Utilization of the PHQ-9 to Monitor Depression Symptoms for Adolescents and Adults

Section 1. Basic Measure Information

1.A. Measure Name

Utilization of the PHQ-9 to Monitor Depression Symptoms for Adolescents and Adults

1.B. Measure Number

0245

1.C. Measure Description

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

This measure, Utilization of the PHQ-9 to Monitor Depression Symptoms for Adolescents and Adults, assesses the percentage of patients 12 years of age and older with a diagnosis of major depression or dysthymia, who had an outpatient encounter with a PHQ-9 score present in their record in the same assessment period as the encounter.

1.D. Measure Owner

The measure owner is the National Committee for Quality Assurance (NCQA). The measure was developed through the National Collaborative for Innovation in Quality Measurement (NCINQ).

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

This measure was adapted from NQF #0712, with permission from Minnesota Community Measurement (MNCM).

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.**

   Not applicable.

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.**

   HEDIS® Depression measures. See more information here: https://www.ncqa.org/hedis/the-future-of-hedis/hedis-depression-measures-specified-for-electronic-clinical-data/

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

To assess if patients with depression are receiving routine monitoring using the PHQ-9 tool, the measurement year is divided into three separate assessment periods, each with a numerator:

- Numerator 1 includes documented results of a PHQ-9 score in the patient's record during assessment period one: first 4-month period of the measurement year (January 1 through April 30).
- Numerator 2 includes documented results of a PHQ-9 score in the patient's record during assessment period two: second 4-month period of the measurement year (May 1 through August 31).
- Numerator 3 includes documented results of a PHQ-9 score in the patient's record during assessment period 3: third 4-month period of the measurement year (September 1 through December 31).

1.H. Numerator Exclusions

Not applicable.

1.I. Denominator Statement

To assess if patients with depression are receiving routine monitoring using the PHQ-9 tool, the measurement year is divided into three separate assessment periods, each with a denominator:

- Denominator 1 includes patients 12 to 17 years of age with at least one interactive outpatient encounter with a diagnosis of major depression or dysthymia during assessment period one: first 4-month period of the measurement year (January 1 through April 30).
- Denominator 2 includes patients 12 to 17 years of age with at least one interactive outpatient encounter with a diagnosis of major depression or dysthymia during assessment period two: second 4-month period of the measurement year (May 1 through August 31).
- Denominator 3 includes patients 12 to 17 years of age with at least one interactive outpatient encounter with a diagnosis of major depression or dysthymia during assessment period three: third 4-month period of the measurement year (September 1 through December 31).

1.J. Denominator Exclusions

The measure excludes patients with any of the following at any time during the measurement year:

- Bipolar Disorder
- Personality Disorder
- Psychotic Disorder
- Autism Spectrum Disorder
- Hospice

1.K. Data Sources

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

Administrative data (e.g., claims data); electronic health records (EHRs); health information exchange (HIE)/clinical registry; case management.

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

Not applicable.

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).**

The purpose of this measure is to improve monitoring of symptoms and treatment progress for adolescents and adults with major depressive disorder or dysthymia through use of a standardized tool. Major depressive disorder (MDD) is a disabling condition that is associated with many long-term complications and may lead to suicide (Williams, O'Connor, Eder, et al., 2009). Dysthymia is a mild but long-term (chronic) form of depression, and it often responds to the same treatments provided for major depression.

Data from the 2010–2011 National Survey on Drug Use and Health and the National Health and Nutrition Examination Survey indicate that the prevalence of depression among adolescents aged 12 to 17 was 12.8 percent over the lifetime and 8.1 percent over the past year (Substance Abuse and Mental Health Services Administration [SAMHSA], 2014).

Adolescent-onset depression increases the risk of attempted suicide by five-fold in comparison to non-depressed adolescents (Williams, et al., 2009). Suicide is the third leading cause of death among 15-to 24-year-olds, and most adolescents who took their own lives had a previous history of depression (U.S. Preventive Services Task Force [USPSTF], 2009; Williams, et al., 2009). Depressive symptoms can be both prolonged and episodic, recurring over weeks and months (National Research Council [NRC] and Institute of Medicine [IOM], 2009). The Centers for Disease Control and Prevention (CDC) noted that if an individual experiences just one episode of depression, he or she has a 50 percent higher risk of experiencing further episodes (CDC, 2011). If such symptoms and/or episodes persist in a manner that significantly impacts day-to-day life, a single episode of depression can develop into MDD (NRC and IOM, 2009).

An important element of depression care is monitoring symptoms and receiving feedback about treatment response (Lewandowski, Acri, Hoagwood, et al., 2013). In a recent study comparing the effects of a collaborative care model (intervention) versus usual care (control) among adolescents with depression in primary care settings, results indicated that the intervention group showed higher rates of response and remission compared to those in the control group. The collaborative care model included monitoring of symptoms and adjusting care based on PHQ-9 results, and usual care was considered depression screening only, with adolescents seeking care on their own. Overall, 86 percent of patients in the intervention group received either psychotherapy or medications that met study quality standards, compared with 27 percent of the control group. Intervention adolescents were significantly more likely than control adolescents to receive four or more psychotherapy sessions in the first 6 months of the study. The overall rate of depression remission at 12 months was 50.4 percent for the intervention group, compared with 20.7 percent for the control group (Richardson, Ludman, McCauley, et al., 2014). The Guideline for Adolescent Depression in Primary Care (GLAD-PC) recommends systematic and regular tracking of treatment goals and outcomes, including assessment of depressive symptoms and functioning, monitoring for adverse events during antidepressant treatment, and reassessment of diagnosis and treatment if no improvement is noted (Cheung, Zuckerbrot, Jensen, et al., 2007). The National Institute for Health and Clinical Excellence (NICE) recommends that structured questionnaires be used in addition to clinical judgment in symptom monitoring (NICE, 2005).

Prevalence

Depressive disorders are common mental disorders that occur in people of all ages. Major depressive disorder (MDD) is the second leading cause of disability worldwide, affecting an estimated 120 million people (Murray, Vos, Lozano, et al., 2013). The lifelong prevalence is estimated to range from 10-15 percent (Lépine, Briley, 2011). In the United States, 15.7 percent of people report that at some point in their lifetime they were told by a healthcare professional that they had depression (CDC, 2009).

Youth and Adolescents

A nationally representative survey by SAMHSA in 2015 found that 11.4 percent of adolescents (12-17 years of age) had at least one major depressive episode in 2014, and 8.2 percent had an episode with severe impairment (SAMHSA, 2015). The same survey found that only 41.2 percent of those who had a major depressive episode received treatment in the past year. Prevalence of depression among adolescents and young adults in the United States increased between 2005 and 2014, with little change observed in mental health treatments, ultimately leading to a growing number of youths with untreated or undertreated depression (Mojtabai, Olfson, Han, 2016).

Lifetime prevalence of depression and dysthymia increases from 8.4 percent, for ages 13-14, to 15 percent for ages 17-18 (Merikangas, He, Burstein, et al., 2010). Female adolescents are more likely to be diagnosed with depression than males (NRC and IOM, 2009). One study found that female adolescents are also more likely than males to experience recurrence (57.6 percent vs. 32.9 percent, respectively) (Curry, Silva, Rohde, et al., 2011). Depression during adolescence has a strong correlation to chronic and recurring depression in adulthood (Garber, Clarke, Weersing, et al., 2009).

Health Importance

Depression—an overwhelming feeling of sadness and hopelessness that can last for months or years—can make people feel that life is no longer worth living. People affected by depression lose interest in activities they used to enjoy, and they also can be affected by physical symptoms that interfere with their ability to participate in normal daily activities. For adolescents, depression can also have a major impact, disrupting daily life at home, at school, or in the community.

Depression can complicate and exacerbate other chronic medical conditions and lead to increased morbidity and mortality. The mortality risk for suicide in depressed patients is more than 20-fold greater than in the general population (Bostwick, Pankratz, 2000). In terms of other chronic conditions, depression is associated with a 60 percent increased risk of type 2 diabetes (Mezuk, Eaton, Albrecht, et al., 2008), and it has been identified as a risk factor for development of cardiovascular disease (Van der Kooy, van Hout, Marwijk, et al., 2007). In addition, depression adversely affects the course, complications, and management of other chronic medical illnesses (Katon, 2011). In adolescents, depression can result in serious long-term morbidities, such as generalized anxiety disorder and panic disorder, or lead to engagement in risky behaviors such as substance use (Foley, Carlton, Howell, 1996; Friedman, Katz-Levey, Manderscheid, et al., 1996; NRC and IOM, 2009; Taylor, Chadwick, Heptinstall, et al., 1996). Adolescent-onset depression increases the risk of attempted suicide five-fold in comparison with non-depressed adolescents (Garber, Clarke, Weersing, et al., 2009). Most adolescents who commit suicide have a history of depression (NRC and IOM, 2009; Williams, et al., 2009).

Depression has long been recognized as a major contributor to disease burden (Murray, Lopez, 1997; Üstün, Ayuso-Mateos, Chatterji, et al., 2004). The Global Burden of Disease Study of 2010 identified depression as a leading cause of disease burden in the world. Depressive disorders were the second largest contributor to years lived with disability, an indicator of the impact of disease burden (Ferrari, Charlson, Norman, et al., 2013). This accounts for an

estimated 10 percent of years lived with disability worldwide, which is triple the impact of diabetes, eight times the impact of heart disease, and 40 times the impact of cancer (Murray, et al., 2013). These findings underscore the need for attention to depressive disorders and to the implementation of effective interventions to reduce their disease burden.

Financial Importance and Cost-Effectiveness

Depression has a large effect on healthcare costs and on productivity. Adolescents with MDD have higher medical expenditures, including those related to general and mental healthcare, than children without an MDD diagnosis (USPSTF, 2009).

Even minor levels of depression symptoms are associated with decreases in work function (Beck, Crain, Solberg, et al., 2011). In a survey study, Birnbaum and colleagues found that major depressive disorder severity is significantly associated with increased treatment usage and costs, unemployment, disability and reduced work performance. When the results of the study were projected to the U.S. workforce, it was estimated that monthly depression-related worker productivity losses had human capital costs of nearly $2 billion (Birnbaum, Kessler, Kelley, et al., 2010).

The Patient Protection and Affordable Care Act (ACA) includes a provision that essential health benefits, which cover behavioral health treatment and prevention among others, are required to be provided without any cost-sharing by the patients. Specifically, prevention services that "are recommended with a grade of A or B by the United States Preventive Services Task Force for any indication or population, the amount paid shall be 100 percent of the lesser of the actual charge for the services or the amount determined under the fee schedule that applies to such services under this part'' (U.S. Congress, 2010). This enables individuals to access the essential healthcare they need without increased cost to themselves.

Gaps in Care

In a recent study comparing the effects of a collaborative care model (intervention) versus usual care (control) among adolescents with depression in primary care settings, results indicated that the intervention group showed higher rates of response and remission compared to those in the control group. The collaborative care model included monitoring of symptoms and adjusting care based on PHQ-9 results, and usual care was considered depression screening only, and adolescents sought care on their own. The overall rate of depression remission at 12 months was 50.4 percent for the intervention group compared with 20.7 percent for the control group. Overall, 86 percent of patients in the intervention group received either psychotherapy or medications that met study quality standards, compared with 27 percent of the control group. Intervention adolescents were significantly more likely than control adolescents to receive four or more psychotherapy sessions in the first 6 months of the study (Williams, et al., 2009).

In a survey asking primary care pediatricians about their roles and perceived responsibilities for depression care, pediatricians cited several factors as impeding their ability to diagnose mental health problems, including lack of time during the visit to provide mental health counseling or collect a patient history and lack of knowledge about the causes, signs, symptoms, and management of pediatric mental illness (SAMHSA, 2014; Williams, et al., 2009). Patient and

parent or caregiver knowledge and cooperation due to stigma or comfort can also contribute to the lack of early detection in primary care. In particular, adolescent patients might have issues with discussing the topic in person, with research noting that adolescents prefer a non-human interface to reveal personal information (USPSTF, 2009).

A survey of pediatricians found that only 25 percent believed it was their responsibility to treat depression in adolescents. Those surveyed also had concerns about treatment options, with 86 percent of those surveyed expressing concern with prescribing medications, and 90 percent expressing concern with counseling (Williams, et al., 2009). Some estimates suggest that only 25 percent of adolescents diagnosed with depression receive treatment; among those who go undetected, 20 percent develop recurrent or chronic depression (Foley, et al., 1996; Taylor, et al., 1996).

Health Disparities

Studies suggest gender disparities exist. Female adolescents are more likely to be diagnosed with depression than males (Friedman, et al., 1996; NRC and IOM, 2009). One study also found that female adolescents are also more likely to experience recurrence than males (57.6 versus 32.9 percent, respectively) (NRC and IOM, 2009).

Disparities in care also exist for minority racial/ethnic groups. Algeria and colleagues discovered that among people with a diagnosed depressive disorder, 63.7 percent of Latinos and 58.8 percent of African Americans did not access any mental health treatment in the past year, compared with 40.2 percent of non-Latino whites (Algeria, Chatterji, Wells, et al., 2008). Hispanic and uninsured children have especially high rates of unmet need for mental health services relative to other children (Kataoka, Zhang, Wells, 2002). Additionally, minority adolescents may present depressive symptoms differently from non-Latino whites, which can be challenging for providers who are trained to only recognize certain symptoms when screening for depression (Algeria, et al., 2008).

Minority children and adolescents are 50 to 60 percent less likely than their white counterparts to receive mental healthcare services, despite a similar overall prevalence of disease. Hispanic youth are the least likely to receive treatment, and a smaller but similar disparity has been found for Asian/Pacific Islander and African American youth. Moreover, of those who do receive care, these minority groups are less likely to complete services and are more likely to receive treatment that is inappropriate, fragmented, or inadequate (Child Mind Institute, 2011).

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).**

This measure addresses monitoring for adolescents identified with depression, which is relevant to receiving treatment for depression, covered under the EPSDT benefit.

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 Utilization of the PHQ-9 to Monitor Depression Symptoms for Adolescents and Adults measure is part of the HEDIS® Depression Measures set. The measure was adapted from the Minnesota Community Measurement (MNCM) adult depression monitoring measure, which has been endorsed by the National Quality Forum [#0712]. By including the adolescent population, this measure complements the adult-focused measure. The measure is also specified at the health plan level, complementing MNCM's provider level measure.

This process measure assesses the routine monitoring of symptoms among those with depression. Routinely monitoring symptoms using the PHQ-9 supports the ability to assess patient outcomes (i.e., if patients with depression are improving over time).

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:** No.
c. **Care Setting – other – please specify:** Not applicable.
d. **Service – preventive health, including services to promote healthy birth:** No.
e. **Service – care for acute conditions:** No.
f. **Service – care for children with special health care needs/chronic conditions:** Yes.
g. **Service – other (please specify):** Not applicable.
h. **Measure Topic – duration of enrollment:** No.
i. **Measure Topic – clinical quality:** Yes.

j. **Measure Topic – patient safety:** No.
k. **Measure Topic – family experience with care:** No.
l. **Measure Topic – care in the most integrated setting:** No.
m. **Measure Topic other (please specify):** Not applicable.
n. **Population – pregnant women:** Yes.
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):** No.
r. **Population – school-aged children (6 years through 10 years) (specify age range):** No.
s. **Population – adolescents (11 years through 20 years) (specify age range):** Yes; 12 years and older.
t. **Population – other (specify age range):** Not applicable.
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.

The use of standardized tools is essential for tracking depressive symptoms and monitoring patient response to treatment. Standardized instruments are useful in identifying meaningful change in clinical outcomes over time. Guidelines recommend that providers establish and maintain regular follow-up with patients diagnosed with depression and use a standardized tool to track symptoms (Mitchell, Trangle, Degnan, et al., 2013).

Meta-analyses of studies in adults indicate that formally monitoring patient progress improves patient outcomes (Knaup, Koesters, Schoefer, et al., 2009; Lambert, Whipple, Hawkings, et al., 2003; Shimokawa, Lambert, Smart, 2010). For adolescents, the Guideline for Adolescent Depression in Primary Care (GLAD-PC) recommends systematic and regular tracking of treatment goals and outcomes, including assessing depressive symptoms and function, monitoring for adverse events during antidepressant treatment, and reassessing diagnosis and treatment if no improvement is noted after 6-8 weeks. One study found that youths with a range of symptoms improve more quickly when clinicians receive feedback from assessments every other week instead of every 3 months (Bickman, Kelley, Breda, et al., 2011).

Existing "gold standard" instruments, such as the Hamilton Depression Rating Scale, can be time-consuming and require a specially trained interviewer. The brief PHQ-9 questionnaire can be self-administered by the patient and has been validated for measuring depression severity and treatment response (Kroenke, Spitzer, Williams, 2001). The PHQ-9 has also been validated for use in psychiatric settings as a severity measure and a measure of treatment outcome (Beard, Hsu, Rifkin, et al., 2016). The PHQ-9 tool assesses the nine DSM, Fourth Edition, Text Revision (DSM-IV-TR) criterion symptoms and effects on functioning and has been shown to be highly accurate in discriminating patients with persistent major depression, partial remission, and full remission (Gilbody, Richards, Brealey, et al., 2007; Lowe, Kroenke, Herzog, et al., 2004; Martin, Rief, Klaiberg, et al., 2006).

The PHQ-9 Modified for Teens is the PHQ-9 tool with slight wording adjustment in three questions in order to tailor the tool for the adolescent population with age-appropriate terms. The copyright statement on the PHQ-9M tool is stated: "Modified with permission by the GLAD-PC team from the PHQ-9 (Spitzer, Williams, Kroenke, 1999), Revised PHQ-A [Johnson, 2002], and the CDS [Shaffer, Fisher, Lucas, et al., 2000])." Although widely used in pediatric practices and endorsed by the AAP, APA and AACAP, the modified version of the PHQ-9 tool has not had separate validation studies, as the nine questions are essentially the same as the original PHQ-9, which has been validated for the adolescent population (ages 13 and older). The APA recommends using the modified version of the PHQ-9 for children ages 11 to 17 to assess depression symptom severity (APA, 2015).

Benefits of the PHQ-9 tool are numerous: it is non-proprietary and widely accepted by primary care providers and in general medical settings, it can be completed by the patient in-person or over the telephone, it is translated into many languages, and it is easy for the patient to complete and the provider to score. Widespread use of the PHQ-9, within a collaborative care model, would allow organizations to systematically assess their effectiveness in helping individuals reduce depressive symptoms with appropriate treatment.

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.

Measuring and tracking depressive symptoms is important to improve patient outcomes (Knaup, et al., 2009; Lambert, et al., 2003; Shimokawa, et al., 2010). As discussed in the research evidence, the Guideline for Adolescent Depression in Primary Care (GLAD-PC) recommends systematic assessment of symptoms and function, monitoring for adverse events during antidepressant treatment, and reassessing diagnosis and treatment if no improvement is noted after 6-8 weeks. Regular tracking and monitoring of symptoms can help clinicians better understand the effectiveness of current treatment and modify their recommendations according to the prevalence and severity of symptoms.

This measure encompasses adolescents aged 12-17 for whom clinical guidelines also recommend a stepped-care approach to depression treatment, beginning with the least-intrusive intervention and stepping up to more intensive care if the patient does not respond to or benefit from the first intervention (Mitchell, et al., 2013; National Collaborating Centre for Mental Health, 2009; Trangle, Gursky, Haight, et al., 2013). For mild and moderate depression, psychotherapy alone may be the preferred initial treatment, to be followed by the use of medication if symptoms persist (APA, 2010). This stepped-care approach includes providing assessment, support, psychoeducation, and monitoring of symptoms as a first step, followed by psychosocial, psychological, and pharmacologic interventions, and then combined treatments for those with inadequate response.

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.

The field testing of the this measure included an alpha phase, to assess the feasibility and reliability of collecting key data elements for the adolescent population, and beta testing, focused on documenting performance results for adolescents along with the reliability and validity of the measure specifications at the provider level. The measure was also tested at the health plan level for the entire age population, age 12 years and older with a specific stratification for the adolescent population. The reliability and validity results of testing this measure for the adolescent population at the provider level are described here. Please see the Depression Measures Testing Report (see Supporting Documents) for details on the health plan testing results.

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.

Testing was completed at two integrated delivery systems and one network of community health centers that met the following participation criteria: had established clinical workflows for using the PHQ-9 or PHQ-9 Modified for teens, used searchable coded fields for documenting PHQ results in electronic medical records, and had at least 500 adolescents who had a diagnosis of depression in 2012. The sites were from different geographic regions in the United States and served both urban and rural populations.

The testing included 3,394 adolescent patients, who met the following inclusion criteria: (1) age 12 to 17 years as of January 1, 2013 and (2) at least one face-to-face visit for a diagnosis of depression during the measure year (January 1, 2013 through December 31, 2013). Adolescents with bipolar, psychotic, autism spectrum, and personality disorders were excluded. No sampling was used: all patients that met the inclusion criteria were included in the testing.

To test reliability of data elements, we obtained manually abstracted data from the electronic record for a sample of charts. We used parallel-forms reliability testing to evaluate the agreement between the electronic extract and the manual review in identifying the numerator, denominator, and exclusions for the measure. Agreement was measured using the kappa statistic (a measure of agreement adjusted for agreement that can occur by chance). Results for parallel-forms reliability testing were as follows: Agreement for the denominators and numerators (before and after exclusions) was high (Kappa of at least 0.73). The exclusions showed low agreement using the Kappa statistic; however, the lowest percentage agreement was 79 percent for time period 1 for the monitoring measure. Comparing results based on electronic extract to full data from the electronic extract plus manual review, we found that sensitivity was at or near 100 percent for all data elements, showing that the electronic extract captures nearly all the information in the manual review. Specificity was 100 percent for all data elements.

Inter-rater reliability was assessed in order to determine if key measure components could be reliably abstracted by reviewers from an electronic health record (HER). Inter-rater agreement among abstractors was high, with a Kappa of 0.80 or higher for all data elements except for the exclusion data elements. While agreement was 92 percent or higher for all exclusion data elements and “any exclusion,” the overall Kappa for “any exclusion” was 0.38, in part because there were few observations of exclusions.

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).

Face validity refers to whether the measure plausibly represents the concept being evaluated in the judgment of likely users of the measure. Throughout the measure development process, multi-stakeholder panels provided input on the importance, face validity, and usefulness of the measure for use by States in programs such as the Core Set of Children's Health Care Quality Measures for Medicaid and CHIP, as well as in the context of reporting at the health plan level. These panels consisted of representatives from States, health plans, pediatricians, behavioral health clinicians, and consumers. Additionally, we posted the measure for public comment, a 30-day period of review that allowed interested parties to offer feedback about the measure. NCQA panels considered all comments and advised NCQA staff on appropriate recommendations. This process ensures measures are reasonable and important to those using them. Our advisory panels concluded this measure is a valid way to assess routine monitoring of depression symptoms for adolescents and adults.

Results

Step 1: This measure was adapted for the adolescent population from the existing Minnesota Community Measurement measure. NCQA and numerous expert panels worked together in 2013 and 2014 to identify the most appropriate method for assessing depression outcome among the adolescent patient population. Across the multiple expert panels that reviewed the measure, all panels concluded this measure was specified appropriately for adolescents.

Step 2: The measure was field-tested for the adolescent population in 2013 and 2014. It was first posted for a public comment period in October 2014. After reviewing these initial public comment results along with field test results, the health plan version of the measure was posted to the HEDIS® public comment period in February 2015. This measure was rated a high priority by many commenters. NCQA's Committee on Performance Measurement recommended moving this measure to first year data collection by a majority vote in May 2015.

Step 4: The measure was introduced in HEDIS® 2016 (representing measurement year 2015). Organizations voluntarily report this measure each year in June, and the results are analyzed.

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

HEDIS data are stratified by type of insurance (e.g., commercial, Medicaid, Medicare). NCQA does not currently collect performance data stratified by race, ethnicity, or language. Escarce and

colleagues have described in detail the difficulty of collecting valid data on race, ethnicity, and language at the health plan level (Escarce, Carreon, Veselovskiy, et al., 2011). While not specified in the measure, this measure can also be stratified by demographic variables, such as race/ethnicity or socioeconomic status, in order to assess the presence of healthcare disparities. The HEDIS Health Plan Measure Set contains two measures that can assist with stratification to assess healthcare disparities. The Race/Ethnicity Diversity of Membership and the Language Diversity of Membership measures were designed to promote standardized methods for collecting these data, and they follow Office of Management and Budget and Institute of Medicine guidelines for collecting and categorizing race/ethnicity and language data. In addition, NCQA's Multicultural Health Care Distinction Program outlines standards for collecting, storing, and using race/ethnicity and language data to assess healthcare disparities. Based on extensive work by NCQA to understand how to promote culturally and linguistically appropriate services among plans and providers, we have many examples of how health plans have used HEDIS measures to design quality improvement programs to decrease disparities in care.

7.B. Special Health Care Needs

Not available.

7.C. Socioeconomic Status

See section 7.A, above.

7.D. Rurality/Urbanicity

Not available.

7.E. Limited English Proficiency (LEP) Populations

See section 7.A, above.

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?

In reporting of HEDIS® Electronic Clinical Data System (ECDS) measures, health plans report to NCQA which type of data source was used for each key data element in the measure. Four categories of data sources are defined, and a hierarchy is used so data elements are not reported

multiple times across data sources. For more information on the data source categories, see NCQA's website at https://www.ncqa.org/hedis/the-future-of-hedis/hedis-electronic-clinical-data-system-ecds-reporting/. Testing and reporting of this measure has demonstrated that key data elements needed for calculating the measure are available in electronic clinical data systems such as EHRs, case management systems, registries, and administrative claims. HEDIS® ECDS allows all of these data sources to be used to contribute to the measure report.

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.

This measure is currently reported as an ECDS measure as part of NCQA's Healthcare Effectiveness Data and Information Set (HEDIS®). The measure is reported annually by Medicaid and by commercial and Medicare plans.

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?

Data are collected by health plans and reported annually to NCQA through its Interactive Data Submission System.

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

This measure had been reported to HEDIS® for 4 years as of September 2019. NCQA receives feedback and questions related to measure reporting through its Policy Clarification Support System. Since HEDIS® reporting began for this measure, questions received for the measure have allowed NCQA to make refinements to the specifications to ensure they are clear and easily implemented by health plans.

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?**
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?**
By reporting at the State level, the results are more highly aggregated, and that could mask key differences in rates, especially regarding disparities that could be found at lower levels of aggregation.

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?**
Not applicable.

Reliability & Validity: **Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?**
Not applicable.

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)**
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?**
Not applicable.

Reliability & Validity: **Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?**
Not applicable.

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?**
Yes.

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?**
From 2019 health plan reporting of this measure for HEDIS, the median denominator size across 17 Medicaid plans was 6,783.

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

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?**
None.

Provider Level
Individual practitioner: Can compare individual health care professionals

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?**
Not applicable.

Reliability & Validity: **Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?**
Not applicable.

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)**
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?**
Not applicable.

Reliability & Validity: **Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?**
Not applicable.

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)**
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?**
Not applicable.

Reliability & Validity: **Is there published evidence about the reliability and validity of the measure when reported at this level of aggregation?**
Not applicable.

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).

This measure was prioritized as an important measure, both through public comment and by NCQA advisory panels. Stakeholders noted the measure topic is of importance for the adolescent population and addresses a known quality issue. This process measure supports our ability to measure important patient outcomes. With routine symptom monitoring using a standardized tool, we are able to assess if patients with depression are responding to treatment and get better over time. Final measure specifications were informed by commenters' and advisory panel feedback. Stakeholders expressed that the measure as specified is an understandable and sensible approach to assessing depression symptom monitoring for adolescents. Since the measure was adapted from a provider level measure that is used in various reporting programs, healthcare providers understand the construct of the 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.

Routine symptom monitoring using the PHQ-9 tool can be built into provider workflows and the results captured in EHRs, case management systems, and clinical registries. These systems can then be used for population health monitoring for those with depression. Research has demonstrated the importance of routinely monitoring symptoms for depression and making adjustments to treatment to help get patients to remission. Electronic caseload tracking tools (e.g., https://aims.uw.edu/resource-library/aims-caseload-tracker), EHRs, and registry systems can provide clinicians with the information needed to manage their patients and the key data elements needed to calculate the quality measure.

11.B. Health IT Testing

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

Yes.

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

Please see the companion measure, Adolescent Depression Monitoring, available at https://www.ahrq.gov/pqmp/measures/management-of-chronic-conditions.html.

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.

Results from the PHQ-9 tool can be collected electronically outside of office visits through patient portals connected with the EHR or through collection during a visit (e.g., by filling out the tool on a tablet). Additionally, a provider at the visit can verbally ask the patient the questions following prompts in the EHR and document the results in the system. Once results are calculated, the score should be saved in a structured field in the EHR that can be easily queried and tracked over time in the system. The measure can then be easily calculated based on querying patients with a depression diagnosis, their visits for depression, and whether PHQ-9 results are documented.

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)?

Yes.

If yes, please describe.

Both Stage 2 of Meaningful Use and the 2014 edition of the ONC Certification of EHR Technology (ONC, 2010; updated to 2014) require the electronic capture of patient demographics, diagnosis, and visit information data in ambulatory settings that are necessary to calculate this measure. The data elements used in this measure (diagnoses, encounters, results of the PHQ-9 tool) are all supported by existing health IT standards and available coding systems such as ICD-10, CPT, and LOINC.

11.E. Health IT Calculation

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

The likelihood of calculation errors is low.

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?

Decision support systems could easily be built to enhance performance on this measure. For example, EHR systems can flag patients with a depression diagnosis as needing assessment with the PHQ-9 tool at each visit. Such a simple flag would enhance performance on this measure. But even more importantly, systems can then track PHQ-9 scores over time and flag patients that are not improving so clinicians can reassess treatment options.

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).

One limitation of this measure is the feasibility of collecting electronic clinical data at the health plan level. To assess routine symptom monitoring, the measure looks for documented results of the PHQ-9 tool at regular intervals. These results may be captured in various systems, such as EHRs, case management registries, or health information exchanges. Health plans have various strategies to leverage clinical data for HEDIS® reporting using the ECDS reporting method, however some health plans are still developing their methods and strategies. NCQA has several ongoing and planned efforts to support implementation of the ECDS measures and help health plans feasibly collect electronic clinical data, including posting FAQs on the NCQA website (https://www.ncqa.org/hedis/the-future-of-hedis/ecds-frequently-asked-questions/), convening learning collaboratives with health plans, and a new Digital Measurement Community launching in 2020 (https://www.ncqa.org/hedis/the-future-of-hedis/the-digital-measurement-community/).

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.

This measure assesses whether adults and adolescents with depression are routinely monitored for their symptoms of depression using a standardized tool. Depression is a prevalent condition that has impacts on mood, wellbeing, function, participation in school or work activities, and management of other chronic medical conditions. There are effective treatments and models of care to manage and improve depression. Routine monitoring of symptoms during treatment is essential to assess patients' response to treatment and improvement over time. This process measure supports the ability to measure important patient-reported outcomes for depression. The measure is specified to use data from electronic clinical data systems and is intended for use by

health plans. Testing results suggest there is much room for improvement in the management of depression and depression outcomes, and extensive feedback from multiple and varied stakeholders found the measure to be understandable, meaningful, and important.

References

Algeria M, Chatterji P, Wells K, et al. Disparity in depression treatment among racial and ethnic minority populations in the United States. Psychiatr Serv 2008; 59(11):1264-72.

American Psychiatric Association. 2015. Online Assessment Measures. Severity Measure for Depression, Child Age 11 to 17 (PHQ-9 modified for Adolescents [PHQ-A], Adapted). Washington, DC: APA; 2015. Available at https://www.psychiatry.org/psychiatrists/practice/dsm/dsm-5/online-assessment-measures. Accessed October 15, 2019.

American Psychiatric Association (APA) Practice Guideline for the Treatment for Patients with Major Depressive Disorder, 3rd Ed. Arlington, VA: Workgroup on Major Depressive Disorder, American Psychiatric Association; 2010.

Beard C, Hsu KJ, Rifkin LS, et at. Validation of the PHQ-9 in a psychiatric sample. J Affect Disorders 2016; 193:267–73.

Beck A, Crain AL, Solberg LI, et al. Severity of depression and magnitude of productivity loss. Ann Fam Med 2011; 9:305-11.

Bickman L, Kelley SD, Breda C, et al. Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. Psychiatr Serv 2011; 62(12):1423-9.

Birnbaum HG, Kessler RC, Kelley D, et al. Employer burden of mild, moderate, and severe major depressive disorder: mental health services utilization and costs, and work performance. Depress Anxiety 2010; 27(1):78-89.

Bostwick JM, Pankratz VS. Affective Disorders and Suicide Risk: a Reexamination. Am J Psychiatry 2000; 157:1925-32.

Centers for Disease Control and Prevention (CDC). Effective treatments exist: People with depression and anxiety should seek help as early as possible to reduce health effects and improve quality of life. Based on the 2006 Behavioral Risk Factor Surveillance System. Atlanta, GA: CDC; 2009.

Centers for Disease Control and Prevention (CDC). Depression. Atlanta, GA: CDC; 2011.

Cheung AH, Zuckerbrot RA, Jensen PS, et al. Guidelines for Adolescent Depression in Primary Care (GLAD-PC): II. Treatment and ongoing management. Pediatrics 2007; 120(5):e1313-26.

Child Mind Institute. Major depressive disorder basics. New York, NY: Child Mind Institute; 2011. Available at http://www.childmind.org/en/health/disorder-guide/major-depressive-disorder. Accessed October 9, 2019.

Curry J, Silva S, Rohde P, et al. Recovery and recurrence following treatment for adolescent major depression. Arch Gen Psychiatry 2011; 68(3):263-9.

Escarce JJ, Carreon R, Veselovskiy G, et al. Collection of race and ethnicity data by health plans has grown substantially, but opportunities remain to expand efforts. Health Aff 2011; 30(10):1984-91.

Ferrari AJ, Charlson FC, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: Findings from the Global Burden of Disease Study 2010. PLoS Medicine 2013; 10(11):e1001547.

Foley HA, Carlton CO, Howell RJ. The relationship of attention deficit hyperactivity disorder and conduct disorders to juvenile delinquency: Legal implications. Bull Am Acad Psychiatry Law 1996; 24:33345.

Friedman RM, Katz-Levey JW, Manderscheid RW, et al. Prevalence of serious emotional disturbance in children and adolescents. In Manderscheid RW, Sonnenschein MA (eds), Mental health, United States. Rockville, MD: Center for Mental Health Services; 1996.

Garber J, Clarke GN, Weersing VR, et al. Prevention of depression in at-risk adolescents: A randomized controlled trial. JAMA 2009; 301(21):2215-24.

Gilbody SC, Richards S, Brealey C, et al. Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): A diagnostic meta-analysis. J Gen Intern Med 2007; 22(11):1596-1602.

Johnson JG, Harris ES, Spitzer RL, et al. The Patient Health Questionnaire for Adolescents: Validation of an instrument for the assessment of mental disorders among adolescent primary care patients. J Adolesc Health 2002; 30:196-204.

Kataoka SH, Zhang L, Wells KB. Unmet need for mental health care among U.S. Children: Variation by ethnicity and insurance status. Am J Psychiatry 2002; 159(9):1548-55.

Katon WJ. Epidemiology and treatment of depression in patients with chronic medical illness. Dialogues Clin Neurosci 2011; 13(1):7-23.

Knaup C, Koesters M, Schoefer D, et al. Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. Br J Psychiatry 2009; 195(1):15-22.

Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med 2001; 16(9):606-13.

Lambert MJ, Whipple JL, Hawkins EJ, et al. Is it time for clinicians to routinely track patient outcome? A meta-analysis. Clin Psychol Sci Practice 2003; 10(3):288-301.

Lépine JP, Briley M. The increasing burden of depression. Neuropsychiatr Dis Treatment 2011; 7(suppl 1):3-7.

Lewandowski RE, Acri MC, Hoagwood KE, et al. Evidence for the management of adolescent depression. Pediatrics 2013; 132(4):e996-1009.

Lowe B, Kroenke K, Herzog W, et al. Measure depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). J Affect Disord 2004; 81(1):61-6.

Martin A, Rief W, Klaiberg A, et al. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. 2006; Gen Hosp Psychiatry 28(1);71-7.

Merikangas KR, He JP, Burstein M, et al. Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication— Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2010; 49(10):980-9.

Mezuk B, Eaton WW, Albrecht S, et al. Depression and Type 2 diabetes over the lifespan: A meta-analysis. Diabetes Care 2008; 31:2383-90.

Mitchell J, Trangle M, Degnan B, et al. Adult depression in primary care. Clinical Guideline. Bloomington, MN: Institute for Clinical Systems Improvement; 2013, updated 2016. Available at https://www.icsi.org/guideline/depression/. Accessed October 9, 2019.

Mojtabai R, Olfson M, Han B. national trends in the prevalence and treatment of depression in adolescents and young adults. Pediatrics 2016 138(6). Available at http://pediatrics.aappublications.org/content/early/2016/11/10/peds.2016-1878. Accessed October 9, 2019.

Murray C, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 1997; 349(9063):1436-42.

Murray C, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet 2013; 380(9859):2197-223.

National Collaborating Centre for Mental Health. Depression: The treatment and management of depression in adults. Clinical guideline No. 90. London: National Institute for Health and Clinical Excellence (NICE); 2009.

National Institute for Health and Clinical Excellence. Clinical Guideline 28: Depression in Children and Young People: Identification and Management in Primary, Community, and Secondary Care. London, UK: NHS; 2005.

National Research Council and Institute of Medicine. Adolescent health services: Missing opportunities. Lawrence RS, Gootman JA, Sim LJ, Eds. Washington, DC: National Academies Press; 2009.

Office of the National Coordinator for Health IT Standards and Certification (ONC). Health information technology: Initial set of standards, implementation specifications, and certification criteria for electronic health record technology. Fed Regist 2010; 75(8):2013-47.

Richardson L, Ludman E, McCauley E, et al. Collaborative care for adolescents with depression in primary care: A randomized clinical trial. JAMA 2014; 312(8):809-16.

Shaffer D, Fisher P, Lucas CP, et al. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. J Am Acad Child Adolesc Psychiatry 2000; 39(1):28-38.

Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. J Consult Clin Psychol 2010; 78(3):298-311.

Spitzer RL, Williams JB, Kroenke K. Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. JAMA 1999; 282(18):1737-44.

Substance Abuse and Mental Health Services Administration (SAMHSA). Behavioral health trends in the United States: Results from the 2014 National Survey on Drug Use and Health. Rockville, MD: SAMHSA; 2015. Available at http://www.samhsa.gov/data/sites/default/files/NSDUH-FRR1-2014/NSDUH-FRR1-2014.pdf. Accessed October 14, 2019.

Substance Abuse and Mental Health Services Administration (SAMHSA). Mental health and substance use disorders. Rockville, MD: SAMHSA; 2014. Available at http://www.samhsa.gov/disorders/mental. Accessed October 9, 2019.

Taylor E, Chadwick O, Heptinstall E, et al. Hyperactivity and conduct problems as risk factors for adolescent development. J Am Acad Child Adolesc Psychiatry 1996; 35(9):1213-26.

Trangle M, Gursky J, Haight R, et al. Depression, Adult in Primary Care. Clinical guideline. Bloomington, MN: Institute for Clinical Systems Improvement (ICSI); 2013 (updated 2016). Available at https://www.icsi.org/guideline/depression/. Accessed October 9, 2019.

U.S. Congress. An Act: The Patient Protection and Affordable Care Act; 2010. Available at https://www.gpo.gov/fdsys/pkg/BILLS-111hr3590enr/pdf/BILLS-111hr3590enr.pdf. Accessed October 9, 2019.

U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents. U.S. Preventive Services Task Force Recommendation Statement. Pediatrics 2009; 123(6):1223-8.

Üstün TB, Ayuso-Mateos JL, Chatterji S, et al. Global burden of depressive disorders in the year 2000. Br J Psychiatry 2004; 184(5):386-92.

Van der Kooy K, van Hout H, Marwijk H, et al. Depression and the risk for cardiovascular diseases: Systematic review and meta-analysis. Int J Geriatr Psychiatry 2007; 2:613-26.

Williams SB, O'Connor E, Eder M, et al. Screening for child and adolescent depression in primary care settings: A systematic evidence review for the U.S. Preventive Services Task Force. Evidence synthesis no. 69 (AHRQ Publication No. 09-05130-EF-1). Rockville, MD: Agency for Healthcare Research and Quality; 2009. Available at https://www.uspreventiveservicestaskforce.org/Page/Document/evidence-summary/depression-in-children-and-adolescents-screening1. Accessed October 9, 2019.

Section 14: Identifying Information for the Measure Submitter

First Name:	Sarah Hudson
Last Name:	Scholle
Title:	Vice President, Research & Analysis
Organization:	National Committee for Quality Assurance
Mailing Address:	1100 13th Street, N.W., Third Floor
City:	Washington
State:	DC
Postal Code:	20005
Telephone:	202-955-1726
Email:	scholle@ncqa.org

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. 20-0002
December 2019