, HRR)

Intended use: Is measure intended to support meaningful comparisons at this level? (Yes/No)

Yes.

Data Sources: Are data sources available to support reporting at this level?

Yes; all 50 States by race, ethnicity, private/public coverage, SES, chronic disease status, geography.

Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Not determined.

In Use: Have measure results been reported at this level previously?

No.

Reliability & Validity: Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

Unintended consequences: What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

Medicaid or CHIP Payment model: Can compare payment models (e.g., managed care, primary care case management, FFS, and other models)

Intended use: Is measure intended to support meaningful comparisons at this level? (Yes/No)

Yes.

Data Sources: Are data sources available to support reporting at this level?

Yes, public vs. private.

Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Not determined.

In Use: Have measure results been reported at this level previously?

No.

Reliability & Validity: Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

Unintended consequences: What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

Health plan*: Can compare quality of care among health plans.

Intended use: Is measure intended to support meaningful comparisons at this level? (Yes/No)

Yes.

Data Sources: Are data sources available to support reporting at this level?

No.

Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Not determined.

In Use: Have measure results been reported at this level previously?

No.

Reliability & Validity: Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

Unintended consequences: What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

Provider Level

Individual practitioner: Can compare individual health care professionals

Intended use: Is measure intended to support meaningful comparisons at this level?
(Yes/No)

Yes.

Data Sources: Are data sources available to support reporting at this level?

No.

Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Not determined.

In Use: Have measure results been reported at this level previously?

No.

Reliability & Validity: Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

Unintended consequences: What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

Provider Level

Hospital: Can compare hospitals

Intended use: Is measure intended to support meaningful comparisons at this level?

(Yes/No)

Yes.

Data Sources: Are data sources available to support reporting at this level?

No.

Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Not determined.

In Use: Have measure results been reported at this level previously?

No.

Reliability & Validity: Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

Unintended consequences: What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

Provider Level

Practice, group, or facility:** Can compare: (i) practice sites; (ii) medical or other professional groups; or (iii) integrated or other delivery networks

Intended use: Is measure intended to support meaningful comparisons at this level?

(Yes/No)

Yes.

Data Sources: Are data sources available to support reporting at this level?

No.

Sample Size: What is the typical sample size available for each unit at this level? What proportion of units at this level of aggregation can achieve an acceptable minimum sample size?

Not determined.

In Use: Have measure results been reported at this level previously?

No.

Reliability & Validity: Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?

No.

Unintended consequences: What are the potential unintended consequences of reporting at this level of aggregation?

Unknown.

Section 10. Understandability

CHIPRA states that the core set should allow purchasers, families, and health care providers to understand the quality of care for children. Please describe the usefulness of this measure toward achieving this goal. Describe efforts to assess the understandability of this measure (e.g., focus group testing with stakeholders).

State Medicaid programs, CHIP programs, and other key stakeholders have identified child health outcomes as an unmet need in pediatric health and healthcare quality measurement and expressed significant interest in identifying meaningful outcome measures that assess children's well-being and functioning. In addition to being validated as an important and meaningful measure concept throughout a large body of literature, the significance of oral health pain as a PRO concept was confirmed by dental clinicians and children through in-depth interviews. Assessment of pain, distress, and well-being by clinicians is already an integral and standard part of the dental clinical examination and provides important clues around a child's oral health. These include characteristics of pain regarding the location, quality, nature, and impact. Among patient reportable outcomes, clinicians highlighted pain and distress as a key concept to target. Furthermore, during the conduct of children and parent concept elicitation interviews, the pain/inflammation concept was clearly mapped to the extensive oral health experiences, including well-being and distress, described from the perspectives of the intended PROM respondents, highlighting both the relevance and understandability of these measures and of the complementary oral health pain/inflammation measure.

Section 11. Health Information Technology

Please respond to the following questions in terms of any health information technology (health IT) that has been or could be incorporated into the measure calculation.

11.A. Health IT Enhancement

Please describe how health IT may enhance the use of this measure.

The utility and availability of the POHR-WB and POHR-D will be greatly enhanced when EHRs widely incorporate patient-reported outcomes in their patient portals. Epic has developed PRO applications in which PROMIS measures are part of the module and freely available to any user. The EHR module enables patients to complete the instrument at home, on a hand held device, or in the office waiting area. Data are stored in the EHR, which permits linkage with other clinical variables. Registries are another health IT application that will increase the availability of PROs.

The arrival and brevity of measure short forms will further increase the likelihood of adoption of these measures as an evaluation of children's health.

11.B. Health IT Testing

Has the measure been tested as part of an electronic health record (EHR) or other health IT system?

No.

If so, in what health IT system was it tested and what were the results of testing?

Not applicable.

11.C. Health IT Workflow

Please describe how the information needed to calculate the measure may be captured as part of routine clinical or administrative workflow.

See Section 11.A, above. In general, the information cannot be captured as part of currently routine workflow.

11.D. Health IT Standards

Are the data elements in this measure supported explicitly by the Office of the National Coordinator for Health IT Standards and Certification (ONC) criteria (see healthit.hhs.gov/portal/server.pt/community/healthit_hhs_gov__standards_ifr/1195)?

No.

If yes, please describe.

Not applicable.

11.E. Health IT Calculation

Please assess the likelihood that missing or ambiguous information will lead to calculation errors.

Not applicable.

11.F. Health IT Other Functions

If the measure is implemented in an EHR or other health IT system, how might implementation of other health IT functions (e.g., computerized decision support systems in an EHR) enhance performance characteristics on the measure?

Not applicable.

Section 12. Limitations of the Measure

Describe any limitations of the measure related to the attributes included in this CPCF (i.e., availability of measure specifications, importance of the measure, evidence for the focus of the measure, scientific soundness of the measure, identification of disparities, feasibility, levels of aggregation, understandability, health information technology).

Age limitations: These measures do not address the health of children ages 0-4 years.

More work required to assess responsiveness: These measures require longitudinal evaluation to assess their responsiveness to change.

Implementation: Activities to encourage PRO adoption are necessary in order to generate data that can be repurposed for performance measurement and other applications.

Section 13. Summary Statement

Provide a summary rationale for why the measure should be selected for use, taking into account a balance among desirable attributes and limitations of the measure. Highlight specific advantages that this measure has over alternative measures on the same topic that were considered by the measure developer or specific advantages that this measure has over existing measures. If there is any information about this measure that is important for the review process but has not been addressed above, include it here.

The Pediatric Oral Health-Related Well-Being and Distress measures are evidence-based, patient reported outcome measures of pediatric oral health. The POHR-WB and POHR-D item banks comprise 14 and 11 questions, respectively. They reflect the complementary and inter-related concepts of one's satisfaction with appearance and the social-emotional impact of these perceptions. Child-report versions are appropriate for children aged 8-17 years, and parent-proxy versions are appropriate for parents of children aged 5-17 years. The measures use a 7-day report period.

The POHR-WB and POHR-D item banks were developed according to the scientific standards of the Patient Reported Outcome Measurement Information System (PROMIS; www.nihpromis.org). PROMIS has developed more than 60 measures of patient-reported health. These measures have excellent reliability, which indicates that they yield very precise measurements of a child's oral health well-being and distress. Likewise, their validity is excellent, which suggests that indeed they do evaluate a child's direct report of well-being and distress. The comprehensive and rigorous PROM-development process ensures that these tools will be useful for clinical, population health, and research applications that require precise and valid measures of children's oral health experiences.

PRO measures may be integrated into existing data collection platforms, such as EHRs, ongoing surveys, and registries, or included in new survey modules. These instruments may be administered in a patient/ provider friendly manner (i.e., as an intake form or at the point of

healthcare program initiation or as part of existing consumer survey administration) and subsequently coded in to an EHR.

The POHR-WB and POHR-D item banks are responsive to the voices of State leaders who advocated for greater attention to child outcomes; they represent an opportunity to address a topic of great public health and clinical significance. The tools summarize children's oral health experiences and the impact of these experiences. Patient reported outcome measures such as the POHR-WB and POHR-D item banks, provide information about health and the impact of healthcare that is unique and complementary to information derived from measures of clinical outcomes and healthcare processes. Such measures have the potential to improve pediatric healthcare quality assessment at the point of care and through system-level performance evaluations.

References

American Dental Association (ADA). Pediatric oral health quality and performance measures concept set: Achieving standardization and alignment. Washington, DC: ADA; 2012.

Bjorner JB, Chang CH, Thissen D, et al. Developing tailored instruments: item banking and computerized adaptive assessment. *Qual Life Res* 2007; 16(Suppl 1):95-108.

Centers for Medicare & Medicaid Services (CMS). National Health Expenditure Data 1960-2014. Baltimore, MD: CMS; 2014. Available at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/index.html>. Accessed July 23, 2019.

Dye BA, Tan S, Smith V, et al. Trends in oral health status: United States, 1988-1994 and 1999-2004. *Vital Health Stat* 11 2007; (248):1-92.

Eli I, Bar-Tal Y, Kostovetzki I. At first glance: Social meanings of dental appearance. *J Public Health Dent* 2001; 61:150-4.

Filstrup SL, Briskie D, da Fonseca M, et al. Early childhood caries and quality of life: Child and Parent Perspectives. *Pediatr Dent* 2003; 25(5):431-40.

Health Management Associates. Managing California's Medicaid Dental Program: Lessons from Other States. Oakland, CA: California Healthcare Foundation; 2009. Available at: <http://www.chcf.org/publications/2009/07/managing-californias-medicaid-dental-program-lessons-from-other-states>. Accessed July 23, 2019.

Jackson SL, Vann WF, Kotch J, et al. Impact of poor oral health of children's school attendance and performance. *Am J Public Health* 2011; 101:1900-6.

Klem M, Saghafi E, Abromitis R, et al. Building PROMIS item banks: Librarians as co-investigators. *Qual Life Res* 2009; 18(7):881-8.

Mann M, Hosman CM, Schaalma HP, et al., Self-esteem in broad-spectrum approach for mental health promotion. Health Educ Res 2004; 19(4):357-72.

Mouradian WE, Wehr E, Crall JJ. Disparities in children's oral health and access to dental care. JAMA 2000; 284(20):2625-31.

Teresi JA, Fleishman JA. Differential item functioning and health assessment. Qual Life Res 2007; 16(Suppl 1):33-42.

Tin-Oo MM, Saddki N, Hassan N. Factors influencing patient satisfaction with dental appearance and treatments they desire to improve aesthetics. BMC Oral Health 2011; 11(6).

U.S. Department of Health and Human Services. Healthy People 2010. Washington, DC: 2010. Available at <http://www.healthypeople.gov/2010/>. Accessed July 23, 2019.

U.S. Department of Health and Human Services (HHS). Oral Health in America: A Report of the Surgeon General. Executive Summary. Rockville, MD: HHS; 2000.

Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life: A conceptual model of patient outcomes. JAMA 1995; 273(1):59-65.

Wu AW. Advances in the use of patient reported outcome measures in electronic health records, including case studies. Technical report in support of the PCORI National Workshop to Advance the Use of PRO measures in Electronic Health Records, Atlanta, GA, November 19-20, 2013. Baltimore, MD: Bloomberg School of Public Health; 2013. Available at <http://www.pcori.org/assets/2013/11/PCORI-PRO-Workshop-EHR-Landscape-Review-111913.pdf>. Accessed July 24, 2019.

Section 14: Identifying Information for the Measure Submitter

First Name: Christopher B.
Last Name: Forrest, MD, PhD
Title: Professor of Pediatrics
Organization: Children's Hospital of Philadelphia
Mailing Address: 3535 Market Street, 16th Floor
City: Philadelphia
State: PA
Postal Code: 19104
Telephone: 267-426-6917

Email: forrest@email.chop.edu

The CHIPRA Pediatric Quality Measures Program (PQMP) Candidate Measure Submission Form (CPCF) was approved by the Office of Management and Budget (OMB) in accordance with the Paperwork Reduction Act.

The OMB Control Number is 0935-0205 and the Expiration Date is December 31, 2015.

Public Disclosure Requirements

Each submission must include a written statement agreeing that, should U.S. Department of Health and Human Services accept the measure for the 2014 and/or 2015 Improved Core Measure Sets, full measure specifications for the accepted measure will be subject to public disclosure (e.g., on the Agency for Healthcare Research and Quality [AHRQ] and/or Centers for Medicare & Medicaid Services [CMS] websites), except that potential measure users will not be permitted to use the measure for commercial use. In addition, AHRQ expects that measures and full measure specifications will be made reasonably available to all interested parties. "Full measure specifications" is defined as all information that any potential measure implementer will need to use and analyze the measure, including use and analysis within an electronic health record or other health information technology. As used herein, "commercial use" refers to any sale, license or distribution of a measure for commercial gain, or incorporation of a measure into any product or service that is sold, licensed or distributed for commercial gain, even if there is no actual charge for inclusion of the measure. This statement must be signed by an individual authorized to act for any holder of copyright on each submitted measure or instrument. The authority of the signatory to provide such authorization should be described in the letter.

AHRQ Publication No. 19-0051
August 2019