Pediatric Global Health-7 Measure

Section 1. Basic Measure Information

1.A. Measure Name

Pediatric Global Health-7 (PGH-7) Measure

1.B. Measure Number

0081

1.C. Measure Description

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

The Pediatric Global Health-7 (PGH-7) comprises seven questions about a child's overall health. The questions can be answered by children ages 8-17 years or the parents of children 5-17 years old. The measure produces a summary score of a child's general health status from his or her perspective. The questions ask about children's perceptions of their health in general, physical health, mental health, pain, friendships, family life, self-esteem, and feelings of worry and sadness.

The PGH-7 was developed according to the scientific standards of the Patient Reported Outcome Measurement Information System (PROMIS—www.nihpromis.org). PROMIS is a federally funded initiative that has developed state-of-the-science criteria for creating patient- reported outcome measures. PROMIS now includes over 60 measures of patient-reported health, including the Adult Global Health Measure. The PGH-7 was developed to be conceptually compatible with the adult version.

Development of the PGH-7 was conducted with 6,072 children and youth (8-17 years old) and parents (of children 5-17 years old). Participants resided in all 50 States. The measure has excellent reliability, which indicates that it yields very precise measurements of a child's global health. Likewise, the validity is excellent, which suggests that indeed it evaluates a child's overall health. The PGH-7 has also been validated with other measures of children's global health and has been shown to be responsive to changes in asthma symptoms among children seen in an emergency room for an asthma exacerbation.

The PGH-7 is an outcome measure. That is, it assesses health and well-being. As a general health measure for children 5-17 years of age, it will be useful across all diseases and in a wide variety of applications, from evaluations of care for specific diseases to assessments of changes in delivery system models, across most pediatric populations.

1.D. Measure Owner

The Children's Hospital of Philadelphia (CHOP).

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

The PGH-7 is not currently NQF-endorsed. In our interviews with State stakeholders, they expressed the value they see in NQF endorsement. They stated that the endorsement increases the chances that they will select and prioritize PGH-7 as a measurement tool.

Notably, NQF was recently commissioned by the Department of Health and Human Services to develop recommendations for the development and selection of patient-reported outcome measures that can be used to encourage provider accountability and inform health care quality improvement initiatives.

We believe the PGH-7 is a strong candidate for NQF endorsement based on the organizations existing evaluation criteria, the similarities between the CPCF and NQF criteria, and because the instrument was developed in compliance with the rigorous NIH PROMIS standards (for PRO measures in particular).

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 PGH-7 has been incorporated into the PROMIS measure collection. The domain framework, which is hierarchically arranged, is shown in Figure 1 (see Supporting Documents). Note that the overall concept that PROMIS measures assess is self-reported health, and the highest level dimensions are physical, mental, and social health. This structure of health is reflective of the WHO tripartite conceptualization of health. The global health measure is an evaluation of overall health, which integrates one's physical, mental, and social health. The questions asked in the global health measure, therefore, evaluate all three dimensions of one's 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.

We developed national normative values with a large sample of the general population. The sum-score to T-score conversion table is shown in the Technical Specifications, Table 3 (see Supporting Documents).

1.H. Numerator Exclusions

For the child self-report measure, we recommend exclusion of children with cognitive limitations or developmental delays, including: autism, intellectual developmental disability, mental retardation, or severe learning disabilities, which affect a child's language capabilities and capacity to self-report.

For the parent proxy-report measure, the same cognitive limitations exclusions for children apply to parent respondents. The child-report form is applicable to children 8-17 years old. Ample research has demonstrated the ability of children as young as age 8 to provide reliable and valid self-reports of their health.

The parent proxy-report form is applicable to parents of children 5-17 years old. Age 5 has been chosen as the floor because the questions used for the measure are not relevant to younger children.

Both age cut-points are consistent with PROMIS standards.

1.I. Denominator Statement

For mean scores, the denominator is the sum of the number of individuals in the target population.

1.J. Denominator Exclusions

These are the same as the numerator exclusions:

For the child self-report measure, we recommend exclusion of children with cognitive limitations or developmental delays, including: autism, intellectual developmental disability, mental retardation, or severe learning disabilities, which affect a child's language capabilities and capacity to self-report.

For the parent proxy-report measure, the same cognitive limitations exclusions for children apply to parent respondents. The child-report form is applicable to children 8-17 years old. Ample research has demonstrated the ability of children as young as age 8 to provide reliable and valid self-reports of their health.

The parent proxy-report form is applicable to parents of children 5-17 years old. Age 5 has been chosen as the floor because the questions used for the measure are not relevant to younger children.

Both age cut-points are consistent with PROMIS standards.

1.K. Data Sources

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

Survey, parent/caregiver report; survey, child report; electronic medical record.

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

Patient registry; personal health record.

Note: The PGH-7 requires obtaining self-reports from children or proxy reports from their caregivers. This can be done using paper-and-pencil questionnaires, web-based forms that store data in electronic health records (e.g., EHRs via patient portals), free-standing web-based data collection systems, patient-reported registries, and personal health records. In ongoing research, we have demonstrated that there are no mode of administration effects (computer, web-based = school, paper-and-pencil = clinic, iPad) for pediatric reported health outcome measures.

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.

See 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 Global Health-7 (PGH-7) is a generic outcome measure that will be useful for evaluating quality within and across any acute or chronic conditions. It may provide an evaluation of how an intervention benefits the health of the patient or patient population and, thus, could add to the meaningfulness of changes in processes of care, particularly when a change in process is measured at the same time as global health.

Several studies have examined the associations of global health with socioeconomic status, chronic disease, and insurance status. Two studies are exemplary. In the first, Newacheck and colleagues found significantly lower general health among poor versus non-poor adolescents (Newacheck, Hung, Park, et al., 2003). Another study found similar results, including that disparities in general health remained between poor and non-poor children even in models that

5

included controls for health insurance status (Larson, Halfon, 2010). Among patients with juvenile idiopathic arthritis, those with publicly financed health insurance had lower self-reported general health than their counterparts with privately financed insurance (Brunner, Taylor, Britto, et al., 2006). Our known-group validation work (see section 6.B Validity in this report) confirms and expands these findings.

Potential for Quality Improvement

The goal of most quality improvement (QI) activities is to improve the reliability of processes of care that have likely or established associations with patient outcomes. Very little QI actually measures the impact of the intervention on health. The PGH-7 can fill this gap. Quality improvement activity that reduces symptom burden or improves children's functional status will have a positive effect on global health (see the Measure Specifications, Figure 1, in the Supporting Documents for our conceptual model). For example, Levy and colleagues found that improving the home environment for children with asthma resulted in better asthma control and general health (Levy, Brugge, Peters, et al., 2006). The PGH-7 has been cited in similar work.

Prevalence of Condition Among Children or Pregnant Women

Not applicable; the PGH-7 is a domain-specific rather than disease-specific measure. Global health is the domain measured. It is applicable to all children and pregnant women.

Severity and Burden of Condition on Children, Family, and Society

Not applicable; the PGH-7 is a domain-specific rather than disease-specific measure. Global health is the domain measured. It is applicable to all children and pregnant women.

Rarity of Condition

The PGH-7 is a domain-specific rather than disease-specific measure. It is applicable to all children and all medical conditions, including rare diseases, defined as those affecting fewer than 200,000 individuals in the U.S. population or an incidence rate of less than 1 in 1,500 individuals.

Fiscal Burden of the Condition Currently and Over the Lifespan

Few studies have examined the fiscal burden of low global health. However, Seid and colleagues found in a 2-year prospective study that global health accounted for a large share in the variance in health care costs (Seid, Varni, Segall, et al., 2004). Children's global health has been associated with future health care utilization, even after controlling for patient morbidity (Forrest, Riley, Vivier, et al., 2004).

Association of Measure with Children's Future Health

In a large meta-analysis of longitudinal studies, global health was consistently found to be a strong predictor of future mortality (Idler, Benyamini, 1997).

Developmental Change of Measure

A key strength of the PGH-7 is its ability to measure change across developmental stages. We have developed the PGH-7 to provide an assessment of global health across the life course, from 5 to 85 years of age.

In our national sample of 3,635 children (see Section 6 in this report for a full description of the development sample), we found that global health decreases with age. Mean sum scores were 29.44 for those aged 8-10 years, 28.32 for those aged 11-13 years, and 27.79 for those 14-17 years of age. Thus, the measure shows age-related differences.

Importantly, in psychometric testing of the PGH-7, we found no age-related differential item functioning (see section 6.B Validity in this report). 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, although older children have lower levels of global 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).**

The PGH-7 represents a tremendous opportunity to expand the scope of pediatric quality measurement by providing an assessment of health evaluated from the perspective of children themselves. Among the limited number of patient-reported outcomes currently in use, most focus on disease states. Disease-oriented outcome frameworks present a challenge for pediatric health outcome assessment because, compared to adults, relatively few children have chronic health conditions. Furthermore, there has been a general failure to consider positive child health outcomes that are essential contributors to a child's growth and thriving in family, school, and social environments. Lastly, patient-reported outcome assessments enhance the patient-centeredness of health care quality because they ensure that health care interventions are evaluated for their impact on health from patients' own perspectives.

Although the PGH-7 would not be informative in terms of determining Early Periodic Screening, Diagnostic, and Treatment (EPSDT) needs, we believe that the measure, given its sensitivity to activities that reduce symptom burden or children's functional status, will be applicable to assessing outcomes related to the provision of EPSDT services (e.g., physical, dental, auditory, vision).

We have compared global health between patients with Medicaid/CHIP versus those with private insurance or no insurance. Results are presented in Table 2 (see Supporting Documents).

Results indicate a consistent finding of lower self-reported general health for publicly insured children, compared with both privately insured and uninsured children.

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).

The PGH-7 is a pediatric extension of the PROMIS Adult Global Health Measure (Hays, Bjorner, Revicki, et al., 2009). It was designed to be conceptually equivalent. In addition, we completed work to compare the PGH-7 with existing legacy measures of children's general health (Forrest, Tucker, Ravens-Sieberer, et al., 2016).

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.
    d. **Service – preventive health, including services to promote healthy birth:** Yes.
    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):** Health promotion; healthy birth; behavioral health; oral health. The PGH-7 is a domain-specific measure, meaning it cuts across service types.
    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; the parent-report version extends to children age 5, and both parent-report and child-report versions cover school-age children 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):** Yes; when the PGH-7 is linked with the adult version, it will also be applicable to women and men.
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 literature to gain a better understanding of the empirical evidence linking health insurance to pediatric health outcomes. The following text summarizes our methods, results, and conclusions.

Methods

The search strategy was based on the procedures outlined by Klem, Saghafi, Abromitis, et al. (2009). Biomedical librarians at the University of Pennsylvania worked with our team to develop a controlled vocabulary that included insurance coverage and access to health care as key variables, in addition to sets of terms related to global health, social indicators, measurement, and self/proxy report. The final iteration of the key word search linked the following categories of terms with the applicable Boolean operators: (Global Health Terms OR Social Indicator Terms) AND Measurement Terms AND Self/ Proxy Report Terms AND (Insurance Terms OR Access Terms). The terms were applied to both the Medline and Health STAR databases to identify

articles that described or applied PRO instruments to assess children's global, system-specific, or disease-specific health characteristics.

Studies were eligible for inclusion if they were (1) written in the English language, (2) published in peer-reviewed journals or government reports dated 1985 or later, (3) included specific measure(s) of child health outcomes, (4) enrolled children and youth to 21 years of age, and (5) included related measures of access to care, regular source of care, health care utilization and costs, or child health outcomes as dependent variables. Studies were excluded if the study population included only adults age 21 years and older; contained measures of the continuity of care, source of care, or initial enrollment but not insurance coverage; reported exclusively on adult health outcomes; or were conducted in a country outside the Organization for Economic Cooperation and Development (OECD). Reports that included individuals under 21 years of age were included due to public insurance standards surrounding age eligibility (which includes individuals up to age 21 years).

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. Eligible papers were abstracted using a standardized electronic abstraction form that had been piloted prior to study initiation. Abstraction variables are provided in Table 3 (see Supporting Documents).

Information abstracted from all relevant reports included characteristics of the study population (age ranges, race, ethnicity, geographical location, etc.), and study design, as well as variables related to demographics, insurance coverage, health services access, and health care utilization variables reported. Since studies may contain more than one relevant measure of general health, we abstracted information on all measures contained in each eligible paper. In addition, we retrieved information on each study's design (cohort, panel, cross-sectional, controlled trial, or descriptive), population, study aims, data source (administrative data, surveys, chart abstraction), sample size, and outcomes (access to care, regular source of care, health services utilization and/or costs, and health outcomes).

Literature Review Results

Following electronic database searches, we identified 880 potentially eligible papers, of which 424 were unique articles. Our search in Health STAR yielded no new articles, as all results were duplicates of the Medline search. After review of the abstracts, we excluded 308 papers that did not fully meet the eligibility criteria. We reviewed the full text of 116 papers and excluded an additional nine papers primarily because those studies were conducted in countries outside the OECD.

The overall agreement rate among reviewers for paper eligibility was 97 percent with $\kappa = 0.9287$ ($p < .0001$). The final adjudicated 107 papers contained 83 separate measures of child health outcomes. Table 4 (see Supporting Documents) compares the number of studies in which child health was reported as an outcome in relation to health insurance status, access to health care services, or utilization of health care services.

The most common study design evaluating the measures was a cohort study design (67 studies), although many relied on secondary analysis of national (e.g., NHANES, CSHCN, NHIS, etc.) or State (Medicaid) health surveys. Most measures were assessed in populations of children and adolescents under age 18 years of age, with few measures exclusively of older adolescents and young adults, ages 18-21 years (four studies).

Many reports (68 percent, 73 studies) included measures that assessed whether insurance coverage was associated with health outcomes, although measures of access to care and utilization of care were more often included within reports (76 percent for access and 90 percent for utilization). The measures of child health consisted of three major categories (1) assessment of general or global health; (2) assessments of specific body system health (e.g., visual, dental); or (3) diagnostic or condition-specific health surveys that may also include assessment of child general health. Several studies included proprietary standardized outcomes of child health including the PedsQL, Kid Screen, CHIP, and CHQ among others.

Discussion of Literature Review Findings

Children with Medicaid were less likely to be in excellent health than those with private insurance or those with no insurance (Blackwell, Tonthat, 2002). Lower PedsQL scores were reported for children with public versus private health insurance (Wade, Mansour, Line, et al., 2008). Medicaid patients were significantly more disabled than privately insured patients; severe disease was observed more often overall and at the time of diagnosis in Medicaid patients compared with those who had private insurance or no insurance (Brunner, et al., 2006). For those who reported forgoing care before enrolling in CHIP, improved access after CHIP enrollment was associated with higher global health (Seid, Varni, Cummings, et al., 2006). Gaining insurance increased the percentage reporting "no unmet health care need due to cost" and the percentage reporting excellent, very good, or good health (Busch, Vigdor, 2008). Children on Medicaid were more likely than the uninsured and privately insured to miss days from school (Dey, Bloom, 2005).

Children with Medicaid insurance self-report poorer health than those with private insurance. The few longitudinal studies that examined the impact of insurance on health found beneficial effects. These findings may be due to increased likelihood among children with insurance of having a medical home and overall better access to care. The literature review provides support for a linkage between insurance and self-assessed health. Interestingly, the few studies that focused on dental outcomes indicate that children's dental health is highly impacted by insurance coverage. The inclusion of oral health may be warranted for inclusion in a health outcome inventory. Similarly, visual health screening and eye exams are also sensitive to health insurance coverage and may help improve the sensitivity of global health measures to differences in health insurance coverage.

NHIS Analyses

Our literature review identified the annual National Health Interview Survey (NHIS) as the most common source for the reports that included secondary analyses. Because NHIS is publicly available, we were able to download these data and conduct independent analyses.

We extracted the variables of health insurance status and the general health item (an indicator of global health) to calculate an effect size for insurance coverage type on health status.

We limited the data extraction to children 5-21 years of age for NHIS calendar years 1997-2009. Data were used to estimate effect sizes (Cohen's D) for the association between insurance type and general health. Table 5 (see Supporting Documents) summarizes our findings.

These NHIS data analyses confirm the findings from the literature: children with publicly financed insurance have lower global health compared with those who have private or no insurance.

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.

Response to Stakeholders

In stakeholder discussions that helped to identify the measure priorities for PQMP, several State leaders advocated for greater attention to outcomes such as global health, functioning, and children's well-being. The PGH-7 fills this need.

Patient-Centeredness

With the emergence of the Patient-Centered Outcomes Research Institute (PCORI) and renewed emphasis within AHRQ and across the public sector on patient-centered care, new measures of outcomes from the person's perspective are needed. The PGH-7 is such a measure.

Domain-Specific

The PGH-7 is similar to all PROMIS outcome measures and is intended to be used across conditions. Thus, the same outcomes data can be obtained for QI activities that are disease-specific, service-specific, or more generic.

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.

Overview of Development and Validation Samples

We collected data from five samples for the development and evaluation of the PGH-7, which are described here. Approximately 8,000 child and parent participants were included in these studies. The Children's Hospital of Philadelphia's institutional review board approved study protocols, parental consent forms, and child assent forms.

1. National Sample of Children and Youth 8-17 years of age (n =3,635) and Parents of Children 5-17 years of age (n =1,807): "Sample 1." Children and their parents were recruited by the internet panel company, Op 4G, over a 10-day period in December 2011. This not-for-profit firm maintains a panel of several hundred thousand individuals who are given the option of participating in surveys. With each completed survey, participants receive a small stipend ($5) which is divided into an account for the charity of their choice and personal use. Children were recruited via their parents, who were sent an email invitation. Consent and assent were completed online. Children were instructed to complete the questionnaire alone. For 30 percent of 8 and 9 year-olds, parents helped by explaining or reading questions to the children. Sensitivity analyses comparing children who had help versus those who did not found no differences in the psychometric properties of the PGH-7. Data from this sample were used for all psychometric testing. The sociodemographic distribution of the samples of children and their parents is shown in Table 6.

2. National Sample of Children and Youth 8-17 years of age (n=1,001) and Parents of Children 5-17 years of age (n=1,001): "Sample 2." Children and their parents were recruited from Op 4G from August-September 2012. Children were recruited via their parents, who were sent an email invitation. Consent and assent were completed online. Children were instructed to complete the questionnaire alone. Data from this sample were used for concurrent validation of the PGH-7 with other legacy instruments of children's self-reported health, as well as evaluation of known group validity for variables only available in this sample. The socio-demographic distribution of the samples of children and their parents is shown in Table 6.

3. National Sample of Children for a Test/Retest Administration (n=334 children and n=296 parents). A separate sample of children was recruited from Op4G between December 2011 and January 2012 to participate in a 2-week test/retest administration. The sociodemographic distribution of the samples of children and their parents is shown in Table 7.

4. Locally Recruited Cognitive Interview Samples (n=20 children, 15 parents). We conducted 20 child and 15 parent cognitive interviews on the 10 pediatric global health items retained after calibration between May-June 2012. CHOP partnered with the Public Citizens for Children and Youth, a local advocacy group focused on helping children and families access quality health care, to assist in the recruitment of subjects for PGH-7 cognitive interviews. PCCY staff assisted by identifying eligible participants via telephone and local PCCY

functions and assessing interest in participation. PCCY provided CHOP study staff with a list of interested contacts. Study participants were met in person by research staff either at the PCCY office or event where consent was obtained and the interview was conducted. Child assent was also obtained before conducting the child interview.

Children were between the ages of 8 and 18, inclusive. There was equal gender representation. Black/African American children constituted 95 percent of the sample. We oversampled children with respect to public insurance coverage and racial/ ethnic group given the relevance of these characteristics to the Medicaid/ CHIP population: 10 children were 8-11 years old, 3 were 12-14 years old, and 7 were 15-18 years old.

5. Locally recruited children presenting to the emergency department (ED) with asthma exacerbations. We conducted a longitudinal clinical validation study by recruiting children age 5-17 presenting to the ED with an acute asthma exacerbation. Data were collected from these child- parent dyads in order to evaluate the construct validity of the PGH-7 as well as the responsiveness of the measure to clinical change among children with asthma.

Recruitment procedures involved consents, questionnaires, and inclusion/exclusion criteria set forth by CHOP investigators.

Data were collected three times. First, data were collected in person at the Children's Hospital of Philadelphia ED by research staff and at two subsequent time points by the Temple University Institute for Survey Research (Temple ISR). At the second and third survey rounds, patients (and parents) were contacted by Temple ISR staff via a mailed latter, e-mail, or telephone, and were provided two options for survey completion: (1) a web link or (2) a toll free number from which a computer-assisted telephone interview was conducted. Details for the second and third rounds of follow-up were outlined in an information sheet given to the patients by research staff while in the ED.

Comprehensive clinical and patient-reported assessments were performed at baseline in the hospital by CHOP staff; person-reported assessments were collected at 2-4 weeks and 6-12 weeks after the ED visit by Temple ISR. CHOP staff reviewed medical records for all participating children. Participants were given gift cards of $15 for each round of participation (see Table 8 in the Supporting Documents for this sample's demographics).

Reliability

We examined reliability in three ways:

- Internal consistency: a measure of the homogeneity of the items.

- Test/retest: a measure of the stability of item responses.

- Error associated with the full range of the latent trait: obtained from IRT analyses of the information function.

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, global health (see Table 9 in the Supporting Documents).

Test-Retest Reliability

To evaluate the instrument's stability, we conducted a 2-week retest reliability evaluation among children and parents. We used the intraclass correlation coefficient to assess the reliability coefficient. Results are shown separately for children and parents in 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, indeed likely, that the precision of a measure will differ by level of a latent trait.

To address this attribute of reliability, we examined the test information function from a two-parameter graded response IRT model. This analysis produced a plot of the standard error of the measure, which is the reciprocal of the information function, by the level of the latent trait, called "theta." A theta of zero is the average level of global health (see Figure 3 in the Supporting Documents). Lower standard error means greater reliability and precision of the estimate.

Reliability Summary

The PGH-7 has excellent reliability and stability. The precision of the measure is superior for individuals with average and low levels of global health compared to those with high levels. This is a reasonable limitation because in QI, we are most interested in moving children with poor health to higher levels.

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).

The PGH-7 has undergone five types of validation testing:

1. Structural validity.
2. Content validity.
3. Concurrent validity.
4. Construct validity.
5. Differential item functioning.

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 item response theory 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 to the pediatric data. These analyses were conducted to assess the extent to which the 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 for all children, the PGH-7 is unidimensional (Table 12, see Supporting Documents).

The ratio of eigenvalues supports factor invariance for 8-17 year-old children and youth (see Table 13 in the Supporting Documents). However, it is intriguing that the global health measure for adults 18 years of age verges on a two-factor solution. This is consistent with the adult version of the global health measure, which can be scored as two factors: physical health and mental health (Hays, et al., 2009).

Once the PGH-7 was shown to be unidimensional, we ran confirmatory factor analysis to assess model fit, overall and for each year of age. The comparative fit index (CFI) was used to assess model fit. A CFI value of 0.95 or above is considered to be excellent fit.

These findings indicate adequate unidimensionality of the PGH-7 across all age groups.

Content Validity

The conceptual breadth of the PGH-7 is based on the adult PROMIS measure, which involved explication of the construct by subject matter experts, and the PROMIS domain framework (Figure 1; see Supporting Documents).

We conducted concept elicitation on the global health concept with children and parents during the cognitive interviews. When asked about the kinds of things parents thought about regarding their child's overall health, parents tended to focus on disease (i.e., child is healthy because he doesn't get sick very often/doesn't have a chronic illness) and preventive care (i.e., it's important to go to the doctor for vaccinations or the dentist for a cleaning to maintain good health). A majority of parents also mentioned the importance of healthy eating and balanced nutrition.

Parents subsequently were given an opportunity to review the PGH-7 and were asked to comment on their reactions to the questionnaire—that is, whether any concepts were missing. Several parents were concerned that children wouldn't understand what "mental health" means, although in empirical analyses this concern was not supported. A couple of parents were interested in seeing a family life question added (e.g., "How often do you have fun with your family?"). Parents also mentioned that asking about sexual health and risky behavior (e.g., getting tattoos or piercings) might be useful for older children. Although both groups identified a limited number of concepts beyond those included in the PGH-7, they affirmed the meaningfulness of those topics included.

We assessed the comprehensibility of the items in cognitive interviews and literacy analyses, as described here.

Content Validity/Cognitive Interviews

The primary purpose of cognitive interviews is to test the comprehension and meaning of items, such that the meaning of each question, as written by the item developer, is consistent with the children's interpretations. Using the cognitive interview procedures of cognitive debriefing, verbal probing, concurrent probing, and retrospective probing, we aimed to solicit important information about how participants interpret questions.

Child interviews lasted approximately 45 minutes. Children were provided a paper questionnaire upon which to record their answers prior to the interview. They were asked to read each item aloud and verbalize their thoughts on the meaning of the questions and their rationale for the selected response. Open-ended probes were used to elicit further explanation of perceptions and experiences. Interviews were audio-taped and scored by interviewers for comprehension. Transcriptions were summarized to identify the concepts/ interpretations being elicited, reading and comprehension difficulties for items, stems, and response options.

Results

See Table 14 in the Supporting Documents for a summary of cognitive interview results.

Based on these interviews, all items were understood by the majority of children. None of the items in the PGH-7 performed poorly on a consistent basis. The items for which greater variation was observed in terms of understanding were those originating from the PROMIS Adult General Health Measure. However, we believe that some of these issues may be related to differences according to age levels. In Table 14 (see Supporting Documents), we outline general themes at the item level, as well as hypotheses regarding the psychometric properties of these items that we tested in the quantitative analyses.

Content Validity / Reading Level Analyses

We sought to create items at the 4th grade reading level. Reading level analysis is based on the Flesch Kincaid Readability Test, which computes a reading level score based on total words, sentences, and syllables. Thus, the measure is written at approximately a 3rd to 4th grade reading level (see Table 15 in the Supporting Documents).

Concurrent Validity

We have conducted a set of analyses that correlated the PGH-7 with other measures that we expected to be strongly related.

To examine convergent and discriminant validity of the PGH-7, we evaluated associations with 15 PROMIS pediatric measures of physical, mental, and social health. We also evaluated correlations with two legacy measures of children's health-related quality of life: KIDSCREEN-10, which assesses positive health, and PedsQL, which assesses problems with a child's health.

Study sample 1 was concurrently validated with the following PROMIS pediatric measures: family involvement, family belonging, life satisfaction, positive affect, meaning and purpose, psychological stress experiences, and physical stress experiences. Study sample 2 was concurrently validated with questionnaires KIDSCREEN-10 and PedsQL, as well as the following PROMIS pediatric measures: mobility, upper extremity physical function, pain interference, fatigue, anxiety, anger, depressive symptoms, and peer relationships.

H 1: Measures of psychological symptoms (PROMIS Anxiety, Anger, Depressive Symptoms, and Psychological Stress; PedsQL Psychosocial Health and Emotional Function) will be negatively associated with the PGH-7.
H 2: Measures of children's positive psychological experiences (PROMIS Positive Affect, Life Satisfaction, Meaning and Purpose) and relationships (Peer Relationships, Family Belonging, and Family Involvement) will positively correlate with the PGH-7.
H 3: Scores on legacy instruments will correlate most strongly with scores on conceptually similar child and parent-proxy PGH-7 items.

Conclusion

These findings indicate that global health increases in the same direction as children's sense of belonging and connectedness with their family and their subjective well-being, while moving in the opposite direction of children's stress experiences. The PGH-7 is strongly correlated with the KIDSCREEN-10, which assesses positive health, and moderately correlated with the Peds QL, which assesses problems with a child's health (Tables 16 and 17, see Supporting Documents).

Construct Validity

We have tested differences in global health that we expected to vary, a priori, by sociodemographic and health status characteristics. These known-group validation results are presented here. Because the PGH-7 is scored on a scale with a mean of 0 and standard deviation of 1, beta coefficients represent differences between groups in standard deviation units. Only children whose parents also participated in the surveys were included in the analyses.

Race/Ethnicity

H1: Black children have lower global health than white children.
H2: Hispanic children have lower global health than non-Hispanic children.

For race determination, participants were asked "What is your ethnicity?" They chose between two options: "Not Hispanic/Latino" or "Hispanic/Latino."

Conclusion

There are no substantive differences by race. However, our results show that children with Hispanic ethnicity have lower global health than their non-Hispanic counterparts (Table 18, see Supporting Documents).

Chronic Conditions/Special Health Care Needs

H1: Children with a chronic condition or special health care needs will have lower global health than children without them.

Participants were asked "In the past 6 months, has the child had a chronic condition? (A chronic condition is a physical or mental condition that has lasted or is expected to last at least 12 months, AND interferes with the child's activities.)" Response options were "No" or "Yes." If "Yes" was selected, the respondent was given the option to specify the chronic condition. A physician on the project team reviewed all open-ended responses to ensure that the response was in fact a chronic disease. Parents were also administered the Children with Special Healthcare Needs Screener, a measure of chronic health problems that require health services or cause functional limitations. Chronic conditions were present in 24 percent of study sample 1, while special health care needs were present in 26 percent of study sample 2. Results are shown in Table 19 and Table 20 (see Supporting Documents).

Conclusion

The findings support our hypothesis that children with chronic conditions or special health care needs will have lower global health. Children with chronic conditions had PGH-7 scores that were from 0.53 to 1.30 standard deviation units lower than children without them. Children with special health care needs had PGH-7 scores that were 0.78 standard deviation units lower than their counterparts without special health care needs.

Family Income

H1: Children with low family income will have lower global health than those with higher family income.

Participants were asked to "Please check your child's total family income before taxes for last year." Categories were: "Less than $10,000," "$10,000 to $14,999," "$15,000 to $19,999," "$20,000 to $29,999," "$30,000 to $39,999," "$40,000 to $79,999," "$80,000 to $119,999," or "$120,000 or more." We selected $40,000 as a logical income cut (30 percent of study sample 1 and 41 percent of study sample 2 reported an income below $40,000).

Conclusion

The findings support our hypothesis that children with low household income will have lower global health than those with higher household income. Low income children in our sample had PGH-7 scores that were 0.21 standard deviation units lower than children with higher family income (Table 21, see Supporting Documents).

Geographic Residence

H1: Global health will not be affected by geographic residence.

Rurality and urbanicity were determined through the participant's 3-digit zip codes and matched to the Department of Agriculture Rural-Urban Continuum Codes. Rural participants accounted for 15 percent of the sample. We also stratified the urban/suburban and rural categories by presence of a chronic condition.

Conclusion

The findings support our hypothesis that geographic residence will have no effect on global health (Table 22, see Supporting Documents).

Differential Item Functioning

In the context of IRT, differential item functioning (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 global health level (theta) and a putative DIF characteristic (e.g., gender). We tested the following characteristics for DIF across each of the seven global health questions: age, gender, receiving help from a parent to complete questionnaire, and presence of a chronic condition.

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.

Construct Validity in a Clinical Population

We conducted analyses in a clinical population of children presenting to the ED with acute asthma exacerbations in order to evaluate the predictive validity of the instrument as well as the responsiveness of the measure to clinical change. The association was examined between scores on the PGH-7 and other quality of life instruments and conventional clinical markers of asthma activity.

Sample Characteristics

Out of 328 parents of children age 5-7 years recruited at baseline, 178 completed all follow-up surveys. Out of 182 children age 8-17 years recruited at baseline, 74 completed all follow-up surveys. For parent/child pairs, 173 dyads were recruited at baseline, and 74 dyads completed all follow-up surveys. For any analyses involving baseline data only, all baseline cases were used. Whenever 2-4 week follow-up data were analyzed, all data at follow up 1 were used. For any analyses including 6-12 week follow-up data, all follow-up 2 data were used (Figure 4; see Supporting Documents).

Characteristics of the baseline, follow-up 1, and follow-up 2 samples are summarized in Table 23 (see Supporting Documents). The sample was mostly made for both child report and parent report during all stages of data collection (55.4-63.1 percent). Patients were mostly black/African American (75-84.3 percent) and non-Hispanic (90-94 percent). For parent report at all three data

collection stages, the largest age group represented comprised children ages 5-7 years (47.3-58.5 percent); for child report, children ages 8-12 years were most represented (60.8-69.2 percent).

The mean follow-up time in days between the patient's baseline visit and follow-up 1 was 20.5 and 21.7 days for parent and child, respectively. Between the patient's baseline visit and follow-up 2, the mean number of days was 57.1 and 60.1 days for parent and child, respectively (Table 23; see Supporting Documents).

Results

In the ED sample, 61-68 percent of children improved, while just 5-15 percent reported worsening in their asthma symptoms after 8 weeks. Lower percentages of children reported improvements in global health: 34 percent improved, 22 percent worsened, and 45 percent remained the same as measured by PGH-7 scores. The PGH-7 was sensitive to change in asthma symptoms. It increased over time among those whose asthma symptoms improved and decreased among those whose asthma symptoms worsened. Moreover, it was responsive to a change in asthma symptoms: the difference in the PGH-7 among children who improved versus those who stayed the same/worsened was approximately 0.7 standard deviation units higher. These findings provide support for the clinical validity of the PGH-7, demonstrating that it is responsive to changes that result from health care (see Technical Specifications [Tables 2 and 3]; also see Tables 2-11 and 24-36; Figures 3-7, in the Supporting Documents).

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

See Section 6, Construct Validity.

7.B. Special Health Care Needs

See Section 6, Construct Validity.

7.C. Socioeconomic Status

See Section 6, Construct Validity.

7.D. Rurality/Urbanicity

See Section 6, Construct Validity.

7.E. Limited English Proficiency (LEP) Populations

See Section 6, Construct Validity.

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 that has not been used. The PGH-7 is a survey tool, and data for it are not available in any existing data collection system.

Pediatric General Health Measure data are acquired through self-report and parent-proxy report. Person-reported outcome (PRO) measures, like the PGH-7, 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.

As a seven-item scale, the brevity of the PGH-7 makes any of these administration methods feasible. However, within the context of performance measurement activities at the State, health plan, or other level of the health care delivery system, implementation of these methods must be considered in terms of their sustainability over time.

We engaged Medicaid and CHIP leaders from the States of New Jersey, Pennsylvania, and Massachusetts to gather feedback on alternative approaches for obtaining global health assessments. State leaders acknowledged that scarcity in resources poses barriers to any data collection effort. However, they generally felt that PGH-7 administration is feasible and desirable given multiple avenues for survey administration/data collection.

The following represent some potential PGH-7 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 health care and to provide feedback at the health plan, provider group, and individual provider levels. The family of CAHPS surveys comprises tools that are widely used

throughout the health care industry. It would be straightforward to add the PGH-7 to an existing survey.

2. **Independent population-based survey.** An alternative to integrating questions directly into another survey is to administer the PGH-7 separately, potentially even parallel, and based on the same sampling procedures. As an independent survey, the PGH-7 could be Internet-based or administered via telephone or paper and pencil. It could be included in a new PQMP survey module that would assess multiple quality indicators. Health plans could integrate the PGH-7 as part of the new beneficiary on-boarding process, whereby new members may participate in health assessment surveys. States or plans could also focus specifically on the chronically ill population, administering the tool via case management systems that are already in place.

3. **Patient portal integrated with an electronic health record (EHR).** The use of patient portals as a mechanism for engaging patients and collecting important information will likely increase. As States and plans continue to evolve in terms of their health information technology 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 (clinic, 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, which 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 the early stages of adoption. As described in Section 3.B, Evidence for Importance of the Measure to Medicaid and/ or CHIP, more than one State voiced the willingness to collaborate in demonstration work that would not only help build empirical evidence of the responsiveness and sensitivity of the PGH-7 to events (e.g., disenrollment) or interventions, but could serve as a proof of concept for implementation.

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?

Not applicable.

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.

Note: the PGH-7 is a survey tool. To complete the self-report PGH-7, children must be 8-17 years of age and without any cognitive limitations or developmental delays that prevent them from responding to a questionnaire. Parents may report on any child age 5-17 regardless of that child's literacy. The age range can be extended to 21 for both editions. See Section 6, Scientific Soundness of the Measure, for comprehensive testing results on the eligible population.

At a power of 0.80 a critical value of 0.05, 100 subjects per group are needed to detect a difference of 4 points (i.e., a clinically meaningful difference). Table 24 (see Supporting Documents) summarizes the total sample sizes required to detect small (.10), medium (.25), and large (.40) differences among two through eight groups, assuming power = .80 and alpha = .05.

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; public and private coverage in States that have sufficient sample sizes.

Data Sources: **Are data sources available to support reporting at this level?**
Not applicable.

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 applicable.

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?**
Not applicable.

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)**
No.

Data Sources: **Are data sources available to support reporting at this level?**
Not applicable.

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 applicable.

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?**
Not applicable.

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?**
Not applicable.

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 applicable.

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?**
Not applicable.

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?**
Not applicable.

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 applicable.

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?**
Not applicable.

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?**
Not applicable.

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 applicable.

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?**
Not applicable.

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?**
Not applicable.

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 applilcable.

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?**
Not applicable.

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?**
Not applicable.

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?**
No.

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?**
Not applicable.

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 health care quality measurement, and they have expressed significant interest in identifying meaningful outcome measures that assess children's well-being and functioning. We engaged leadership from New Jersey, Pennsylvania, and Massachusetts Medicaid/CHIP programs to gather stakeholder feedback on the PGH-7.

The PGH-7 was viewed as an efficient (low respondent burden) tool for measuring person-reported global health from the perspectives of either children or parents. Stakeholders

emphasized that because pediatric populations eventually reach age 21, an instrument that accounts for developmental changes in physical, mental, and social health is necessary. One of the greatest strengths of the PGH-7 is that it is a highly inclusive measure that allows State Medicaid and CHIP programs to assess the majority of their pediatric populations.

The greatest challenge for States is in being able to attribute changes in global health, both at the individual and population levels, to specific points/processes in the delivery system. To interpret changes in global health, a program may want to give the measure along with process measures and associate change in one with change in the other.

We believe there is a strong conceptual basis for improvements in global health resulting from any QI activity that reduces symptom burden or improves children's participation in desired activities, and that, as outlined in Section 3 of this report, there is existing evidence that global health measures are indeed sensitive to health care interventions. Despite any concerns they might have, States voiced an interest and willingness to collaborate in demonstration work that would help build empirical evidence of the responsiveness and sensitivity of the PGH-7.

One State referred to the PGH-7 as a potential means of "triaging" children at the individual level, whereby endorsement of particular items could serve as a flag or indicator for further investigation. At the population level, plans could allocate resources to areas identified as higher need. In our analyses, we observed variations in levels of global health across States (see Figure 8 in the Supporting Documents). Similar analyses can be conducted at any geographic unit or unit of the health care delivery system. State leaders were interested in further discussing the relationship of global health with other aspects of health and perhaps most importantly implications for health care service delivery.

One of our State partners further sought the input of seven of its contracting Medicaid managed care plans. The MCOs provided comments about the usability and feasibility of the measure, with only one expressing disinterest in the tool. Several plans expressed support for the PGH-7 as a potential CHIPRA core measure, and others acknowledged its usefulness even as an autonomous tool. Enthusiasm for the PGH-7 stemmed from its potential efficacy in identifying at-risk populations, providing information that would allow plans to implement outreach programs and conduct existing program evaluations. Plans shared a general sense that the PGH-7 could potentially be a useful, feasible measurement device, and that they would be interested in exploring the measure further.

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 PGH-7 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 brevity of the measure further increases the likelihood of adoption 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?

The PGH-7 has not been tested as part of an EHR system.

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, at this time, be captured as part of 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 criteria (see healthit.hhs.gov/portal/server.pt/community/healthit_hhs_gov__standards_ifr/1195)?

Although the PGH-7 does not use data elements supported by the ONC, it is likely that PROs will soon become part of these standards.

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).

1. **Does not measure specific needs:** The PGH-7 provides a summary measure of outcome. A PGH-7 score does not indicate the need for a specific service. It is best used as an assessment of the end results of a quality improvement activity and will not guide the design of specific (versus global) interventions.

2. **Measures health at a high level.** The measure evaluates global health, which is the highest level of self-reported health possible. In some applications, a user may prefer to evaluate physical, mental, or social dimensions more specifically.

3. **More work required to assess responsiveness.** The measure requires longitudinal evaluation to assess its responsiveness to change.

4. **Age limitations.** The measure does not address the health of children ages 0-4 years.

5. **National norms:** We are working to develop national norms, so that a score of 50 is set as the U.S. average based on the 2010 census.

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 Global Health-7 (PGH-7) is a seven-question outcome measure that assesses a child's overall health. The questions can be answered by children ages 8-17 years or the parents of children 5-17 years old. Questions ask about children's perceptions of their health in general, physical health, mental health, pain, friendships, family life, self-esteem, and feelings of worry and sadness.

The PGH-7 was developed according to the scientific standards of the Patient Reported Outcome Measurement Information System (PROMIS—www.nihpromis.org). PROMIS has developed over 60 measures of patient-reported health, including an Adult Global Health Measure. The PGH-7 is conceptually comparable with the adult version. Development and validation was conducted with over 7,000 children and parents located in all 50 States. The measure has excellent reliability, which indicates that it yields very precise measurements of a child's global health. Likewise, the validity is excellent, which suggests that indeed it evaluates a child's overall health.

The PGH-7 is an outcome measure. That is, it evaluates the results of health care, rather than the specific services that were delivered. As a general health measure for children 5-17 years old, it will be useful in a wide variety of quality applications, from evaluations of care for specific diseases to assessments of changes in delivery system models.

Potential end-users indicated that the PGH-7 (and other PRO measures) can be integrated into existing data collection platforms, such as EHRs, ongoing surveys, and registries, or included in a new PQMP survey module. The instrument may be administered in a patient/provider friendly manner (i.e., as an intake form or at the point of health care program initiation or as part of existing consumer survey administration) and subsequently coded into an EHR.

State Medicaid/CHIP leaders and managed care plans see the potential value of the PGH-7 and are interested in exploring further its use as a tool for QI. They identified the ability of the PGH-7 to assess the majority of its pediatric population to be one of its greatest strengths.

References

Blackwell DL, Tonthat L. Summary health statistics for U.S. Children: National Health Interview Survey, 1998. Vital Health Stat 2002; 10(208):1-53.

Brunner HI, Taylor J, Britto MT, et al. Differences in disease outcomes between Medicaid and privately insured children: Possible health disparities in juvenile rheumatoid arthritis. Arthritis Care Res 2006; 55(3):378-84.

Busch SH, Vigdor ER. Are adults in poor health more likely to enroll in public insurance? Inquiry 2008; 45(4):380-94.

Dey AN, Bloom B. Summary health statistics for U.S. children: National Health Interview Survey, 2003. Vital Health Stat 2005; 10(223):1-78.

Forrest CB, Riley AW, Vivier PM, et al. Predictors of children's healthcare use: The value of child versus parental perspectives on healthcare needs. Med Care 2004; 42(3):223-8.

Forrest CB, Tucker CA, Ravens-Sieberer U, et al. Concurrent validity of the PROMIS® pediatric global health measure. Qual Life Res 2016; 25(3):739-51.

Hays RD, Bjorner JB, Revicki DA, et al. Development of physical and mental health summary scores for the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009; 18(7):873-80.

Idler EL, Benyamini Y. Self-rated health and mortality: A review of twenty-seven community studies. J Health Social Behav 1997; 38(1):21-37.

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

Larson K, Halfon N. Family income gradients in the health and health care access of US children. Matern Child Health J 2010; 14(3):332-42.

Levy JI, Brugge D, Peters JL, et al. A community-based participatory research study of multifaceted in-home environmental interventions for pediatric asthmatics in public housing. Soc Sci Med 2006; 63(8):2191-203.

Newacheck PW, Hung YY, Park MJ, et al. Disparities in adolescent health and health care: Does socioeconomic status matter? Health Serv Res 2003; 38(5):1235-52.

Seid M, Varni JW, Segall D, et al. Health-related quality of life as a predictor of pediatric healthcare-costs: A two-year prospective cohort analysis. Health Qual Life Outcomes 2004; 2(48):epub.

Seid M, Varni JW, Cummings L, et al. The impact of realized access to care on health-related quality of life: a two-year prospective cohort study of children in the California State Children's Health Insurance Program. J Pediatr 2006; 149(3):354-61.

Wade TJ, Mansour ME, Line K, et al. Improvements in health-related quality of life among school-based health center users in elementary and middle school. Ambul Pediatr 2008; 8(4):241-9.

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: forrestc@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.