NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 1394 **NQF Project:** Child Health Quality Measures 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Depression Screening By 13 years of age
De.2 **Brief description of measure:** The percentage of adolescents who turn 13 years of age in the measurement year who had a screening for depression using a standardized tool.
1.1-2 **Type of Measure:** Process De.3 **If included in a composite or paired with another measure, please identify composite or paired measure** This measure appears in the composite Comprehensive Well Care by Age 13 Years.
De.4 **National Priority Partners Priority Area:** Population health De.5 **IOM Quality Domain**: Effectiveness, Timeliness De.6 **Consumer Care Need:** Staying healthy

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	**NQF Staff**
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** Proprietary measure A.3 **Measure Steward Agreement:** Agreement will be signed and submitted prior to or at the time of measure submission A.4 **Measure Steward Agreement attached:**	**A** Y☐ N☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and	**B**

update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	Y☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement Accountability	**C** Y☐ N☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** Yes, fully developed and tested D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** Y☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. ***Measures must be judged to be important to measure and report*** *in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** High resource use, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Major depressive disorder (MDD) affects more than 7 percent of adolescents in the U.S. In 2006, around 2.3 million 12-17 year-old adolescents had a major depressive episode in their life. Depression is much less common in children under the age of 11 (Williams, 2009); MDD occurs in about 2.8 percent of children younger than 13 years old (USPSTF, 2009). Signs of major depressive disorder include: sadness, irritability, isolation, trouble completing work, problems sleeping, and unexplained body pains. These MDD symptoms "cluster" and can last for two weeks or longer (USPSTF, 2009). Depression, which can vary in severity, can have a major impact on people's lives, including serious long-term morbidities (USPSTF, 2009). It can disrupt daily life at home, at school or in the community and can lead to drug use and other risky behavior, even suicide (Taylor, 1996; Foley, 1996; Friedman, 1996; NRCIM, 2009). Most adolescents that committed suicide, which is the third leading cause of death in 15 to 24 year olds and the sixth leading for children 5 to 14 years, had a history of depression or long-term MDD (NRCIM, 2009; Williams, 2009). The adolescent-onset depressed have upwards of a five-fold increase in attempting suicide risk compared to non-depressed adolescents (Williams SB, 2009). Children with MDD have higher medical expenditures, including general health care and mental health care, than children without (USPSTF, 2009). Outpatient care is the most common treatment; it accounts for	**1a** C☐ P☐ M☐ N☐

nearly 60 percent of all mental health expenditures, including major depressive disorder, for young people, a large portion of which is from school-based programs (MHCY, 2001). Inpatient care accounts for about 33 percent of all mental health expenditures, and the remaining seven percent is for medications and other mental health services related to mental health (MHCY, 2001).

1a.4 Citations for Evidence of High Impact: Foley, H.A.; Carlton, C.O.; and Howell, R.J. The relationship of attention deficit hyperactivity disorder and conduct disorders to juvenile delinquency: Legal implications. Bulletin of the American Academy of Psychiatry Law 24:333 345, 1996.

Friedman, R.M.; Katz-Levey, J.W.; Manderschied, R.W.; and Sondheimer, D.L. Prevalence of serious emotional disturbance in children and adolescents. In: Manderscheid, R.W., and Sonnenschein, M.A. (eds.). Mental Health, United States, 1996. Rockville, MD: Center for Mental Health Services, 1996, 71-78.

National Research Council and Institute of Medicine. (2009). Adolescent Health Services: Missing Opportunities. Committee on Adolescent Health Care Services and Models of Care for Treatment, Prevention, and Healthy Development, R.S. Lawrence, J. Appleton Gootman, and L.J. Sim, Editors. Board on Children, Youth, and Families. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

RAND Health. Mental Health Care for Youth: Who Gets It? How Much Does It Cost? Who Pays? Where Does the Money Go? http://www.rand.org/pubs/research_briefs/RB4541/index1.html . Updated 2001.

Surgeon General report. http://www.surgeongeneral.gov/library/mentalhealth/pdfs/c3.pdf

Taylor, E.; Chadwick, O.; Heptinstall, E; et al. Hyperactivity and conduct problems as risk factors for adolescent development. Journal of the American Academy of Child and Adolescent Psychiatry 35:1213 1226, 1996.

U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. Pediatrics 2009;123:1223–1228

Williams SB, O'Connor, E, Eder M, Whitlock E. 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. April 2009.

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: This measure highlights the need for screening of major depressive disorder in adolescents. Early intervention in adolescents diagnosed with depression can lead to needed treatment. Once depression is diagnosed, around 95 percent of physicians report further assessment of specific symptoms and contributing factors. Another study found that 52 percent of the times that depression was reported in adolescent primary care visits, antidepressants were prescribed; 68 percent of cases led to psychotherapy or counseling (Williams SB, 2009).

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
Despite the prevalence of mental health concerns, most adolescents are undiagnosed and untreated (USPSTF, 2009). Documentation from community health centers shows screening for only 3 percent of patients. HMO providers screen around 40 percent of their patients for depression. Those physicians that do screen for depression report not systematically using a standardized tool or the DSM-IV criteria (Williams, 2009).

1b.3 Citations for data on performance gap:
U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. Pediatrics 2009;123:1223–1228

Williams SB, O'Connor, E, Eder M, Whitlock E. 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. April 2009.

1b
C☐
P☐
M☐
N☐

1b.4 Summary of Data on disparities by population group:
MDD can appear in both males and females during childhood or adolescence. However, young female adolescents are more likely to be diagnosed with depression than males (National Research Council and Institute of Medicine, 2009). Minority racial/ethnic groups are at an even further disadvantage. Minority children are 50 to 60 percent less likely to receive mental health care as white children, despite a similar overall prevalence of disease. Hispanic/Latino youth are the least likely to receive treatment, and a smaller, similar disparity has been found for Asian/Pacific Islander as well as African American youth. Moreover, of those who do receive care, these minority groups are less likely to receive complete services and are more likely to receive treatment that is inappropriate, fragmented, or inadequate (Cheryl Holm-Hansen, 2006).

1b.5 Citations for data on Disparities:
Cheryl Holm-Hansen. Racial and ethnic disparities in children's mental health. http://www.wilder.org/reportsummary.0.html?tx_ttnews percent5Btt_news percent5D=1964. Updated 2006

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** The U.S. Preventive Services Task Force (USPSTF) found no studies that directly examined the health outcomes of screening children and adolescents for depression.

http://www.ahrq.gov/clinic/uspstf09/depression/chdeprart.htm

1c.2-3. Type of Evidence: Evidence-based guideline, Expert opinion

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**
The U.S. Preventives Services Task Force (USPSTF) recommends that adolescents aged 12-18 years old be screened for major depressive disorder. The USPSTF found adequate evidence that screening tests can accurately identify MDD in adolescents. Adequate evidence also supports beneficial decreases in MDD symptoms associated with treatment of adolescents with SSRIs, psychotherapy, and therapy combining SSRIs with psychotherapy. The USPSTF found inadequate evidence of harms of screening adolescents. There is adequate evidence on the harms of SSRIs (risk of suicidality), but there is no evidence on the harms of psychotherapy or combined treatment of adolescents with psychotherapy and SSRIs (fluoxetine), which is bounded to be low. The USPSTF found moderate certainty that the net benefit is moderate for screening followed by treatment with psychotherapy in adolescents.

The USPSTF concluded that co-morbid mental health problems, chronic conditions, parental depression, along with major life-changing events are risk factors of depression that can be assessed accurately and reliably. Similarly, external risk factors such as poverty, deprivation, abuse and neglect, unsatisfactory relationships, or exposure to traumatic events may also play a role in depression (Surgeon General report).

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**
Good

1c.6 Method for rating evidence: Expert consensus based on evidence review

1c.7 Summary of Controversy/Contradictory Evidence: None

1c.8 Citations for Evidence (*other than guidelines***):** U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. Pediatrics 2009;123:1223–1228

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number***):**
U.S. Preventive Services Task Force (2009)
The USPSTF recommends that adolescents aged 12-18 years old be screened for major depressive disorder

1c
C☐
P☐
M☐
N☐

when there are systems in place to ensure accurate diagnosis. The USPSTF recommends using the Patient Health Questionnaire for Adolescents (PHQ-A) or the Beck Depression Inventory-Primary Care Version (BDI-PC). (B Recommendation) American Academy of Family Physicians (AAFP) (2009) The AAFP endorses the USPSTF recommendation. Michigan Quality Improvement Consortium (2007) The Michigan Quality Improvement Consortium recommends that health care professionals screen adolescents age 13-18 years. Parent/Child education and counseling should include: depression, suicide threats, alcohol/drug abuse, anxiety, stress reduction, coping skills. (Expert Consensus) Bright Futures (2008) Bright Futures states that health care professionals should screen adolescents 15 to 21 years of age. Discussion topics should include coping, mood regulation and mental health sexuality. (Expert Consensus) **1c.10 Clinical Practice Guideline Citation:** American Academy of Family Physicians (AAFP). Summary of recommendations for clinical preventive services. Revision 6.4. Leawood (KS): American Academy of Family Physicians (AAFP); 2008 * The AAFP "clinical considerations" link goes to USPSTF 2009 updated recommendation Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics Michigan Quality Improvement Consortium. Routine preventive services for infants and children (ages 2-18). May 2007 U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. Pediatrics 2009;123:1223–1228 **1c.11 National Guideline Clearinghouse or other URL:** Screening and treatment for major depressive disorder in children and adolescents: U.S. Preventive Services Task Force recommendation statement. http://www.guideline.gov/content.aspx?id=14294&search=depression+screening **1c.12 Rating of strength of recommendation (***also provide narrative description of the rating and by whom***):** USPSTF B Recommendation **1c.13 Method for rating strength of recommendation** (*If different from USPSTF system, also describe rating and how it relates to USPSTF)*: USPSTF Based **1c.14 Rationale for using this guideline over others:** In general, guidelines from major clinical bodies are in alignment with the USPSTF Recommendation.	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	**Eval Rating**
2a. MEASURE SPECIFICATIONS	

S.1 **Do you have a web page where current detailed measure specifications can be obtained?**
S.2 **If yes, provide web page URL:**

2a. Precisely Specified

2a.1 **Numerator Statement** (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome*)**:**
Children who had a screening for depression using a standardized tool by age 13 years

2a.2 **Numerator Time Window** (*The time period in which cases are eligible for inclusion in the numerator*)**:**
2 years

2a.3 **Numerator Details** (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*)**:**
Documentation of depression screening using a standardized tool. Any of the following qualifies as a standardized tool:
- Patient Health Questionnaire for Adolescents (PHQ-A).
- Beck Depression Inventory-Primary Care Version (BDI-PC).
- PHQ-2—Patient Health Questionnaire-2 Item
- PHQ-9—Patient Health Questionnaire-9 Item
- Columbia Depression Scale - Teen Version
- Kutcher Adolescent Depression Scale (KADS) 6-item

2a.4 **Denominator Statement** (*Brief, text description of the denominator - target population being measured*)**:**
Children with a visit who turned 13 years in the measurement year

2a.5 **Target population gender:** Female, Male
2a.6 **Target population age range:** 11 years-13 years

2a.7 **Denominator Time Window** (*The time period in which cases are eligible for inclusion in the denominator*)**:**
1 year

2a.8 **Denominator Details** (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*)**:**
Children who turned 13 years of age between January 1 of the measurement year and December 31 of the measurement year and who had documentation of a face-to-face visit between the clinician and the child that predates the child's birthday by at least 12 months.

2a.9 **Denominator Exclusions** (*Brief text description of exclusions from the target population*)**:** None

2a.10 **Denominator Exclusion Details** (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*)**:**
NA

2a.11 **Stratification Details/Variables** (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*)**:**
None

2a.12-13 **Risk Adjustment Type:** No risk adjustment necessary

2a.14 **Risk Adjustment Methodology/Variables** (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*)**:**
NA

2a.15-17 **Detailed risk model available Web page URL or attachment:**

2a.18-19 **Type of Score:** Rate/proportion
2a.20 **Interpretation of Score:** Better quality = Higher score
2a.21 **Calculation Algorithm** (*Describe the calculation of the measure as a flowchart or series of steps*)**:**

2a-specs
C ☐
P ☐
M ☐
N ☐

Step 1: Determine the denominator
Children who turned the requisite age in the measurement year, AND
Who had a visit within the past 12 months of the child´s birthday
Step 2: Determine the numerator
Children who had documentation in the medical record of the screening or service during the measurement year or the year previous to the measurement year.

2a.22 **Describe the method for discriminating performance** *(e.g., significance testing)*:
Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance.

2a.23 **Sampling (Survey) Methodology** *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:
For this physician-level measure, we anticipate the entire population will be used in the denominator. If a sample is used, a random sample is ideal. NCQA's work has indicated that a sample size of 30-50 patients would be necessary for a typical practice size of 2000 patients.

2a.24 **Data Source (***Check the source(s) for which the measure is specified and tested)*
Paper medical record/flow-sheet, Electronic clinical data, Electronic Health/Medical Record

2a.25 **Data source/data collection instrument (***Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**
Medical Record

2a.26-28 **Data source/data collection instrument reference web page URL or attachment:**

2a.29-31 **Data dictionary/code table web page URL or attachment:**

2a.32-35 **Level of Measurement/Analysis** (*Check the level(s) for which the measure is specified and tested)*
Clinicians: Individual, Clinicians: Group, Population: national, Population: regional/network

2a.36-37 **Care Settings (***Check the setting(s) for which the measure is specified and tested)*
Ambulatory Care: Office, Ambulatory Care: Clinic, Ambulatory Care: Hospital Outpatient, Behavioral health/psychiatric unit

2a.38-41 **Clinical Services** (*Healthcare services being measured, check all that apply*)
Behavioral Health: Mental Health, Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO), Clinicians: Psychologist/LCSW

TESTING/ANALYSIS

2b. Reliability testing

2b.1 **Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure)

2b.2 **Analytic Method** *(type of reliability & rationale, method for testing)*:
We calculated 95% confidence intervals, which speak to the precision of the rates obtained from field testing.

2b.3 **Testing Results** *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
Rate (Upper Confidence Interval, Lower Confidence Interval):
0.520 (0.45, 0.59)

2b
C☐
P☐
M☐
N☐

2c. Validity testing

2c.1 **Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure)

2c
C☐
P☐
M☐
N☐

2c.2 Analytic Method *(type of validity & rationale, method for testing)*:
NCQA tested the measure for face validity using a panel of stakeholders with specific expertise in measurement and child health care. This panel included representatives from key stakeholder groups, including pediatricians, family physicians, health plans, state Medicaid agencies and researchers. Experts reviewed the results of the field test and assessed whether the results were consistent with expectations, whether the measure represented quality care, and whether we were measuring the most important aspect of care in this area. This measure was deemed valid by the expert panel. In addition, this measure does not utilize administrative data sources; data recorded in the chart is considered the gold standard.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
NA

2d. Exclusions Justified

2d.1 Summary of Evidence supporting exclusion(s):
None

2d.2 Citations for Evidence:
NA

2d.3 Data/sample *(description of data/sample and size)*: NA

2d.4 Analytic Method *(type analysis & rationale)*:
NA

2d.5 Testing Results *(e.g., frequency, variability, sensitivity analyses)*:
NA

2d
C☐
P☐
M☐
N☐
NA☐

2e. Risk Adjustment for Outcomes/ Resource Use Measures

2e.1 Data/sample *(description of data/sample and size)*: NA

2e.2 Analytic Method *(type of risk adjustment, analysis, & rationale)*:
NA

2e.3 Testing Results *(risk model performance metrics)*:
NA

2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: The measure assesses prevention and wellness in a general population; risk adjustment is not indicated.

2e
C☐
P☐
M☐
N☐
NA☐

2f. Identification of Meaningful Differences in Performance

2f.1 Data/sample from Testing or Current Use *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure)

2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance *(type of analysis & rationale)*:
Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance

2f.3 Provide Measure Scores from Testing or Current Use *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:
Elig Population: 179
We tested several different numerator options during field test:
Documentation of a Depression Screening: 52.0
Documentation of Results of the Screening: 51.4%
Documentation that Screening was Done Using a Standardized Tool: 17.3%

2f
C☐
P☐
M☐
N☐

2g. Comparability of Multiple Data Sources/Methods 2g.1 **Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure) 2g.2 **Analytic Method** *(type of analysis & rationale)*: This measure is chart review only; no other sources were identified by the expert panel; this measure does not utilize administrative data 2g.3 **Testing Results** *(e.g., correlation statistics, comparison of rankings)*: NA	**2g** C☐ P☐ M☐ N☐ NA☐
2h. Disparities in Care 2h.1 **If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: The measure is not stratified to detect disparities. 2h.2 **If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** NA	**2h** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	***2*** C☐ P☐ M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	**Eval Rating**
3a. Meaningful, Understandable, and Useful Information 3a.1 **Current Use:** Not in use but testing completed 3a.2 **Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*): This measure is not currently publicly reported. NCQA is exploring the feasibility of adding this measure and its related measures into a physician-level program and/or the HEDIS® measurement set as appropriate. 3a.3 **If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):** This measure is not currently used in QI. NCQA is exploring the feasibility of adding this measure and its related measures into a physician-level program and/or the HEDIS® measurement set as appropriate. NCQA anticipates that after we release these measures, they will become widely used, as all our measures do. **Testing of Interpretability** (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) 3a.4 **Data/sample** *(description of data/sample and size)*: NA 3a.5 **Methods** *(e.g., focus group, survey, QI project)*: NCQA vetted the measures with its expert panel. In addition, throughout the development process, NCQA vetted the measure concepts and specifications with other stakeholder groups, including the National Association of State Medicaid Directors, NCQA's Health Plan Advisory Council, NCQA's Committee on Performance Measurement, and the American Academy of Pediatrician's Quality Improvement Innovation Network.	**3a** C☐ P☐ M☐ N☐

After field testing, NCQA also conducted a debrief call with field test participants. In the form of a group interview, NCQA systematically sought feedback on whether the measures were understandable, feasible, important, and had face validity. 3a.6 **Results** *(qualitative and/or quantitative results and conclusions)*: NCQA received feedback that the measure is understandable, feasible, important and valid.	
3b/3c. Relation to other NQF-endorsed measures 3b.1 **NQF # and Title of similar or related measures:**	
(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 **Are the measure specifications harmonized? If not, why?**	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value 3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** 5.1 **If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:**	**3c** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐ P☐ M☐ N☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	**Eval Rating**
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition), Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	**4a** C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) No 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.** NCQA plans to eventually adopt this measure in electronic health records.	**4b** C☐ P☐ M☐ N☐
4c. Exclusions	**4c** C☐

4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications? No **4c.2 If yes, provide justification.**	**P☐ M☐ N☐ NA☐**
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences **4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** During the measure development process the Child Health MAP and measure development team worked with NCQA's certified auditors and audit department to ensure that the measure specifications were clear and auditable. The denominator, numerator and optional exclusions are concisely specified and align with our audit standards.	**4d C☐ P☐ M☐ N☐**
4e. Data Collection Strategy/Implementation **4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:** Based on field test results, we have specified the measure to assess whether screening was documented and whether use of a standardized tool was documented. Our field test results showed that these data elements are available in the medical record. In addition, our field test participants noted that many were able to program these requirements into their electronic health record systems, and several implemented point-of-service physician reminders for this measure. **4e.2 Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** Collecting measures from medical charts is time-consuming and can be burdensome. Adapting this measure in electronic health records may relieve some of this burden. **4e.3 Evidence for costs:** Based on field test participant feedback and other stakeholder input **4e.4 Business case documentation:**	**4e C☐ P☐ M☐ N☐**
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility?*	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	**4 C☐ P☐ M☐ N☐**
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited ☐**
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y☐ N☐ A☐**
CONTACT INFORMATION	
Co.1 Measure Steward (Intellectual Property Owner) **Co.1 Organization** National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005	

Co.2 **Point of Contact**
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-

Measure Developer If different from Measure Steward
Co.3 **Organization**
National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005

Co.4 **Point of Contact**
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-

Co.5 **Submitter If different from Measure Steward POC**
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-, National Committee for Quality Assurance

Co.6 **Additional organizations that sponsored/participated in measure development**

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 **Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.**
Child Health Measurement Advisory Panel:
Jeanne Alicandro
Barbara Dailey
Denise Dougherty, PhD
Ted Ganiats, MD
Foster Gesten, MD
Nikki Highsmith, MPA
Charlie Homer, MD, MPH
Jeff Kamil, MD
Elizabeth Siteman
Mary McIntyre, MD, MPH
Virginia Moyer, MD, MPH, FAAP
Lee Partridge
Xavier Sevilla, MD, FAAP
Michael Siegal
Jessie Sullivan

Ad.2 **If adapted, provide name of original measure:** NA
Ad.3-5 **If adapted, provide original specifications URL or attachment**

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 **Year the measure was first released:**
Ad.7 **Month and Year of most recent revision:**
Ad.8 **What is your frequency for review/update of this measure?**
Ad.9 **When is the next scheduled review/update for this measure?**

Ad.10 **Copyright statement/disclaimers:**

Ad.11 -13 **Additional Information web page URL or attachment:**

Date of Submission (*MM/DD/YY*)**:** 01/06/2011

NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 1515 **NQF Project:** Child Health Quality Measures 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Depression Screening By 18 years of age
De.2 **Brief description of measure***:* The percentage of adolescents who turn 18 years of age in the measurement year who had a screening for depression using a standardized tool.
1.1-2 **Type of Measure:** Process De.3 **If included in a composite or paired with another measure, please identify composite or paired measure** This measure appears in the composite Comprehensive Well Care by Age 18 Years.
De.4 **National Priority Partners Priority Area:** Population health De.5 **IOM Quality Domain**: Effectiveness, Timeliness De.6 **Consumer Care Need:** Staying healthy

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	**NQF Staff**
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** Proprietary measure A.3 **Measure Steward Agreement:** Agreement will be signed and submitted prior to or at the time of measure submission A.4 **Measure Steward Agreement attached:**	**A** Y☐ N☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and	**B**

update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	Y☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement Accountability	**C** Y☐ N☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** Yes, fully developed and tested D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** Y☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. ***Measures must be judged to be important to measure and report*** *in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** High resource use, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Major depressive disorder (MDD) affects more than 7 percent of adolescents in the U.S. In 2006, around 2.3 million 12-17 year-old adolescents had a major depressive episode in their life. Depression is much less common in children under the age of 11 (Williams, 2009); MDD occurs in about 2.8 percent of children younger than 13 years old (USPSTF, 2009). Signs of major depressive disorder include: sadness, irritability, isolation, trouble completing work, problems sleeping, and unexplained body pains. These MDD symptoms "cluster" and can last for two weeks or longer (USPSTF, 2009). Depression, which can vary in severity, can have a major impact on people's lives, including serious long-term morbidities (USPSTF, 2009). It can disrupt daily life at home, at school or in the community and can lead to drug use and other risky behavior, even suicide (Taylor, 1996; Foley, 1996; Friedman, 1996; NRCIM, 2009). Most adolescents that committed suicide, which is the third leading cause of death in 15 to 24 year olds and the sixth leading for children 5 to 14 years, had a history of depression or long-term MDD (NRCIM, 2009; Williams, 2009). The adolescent-onset depressed have upwards of a five-fold increase in attempting suicide risk compared to non-depressed adolescents (Williams SB, 2009). Children with MDD have higher medical expenditures, including general health care and mental health care, than children without (USPSTF, 2009). Outpatient care is the most common treatment; it accounts for	**1a** C☐ P☐ M☐ N☐

nearly 60 percent of all mental health expenditures, including major depressive disorder, for young people, a large portion of which is from school-based programs (MHCY, 2001). Inpatient care accounts for about 33 percent of all mental health expenditures, and the remaining seven percent is for medications and other mental health services related to mental health (MHCY, 2001).

1a.4 Citations for Evidence of High Impact: Foley, H.A.; Carlton, C.O.; and Howell, R.J. The relationship of attention deficit hyperactivity disorder and conduct disorders to juvenile delinquency: Legal implications. Bulletin of the American Academy of Psychiatry Law 24:333 345, 1996.

Friedman, R.M.; Katz-Levey, J.W.; Manderschied, R.W.; and Sondheimer, D.L. Prevalence of serious emotional disturbance in children and adolescents. In: Manderscheid, R.W., and Sonnenschein, M.A. (eds.). Mental Health, United States, 1996. Rockville, MD: Center for Mental Health Services, 1996, 71-78.

National Research Council and Institute of Medicine. (2009). Adolescent Health Services: Missing Opportunities. Committee on Adolescent Health Care Services and Models of Care for Treatment, Prevention, and Healthy Development, R.S. Lawrence, J. Appleton Gootman, and L.J. Sim, Editors. Board on Children, Youth, and Families. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

RAND Health. Mental Health Care for Youth: Who Gets It? How Much Does It Cost? Who Pays? Where Does the Money Go? http://www.rand.org/pubs/research_briefs/RB4541/index1.html . Updated 2001.

Surgeon General report. http://www.surgeongeneral.gov/library/mentalhealth/pdfs/c3.pdf

Taylor, E.; Chadwick, O.; Heptinstall, E; et al. Hyperactivity and conduct problems as risk factors for adolescent development. Journal of the American Academy of Child and Adolescent Psychiatry 35:1213 1226, 1996.

U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. Pediatrics 2009;123:1223–1228

Williams SB, O'Connor, E, Eder M, Whitlock E. 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. April 2009.

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: This measure highlights the need for screening of major depressive disorder in adolescents. Early intervention in adolescents diagnosed with depression can lead to needed treatment. Once depression is diagnosed, around 95 percent of physicians report further assessment of specific symptoms and contributing factors. Another study found that 52 percent of the times that depression was reported in adolescent primary care visits, antidepressants were prescribed; 68 percent of cases led to psychotherapy or counseling (Williams SB, 2009).

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
Despite the prevalence of mental health concerns, most adolescents are undiagnosed and untreated (USPSTF, 2009). Documentation from community health centers shows screening for only 3 percent of patients. HMO providers screen around 40 percent of their patients for depression. Those physicians that do screen for depression report not systematically using a standardized tool or the DSM-IV criteria (Williams, 2009).

1b.3 Citations for data on performance gap:
U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. Pediatrics 2009;123:1223–1228

Williams SB, O'Connor, E, Eder M, Whitlock E. 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. April 2009.

1b
C☐
P☐
M☐
N☐

1b.4 Summary of Data on disparities by population group:
MDD can appear in both males and females during childhood or adolescence. However, young female adolescents are more likely to be diagnosed with depression than males (National Research Council and Institute of Medicine, 2009). Minority racial/ethnic groups are at an even further disadvantage. Minority children are 50 to 60 percent less likely to receive mental health care as white children, despite a similar overall prevalence of disease. Hispanic/Latino youth are the least likely to receive treatment, and a smaller, similar disparity has been found for Asian/Pacific Islander as well as African American youth. Moreover, of those who do receive care, these minority groups are less likely to receive complete services and are more likely to receive treatment that is inappropriate, fragmented, or inadequate (Cheryl Holm-Hansen, 2006).

1b.5 Citations for data on Disparities:
Cheryl Holm-Hansen. Racial and ethnic disparities in children's mental health. http://www.wilder.org/reportsummary.0.html?tx_ttnews percent5Btt_news percent5D=1964. Updated 2006

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** The U.S. Preventive Services Task Force (USPSTF) found no studies that directly examined the health outcomes of screening children and adolescents for depression.

http://www.ahrq.gov/clinic/uspstf09/depression/chdeprart.htm

1c.2-3. Type of Evidence: Evidence-based guideline, Expert opinion

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**
The U.S. Preventives Services Task Force (USPSTF) recommends that adolescents aged 12-18 years old be screened for major depressive disorder. The USPSTF found adequate evidence that screening tests can accurately identify MDD in adolescents. Adequate evidence also supports beneficial decreases in MDD symptoms associated with treatment of adolescents with SSRIs, psychotherapy, and therapy combining SSRIs with psychotherapy. The USPSTF found inadequate evidence of harms of screening adolescents. There is adequate evidence on the harms of SSRIs (risk of suicidality), but there is no evidence on the harms of psychotherapy or combined treatment of adolescents with psychotherapy and SSRIs (fluoxetine), which is bounded to be low. The USPSTF found moderate certainty that the net benefit is moderate for screening followed by treatment with psychotherapy in adolescents.

The USPSTF concluded that co-morbid mental health problems, chronic conditions, parental depression, along with major life-changing events are risk factors of depression that can be assessed accurately and reliably. Similarly, external risk factors such as poverty, deprivation, abuse and neglect, unsatisfactory relationships, or exposure to traumatic events may also play a role in depression (Surgeon General report).

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**
Good

1c.6 Method for rating evidence: Expert consensus based on evidence review

1c.7 Summary of Controversy/Contradictory Evidence: None

1c.8 Citations for Evidence (*other than guidelines***):** U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. Pediatrics 2009;123:1223–1228

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number***):**
U.S. Preventive Services Task Force (2009)
The USPSTF recommends that adolescents aged 12-18 years old be screened for major depressive disorder

1c
C☐
P☐
M☐
N☐

when there are systems in place to ensure accurate diagnosis. The USPSTF recommends using the Patient Health Questionnaire for Adolescents (PHQ-A) or the Beck Depression Inventory-Primary Care Version (BDI-PC). (B Recommendation) American Academy of Family Physicians (AAFP) (2009) The AAFP endorses the USPSTF recommendation. Michigan Quality Improvement Consortium (2007) The Michigan Quality Improvement Consortium recommends that health care professionals screen adolescents age 13-18 years. Parent/Child education and counseling should include: depression, suicide threats, alcohol/drug abuse, anxiety, stress reduction, coping skills. (Expert Consensus) Bright Futures (2008) Bright Futures states that health care professionals should screen adolescents 15 to 21 years of age. Discussion topics should include coping, mood regulation and mental health sexuality. (Expert Consensus) **1c.10 Clinical Practice Guideline Citation:** American Academy of Family Physicians (AAFP). Summary of recommendations for clinical preventive services. Revision 6.4. Leawood (KS): American Academy of Family Physicians (AAFP); 2008 * The AAFP "clinical considerations" link goes to USPSTF 2009 updated recommendation Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics Michigan Quality Improvement Consortium. Routine preventive services for infants and children (ages 2-18). May 2007 U.S. Preventive Services Task Force. Screening and Treatment for Major Depressive Disorder in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. Pediatrics 2009;123:1223–1228 **1c.11 National Guideline Clearinghouse or other URL:** Screening and treatment for major depressive disorder in children and adolescents: U.S. Preventive Services Task Force recommendation statement. http://www.guideline.gov/content.aspx?id=14294&search=depression+screening **1c.12 Rating of strength of recommendation (***also provide narrative description of the rating and by whom***):** USPSTF B Recommendation **1c.13 Method for rating strength of recommendation** (*If different from USPSTF system, also describe rating and how it relates to USPSTF)*: USPSTF Based **1c.14 Rationale for using this guideline over others:** In general, guidelines from major clinical bodies are in alignment with the USPSTF Recommendation.	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	**Eval Rating**
2a. MEASURE SPECIFICATIONS	

S.1 **Do you have a web page where current detailed measure specifications can be obtained?**
S.2 **If yes, provide web page URL:**

2a. Precisely Specified

2a.1 **Numerator Statement** (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome*)**:**
Adolescents who had a screening for depression using a standardized tool by age 18 years

2a.2 **Numerator Time Window** (*The time period in which cases are eligible for inclusion in the numerator*)**:**
2 years

2a.3 **Numerator Details** (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*)**:**
Documentation of depression screening using a standardized tool. Any of the following qualifies as a standardized tool:
- Patient Health Questionnaire for Adolescents (PHQ-A).
- Beck Depression Inventory-Primary Care Version (BDI-PC).
- PHQ-2—Patient Health Questionnaire-2 Item
- PHQ-9—Patient Health Questionnaire-9 Item
- Columbia Depression Scale - Teen Version
- Kutcher Adolescent Depression Scale (KADS) 6-item

2a.4 **Denominator Statement** (*Brief, text description of the denominator - target population being measured*)**:**
Adolescents with a visit who turned 18 years in the measurement year

2a.5 **Target population gender:** Female, Male
2a.6 **Target population age range:** 16 years-18 years

2a.7 **Denominator Time Window** (*The time period in which cases are eligible for inclusion in the denominator*)**:**
1 year

2a.8 **Denominator Details** (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*)**:**
Adolescents who turned 18 years of age between January 1 of the measurement year and December 31 of the measurement year and who had documentation of a face-to-face visit between the clinician and the adolescent that predates the adolescent's birthday by at least 12 months.

2a.9 **Denominator Exclusions** (*Brief text description of exclusions from the target population*)**:** None

2a.10 **Denominator Exclusion Details** (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*)**:**
NA

2a.11 **Stratification Details/Variables** (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*)**:**
None

2a.12-13 **Risk Adjustment Type:** No risk adjustment necessary

2a.14 **Risk Adjustment Methodology/Variables** (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*)**:**
NA

2a.15-17 **Detailed risk model available Web page URL or attachment:**

2a.18-19 **Type of Score:** Rate/proportion
2a.20 **Interpretation of Score:** Better quality = Higher score
2a.21 **Calculation Algorithm** (*Describe the calculation of the measure as a flowchart or series of steps*)**:**

2a-specs
C ☐
P ☐
M ☐
N ☐

Step 1: Determine the denominator
Adolescents who turned 18 years in the measurement year, AND
Who had a visit within the past 12 months of the adolescent´s birthday
Step 2: Determine the numerator
Children who had documentation in the medical record of the screening or service during the measurement year or the year previous to the measurement year.

2a.22 **Describe the method for discriminating performance** *(e.g., significance testing)*:
Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance.

2a.23 **Sampling (Survey) Methodology** *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:
For this physician-level measure, we anticipate the entire population will be used in the denominator. If a sample is used, a random sample is ideal. NCQA's work has indicated that a sample size of 30-50 patients would be necessary for a typical practice size of 2000 patients.

2a.24 **Data Source (***Check the source(s) for which the measure is specified and tested)*
Paper medical record/flow-sheet, Electronic clinical data, Electronic Health/Medical Record

2a.25 **Data source/data collection instrument (***Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**
Medical Record

2a.26-28 **Data source/data collection instrument reference web page URL or attachment:**

2a.29-31 **Data dictionary/code table web page URL or attachment:**

2a.32-35 **Level of Measurement/Analysis** (*Check the level(s) for which the measure is specified and tested)*
Clinicians: Individual, Clinicians: Group, Population: national, Population: regional/network

2a.36-37 **Care Settings (***Check the setting(s) for which the measure is specified and tested)*
Ambulatory Care: Office, Ambulatory Care: Clinic, Ambulatory Care: Hospital Outpatient, Behavioral health/psychiatric unit

2a.38-41 **Clinical Services** (*Healthcare services being measured, check all that apply*)
Behavioral Health: Mental Health, Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO), Clinicians: Psychologist/LCSW

TESTING/ANALYSIS

2b. Reliability testing

2b.1 **Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure)

2b.2 **Analytic Method** *(type of reliability & rationale, method for testing)*:
We calculated 95% confidence intervals, which speak to the precision of the rates obtained from field testing.

2b.3 **Testing Results** *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
Rate (Upper Confidence Interval, Lower Confidence Interval):
0.497 (0.42, 0.57)

2b
C☐
P☐
M☐
N☐

2c. Validity testing

2c.1 **Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure)

2c
C☐
P☐
M☐
N☐

2c.2 Analytic Method *(type of validity & rationale, method for testing)*:
NCQA tested the measure for face validity using a panel of stakeholders with specific expertise in measurement and child health care. This panel included representatives from key stakeholder groups, including pediatricians, family physicians, health plans, state Medicaid agencies and researchers. Experts reviewed the results of the field test and assessed whether the results were consistent with expectations, whether the measure represented quality care, and whether we were measuring the most important aspect of care in this area. This measure was deemed valid by the expert panel. In addition, this measure does not utilize administrative data sources; data recorded in the chart is considered the gold standard.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
NA

2d. Exclusions Justified

2d.1 Summary of Evidence supporting exclusion(s):
None

2d.2 Citations for Evidence:
NA

2d.3 Data/sample *(description of data/sample and size)*: NA

2d.4 Analytic Method *(type analysis & rationale)*:
NA

2d.5 Testing Results *(e.g., frequency, variability, sensitivity analyses)*:
NA

2d
C☐
P☐
M☐
N☐
NA☐

2e. Risk Adjustment for Outcomes/ Resource Use Measures

2e.1 Data/sample *(description of data/sample and size)*: NA

2e.2 Analytic Method *(type of risk adjustment, analysis, & rationale)*:
NA

2e.3 Testing Results *(risk model performance metrics)*:
NA

2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: The measure assesses prevention and wellness in a general population; risk adjustment is not indicated.

2e
C☐
P☐
M☐
N☐
NA☐

2f. Identification of Meaningful Differences in Performance

2f.1 Data/sample from Testing or Current Use *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure)

2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance *(type of analysis & rationale)*:
Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance

2f.3 Provide Measure Scores from Testing or Current Use *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:
Elig Population: 163
We tested several different numerator options during field test:
Documentation of a Depression Screening: 49.7
Documentation of Results of the Screening: 49.7
Documentation that Screening was Done Using a Standardized Tool: 10.4

2f
C☐
P☐
M☐
N☐

2g. Comparability of Multiple Data Sources/Methods 2g.1 **Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure) 2g.2 **Analytic Method** *(type of analysis & rationale)*: This measure is chart review only; no other sources were identified by the expert panel; this measure does not utilize administrative data 2g.3 **Testing Results** *(e.g., correlation statistics, comparison of rankings)*: NA	**2g** C☐ P☐ M☐ N☐ NA☐
2h. Disparities in Care 2h.1 **If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: The measure is not stratified to detect disparities. 2h.2 **If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** NA	**2h** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	***2*** C☐ P☐ M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	**Eval Rating**
3a. Meaningful, Understandable, and Useful Information 3a.1 **Current Use:** Not in use but testing completed 3a.2 **Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*): This measure is not currently publicly reported. NCQA is exploring the feasibility of adding this measure and its related measures into a physician-level program and/or the HEDIS® measurement set as appropriate. 3a.3 **If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):** This measure is not currently used in QI. NCQA is exploring the feasibility of adding this measure and its related measures into a physician-level program and/or the HEDIS® measurement set as appropriate. NCQA anticipates that after we release these measures, they will become widely used, as all our measures do. **Testing of Interpretability** (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) 3a.4 **Data/sample** *(description of data/sample and size)*: NA 3a.5 **Methods** *(e.g., focus group, survey, QI project)*: NCQA vetted the measures with its expert panel. In addition, throughout the development process, NCQA vetted the measure concepts and specifications with other stakeholder groups, including the National Association of State Medicaid Directors, NCQA's Health Plan Advisory Council, NCQA's Committee on Performance Measurement, and the American Academy of Pediatrician's Quality Improvement Innovation Network.	**3a** C☐ P☐ M☐ N☐

After field testing, NCQA also conducted a debrief call with field test participants. In the form of a group interview, NCQA systematically sought feedback on whether the measures were understandable, feasible, important, and had face validity. 3a.6 **Results** *(qualitative and/or quantitative results and conclusions)*: NCQA received feedback that the measure is understandable, feasible, important and valid.	
3b/3c. Relation to other NQF-endorsed measures 3b.1 **NQF # and Title of similar or related measures:**	
(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 **Are the measure specifications harmonized? If not, why?**	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value 3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** 5.1 **If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:**	**3c** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐ P☐ M☐ N☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	**Eval Rating**
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition), Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	**4a** C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) No 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.** NCQA plans to eventually adopt this measure in electronic health records.	**4b** C☐ P☐ M☐ N☐
4c. Exclusions	**4c** C☐

4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications? No **4c.2 If yes, provide justification.**	P☐ M☐ N☐ NA☐
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences **4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** During the measure development process the Child Health MAP and measure development team worked with NCQA's certified auditors and audit department to ensure that the measure specifications were clear and auditable. The denominator, numerator and optional exclusions are concisely specified and align with our audit standards.	4d C☐ P☐ M☐ N☐
4e. Data Collection Strategy/Implementation **4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:** Based on field test results, we have specified the measure to assess whether screening was documented and whether use of a standardized tool was documented. Our field test results showed that these data elements are available in the medical record. In addition, our field test participants noted that many were able to program these requirements into their electronic health record systems, and several implemented point-of-service physician reminders for this measure. **4e.2 Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** Collecting measures from medical charts is time-consuming and can be burdensome. Adapting this measure in electronic health records may relieve some of this burden. **4e.3 Evidence for costs:** Based on field test participant feedback and other stakeholder input **4e.4 Business case documentation:**	4e C☐ P☐ M☐ N☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility?*	4
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	4 C☐ P☐ M☐ N☐
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	Time-limited ☐
Steering Committee: Do you recommend for endorsement? **Comments:**	Y☐ N☐ A☐
CONTACT INFORMATION	
Co.1 Measure Steward (Intellectual Property Owner) **Co.1 Organization** National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005	

Co.2 **Point of Contact**
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-

Measure Developer If different from Measure Steward
Co.3 **Organization**
National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005

Co.4 **Point of Contact**
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-

Co.5 **Submitter If different from Measure Steward POC**
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-, National Committee for Quality Assurance

Co.6 **Additional organizations that sponsored/participated in measure development**

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 **Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.**
Child Health Measurement Advisory Panel:
Jeanne Alicandro
Barbara Dailey
Denise Dougherty, PhD
Ted Ganiats, MD
Foster Gesten, MD
Nikki Highsmith, MPA
Charlie Homer, MD, MPH
Jeff Kamil, MD
Elizabeth Siteman
Mary McIntyre, MD, MPH
Virginia Moyer, MD, MPH, FAAP
Lee Partridge
Xavier Sevilla, MD, FAAP
Michael Siegal
Jessie Sullivan

Ad.2 **If adapted, provide name of original measure:** NA
Ad.3-5 **If adapted, provide original specifications URL or attachment**

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 **Year the measure was first released:**
Ad.7 **Month and Year of most recent revision:**
Ad.8 **What is your frequency for review/update of this measure?**
Ad.9 **When is the next scheduled review/update for this measure?**

Ad.10 **Copyright statement/disclaimers:**

Ad.11 -13 **Additional Information web page URL or attachment:**

Date of Submission (*MM/DD/YY*)**:** 01/06/2011

NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 1364 **NQF Project:** Child Health Quality Measures 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Child and Adolescent Major Depressive Disorder: Diagnostic Evaluation
De.2 **Brief description of measure***:* Percentage of patients aged 6 through 17 years with a diagnosis of major depressive disorder with documented evidence that they met the DSM-IV criteria [at least 5 elements with symptom duration of two weeks or longer, including 1) depressed mood (can be irritable mood in children and adolescents) or 2) loss of interest or pleasure] during the visit in which the new diagnosis or recurrent episode was identified
1.1-2 **Type of Measure:** Process De.3 **If included in a composite or paired with another measure, please identify composite or paired measure**
De.4 **National Priority Partners Priority Area:** Population health De.5 **IOM Quality Domain**: Effectiveness, Patient-centered De.6 **Consumer Care Need:** Getting better

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** A.3 **Measure Steward Agreement:** Agreement will be signed and submitted prior to or at the time of measure submission A.4 **Measure Steward Agreement attached:**	**A** Y☐ N☐

B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	**B** Y☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement Accountability	**C** Y☐ N☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** No, testing will be completed within 12 months D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** Y☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. ***Measures must be judged to be important to measure and report*** *in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** Affects large numbers, Leading cause of morbidity/mortality, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** "Major depressive disorder (MDD) is a debilitating condition that has been increasingly recognized among youth, particularly adolescents. The prevalence of current or recent depression among children is 3% and among adolescents is 6%.1 The lifetime prevalence of MDD among adolescents may be as high as 20%.2–4 Adolescent-onset MDD is associated with an increased risk of death by suicide, suicide attempts, and recurrence of major depression by young adulthood.5–7 MDD is also associated with early pregnancy, decreased school performance, and impaired work, social, and family functioning during young adulthood.6–8" 1a.4 **Citations for Evidence of High Impact:** Williams SB, O'Connor EA, Eder M, Whitlock EP. Screening for Child and Adolescent Depression in Primary Care Settings: A Systematic Evidence Review for the US Preventive Services Task Force. Pediatrics 2009;123:e716–e735. Citing: 1. Jane Costello E, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression? J Child Psychol Psychiatry. 2006; 47(12):1263–1271 2. Lewinsohn PM, Rohde P, Seeley JR. Major depressive disorder in older adolescents: prevalence, risk factors, and clinical implications. Clin Psychol Rev. 1998;18(7):765–794 3. Cheung A. Canadian community health survey: major depressive disorder and suicidality in adolescents.	**1a** C☐ P☐ M☐ N☐

Healthc Policy. 2006; 2(2):76–89
4. Whitaker A, Johnson J, Shaffer D, et al. Uncommon troubles in young people: prevalence estimates of selected psychiatric disorders in a nonreferred adolescent population. Arch Gen Psychiatry. 1990;47(5):487–496
5. Shaffer D, Gould MS, Fisher P, et al. Psychiatric diagnosis in child and adolescent suicide. Arch Gen Psychiatry. 1996;53(4):339–348
6. Weissman MM, Wolk S, Goldstein RB, et al. Depressed adolescents grown up. JAMA. 1999;281(18):1707–1713
7. Fergusson DM, Woodward LJ. Mental health, educational, and social role outcomes of adolescents with depression. Arch Gen Psychiatry. 2002;59(3):225–231
8. Keenan-Miller D, Hammen CL, Brennan PA. Health outcomes related to early adolescent depression. J Adolesc Health. 2007; 41(3):256–262

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: Depression in children and adolescents is often underdiagnosed; one-quarter to one-half of all cases of major depressive disorders are estimated to be properly recognized by primary care and non-psychiatric practitioners. (1)(2)(3)Thorough assessment of depressive symptoms as enumerated by DSM-IV sets the basis for accurate diagnosis and treatment of major depressive disorder. Despite its importance, significant gaps in the knowledge or application of the DSM-IV criteria, even among psychiatrists exist and represent a tremendous opportunity for improvement.

(1)Kerr E. Depression, in Elizabeth McGlynn, Cheryl Damberg, Eve Kerr, and Mark Schuster (eds.), Quality of Care for Children and Adolescents: A Review of Selected Clinical Conditions and Quality Indicators, Santa Monica: RAND, 141-155, 2000.
(2) Depression Guideline Panel. Depression in Primary Care: Volume 1. Detection and Diagnosis. Clinical Practice Guideline, Number 5.AHCPR Publication No. 93-0550. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. April 1993.
(3) Katon WJ, Richardson L, Russo J, Lozano P, McCauley E. Quality of Mental Health Care for Youth With Asthma and Comorbid Anxiety and Depression. Medical Care 2006; 44:12, 1064-1072.

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
A recent survey analyzed psychiatrists' reported use of the DSM-IV criteria for MDD to diagnose depression and compared their use to the use by nonpsychiatrist physicians. Nearly one quarter of the psychiatrists indicated that they usually did not use the DSM-IV criteria when diagnosing depression while nearly half of the nonpsychiatrist physicians indicated that they rarely used the DSM-IV MDD criteria to diagnose depression.(1) A 2003 study reviewed medical records to assess the degree to which providers adhered to depression guidelines in a VA primary care setting. Providers documented review of at least five DSM-IV criteria in 46% of the records.(2)

1b.3 Citations for data on performance gap:
(1) Zimmerman M, Galione J. Psychiatrists´ and Nonpsychiatrist Physicians´ Reported Use of the DSM-IV Criteria for Major Depressive Disorder. J Clin Psychiatry. 2010;71:235-238
(2) Dobscha SK, Gerrity MS, Corson K, Bahr A, Cuilwik NM. Measuring adherence to depression treatment guidelines in a VA primary care clinic. Gen Hosp Psychiatry. 2003;25:230-7.

1b.4 Summary of Data on disparities by population group:
We are not aware of any publications/evidence outlining disparities in this area.

1b.5 Citations for data on Disparities:

1b
C☐
P☐
M☐
N☐

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** Thorough assessment of depressive symptoms as enumerated by DSM-IV sets the basis for accurate diagnosis and treatment of major

1c
C☐
P☐
M☐
N☐

depressive disorder. A variety of treatment strategies have demonstrated efficacy leading to symptomatic remission.

1c.2-3. Type of Evidence: Evidence-based guideline

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**
A diagnostic evaluation should be instituted for all patients with major depressive disorder to determine whether a diagnosis of depression is warranted and to reveal the presence of other conditions that may have an impact on treatment.

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom*)**:**

1c.6 Method for rating evidence:

1c.7 Summary of Controversy/Contradictory Evidence: None

1c.8 Citations for Evidence (*other than guidelines*)**:**

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number*)**:**
If the screening indicates significant depressive symptomatology, the clinician should perform a thorough evaluation to determine the presence of depressive and other comorbid psychiatric and medical disorders [MS]. A comprehensive psychiatric diagnostic evaluation is the single most useful tool currently available to diagnose depressive disorders.(AACAP (1))

The criteria for a major depressive disorder episode include five (or more) of nine specific symptoms which have been present during the same two-week period and represent a change from previous functioning; at least one of the symptoms is either 1) depressed mood or 2) loss of interest or pleasure. In addition, these symptoms do not meet criteria for a mixed episode (e.g., criteria for both a manic episode and for major depressive order are exhibited nearly daily). The symptoms cause clinically significant distress or impairment in social, occupations, or other important areas of functioning. The symptoms are not due to the direct physiological effects of a substance or general medical condition. The symptoms are not due to bereavement and they persist longer than two months. The symptoms may be characterized by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation. (DSM-IV (2))

In children and adolescents, an irritable or cranky mood may develop rather than a sad or dejected mood. (DSM-IV (2))

1c.10 Clinical Practice Guideline Citation: (1) American Academy of Child and Adolescent Psychiatry (AACAP). Practice parameters for the assessment and treatment of children and adolescents with depressive disorders. J. Am. Acad. Child Adolesc. Psychiatry, 2007; 46(11):1503-1526. Available at: http://www.aacap.org/galleries/PracticeParameters/Vol%2046%20Nov%202007.pdf
(2) Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition-TR (DSM-IV). American Psychiatric Association, 2000.
1c.11 National Guideline Clearinghouse or other URL: (1) http://www.guideline.gov/content.aspx?id=11404

1c.12 Rating of strength of recommendation (*also provide narrative description of the rating and by whom*)**:**
(1) Minimal Standard (MS) [see below for narrative description of the rating] (2) Not available [see below for description of revision process]

1c.13 Method for rating strength of recommendation (*If different from USPSTF system, also describe rating and how it relates to USPSTF)***:**

American Academy of Child and Adolescent Psychiatry (AACAP) Grades of Recommendations •Minimal Standard [MS] is applied to recommendations that are based on rigorous empirical evidence (such as randomized, controlled trials) and/or overwhelming clinical consensus. Minimal standards apply more than 95% of the time; i.e., in almost all cases. •Clinical Guideline [CG] is applied to recommendations that are based on strong empirical evidence (such as non-randomized control trials) and/or strong clinical consensus. Clinical guidelines apply approximately 75% of the time; i.e., in most cases. •Option [OP] is applied to recommendations that are acceptable based on emerging empirical evidence (such as uncontrolled trials or reports) or clinical opinion, but lack strong empirical evidence and/or strong clinical consensus. •Not Endorsed [NE] is applied to practices that are known to be ineffective or contraindicated. DSM-IV Revision Process: The Task Force on DSM-IV and its Work Groups conducted a three-stage empirical process that included 1) comprehensive and systematic reviews of the published literature, 2) reanalyses of already-collected data sets and 3) extensive issue-focused field trials. 1c.14 **Rationale for using this guideline over others:** It is the PCPI policy to use guidelines, which are evidence-based, applicable to physicians and other healthcare providers, and developed by a national speciality organization or government agency. In addition, the PCPI has now expanded what is acceptable as the evidence base for measures to included documented quality improvement (QI) initiatives or implementation projects that have demonstrated improvement in the quality of care.	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	**Eval Rating**
2a. MEASURE SPECIFICATIONS	
S.1 **Do you have a web page where current detailed measure specifications can be obtained?** S.2 **If yes, provide web page URL:** **2a. Precisely Specified**	
2a.1 **Numerator Statement (***Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome***):** Patients with documented evidence that they met the DSM-IV criteria [at least 5 elements with symptom duration of two weeks or longer, including 1) depressed mood (can be irritable mood in children and adolescents) or 2) loss of interest or pleasure] during the visit in which the new diagnosis or recurrent episode was identified 2a.2 **Numerator Time Window (***The time period in which cases are eligible for inclusion in the numerator***):** Once per episode (at initial evaluation) within a 12-month period 2a.3 **Numerator Details (***All information required to collect/calculate the numerator, including all codes, logic, and definitions***):** The DSM-IV Criteria for a MDD episode includes five (or more) of nine specific symptoms: - depressed mood (Note: in children and adolescents, can be irritable mood) - marked diminished interest/pleasure; - significant weight loss or gain; (Note: in children, consider failure to make expected weight gains)	**2a-specs** **C☐** **P☐** **M☐** **N☐**

- insomnia or hypersomnia;
- psychomotor agitation/ retardation;
- fatigue or lost of energy;
- feelings of worthlessness;
- diminished ability to concentrate; and
- recurrent suicidal ideation

which have been present during the same two-weeks period and represent a change from previous functioning; at least one of the symptoms is either 1) depressed mood or 2) loss of interest or pleasure.

Note: The essential feature of a major depressive disorder is a period of at least two weeks during which there is either depressed mood or irritability or the loss of interest or pleasure in nearly all activities. In children and adolescents, can be irritable or cranky mood.

2a.4 **Denominator Statement (***Brief, text description of the denominator - target population being measured***):**
All patients aged 6 through 17 years with a diagnosis of major depressive disorder

2a.5 **Target population gender:** Female, Male
2a.6 **Target population age range:** 6 through 17 years

2a.7 **Denominator Time Window (***The time period in which cases are eligible for inclusion in the denominator***):**
12 months

2a.8 **Denominator Details (***All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
See attached Level I EHR Specifications

2a.9 **Denominator Exclusions (***Brief text description of exclusions from the target population***):** None

2a.10 **Denominator Exclusion Details (***All information required to collect exclusions to the denominator, including all codes, logic, and definitions***):**

2a.11 **Stratification Details/Variables (***All information required to stratify the measure including the stratification variables, all codes, logic, and definitions***):**
Stratification by insurance coverage (commercial, Medicare and Medicaid) is recommended by some implementers

2a.12-13 **Risk Adjustment Type:** No risk adjustment necessary

2a.14 **Risk Adjustment Methodology/Variables (***List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method***):**

2a.15-17 **Detailed risk model available Web page URL or attachment:**

2a.18-19 **Type of Score:** Rate/proportion
2a.20 **Interpretation of Score:** Better quality = Higher score
2a.21 **Calculation Algorithm (***Describe the calculation of the measure as a flowchart or series of steps***):**
See attached documents

2a.22 **Describe the method for discriminating performance** *(e.g., significance testing)***:**

2a.23 **Sampling (Survey) Methodology** *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:

2a.24 **Data Source (***Check the source(s) for which the measure is specified and tested)*
Electronic Health/Medical Record

2a.25 Data source/data collection instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**

2a.26-28 Data source/data collection instrument reference web page URL or attachment:

2a.29-31 Data dictionary/code table web page URL or attachment: Attachment MDD 2 Complete.pdf

2a.32-35 Level of Measurement/Analysis (*Check the level(s) for which the measure is specified and tested)*
Clinicians: Individual, Clinicians: Group

2a.36-37 Care Settings (*Check the setting(s) for which the measure is specified and tested)*
Ambulatory Care: Office, Ambulatory Care: Clinic, Ambulatory Care: Hospital Outpatient, Behavioral health/psychiatric unit

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)
Behavioral Health: Mental Health, Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO), Clinicians: Psychologist/LCSW

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample *(description of data/sample and size)***:** Are Claims Data Accurate Enough to Identify Patients for Performance Measures or Quality Improvement? The Case of Diabetes, Heart Disease, and Depression. Leif I. Solberg, Karen I. Engebretson, Joann M. Sperl-Hillen, Mary C. Hroscikoski and Patrick J. O´Connor. American Journal of Medical Quality 2006; 21; 238.

The Challenge of Measuring Quality of Care From the Electronic Health Record. Carol P. Roth, Yee-Wei Lim, Joshua M. Pevnick, Steven M. Asch and Elizabeth A. McGlynn. American Journal of Medical Quality 2009; 24; 385 originally published online May 29, 2009.

Measuring adherence to depression treatment guidelines in a VA primary care clinic. Dobscha SK, Gerrity MS, Corson K, Bahr A, Cuilwik NM. General Hospital Psychiatry 25 (2003) 230–237

2b.2 Analytic Method *(type of reliability & rationale, method for testing)***:**
(Solberg, 2006) The objective of this study was to demonstrate a method to accurately identify patients with specific conditions from claims data for care improvement or performance measurement. Using an iterative process of trial case definitions followed by review of repeated random samples of 10 to 20 cases for newly treated depression, a final identification algorithm was created from claims files of health plan members. A final sample was used to calculate the positive predictive value (PPV).

(Roth 2009) The electronic health record (EHR) is seen by many as an ideal vehicle for measuring quality of health care and monitoring ongoing provider performance. It is anticipated that the availability of EHR-extracted data will allow quality assessment without the expensive and time-consuming process of medical record abstraction. Each quality measure was classified by the anticipated difficulty of satisfying eligibility and scoring statements using an EHR-enhanced data warehouse as the source of data. Measures were considered level 1 if all requisite data elements were accessible. Measures were considered level 2 if the denominator was accessible but the numerator was in some way inaccessible. Measures were considered level 3 if the denominator was difficult to access.

(Dobscha 2003) Researchers created one composite, measure, based on 3 national guidelines.
The DSM-IV Major depression criteria corresponds with our Diagnostic Evaluation measure.
The Evaluate level of safety/suicide history criteria corresponds with our Suicide Risk Assessment measure.
Data was analyzed for internal consistency and inter-rater reliability.

2b.3 Testing Results *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)***:**

2b
C ☐
P ☐
M ☐
N ☐

(Solberg, 2006) MDD had an unacceptably low PPV (0.65) when cases were identified on the basis of only 1 International Classification of Diseases, ninth revision, code per year. Requiring 2 outpatient ICD-9 codes or 1 inpatient ICD-9 code within 12 months (plus consideration of extra criteria for depression) resulted in PPV of 0.95. This approach is feasible and necessary for those wanting to use administrative data for case identification for performance measurement or quality improvement. The PCPI measure utilizes this approach. (Roth 2009) Accurately identifying eligible cases for quality assessment and validly scoring those cases with EHR extracted data will pose challenges but could potentially plummet the cost and therefore expand the use of quality assessment. A review of the data requirements for the depression related indicators in the Quality Assessment Tools system suggests that 41% of measures would be readily accessible from EHR data. Another 29% of the depression-related indicators have denominators that are readily accessible. Accessibility of data used to calculate the measure in an EHR reflects reliability of measure calculation. (Dobscha 2003) Inter-rater reliability was assessed, using the kappa coefficient. The Diagnosis measure (documentation of review of >= 5 DSM-IV criteria or of specific PHQ results) had a kappa = 0.83. The performance rate for this measure was 46.0% (37.0 - 55.2 95%CI).	
2c. Validity testing **2c.1 Data/sample** *(description of data/sample and size)*: **2c.2 Analytic Method** *(type of validity & rationale, method for testing)*: During measure development, the PCPI-convened expert work groups assess the face and content validity of each measure. The groups establish the measure's ability to capture what it is designed to capture using a consensus process that consists of input from multiple stakeholders, including practicing physicians and experts with technical measure expertise, as well as a review of additional input received through a PCPI public comment period. **2c.3 Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:	**2c** C☐ P☐ M☐ N☐
2d. Exclusions Justified **2d.1 Summary of Evidence supporting exclusion(s):** No Exceptions are allowed for this measure. **2d.2 Citations for Evidence:** **2d.3 Data/sample** *(description of data/sample and size)*: **2d.4 Analytic Method** *(type analysis & rationale)*: **2d.5 Testing Results** *(e.g., frequency, variability, sensitivity analyses)*:	**2d** C☐ P☐ M☐ N☐ NA☐
2e. Risk Adjustment for Outcomes/ Resource Use Measures **2e.1 Data/sample** *(description of data/sample and size)*: **2e.2 Analytic Method** *(type of risk adjustment, analysis, & rationale)*: **2e.3 Testing Results** *(risk model performance metrics)*:	**2e** C☐ P☐ M☐ N☐ NA☐

2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:	
2f. Identification of Meaningful Differences in Performance **2f.1 Data/sample from Testing or Current Use** *(description of data/sample and size)*: **2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: **2f.3 Provide Measure Scores from Testing or Current Use** *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:	**2f** C☐ P☐ M☐ N☐
2g. Comparability of Multiple Data Sources/Methods **2g.1 Data/sample** *(description of data/sample and size)*: **2g.2 Analytic Method** *(type of analysis & rationale)*: **2g.3 Testing Results** *(e.g., correlation statistics, comparison of rankings)*:	**2g** C☐ P☐ M☐ N☐ NA☐
2h. Disparities in Care **2h.1 If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: The measure is not stratified by patient groups or cohorts that could potentially be affected by disparities in care, nor are we aware of any existing research identifying disparities in care that may be relevant to this measure. **2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** We are not aware of any relevant disparities that have been identified.	**2h** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	***2*** C☐ P☐ M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	**Eval Rating**
3a. Meaningful, Understandable, and Useful Information **3a.1 Current Use:** In use **3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large)** *(If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years)*: This measure in its adult form is currently utilized in the CMS PQRI Program. **3a.3 If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):**	**3a** C☐ P☐ M☐ N☐

Testing of Interpretability (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) 3a.4 **Data/sample** *(description of data/sample and size)*: 3a.5 **Methods** *(e.g., focus group, survey, QI project)*: 3a.6 **Results** *(qualitative and/or quantitative results and conclusions)*:	
3b/3c. Relation to other NQF-endorsed measures 3b.1 **NQF # and Title of similar or related measures:** 103: Major Depressive Disorder: Diagnostic Evaluation	
(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 **Are the measure specifications harmonized? If not, why?** Yes	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value 3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** 5.1 **If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:**	**3c** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐ P☐ M☐ N☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	**Eval Rating**
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition)	**4a** C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) Yes 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.**	**4b** C☐ P☐ M☐ N☐
4c. Exclusions	**4c**

4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications? No **4c.2 If yes, provide justification.**	**C☐ P☐ M☐ N☐ NA☐**
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences **4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** We are not aware of any unintended consequences related to this measurement.	**4d** **C☐ P☐ M☐ N☐**
4e. Data Collection Strategy/Implementation **4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:** This pediatric MDD measure has a corresponding adult measure, which differs only in having an different age range. Therefore, implementation results for the adult measures are expected to be applicable to the pediatric measures. Through a partnership with the American Medical Association (AMA) and Healthcare Information and Management Systems Society (HIMSS), the Alliance of Chicago Community Health Centers developed the AHRQ-funded 3-year Enhancing Quality in Patient Care (EQUIP) project to augment its EHR implementation. This project implemented all 5 AMA-PCPI Adult MDD measures in the EHR. As part of the AHRQ-funded Effecting Change in Chronic Care: The Tipping Point project, 3 physicians implemented performance measures into existing electronic health record systems. One additional physician implemented a paper flow sheet documentation system where the flow sheet was placed in each chart at the time of the visit. This project found that the adult MDD measures were feasible to collect after the process changes were put into place. Additionally, the adult MDD version of this measure was utilized in the CMS PQRI program, in 2008, 2009, and 2010. The average performance rate for the 2008 PQRI program for the Diagnostic Evaluation measure was 86% with n=1328. **4e.2 Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** Costs to implement this specific measure have not been calculated. **4e.3 Evidence for costs:** **4e.4 Business case documentation:**	**4e** **C☐ P☐ M☐ N☐**
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility?*	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	**4** **C☐ P☐ M☐ N☐**
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited ☐**
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y☐ N☐**

A☐

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner)
Co.1 Organization
American Medical Association, 515 N State St., Chicago, Illinois, 60654

Co.2 Point of Contact
Mark, Antman, DDS, MBA, mark.antman@ama-assn.org, 312-464-5056-

Measure Developer If different from Measure Steward
Co.3 Organization
American Medical Association, 515 N State St., Chicago, Illinois, 60654

Co.4 Point of Contact
Mark, Antman, DDS, MBA, mark.antman@ama-assn.org, 312-464-5056-

Co.5 Submitter If different from Measure Steward POC
Mark, Antman, DDS, MBA, mark.antman@ama-assn.org, 312-464-5056-, American Medical Association

Co.6 Additional organizations that sponsored/participated in measure development
American Psychiatric Association, American Academy of Child and Adolescent Psychiatry

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
Boris Birmaher, MD (child/adolescent psychiatry)
Mary Dobbins, MD, FAAP (pediatrics/psychiatry)
Scott Endsley, MD, MSc (family medicine)
William E. Golden, MD, FACP (internal medicine)
Margaret L. Keeler, MD, MS, FACEP (emergency medicine)
Louis J. Kraus, MD (child/adolescent psychiatry)
Laurent S. Lehmann, MD (psychiatry)
Karen Pierce, MD (child/adolescent psychiatry)
Reed E. Pyeritz, MD, PhD, FACP, FACMG (medical genetics)
Laura Richardson, MD, MPH (internal medicine/pediatrics)
Sam J.W. Romeo, MD, MBA (family medicine)
Carl A. Sirio, MD (critical care medicine)
Sharon Sweede, MD (family medicine)
Scott Williams, PsyD (The Joint Commission)

PCPI measures are developed through cross-specialty, multi-disciplinary work groups. All medical specialties and other health care professional disciplines participating in patient care for the clinical condition or topic under study must be equal contributors to the measure development process. In addition, the PCPI strives to include on its work groups individuals representing the perspectives of patients, consumers, private health plans, and employers. This broad-based approach to measure development ensures buy-in on the measures from all stakeholders and minimizes bias toward any individual specialty or stakeholder group. All work groups have at least two co-chairs who have relevant clinical and/or measure development expertise and who are responsible for ensuring that consensus is achieved and that all perspectives are voiced.

Ad.2 If adapted, provide name of original measure:
Ad.3-5 If adapted, provide original specifications URL or attachment

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 Year the measure was first released: 2008
Ad.7 Month and Year of most recent revision: 09, 2008
Ad.8 What is your frequency for review/update of this measure? Every 3 years or as new evidence becomes available that materially affects the measures
Ad.9 When is the next scheduled review/update for this measure? 09, 2011

Ad.10 Copyright statement/disclaimers: Physician Performance Measures (Measures) and related data specifications are developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement® (PCPI).

These performance Measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications.

The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measures require a license agreement between the user and the AMA, (on behalf of the PCPI). Neither the AMA, the PCPI nor its members shall be responsible for any use of the Measures.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. The AMA, NCQA, the PCPI and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT®) or other coding contained in the specifications.

Ad.11 -13 Additional Information web page URL or attachment: Attachment NQF Aug 2010 Submission Letter-634187846588122861.pdf

Date of Submission (*MM/DD/YY*)**:** 08/30/2010

NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 1406 **NQF Project:** Child Health Quality Measures 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risky Behavior Assessment or Counseling by Age 13 Years
De.2 **Brief description of measure***:* Percentage of children with documentation of a risk assessment or counseling for risky behaviors by the age of 13 Years. Four rates are reported: Risk Assessment or Counseling for Alcohol Use, Risk Assessment or Counseling for Tobacco Use, Risk Assessment or Counseling for Other Substance Abuse, Risk Assessment or Counseling for Sexual Activity
1.1-2 **Type of Measure:** Process De.3 **If included in a composite or paired with another measure, please identify composite or paired measure** This measure appears in the composite Comprehensive Well Care by Age 13 Years
De.4 **National Priority Partners Priority Area:** Care coordination, Population health De.5 **IOM Quality Domain**: Effectiveness, Timeliness De.6 **Consumer Care Need:** Staying healthy

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** Proprietary measure A.3 **Measure Steward Agreement:** Agreement will be signed and submitted prior to or at the time of measure submission A.4 **Measure Steward Agreement attached:**	**A** Y☐ N☐

B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	**B** Y☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement Accountability	**C** Y☐ N☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** Yes, fully developed and tested D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** Y☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. ***Measures must be judged to be important to measure and report*** *in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** Leading cause of morbidity/mortality, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Adolescents are at risk for participating in risky behaviors that include sexual activity and alcohol, tobacco and substance use. Alcohol and drug abuse can have serious consequences for the user: heavy drinking increases one's risk for many forms of cancer and are connected to many injuries, abuse cases, and near-fatal and fatal accidents. Illegal drug use is connected to serious health consequences such as heart failure, convulsions, chronic sexual problems, depression, and societal costs such as increasing crime, loss of familial ties and employment. Adolescents that abuse drugs are more likely to engage in other risky behavior such as stealing, sexual intercourse, and more intense drug abuse (HHS, 2000). Nationwide, 45 percent of students had at least one alcoholic beverage in the past month; 20 percent had used marijuana one or more times in the month; seven percent had used some form of cocaine, four percent had used methamphetamine, two percent had used heroin, and eight percent had used hallucinogenic drugs one or more times in their life (CDC, 2008). The Youth Risk Behavior Surveillance national survey showed that, nationwide, 50 percent of teenagers have smoked at least one puff of a cigarette. Twenty percent of students in grades 9-12 are categorized as "currently smoking," and ten percent smoked ten or more cigarettes a day (CDC, 2008). The annual direct and indirect costs to society due to sexually transmitted diseases (STDs) and the resulting	**1a** C☐ P☐ M☐ N☐

complications are conservatively estimated at $17 billion (HHS, 2000). For example: Many unintended pregnancies receive late to no prenatal care and result in low-birth-weight infants, children with behavioral problems, and child abuse. In 1995, the nation incurred $246 billion in costs due to substance abuse to cover health care, vehicle accidents, crime, and other adverse effects. Direct costs due to tobacco use totaled at least $50 billion per year.

1a.4 Citations for Evidence of High Impact: Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics.

U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, DC: U.S. Government Printing Office, November 2000.

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: This measure promotes counseling to educate adolescents on the dangers of risky behavior (sexual activity and alcohol, tobacco and substance use). The need to prevent tobacco and other substance use early in a child's life is important. Tobacco use and addiction usually begin in adolescence. Of adults that smoke daily, 82 percent tried their first cigarette before age 18, and 53 percent became daily smokers before that age. Age of onset of drinking is connected to the amount of alcohol dependency over a lifetime: 40 percent of people that begin drinking at age 14 or under develop alcohol dependency sometime in their life compared to ten percent of those that begin at age 21 or older (CDC, 2008).

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
Of students grade 9-12 nationwide who have had sexual intercourse at least once, seven percent had sexual intercourse before they were age 13. Of the 35 percent considered sexually active, only 62 percent of students used condoms during the last encounter, and 23 percent had consumed drugs or alcohol before their last sexual encounter (CDC, 2008). Unintended pregnancies and STDs may be the consequences of this behavior. Sexually transmitted diseases remain a large national public health problem despite efforts to curb them.

Approximately one quarter of teenage girls in the United States currently have a sexually transmitted disease (STD), which suggests that an estimated 3.2 million teenagers between the ages of 14 and 19 are infected with HPV, Chlamydia, herpes or trichomoniasis. This is evidence there is a lack of STD screening and counseling in contraceptive services for teens and young women (Hampton, 2008).

In 2008, 1,210,523 Chlamydia trachomatis infection cases were reported to CDC, the largest number of cases ever reported for any condition. This is a 9.7 percent increase from 2007 (CDC, 2008).

1b.3 Citations for data on performance gap:
Centers for Disease Control and Prevention (2009, November) ´Sexually Transmitted Disease Surveillance, 2008´, Atlanta, GA: U.S. Department of Health and Human Services

Tracy Hampton. Researchers Seek Ways to Stem STDs. "Alarming" STD Rates Found in Teenaged Girls. JAMA. 2008;299(16):1888-1889.

University of Texas at Austin (2010, June 6). Adolescent brains biologically wired to engage in risky behavior, study finds. ScienceDaily. Retrieved August 26, 2010, from http://www.sciencedaily.com-/releases/2010/06/100603132458.htm.

1b.4 Summary of Data on disparities by population group:
Overall, the prevalence of sexual intercourse among students in grades nine through 12 was higher among African American and Hispanic males and females than white males and females; among African Americans and Hispanics, prevalence was higher in males than females. Prevalence of sex before age 13 was higher among males than females and higher among African American and Hispanic males and females than white males and females. Prevalence of condom use during last sexual intercourse was higher among African

1b
C☐
P☐
M☐
N☐

Americans than whites and higher among African American male than white male students (CDC, 2008). STDs disproportionally affect adolescents. Overall, women have more serious STDs than men, and African Americans and Hispanics have the highest rates of STDs (CDC, 2008).

Overall, whites and Hispanics are more likely to use alcohol and illicit drugs than African Americans (CDC 2008). Heavy episodic drinking was more common among males than females, in white males and females and Hispanics males and females than in African Americans males and females.

Males are more likely to smoke tobacco than females. American Indians or Alaska Natives are more likely to smoke than other racial/ethnic groups and Hispanics, and Asians are least likely to smoke (JAMA, 2009). Among students, frequent smoking was more common among white students in grades 9-12 (both males and females) than among African American and Hispanic males and females (CDC, 2009).

1b.5 Citations for data on Disparities:

Centers for Disease Control and Prevention. Youth Risk Behavior Surveillance — United States, 2009. Surveillance Summaries, June 4, 2010. MMWR 2010;59(No. SS-5)

State-Specific Prevalence and Trends in Adult Cigarette Smoking—United States, 1998-2007. JAMA. 2009;302(3):250-252. MMWR. 2009;58:221-226.

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** Teens engaging in one form of risk behavior, such as alcohol or drug use, will often times lead them to engage in others like unprotected sex. Unfortunately, the outcomes of taking these risks are not always discussed with the teen. Studies show that simple and brief screenings provided during regular medical visits, known as adolescent risk inventory (ARI), are an important way of identifying teens in trouble (Lifespan, 2007).

Adolescents could benefit greatly through risk behavior counseling. Primary care clinicians are able to identify those at increased risk of participating in risky behavior, including substance abuse and unsafe sexual activities. There is evidence that behavioral counseling targeted at sexually active adolescents could reduce the incidence of sexually transmitted infections (STIs). There is also no evidence of behavioral or biological harms of the counseling (Lin, Whitlock, O'Connor, Bauer, 2008). There are nearly 19 million new STIs diagnosed in the United States each year, occurring in those between the ages of 15 and 24 years.

1c.2-3. Type of Evidence: Evidence-based guideline, Expert opinion

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**
Healthy People 2010, Bright Futures, and other major bodies recommend the following risky behavior topics be discussed with adolescents: sexual activity, substance abuse, and tobacco use and cessation (HHS, 2000; Hagan et al, 2008). However, the evidence is mixed. Currently there is an abundance of evidence supporting the fact that high-intensive counseling can alter adolescent risky behavior trends, however there is not enough evidence to determine the positive outcomes that could result from a lower scale of counseling for youths and parents during regular pediatric and primary care visits.

Counseling for Sexual Activity
Good evidence suggests the effectiveness of moderate- to high-intensity behavioral counseling in reducing the incidence of overall STIs (excluding herpes simplex virus) and common bacterial STIs (such as gonorrhea and Chlamydia). However, evidence is lacking for the effectiveness of low-intensity behavioral counseling interventions, especially in lower-risk populations (Lin, Whitlock, O'Connor, Bauer, 2008).

Counseling for Substance Use, including Alcohol and Tobacco
As part of a larger risk reduction intervention among 13- to 16-year-olds and their parents, intensive counseling demonstrated decreased use of illicit drugs, though no change in alcohol use was reported. (Hagan et al, 2008).

No studies were found that addressed the effectiveness of screening for substance abuse/misuse in the

1c
C☐
P☐
M☐
N☐

primary care setting. In the school setting, mandatory drug testing among athletes decreased the use of body image-changing substances and illicit drugs, but was associated with increased risk factors that are known to be associated with drug misuse. (Hagan et al, 2008)

The USPSTF found limited evidence that screening and counseling children and adolescents in the primary care setting are effective in either preventing initiation or promoting cessation of tobacco use (USPSTF, 2003).

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**
Fair to good

1c.6 Method for rating evidence: Expert consensus

1c.7 Summary of Controversy/Contradictory Evidence: While Bright Futures and other major bodies recommend counseling adolescents on risky behavior topics, the U.S. Preventive Services Task Force concluded the evidence was insufficient to recommend for or against screening for illicit drug use and routine screening and interventions for tobacco use in adolescents. (Hagan et al, 2008)

1c.8 Citations for Evidence (*other than guidelines***):** Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics.

Jennifer S. Lin, MD, MCR; Evelyn Whitlock, MD, MPH; Elizabeth O'Connor, PhD; and Vance Bauer, MA. Behavioral Counseling to Prevent Sexually Transmitted Infections: A Systematic Review for the U.S. Preventive Services Task Force. Ann Intern Med. 2008;149:497-508.

Lifespan (2007, April 30). Teen Risk Behaviors Can Be Identified Through Simple Screening. ScienceDaily. Retrieved August 27, 2010, from http://www.sciencedaily.com- /releases/2007/04/070430102036.htm

U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, DC: U.S. Government Printing Office, November 2000.

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number***):**
Risky Behavior: Risk Reduction, Sexual Activity, Substance Abuse, and Tobacco Use

Bright Futures
Bright Futures recommends that health care providers counsel adolescents age 11-18 years on risk reduction of tobacco, alcohol or other drugs and STIs
Consensus Based

U.S. Preventive Services Task Force
The USPSTF recommends high-intensity behavioral counseling to prevent sexually transmitted infections (STIs) for all sexually active adolescents and for adults at increased risk for STIs.
Grade: B Recommendation.

The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of behavioral counseling to prevent STIs in non-sexually-active adolescents and in adults not at increased risk for STIs.
Grade: I Statement.

The USPSTF concludes that the evidence is insufficient to recommend for or against routine screening for tobacco use or interventions to prevent and treat tobacco use and dependence among children or adolescents.
Grade: I Statement.

The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms

of screening adolescents, adults, and pregnant women for illicit drug use.
Grade: I Statement.

Institute for Clinical Systems Improvement (2009)
ICSI recommends the following discussion topics on alcohol use for adolescents age 7-12:
• Reinforce alcohol abuse prevention and education.
ICSI recommends the following discussion topics for adolescents age 13 and older:
Don´t ride with someone who is under the influence of alcohol.
• Prevent others from driving in this condition: "Friends don´t let friends drive drunk."
• Reinforce not drinking and driving, and the dangers of it.
-Abstinence if driving
-Have a designated driver
• Discuss characteristics of dependency.
• Assess current use of alcohol (by history and/or use of standardized screening questionnaire).
• Advise all females of the harm of alcohol on a fetus, and advise them to limit or cease alcohol intake.
Level III

ICSI recommends the following discussion topics on sexual activity for adolescents age 12 and older, or earlier if sexually active
• Obtain a sexual history from all adolescents.
• Inform adolescents that abstinence is the most effective way to prevent pregnancy and sexually transmitted infections.
• Provide detailed education and written information regarding all contraceptive methods including barrier contraceptives, birth control pills, injectables, implantables, tubal sterilization and vasectomy. Longer-duration methods may improve compliance and efficacy.
• To enhance acceptance of contraceptive methods, health benefits should be discussed:
- Use of oral contraceptives will reduce lifetime risks of ovarian and uterine cancer.
- Use of barrier contraceptives and spermicides will reduce the risk of developing cervical cancer and sexually
transmitted infections.
These messages should also be given as indicated by clinical discretion (e.g., genitourinary symptoms).
Grade: Level III

Bright Futures (2008)
Bright Futures recommends the following topics about sexual activity for adolescents age 11-18 years
At every visit: talk to parent and adolescent: abstinence for those who have not had sex, and as an option to those who are sexually experienced, is the best protection from pregnancy, STIs, and the emotional distress
Provide information and/or role-play on how to resist peer pressure to smoke, drink alcohol, or use drugs
Administer alcohol and drug screening tool
Grade: Expert consensus

AAFP
• Risks for sexually transmitted diseases and how to prevent them.
• Effective sexuality education, pregnancy prevention and sexually transmitted disease prevention programs as those using a comprehensive approach to sexuality education that includes medically accurate information on contraception and abstinence.
• Stress abstinence which, when practiced consistently, is the most effective method of preventing unplanned pregnancy and the transmission of sexually transmitted disease(s).
• Responsible sexual behavior is also an effective method of preventing pregnancy and sexually transmitted diseases.
• Adolescents receiving contraceptive services should be accorded strict patient confidentiality
Work to prevent unintended teenage pregnancies and prevention of STDs, by providing appropriate guidance/ counseling and effective sex education to their adolescent patient population.

1c.10 **Clinical Practice Guideline Citation:** Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics

U.S. Preventive Services Task Force. Behavioral counseling to prevent sexually transmitted infections: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2008 Oct 7;149(7):491-6, W95.

Institute for Clinical Systems Improvement. Health care guideline: Preventive Services for Children and Adolescents. Fifteenth Edition. October 2009.

AAFP. Substance and Alcohol Abuse and Addiction. American Academy of Family Physicians. 2003. http://www.aafp.org/online/en/home/policy/policies/s/substanceabuse.html

Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics Screening and behavioral counseling interventions in primary care to reduce alcohol misuse: recommendation statement. Ann Intern Med 2004 Apr 6;140(7):554-6.

AAP. Kulig JW. Tobacco, alcohol, and other drugs: the role of the pediatrician in prevention, identification, and management of substance abuse. Pediatrics 2005 Mar;115(3):816-21.

1c.11 National Guideline Clearinghouse or other URL: Behavioral counseling to prevent sexually transmitted infections: U.S. Preventive Services Task Force recommendation statement. 1996 (revised 2008 Oct). NGC:006686 http://www.guideline.gov/content.aspx?id=12990&search=at+risk+adolescents

1c.12 Rating of strength of recommendation (*also provide narrative description of the rating and by whom*)**:**
Fair to good

1c.13 Method for rating strength of recommendation (*If different from USPSTF system, also describe rating and how it relates to USPSTF)*:
USPSTF based

1c.14 Rationale for using this guideline over others:
Healthy People 2010, Bright Futures, and other major bodies recommend the following risky behavior topics be discussed with adolescents: sexual activity, substance abuse, and tobacco use and cessation. Based on expert feedback, we based the measure on these guidelines and the body of evidence.

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	**Eval Rating**
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? **S.2 If yes, provide web page URL:** **2a. Precisely Specified**	
2a.1 Numerator Statement (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome*)**:** Children with documentation of a risk assessment or counseling for risky behaviors by the age of 13 Years **2a.2 Numerator Time Window** (*The time period in which cases are eligible for inclusion in the numerator*)**:** 2 years	**2a-specs** **C☐** **P☐** **M☐** **N☐**

2a.3 **Numerator Details** (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):
Documentation must include a note indicating the date and that the provider asked or counseled about the following.
- Sexual activity
- Substance use
- Alcohol use
- Tobacco use

Counseling is any of the following.
- Engagement in discussion of current risky behaviors (e.g., sexual activity or substance use)
- Checklist indicating that risky behavior was addressed
- Counseling or referral for risky behavior education
- Member received educational materials on risky behavior
- Anticipatory guidance for risky behavior

2a.4 **Denominator Statement** (*Brief, text description of the denominator - target population being measured*):
Children with a visit who turned 13 years old in the measurement year

2a.5 **Target population gender:** Female, Male
2a.6 **Target population age range:** 11 years-13 years

2a.7 **Denominator Time Window** (*The time period in which cases are eligible for inclusion in the denominator*):
1 year

2a.8 **Denominator Details** (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):
Children who turned 13 years of age between January 1 of the measurement year and December 31 of the measurement year and who had documentation of a face-to-face visit between the clinician and the child that predates the child's birthday by at least 12 months.

2a.9 **Denominator Exclusions** (*Brief text description of exclusions from the target population*): None

2a.10 **Denominator Exclusion Details** (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*):
NA

2a.11 **Stratification Details/Variables** (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*):
The measure is not stratified

2a.12-13 **Risk Adjustment Type:** No risk adjustment necessary

2a.14 **Risk Adjustment Methodology/Variables** (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*):
NA

2a.15-17 **Detailed risk model available Web page URL or attachment:**

2a.18-19 **Type of Score:** Rate/proportion
2a.20 **Interpretation of Score:** Better quality = Higher score
2a.21 **Calculation Algorithm** (*Describe the calculation of the measure as a flowchart or series of steps*):
Step 1: Determine the denominator
Children who turned the requisite age in the measurement year, AND
Who had a visit within the past 12 months of the child´s birthday
Step 2: Determine the numerator
Children who had documentation in the medical record of the service during the measurement year or the year previous to the measurement year.

2a.22 **Describe the method for discriminating performance** *(e.g., significance testing)*:

Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance	
2a.23 Sampling (Survey) Methodology *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*: For this physician-level measure, we anticipate the entire population will be used in the denominator. If a sample is used, a random sample is ideal. NCQA's work has indicated that a sample size of 30-50 patients would be necessary for a typical practice size of 2000 patients.	
2a.24 Data Source (*Check the source(s) for which the measure is specified and tested)* Paper medical record/flow-sheet, Electronic clinical data, Electronic Health/Medical Record **2a.25 Data source/data collection instrument** (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*)**:** Medical Record **2a.26-28 Data source/data collection instrument reference web page URL or attachment:** **2a.29-31 Data dictionary/code table web page URL or attachment:** **2a.32-35 Level of Measurement/Analysis** (*Check the level(s) for which the measure is specified and tested)* Clinicians: Individual, Clinicians: Group, Population: national, Population: regional/network **2a.36-37 Care Settings** (*Check the setting(s) for which the measure is specified and tested)* Ambulatory Care: Office, Ambulatory Care: Clinic, Behavioral health/psychiatric unit **2a.38-41 Clinical Services** (*Healthcare services being measured, check all that apply*) Behavioral Health: Mental Health, Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO)	
TESTING/ANALYSIS	
2b. Reliability testing **2b.1 Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure) **2b.2 Analytic Method** *(type of reliability & rationale, method for testing)*: We calculated 95% confidence intervals, which speak to the precision of the rates obtained from field testing. **2b.3 Testing Results** *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*: Rate (Upper Confidence Interval, Lower Confidence Interval): Risk Assessment/Counseling for Alcohol Use Rate: 0.737 (0.67, 0.80) Risk Assessment/Counseling for Sexual Activity Rate: 0.704 (0.64, 0.77) Risk Assessment/Counseling for Substance Use Rate: 0.715 (0.65, 0.78) Risk Assessment/Counseling for Tobacco Use Rate: 0.777 (0.72, 0.84)	**2b** C☐ P☐ M☐ N☐
2c. Validity testing **2c.1 Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure) **2c.2 Analytic Method** *(type of validity & rationale, method for testing)*: NCQA tested the measure for face validity using a panel of stakeholders with specific expertise in measurement and child health care. This panel included representatives from key stakeholder groups, including pediatricians, family physicians, health plans, state Medicaid agencies and researchers. Experts reviewed the results of the field test and assessed whether the results were consistent with expectations,	**2c** C☐ P☐ M☐ N☐

whether the measure represented quality care, and whether we were measuring the most important aspect of care in this area. **2c.3 Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*: This measure was deemed valid by the expert panel. In addition, this measure does not utilize administrative data sources; data recorded in the chart is considered the gold standard.	
2d. Exclusions Justified **2d.1 Summary of Evidence supporting exclusion(s):** No exclusions **2d.2 Citations for Evidence:** NA **2d.3 Data/sample** *(description of data/sample and size)*: NA **2d.4 Analytic Method** *(type analysis & rationale)*: NA **2d.5 Testing Results** *(e.g., frequency, variability, sensitivity analyses)*: NA	**2d** C☐ P☐ M☐ N☐ NA☐
2e. Risk Adjustment for Outcomes/ Resource Use Measures **2e.1 Data/sample** *(description of data/sample and size)*: NA **2e.2 Analytic Method** *(type of risk adjustment, analysis, & rationale)*: NA **2e.3 Testing Results** *(risk model performance metrics)*: NA **2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:** The measure assesses prevention and wellness in a general population; risk adjustment is not indicated.	**2e** C☐ P☐ M☐ N☐ NA☐
2f. Identification of Meaningful Differences in Performance **2f.1 Data/sample from Testing or Current Use** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure) **2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance **2f.3 Provide Measure Scores from Testing or Current Use** *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*: Below is eligible population. The eligible population applies to all four rates. Eligible = 179 Below are performance rates for each rate: Rate 1: Sexual Activity 70% Rate 2: Substance Use 72% Rate 3: Alcohol Use 74%	**2f** C☐ P☐ M☐ N☐

Rate 4: Tobacco Use 78%	
2g. Comparability of Multiple Data Sources/Methods 2g.1 **Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure) 2g.2 **Analytic Method** *(type of analysis & rationale)*: This measure is chart review only; no other sources were identified by the expert panel; this measure does not utilize administrative data 2g.3 **Testing Results** *(e.g., correlation statistics, comparison of rankings)*: NA	**2g** C☐ P☐ M☐ N☐ NA☐
2h. Disparities in Care 2h.1 **If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: The measure is not stratified to detect disparities. 2h.2 **If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** NA	**2h** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	***2*** C☐ P☐ M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	**Eval Rating**
3a. Meaningful, Understandable, and Useful Information 3a.1 **Current Use:** Not in use but testing completed 3a.2 **Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*): This measure is not currently publicly reported. NCQA is exploring the feasibility of adding this measure and its related measures into a physician-level program and/or the HEDIS® measurement set as appropriate. 3a.3 **If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):** This measure is not currently used in QI. NCQA is exploring the feasibility of adding this measure and its related measures into a physician-level program and/or the HEDIS® measurement set as appropriate. NCQA anticipates that after we release these measures, they will become widely used, as all our measures do. **Testing of Interpretability** (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) 3a.4 **Data/sample** *(description of data/sample and size)*: Expert panel, other stakeholders, and 19 physician field test participants 3a.5 **Methods** *(e.g., focus group, survey, QI project)*: NCQA vetted the measures with its expert panel. In addition, throughout the development process, NCQA vetted the measure concepts and specifications with other stakeholder groups, including the National	**3a** C☐ P☐ M☐ N☐

Association of State Medicaid Directors, NCQA's Health Plan Advisory Council, NCQA's Committee on Performance Measurement, and the American Academy of Pediatrician's Quality Improvement Innovation Network. After field testing, NCQA also conducted a debrief call with field test participants. In the form of a group interview, NCQA systematically sought feedback on whether the measures were understandable, feasible, important, and had face validity. 3a.6 **Results** *(qualitative and/or quantitative results and conclusions)*: NCQA received feedback that the measure is understandable, feasible, important and valid.	
3b/3c. Relation to other NQF-endorsed measures 3b.1 **NQF # and Title of similar or related measures:**	
(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 **Are the measure specifications harmonized? If not, why?**	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value 3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** 5.1 **If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:** NA	**3c** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐ P☐ M☐ N☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	**Eval Rating**
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition), Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	**4a** C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) No 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.**	**4b** C☐ P☐ M☐ N☐

NCQA plans to eventually adapt this measure for use in electronic health records.	
4c. Exclusions **4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?** No **4c.2 If yes, provide justification.**	**4c** C☐ P☐ M☐ N☐ NA☐
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences **4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** During the measure development process the Child Health MAP and measure development team worked with NCQA's certified auditors and audit department to ensure that the measure specifications were clear and auditable. The denominator, numerator and any exclusions are concisely specified and align with our audit standards.	**4d** C☐ P☐ M☐ N☐
4e. Data Collection Strategy/Implementation **4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:** Based on field test results, we have specified the measure to assess whether physicians assessed OR counseled adolescents on the four risky behavior topics. Our field test results showed that these data elements are available in the medical record. In addition, our field test participants noted that many were able to program these requirements into their electronic health record systems, and several implemented point-of-service physician reminders for this measure. **4e.2 Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** Collecting measures from medical charts is time-consuming and can be burdensome. Adapting this measure in electronic health records may relieve some of this burden. **4e.3 Evidence for costs:** Based on field test participant feedback and other stakeholder input. **4e.4 Business case documentation:**	**4e** C☐ P☐ M☐ N☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility?*	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	**4** C☐ P☐ M☐ N☐
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited** ☐
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y**☐ **N**☐ **A**☐
CONTACT INFORMATION	

Co.1 Measure Steward (Intellectual Property Owner)
Co.1 Organization
National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005

Co.2 Point of Contact
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-

Measure Developer If different from Measure Steward
Co.3 Organization
National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005

Co.4 Point of Contact
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-

Co.5 Submitter If different from Measure Steward POC
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-, National Committee for Quality Assurance

Co.6 Additional organizations that sponsored/participated in measure development

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
Child Health Measurement Advisory Panel:
Jeanne Alicandro
Barbara Dailey
Denise Dougherty, PhD
Ted Ganiats, MD
Foster Gesten, MD
Nikki Highsmith, MPA
Charlie Homer, MD, MPH
Jeff Kamil, MD
Elizabeth Siteman
Mary McIntyre, MD, MPH
Virginia Moyer, MD, MPH, FAAP
Lee Partridge
Xavier Sevilla, MD, FAAP
Michael Siegal
Jessie Sullivan

Ad.2 If adapted, provide name of original measure: NA
Ad.3-5 If adapted, provide original specifications URL or attachment

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 Year the measure was first released:
Ad.7 Month and Year of most recent revision:
Ad.8 What is your frequency for review/update of this measure?
Ad.9 When is the next scheduled review/update for this measure?

Ad.10 Copyright statement/disclaimers:

Ad.11 -13 Additional Information web page URL or attachment:

Date of Submission (*MM/DD/YY*)**:** 01/06/2011

NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 1507 **NQF Project:** Child Health Quality Measures 2010

MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risky Behavior Assessment or Counseling by Age 18 Years
De.2 **Brief description of measure:** Percentage of children with documentation of assessment or counseling for risky behavior. Four rates are reported: assessment or counseling for alcohol use, tobacco use, other substance use, and sexual activity.
1.1-2 **Type of Measure:** Process De.3 **If included in a composite or paired with another measure, please identify composite or paired measure** This measure appears in the composite Comprehensive Well Care by Age 18 Years.
De.4 **National Priority Partners Priority Area:** Care coordination, Population health De.5 **IOM Quality Domain**: Effectiveness, Timeliness De.6 **Consumer Care Need:** Staying healthy

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** Proprietary measure A.3 **Measure Steward Agreement:** Agreement will be signed and submitted prior to or at the time of measure submission A.4 **Measure Steward Agreement attached:**	A Y☐ N☐

B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	**B** Y☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement Accountability	**C** Y☐ N☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** Yes, fully developed and tested D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** Y☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. ***Measures must be judged to be important to measure and report*** *in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** Leading cause of morbidity/mortality, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Adolescents are at risk for participating in risky behaviors that include sexual activity and alcohol, tobacco and substance use. Alcohol and drug abuse can have serious consequences for the user: heavy drinking increases one's risk for many forms of cancer and are connected to many injuries, abuse cases, and near-fatal and fatal accidents. Illegal drug use is connected to serious health consequences such as heart failure, convulsions, chronic sexual problems, depression, and societal costs such as increasing crime, loss of familial ties and employment. Adolescents that abuse drugs are more likely to engage in other risky behavior such as stealing, sexual intercourse, and more intense drug abuse (HHS, 2000). Nationwide, 45 percent of students had at least one alcoholic beverage in the past month; 20 percent had used marijuana one or more times in the month; seven percent had used some form of cocaine, four percent had used methamphetamine, two percent had used heroin, and eight percent had used hallucinogenic drugs one or more times in their life (CDC, 2008). The Youth Risk Behavior Surveillance national survey showed that, nationwide, 50 percent of teenagers have smoked at least one puff of a cigarette. Twenty percent of students in grades 9-12 are categorized as "currently smoking," and ten percent smoked ten or more cigarettes a day (CDC, 2008). The annual direct and indirect costs to society due to sexually transmitted diseases (STDs) and the resulting	**1a** C☐ P☐ M☐ N☐

complications are conservatively estimated at $17 billion (HHS, 2000). For example: Many unintended pregnancies receive late to no prenatal care and result in low-birth-weight infants, children with behavioral problems, and child abuse. In 1995, the nation incurred $246 billion in costs due to substance abuse to cover health care, vehicle accidents, crime, and other adverse effects. Direct costs due to tobacco use totaled at least $50 billion per year.

1a.4 Citations for Evidence of High Impact: Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics.

U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, DC: U.S. Government Printing Office, November 2000.

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: This measure promotes counseling to educate adolescents on the dangers of risky behavior (sexual activity and alcohol, tobacco and substance use). The need to prevent tobacco and other substance use early in a child's life is important. Tobacco use and addiction usually begin in adolescence. Of adults that smoke daily, 82 percent tried their first cigarette before age 18, and 53 percent became daily smokers before that age. Age of onset of drinking is connected to the amount of alcohol dependency over a lifetime: 40 percent of people that begin drinking at age 14 or under develop alcohol dependency sometime in their life compared to ten percent of those that begin at age 21 or older (CDC, 2008).

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
Of students grade 9-12 nationwide who have had sexual intercourse at least once, seven percent had sexual intercourse before they were age 13. Of the 35 percent considered sexually active, only 62 percent of students used condoms during the last encounter, and 23 percent had consumed drugs or alcohol before their last sexual encounter (CDC, 2008). Unintended pregnancies and STDs may be the consequences of this behavior. Sexually transmitted diseases remain a large national public health problem despite efforts to curb them.

Approximately one quarter of teenage girls in the United States currently have a sexually transmitted disease (STD), which suggests that an estimated 3.2 million teenagers between the ages of 14 and 19 are infected with HPV, Chlamydia, herpes or trichomoniasis. This is evidence there is a lack of STD screening and counseling in contraceptive services for teens and young women (Hampton, 2008).

In 2008, 1,210,523 Chlamydia trachomatis infection cases were reported to CDC, the largest number of cases ever reported for any condition. This is a 9.7 percent increase from 2007 (CDC, 2008).

1b.3 Citations for data on performance gap:
Centers for Disease Control and Prevention (2009, November) ´Sexually Transmitted Disease Surveillance, 2008´, Atlanta, GA: U.S. Department of Health and Human Services

Tracy Hampton. Researchers Seek Ways to Stem STDs. "Alarming" STD Rates Found in Teenaged Girls. JAMA. 2008;299(16):1888-1889.

University of Texas at Austin (2010, June 6). Adolescent brains biologically wired to engage in risky behavior, study finds. ScienceDaily. Retrieved August 26, 2010, from http://www.sciencedaily.com-/releases/2010/06/100603132458.htm.

1b.4 Summary of Data on disparities by population group:
Overall, the prevalence of sexual intercourse among students in grades nine through 12 was higher among African American and Hispanic males and females than white males and females; among African Americans and Hispanics, prevalence was higher in males than females. Prevalence of sex before age 13 was higher among males than females and higher among African American and Hispanic males and females than white males and females. Prevalence of condom use during last sexual intercourse was higher among African

1b
C☐
P☐
M☐
N☐

Americans than whites and higher among African American male than white male students (CDC, 2008). STDs disproportionally affect adolescents. Overall, women have more serious STDs than men, and African Americans and Hispanics have the highest rates of STDs (CDC, 2008).

Overall, whites and Hispanics are more likely to use alcohol and illicit drugs than African Americans (CDC 2008). Heavy episodic drinking was more common among males than females, in white males and females and Hispanics males and females than in African Americans males and females.

Males are more likely to smoke tobacco than females. American Indians or Alaska Natives are more likely to smoke than other racial/ethnic groups and Hispanics, and Asians are least likely to smoke (JAMA, 2009). Among students, frequent smoking was more common among white students in grades 9-12 (both males and females) than among African American and Hispanic males and females (CDC, 2009).

1b.5 Citations for data on Disparities:

Centers for Disease Control and Prevention. Youth Risk Behavior Surveillance — United States, 2009. Surveillance Summaries, June 4, 2010. MMWR 2010;59(No. SS-5)

State-Specific Prevalence and Trends in Adult Cigarette Smoking—United States, 1998-2007. JAMA. 2009;302(3):250-252. MMWR. 2009;58:221-226.

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** Teens engaging in one form of risk behavior, such as alcohol or drug use, will often times lead them to engage in others like unprotected sex. Unfortunately, the outcomes of taking these risks are not always discussed with the teen. Studies show that simple and brief screenings provided during regular medical visits, known as adolescent risk inventory (ARI), are an important way of identifying teens in trouble (Lifespan, 2007).

Adolescents could benefit greatly through risk behavior counseling. Primary care clinicians are able to identify those at increased risk of participating in risky behavior, including substance abuse and unsafe sexual activities. There is evidence that behavioral counseling targeted at sexually active adolescents could reduce the incidence of sexually transmitted infections (STIs). There is also no evidence of behavioral or biological harms of the counseling (Lin, Whitlock, O'Connor, Bauer, 2008). There are nearly 19 million new STIs diagnosed in the United States each year, occurring in those between the ages of 15 and 24 years.

1c.2-3. Type of Evidence: Evidence-based guideline, Expert opinion

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**
Healthy People 2010, Bright Futures, and other major bodies recommend the following risky behavior topics be discussed with adolescents: sexual activity, substance abuse, and tobacco use and cessation (HHS, 2000; Hagan et al, 2008). However, the evidence is mixed. Currently there is an abundance of evidence supporting the fact that high-intensive counseling can alter adolescent risky behavior trends, however there is not enough evidence to determine the positive outcomes that could result from a lower scale of counseling for youths and parents during regular pediatric and primary care visits.

Counseling for Sexual Activity
Good evidence suggests the effectiveness of moderate- to high-intensity behavioral counseling in reducing the incidence of overall STIs (excluding herpes simplex virus) and common bacterial STIs (such as gonorrhea and Chlamydia). However, evidence is lacking for the effectiveness of low-intensity behavioral counseling interventions, especially in lower-risk populations (Lin, Whitlock, O'Connor, Bauer, 2008).

Counseling for Substance Use, including Alcohol and Tobacco
As part of a larger risk reduction intervention among 13- to 16-year-olds and their parents, intensive counseling demonstrated decreased use of illicit drugs, though no change in alcohol use was reported. (Hagan et al, 2008).

No studies were found that addressed the effectiveness of screening for substance abuse/misuse in the

1c
C☐
P☐
M☐
N☐

primary care setting. In the school setting, mandatory drug testing among athletes decreased the use of body image-changing substances and illicit drugs, but was associated with increased risk factors that are known to be associated with drug misuse. (Hagan et al, 2008)

The USPSTF found limited evidence that screening and counseling children and adolescents in the primary care setting are effective in either preventing initiation or promoting cessation of tobacco use (USPSTF, 2003).

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**
Fair to good

1c.6 Method for rating evidence: Expert consensus

1c.7 Summary of Controversy/Contradictory Evidence: While Bright Futures and other major bodies recommend counseling adolescents on risky behavior topics, the U.S. Preventive Services Task Force concluded the evidence was insufficient to recommend for or against screening for illicit drug use and routine screening and interventions for tobacco use in adolescents. (Hagan et al, 2008)

1c.8 Citations for Evidence (*other than guidelines***):** Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics.

Jennifer S. Lin, MD, MCR; Evelyn Whitlock, MD, MPH; Elizabeth O'Connor, PhD; and Vance Bauer, MA. Behavioral Counseling to Prevent Sexually Transmitted Infections: A Systematic Review for the U.S. Preventive Services Task Force. Ann Intern Med. 2008;149:497-508.

Lifespan (2007, April 30). Teen Risk Behaviors Can Be Identified Through Simple Screening. ScienceDaily. Retrieved August 27, 2010, from http://www.sciencedaily.com- /releases/2007/04/070430102036.htm

U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, DC: U.S. Government Printing Office, November 2000.

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number***):**
Risky Behavior: Risk Reduction, Sexual Activity, Substance Abuse, and Tobacco Use

Bright Futures
Bright Futures recommends that health care providers counsel adolescents age 11-18 years on risk reduction of tobacco, alcohol or other drugs and STIs
Consensus Based

U.S. Preventive Services Task Force
The USPSTF recommends high-intensity behavioral counseling to prevent sexually transmitted infections (STIs) for all sexually active adolescents and for adults at increased risk for STIs.
Grade: B Recommendation.

The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of behavioral counseling to prevent STIs in non-sexually-active adolescents and in adults not at increased risk for STIs.
Grade: I Statement.

The USPSTF concludes that the evidence is insufficient to recommend for or against routine screening for tobacco use or interventions to prevent and treat tobacco use and dependence among children or adolescents.
Grade: I Statement.

The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms

of screening adolescents, adults, and pregnant women for illicit drug use.
Grade: I Statement.

Institute for Clinical Systems Improvement (2009)
ICSI recommends the following discussion topics on alcohol use for adolescents age 7-12:
• Reinforce alcohol abuse prevention and education.
ICSI recommends the following discussion topics for adolescents age 13 and older:
Don´t ride with someone who is under the influence of alcohol.
• Prevent others from driving in this condition: "Friends don´t let friends drive drunk."
• Reinforce not drinking and driving, and the dangers of it.
-Abstinence if driving
-Have a designated driver
• Discuss characteristics of dependency.
• Assess current use of alcohol (by history and/or use of standardized screening questionnaire).
• Advise all females of the harm of alcohol on a fetus, and advise them to limit or cease alcohol intake.
Level III

ICSI recommends the following discussion topics on sexual activity for adolescents age 12 and older, or earlier if sexually active
• Obtain a sexual history from all adolescents.
• Inform adolescents that abstinence is the most effective way to prevent pregnancy and sexually transmitted infections.
• Provide detailed education and written information regarding all contraceptive methods including barrier contraceptives, birth control pills, injectables, implantables, tubal sterilization and vasectomy. Longer-duration methods may improve compliance and efficacy.
• To enhance acceptance of contraceptive methods, health benefits should be discussed:
- Use of oral contraceptives will reduce lifetime risks of ovarian and uterine cancer.
- Use of barrier contraceptives and spermicides will reduce the risk of developing cervical cancer and sexually
transmitted infections.
These messages should also be given as indicated by clinical discretion (e.g., genitourinary symptoms).
Grade: Level III

Bright Futures (2008)
Bright Futures recommends the following topics about sexual activity for adolescents age 11-18 years
At every visit: talk to parent and adolescent: abstinence for those who have not had sex, and as an option to those who are sexually experienced, is the best protection from pregnancy, STIs, and the emotional distress
Provide information and/or role-play on how to resist peer pressure to smoke, drink alcohol, or use drugs
Administer alcohol and drug screening tool
Grade: Expert consensus

AAFP
• Risks for sexually transmitted diseases and how to prevent them.
• Effective sexuality education, pregnancy prevention and sexually transmitted disease prevention programs as those using a comprehensive approach to sexuality education that includes medically accurate information on contraception and abstinence.
• Stress abstinence which, when practiced consistently, is the most effective method of preventing unplanned pregnancy and the transmission of sexually transmitted disease(s).
• Responsible sexual behavior is also an effective method of preventing pregnancy and sexually transmitted diseases.
• Adolescents receiving contraceptive services should be accorded strict patient confidentiality
Work to prevent unintended teenage pregnancies and prevention of STDs, by providing appropriate guidance/ counseling and effective sex education to their adolescent patient population.

1c.10 **Clinical Practice Guideline Citation:** Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics

U.S. Preventive Services Task Force. Behavioral counseling to prevent sexually transmitted infections: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2008 Oct 7;149(7):491-6, W95.

Institute for Clinical Systems Improvement. Health care guideline: Preventive Services for Children and Adolescents. Fifteenth Edition. October 2009.

AAFP. Substance and Alcohol Abuse and Addiction. American Academy of Family Physicians. 2003. http://www.aafp.org/online/en/home/policy/policies/s/substanceabuse.html

Hagan, JF, Shaw JS, Duncan PM, eds. 2008. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents, Third Edition. Elk Grove, IL: American Academy of Pediatrics Screening and behavioral counseling interventions in primary care to reduce alcohol misuse: recommendation statement. Ann Intern Med 2004 Apr 6;140(7):554-6.

AAP. Kulig JW. Tobacco, alcohol, and other drugs: the role of the pediatrician in prevention, identification, and management of substance abuse. Pediatrics 2005 Mar;115(3):816-21.

1c.11 National Guideline Clearinghouse or other URL: Behavioral counseling to prevent sexually transmitted infections: U.S. Preventive Services Task Force recommendation statement. 1996 (revised 2008 Oct). NGC:006686 http://www.guideline.gov/content.aspx?id=12990&search=at+risk+adolescents

1c.12 Rating of strength of recommendation (*also provide narrative description of the rating and by whom*)**:**
Fair to good

1c.13 Method for rating strength of recommendation (*If different from USPSTF system, also describe rating and how it relates to USPSTF)*:
USPSTF based

1c.14 Rationale for using this guideline over others:
Healthy People 2010, Bright Futures, and other major bodies recommend the following risky behavior topics be discussed with adolescents: sexual activity, substance abuse, and tobacco use and cessation. Based on expert feedback, we based the measure on these guidelines and the body of evidence.

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	**Eval Rating**
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? **S.2 If yes, provide web page URL:** **2a. Precisely Specified**	
2a.1 Numerator Statement (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome*)**:** Children who had documentation in the medical record of a Risky Behavior Assessment or Counseling By Age 18 Years **2a.2 Numerator Time Window** (*The time period in which cases are eligible for inclusion in the numerator*)**:** 2 years	**2a-specs** **C☐** **P☐** **M☐** **N☐**

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):
Documentation must include a note indicating the date and that the provider asked or counseled about the following. Report each rate separately.
- Sexual activity
- Substance use
- Alcohol use
- Tobacco use

Counseling is any of the following.
- Engagement in discussion of current risky behaviors (e.g., sexual activity or substance use)
- Checklist indicating that risky behavior was addressed
- Counseling or referral for risky behavior education
- Member received educational materials on risky behavior
- Anticipatory guidance for risky behavior

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*):
Children with a visit who turned 18 years of age in the measurement year

2a.5 Target population gender: Female, Male
2a.6 Target population age range: 16 years-18 years

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator*):
1 year

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):
Children who turned 18 years of age between January 1 of the measurement year and December 31 of the measurement year and who had documentation of a face-to-face visit between the clinician and the child that predates the child's birthday by at least 12 months.

2a.9 Denominator Exclusions (*Brief text description of exclusions from the target population*): None

2a.10 Denominator Exclusion Details (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*):
NA

2a.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*):
The measure is not stratified

2a.12-13 Risk Adjustment Type: No risk adjustment necessary

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*):
NA

2a.15-17 Detailed risk model available Web page URL or attachment:

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Higher score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps*):
Step 1: Determine the denominator
Children who turned the requisite age in the measurement year, AND
Who had a visit within the past 12 months of the child´s birthday
Step 2: Determine the numerator
Children who had documentation in the medical record of the screening or service during the measurement year or the year previous to the measurement year.

2a.22 Describe the method for discriminating performance *(e.g., significance testing)*:
Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance

2a.23 Sampling (Survey) Methodology *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:
For this physician-level measure, we anticipate the entire population will be used in the denominator. If a sample is used, a random sample is ideal. NCQA's work has indicated that a sample size of 30-50 patients would be necessary for a typical practice size of 2000 patients.

2a.24 Data Source (*Check the source(s) for which the measure is specified and tested)*
Paper medical record/flow-sheet, Electronic clinical data, Electronic Health/Medical Record

2a.25 Data source/data collection instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**
Medical Record

2a.26-28 Data source/data collection instrument reference web page URL or attachment:

2a.29-31 Data dictionary/code table web page URL or attachment:

2a.32-35 Level of Measurement/Analysis (*Check the level(s) for which the measure is specified and tested)*
Clinicians: Individual, Clinicians: Group, Population: national, Population: regional/network

2a.36-37 Care Settings (*Check the setting(s) for which the measure is specified and tested)*
Ambulatory Care: Office, Ambulatory Care: Clinic, Behavioral health/psychiatric unit

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)
Behavioral Health: Mental Health, Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO)

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure)

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:
We calculated 95% confidence intervals, which speak to the precision of the rates obtained from field testing.

2b.3 Testing Results *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
Performance Rate (Upper Confidence Interval, Lower Confidence Interval):
Risk Assessment/Counseling for Alochol Use Rate: 0.810 (0.75, 0.87)
Risk Assessment/Counseling for Sexual Activity Rate: 0.890 (0.84, 0.94)
Risk Assessment/Counseling for Substance Use Rate: 0.785 (0.72,0.85)
Risk Assessment/Counseling for Tobacco Use Rate: 0.791 (0.73, 0.85)

2b C☐ P☐ M☐ N☐

2c. Validity testing

2c.1 Data/sample *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure)

2c.2 Analytic Method *(type of validity & rationale, method for testing)*:
NCQA tested the measure for face validity using a panel of stakeholders with specific expertise in measurement and child health care. This panel included representatives from key stakeholder groups, including pediatricians, family physicians, health plans, state Medicaid agencies and researchers. Experts

2c C☐ P☐ M☐ N☐

reviewed the results of the field test and assessed whether the results were consistent with expectations, whether the measure represented quality care, and whether we were measuring the most important aspect of care in this area. 2c.3 **Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*: This measure was deemed valid by the expert panel. In addition, this measure does not utilize administrative data sources; data recorded in the chart is considered the gold standard.	
2d. Exclusions Justified 2d.1 **Summary of Evidence supporting exclusion(s):** No exclusions 2d.2 **Citations for Evidence:** NA 2d.3 **Data/sample** *(description of data/sample and size)*: NA 2d.4 **Analytic Method** *(type analysis & rationale)*: NA 2d.5 **Testing Results** *(e.g., frequency, variability, sensitivity analyses)*: NA	**2d** C☐ P☐ M☐ N☐ NA☐
2e. Risk Adjustment for Outcomes/ Resource Use Measures 2e.1 **Data/sample** *(description of data/sample and size)*: NA 2e.2 **Analytic Method** *(type of risk adjustment, analysis, & rationale)*: NA 2e.3 **Testing Results** *(risk model performance metrics)*: NA 2e.4 **If outcome or resource use measure is not risk adjusted, provide rationale:** The measure assesses prevention and wellness in a general population; risk adjustment is not indicated.	**2e** C☐ P☐ M☐ N☐ NA☐
2f. Identification of Meaningful Differences in Performance 2f.1 **Data/sample from Testing or Current Use** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure) 2f.2 **Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: Comparison of means and percentiles; analysis of variance against established benchmarks; if sample size is >400, we would use an analysis of variance 2f.3 **Provide Measure Scores from Testing or Current Use** *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*: Below is eligible population Risky Behavior Counseling or Assessment in Adolescents By 18 Years: 163 Performance rates listed by rate. Rate 1: Sexual Activity By 18 Years: 89% Rate 2: Substance Use By 18 Years: 79% Rate 3: Alcohol Use	**2f** C☐ P☐ M☐ N☐

By 18 Years: 81% Rate 4: Tobacco Use By 18 Years: 79%	
2g. Comparability of Multiple Data Sources/Methods **2g.1 Data/sample** *(description of data/sample and size)*: NCQA received data from 18 physician practices who submitted 10 records per measure (total 180 records per measure) **2g.2 Analytic Method** *(type of analysis & rationale)*: This measure is chart review only; no other sources were identified by the expert panel; this measure does not utilize administrative data **2g.3 Testing Results** *(e.g., correlation statistics, comparison of rankings)*: NA	**2g** C☐ P☐ M☐ N☐ NA☐
2h. Disparities in Care **2h.1 If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: The measure is not stratified to detect disparities. **2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** NA	**2h** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	***2*** C☐ P☐ M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	**Eval Rating**
3a. Meaningful, Understandable, and Useful Information **3a.1 Current Use:** Not in use but testing completed **3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*): This measure is not currently publicly reported. NCQA is exploring the feasibility of adding this measure and its related measures into a physician-level program and/or the HEDIS® measurement set as appropriate. **3a.3 If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):** This measure is not currently used in QI. NCQA is exploring the feasibility of adding this measure and its related measures into a physician-level program and/or the HEDIS® measurement set as appropriate. NCQA anticipates that after we release these measures, they will become widely used, as all our measures do. **Testing of Interpretability** (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) **3a.4 Data/sample** *(description of data/sample and size)*: Expert panel, other stakeholders, and 19 physician field test participants **3a.5 Methods** *(e.g., focus group, survey, QI project)*: NCQA vetted the measures with its expert panel. In addition, throughout the development process, NCQA	**3a** C☐ P☐ M☐ N☐

vetted the measure concepts and specifications with other stakeholder groups, including the National Association of State Medicaid Directors, NCQA's Health Plan Advisory Council, NCQA's Committee on Performance Measurement, and the American Academy of Pediatrician's Quality Improvement Innovation Network. After field testing, NCQA also conducted a debrief call with field test participants. In the form of a group interview, NCQA systematically sought feedback on whether the measures were understandable, feasible, important, and had face validity. 3a.6 **Results** *(qualitative and/or quantitative results and conclusions)*: NCQA received feedback that the measure is understandable, feasible, important and valid.	
3b/3c. Relation to other NQF-endorsed measures 3b.1 **NQF # and Title of similar or related measures:**	
(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 **Are the measure specifications harmonized? If not, why?**	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value 3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** 5.1 **If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:** NA	**3c** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐ P☐ M☐ N☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	**Eval Rating**
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition), Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	**4a** C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) No	**4b** C☐ P☐ M☐ N☐

4b.2 If not, specify the near-term path to achieve electronic capture by most providers. NCQA plans to eventually adapt this measure for use in electronic health records.	
4c. Exclusions **4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?** No **4c.2 If yes, provide justification.**	**4c** C☐ P☐ M☐ N☐ NA☐
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences **4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** During the measure development process the Child Health MAP and measure development team worked with NCQA's certified auditors and audit department to ensure that the measure specifications were clear and auditable. The denominator, numerator and any exclusions are concisely specified and align with our audit standards.	**4d** C☐ P☐ M☐ N☐
4e. Data Collection Strategy/Implementation **4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:** Based on field test results, we have specified the measure to assess whether physicians assessed OR counseled adolescents on the four risky behavior topics. Our field test results showed that these data elements are available in the medical record. In addition, our field test participants noted that many were able to program these requirements into their electronic health record systems, and several implemented point-of-service physician reminders for this measure. **4e.2 Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** Collecting measures from medical charts is time-consuming and can be burdensome. Adapting this measure in electronic health records may relieve some of this burden. **4e.3 Evidence for costs:** Based on field test participant feedback and other stakeholder input. **4e.4 Business case documentation:**	**4e** C☐ P☐ M☐ N☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility?*	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	**4** C☐ P☐ M☐ N☐
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited** ☐
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y**☐ **N**☐ **A**☐
CONTACT INFORMATION	

Co.1 Measure Steward (Intellectual Property Owner)
Co.1 Organization
National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005

Co.2 Point of Contact
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-

Measure Developer If different from Measure Steward
Co.3 Organization
National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005

Co.4 Point of Contact
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-

Co.5 Submitter If different from Measure Steward POC
Sepheen, Byron, MHS, byron@ncqa.org, 202-955-3573-, National Committee for Quality Assurance

Co.6 Additional organizations that sponsored/participated in measure development

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
Child Health Measurement Advisory Panel:
Jeanne Alicandro
Barbara Dailey
Denise Dougherty, PhD
Ted Ganiats, MD
Foster Gesten, MD
Nikki Highsmith, MPA
Charlie Homer, MD, MPH
Jeff Kamil, MD
Elizabeth Siteman
Mary McIntyre, MD, MPH
Virginia Moyer, MD, MPH, FAAP
Lee Partridge
Xavier Sevilla, MD, FAAP
Michael Siegal
Jessie Sullivan

Ad.2 If adapted, provide name of original measure: NA
Ad.3-5 If adapted, provide original specifications URL or attachment

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 Year the measure was first released:
Ad.7 Month and Year of most recent revision:
Ad.8 What is your frequency for review/update of this measure?
Ad.9 When is the next scheduled review/update for this measure?

Ad.10 Copyright statement/disclaimers: © 2009 by the National Committee for Quality Assurance
1100 13th Street, NW, Suite 1000
Washington, DC 20005

Ad.11 -13 Additional Information web page URL or attachment:

Date of Submission (*MM/DD/YY*)**:** 01/06/2011

NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 1365 **NQF Project:** Child Health Quality Measures 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Child and Adolescent Major Depressive Disorder: Suicide Risk Assessment
De.2 **Brief description of measure:** Percentage of patient visits for those patients aged 6 through 17 years with a diagnosis of major depressive disorder with an assessment for suicide risk
1.1-2 **Type of Measure:** Process De.3 **If included in a composite or paired with another measure, please identify composite or paired measure**
De.4 **National Priority Partners Priority Area:** Population health De.5 **IOM Quality Domain**: Effectiveness, Patient-centered De.6 **Consumer Care Need:** Getting better

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** A.3 **Measure Steward Agreement:** Agreement will be signed and submitted prior to or at the time of measure submission A.4 **Measure Steward Agreement attached:**	**A** Y☐ N☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least	**B** Y☐

every 3 years. Yes, information provided in contact section	N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement Accountability	**C** Y☐ N☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** No, testing will be completed within 12 months D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** Y☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. ***Measures must be judged to be important to measure and report*** *in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** Affects large numbers, Leading cause of morbidity/mortality, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** "Major depressive disorder (MDD) is a debilitating condition that has been increasingly recognized among youth, particularly adolescents. The prevalence of current or recent depression among children is 3% and among adolescents is 6%.1 The lifetime prevalence of MDD among adolescents may be as high as 20%.2-4 Adolescent-onset MDD is associated with an increased risk of death by suicide, suicide attempts, and recurrence of major depression by young adulthood.5-7 MDD is also associated with early pregnancy, decreased school performance, and impaired work, social, and family functioning during young adulthood.6-8" In 2006, suicide was the third leading cause of death for young people ages 15 to 24, accounting for 12% of all deaths annually. 9 Of every 100,000 young people aged 10-14, 1.3 died by suicide. Of every 100,000 young people aged 15-19, 8.2 died by suicide. 9 Among young adults ages 15 to 24 years old, there are approximately 100-200 attempts for every completed suicide. 9 In 2007, 14.5% of U.S. high school students reported that they had seriously considered attempting suicide during the 12 months preceding the survey; 6.9% of students reported that they had actually attempted suicide one or more times during the same period.9 1a.4 **Citations for Evidence of High Impact:** Williams SB, O'Connor EA, Eder M, Whitlock EP. Screening for Child and Adolescent Depression in Primary Care Settings: A Systematic Evidence Review for the US	**1a** C☐ P☐ M☐ N☐

Preventive Services Task Force. Pediatrics 2009;123:e716–e735. Citing:
1. Jane Costello E, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression? J Child Psychol Psychiatry. 2006; 47(12):1263–1271
2. Lewinsohn PM, Rohde P, Seeley JR. Major depressive disorder in older adolescents: prevalence, risk factors, and clinical implications. Clin Psychol Rev. 1998;18(7):765–794
3. Cheung A. Canadian community health survey: major depressive disorder and suicidality in adolescents. Healthc Policy. 2006; 2(2):76–89
4. Whitaker A, Johnson J, Shaffer D, et al. Uncommon troubles in young people: prevalence estimates of selected psychiatric disorders in a nonreferred adolescent population. Arch Gen Psychiatry. 1990;47(5):487–496
5. Shaffer D, Gould MS, Fisher P, et al. Psychiatric diagnosis in child and adolescent suicide. Arch Gen Psychiatry. 1996;53(4):339–348
6. Weissman MM, Wolk S, Goldstein RB, et al. Depressed adolescents grown up. JAMA. 1999;281(18):1707–1713
7. Fergusson DM, Woodward LJ. Mental health, educational, and social role outcomes of adolescents with depression. Arch Gen Psychiatry. 2002;59(3):225–231
8. Keenan-Miller D, Hammen CL, Brennan PA. Health outcomes related to early adolescent depression. J Adolesc Health. 2007; 41(3):256–262
9. Centers for Disease Control and Prevention (CDC). Suicide: Facts at a Glance. CDC; Summer 2009. http://www.cdc.gov/violenceprevention. Accessed August 25, 2010.

1b. Opportunity for Improvement

1b.1 **Benefits (improvements in quality) envisioned by use of this measure:** Research has shown that patients with major depressive disorder are at a high risk for suicide, which makes this assessment an important aspect of care that should be evaluated at each visit.

1b.2 **Summary of data demonstrating performance gap (variation or overall poor performance) across providers:**
According to a study analyzing the quality of health care in the United States, only about 25.8% of patients with depression had documentation of the presence or absence of suicidal ideation during the first or second diagnostic visit. 76.11% of those patients who have suicidality were asked if they have specific plans to carry out suicide.(1) A 2003 study reviewed medical records to assess the degree to which providers adhered to depression guidelines in a VA primary care setting. Providers documented exploration for suicidal ideation in 57% of the records.(2)

1b.3 **Citations for data on performance gap:**
1. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of health care delivered to adults in the United States. New England Journal of Medicine. 2003;348(26):2635-2645.
2. Dobscha SK, Gerrity MS, Corson K, Bahr A, Cuilwik NM. Measuring adherence to depression treatment guidelines in a VA primary care clinic. Gen Hosp Psychiatry. 2003;25:230-7.

1b.4 **Summary of Data on disparities by population group:**
We are not aware of any publications/evidence outlining disparities in this area.

1b.5 **Citations for data on Disparities:**

1b
C☐
P☐
M☐
N☐

1c. Outcome or Evidence to Support Measure Focus

1c.1 **Relationship to Outcomes** (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** Suicide attempts and completion are among the most significant and devastating sequelae of MDD. Suicide risk should therefore be assessed at each visit and subsequently managed to minimize that risk.

1c.2-3. **Type of Evidence:** Evidence-based guideline

1c.4 **Summary of Evidence** (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**

1c
C☐
P☐
M☐
N☐

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**

1c.6 Method for rating evidence:

1c.7 Summary of Controversy/Contradictory Evidence: None

1c.8 Citations for Evidence (*other than guidelines***):**

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number***):**
The evaluation must include assessment for the presence of harm to self or others (MS). (AACAP (1))

Suicidal behavior exists along a continuum from passive thoughts of death to a clearly developed plan and intent to carry out that plan. Because depression is closely associated with suicidal thoughts and behavior, it is imperative to evaluate these symptoms at the initial and subsequent assessments. For this purpose, low burden tools to track suicidal ideation and behavior such as the Columbia-Suicidal Severity Rating Scale can be used. Also, it is crucial to evaluate the risk (e.g., age, sex, stressors, comorbid conditions, hopelessness, impulsivity) and protective factors (e.g., religious belief, concern not to hurt family) that might influence the desire to attempt suicide. The risk for suicidal behavior increases if there is a history of suicide attempts, comorbid psychiatric disorders (e.g., disruptive disorders, substance abuse), impulsivity and aggression, availability of lethal agents (e.g., firearms), exposure to negative events (e.g., physical or sexual abuse, violence), and a family history of suicidal behavior. (AACAP (1))

1c.10 Clinical Practice Guideline Citation: (1) American Academy of Child and Adolescent Psychiatry (AACAP). Practice parameters for the assessment and treatment of children and adolescents with depressive disorders. J. Am. Acad. Child Adolesc. Psychiatry, 2007; 46(11):1503-1526. Available at: http://www.aacap.org/galleries/PracticeParameters/Vol%2046%20Nov%202007.pdf
1c.11 National Guideline Clearinghouse or other URL: (1) http://www.guideline.gov/content.aspx?id=11404

1c.12 Rating of strength of recommendation (*also provide narrative description of the rating and by whom***):**
Minimal Standard (MS) [see below for narrative description of the rating]

1c.13 Method for rating strength of recommendation (*If different from USPSTF system, also describe rating and how it relates to USPSTF)***:**
American Academy of Child and Adolescent Psychiatry (AACAP) Grades of Recommendations

•Minimal Standard [MS] is applied to recommendations that are based on rigorous empirical evidence (such as randomized, controlled trials) and/or overwhelming clinical consensus. Minimal standards apply more than 95% of the time; i.e., in almost all cases.
•Clinical Guideline [CG] is applied to recommendations that are based on strong empirical evidence (such as non-randomized control trials) and/or strong clinical consensus. Clinical guidelines apply approximately 75% of the time; i.e., in most cases.
•Option [OP] is applied to recommendations that are acceptable based on emerging empirical evidence (such as uncontrolled trials or reports) or clinical opinion, but lack strong empirical evidence and/or strong clinical consensus.
•Not Endorsed [NE] is applied to practices that are known to be ineffective or contraindicated.

1c.14 Rationale for using this guideline over others:
It is the PCPI policy to use guidelines, which are evidence-based, applicable to physicians and other healthcare providers, and developed by a national speciality organization or government agency. In addition, the PCPI has now expanded what is acceptable as the evidence base for measures to included documented quality improvement (QI) initiatives or implementation projects that have demonstrated

improvement in the quality of care.	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	**Eval Rating**
2a. MEASURE SPECIFICATIONS	
S.1 **Do you have a web page where current detailed measure specifications can be obtained?** S.2 **If yes, provide web page URL:** **2a. Precisely Specified**	**2a-specs** **C☐** **P☐** **M☐** **N☐**
2a.1 **Numerator Statement** (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome*)**:** Patient visits with an assessment for suicide risk 2a.2 **Numerator Time Window** (*The time period in which cases are eligible for inclusion in the numerator*)**:** Each patient visit within a 12-month period 2a.3 **Numerator Details** (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*)**:**	
2a.4 **Denominator Statement** (*Brief, text description of the denominator - target population being measured*)**:** All patient visits for those patients aged 6 through 17 years with a diagnosis of major depressive disorder 2a.5 **Target population gender:** Female, Male 2a.6 **Target population age range:** Aged 6 through 17 years 2a.7 **Denominator Time Window** (*The time period in which cases are eligible for inclusion in the denominator*)**:** 12 months 2a.8 **Denominator Details** (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*)**:** See attached Level I EHR Specifications	
2a.9 **Denominator Exclusions** (*Brief text description of exclusions from the target population*)**:** None 2a.10 **Denominator Exclusion Details** (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*)**:**	
2a.11 **Stratification Details/Variables** (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*)**:** Stratification by insurance coverage (commercial, Medicare and Medicaid) is recommended by some implementers	
2a.12-13 **Risk Adjustment Type:** No risk adjustment necessary 2a.14 **Risk Adjustment Methodology/Variables** (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*)**:**	

2a.15-17 Detailed risk model available Web page URL or attachment:

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Higher score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps***):**
See attached documents

2a.22 Describe the method for discriminating performance *(e.g., significance testing)***:**

2a.23 Sampling (Survey) Methodology *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:

2a.24 Data Source (*Check the source(s) for which the measure is specified and tested)*
Electronic Health/Medical Record

2a.25 Data source/data collection instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**

2a.26-28 Data source/data collection instrument reference web page URL or attachment:

2a.29-31 Data dictionary/code table web page URL or attachment: Attachment MDD 3 Complete.pdf

2a.32-35 Level of Measurement/Analysis (*Check the level(s) for which the measure is specified and tested)*
Clinicians: Individual

2a.36-37 Care Settings (*Check the setting(s) for which the measure is specified and tested)*
Ambulatory Care: Office, Ambulatory Care: Clinic, Ambulatory Care: Hospital Outpatient, Behavioral health/psychiatric unit

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)
Behavioral Health: Mental Health, Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO), Clinicians: Psychologist/LCSW

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample *(description of data/sample and size)***:** Are Claims Data Accurate Enough to Identify Patients for Performance Measures or Quality Improvement? The Case of Diabetes, Heart Disease, and Depression. Leif I. Solberg, Karen I. Engebretson, Joann M. Sperl-Hillen, Mary C. Hroscikoski and Patrick J. O´Connor. American Journal of Medical Quality 2006; 21; 238.

The Challenge of Measuring Quality of Care From the Electronic Health Record. Carol P. Roth, Yee-Wei Lim, Joshua M. Pevnick, Steven M. Asch and Elizabeth A. McGlynn. American Journal of Medical Quality 2009; 24; 385 originally published online May 29, 2009.

Measuring adherence to depression treatment guidelines in a VA primary care clinic. Dobscha SK, Gerrity MS, Corson K, Bahr A, Cuilwik NM. General Hospital Psychiatry 25 (2003) 230–237

2b.2 Analytic Method *(type of reliability & rationale, method for testing)***:**
(Solberg, 2006) The objective of this study was to demonstrate a method to accurately identify patients with specific conditions from claims data for care improvement or performance measurement. Using an iterative process of trial case definitions followed by review of repeated random samples of 10 to 20 cases for newly treated depression, a final identification algorithm was created from claims files of health plan members. A final sample was used to calculate the positive predictive value (PPV).

2b
C ☐
P ☐
M ☐
N ☐

(Roth 2009) The electronic health record (EHR) is seen by many as an ideal vehicle for measuring quality of health care and monitoring ongoing provider performance. It is anticipated that the availability of EHR-extracted data will allow quality assessment without the expensive and time-consuming process of medical record abstraction. Each quality measure was classified by the anticipated difficulty of satisfying eligibility and scoring statements using an EHR-enhanced data warehouse as the source of data. Measures were considered level 1 if all requisite data elements were accessible. Measures were considered level 2 if the denominator was accessible but the numerator was in some way inaccessible. Measures were considered level 3 if the denominator was difficult to access.

(Dobscha 2003) Researchers created one composite, measure, based on 3 national guidelines.
The DSM-IV Major depression criteria corresponds with our Diagnostic Evaluation measure.
The Evaluate level of safety/suicide history criteria corresponds with our Suicide Risk Assessment measure.
Data was analyzed for internal consistency and inter-rater reliability.

2b.3 **Testing Results** *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
(Solberg, 2006) MDD had an unacceptably low PPV (0.65) when cases were identified on the basis of only 1 International Classification of Diseases, ninth revision, code per year. Requiring 2 outpatient ICD-9 codes or 1 inpatient ICD-9 code within 12 months (plus consideration of extra criteria for depression) resulted in PPV of 0.95. This approach is feasible and necessary for those wanting to use administrative data for case identification for performance measurement or quality improvement. The PCPI measure utilizes this approach.

(Dobscha 2003)
Inter-rater reliability was assessed, using the kappa coefficient
The Self Harm measure (documentation of past or present suicidal ideation) had a kappa = 0.96. The performance rate for this measure was 56.8% (47.5 - 65.6 95%CI).

(Roth 2009) Accurately identifying eligible cases for quality assessment and validly scoring those cases with EHR extracted data will pose challenges but could potentially plummet the cost and therefore expand the use of quality assessment. A review of the data requirements for the depression related indicators in the Quality Assessment Tools system suggests that 41% of measures would be readily accessible from EHR data. Another 29% of the depression-related indicators have denominators that are readily accessible. Accessibility of data used to calculate the measure in an EHR reflects reliability of measure calculation.

2c. Validity testing

2c.1 **Data/sample** *(description of data/sample and size)*:

2c.2 **Analytic Method** *(type of validity & rationale, method for testing)*:
During measure development, the PCPI-convened expert work groups assess the face and content validity of each measure. The groups establish the measure's ability to capture what it is designed to capture using a consensus process that consists of input from multiple stakeholders, including practicing physicians and experts with technical measure expertise, as well as a review of additional input received through a PCPI public comment period.

2c.3 **Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:

2c
C☐
P☐
M☐
N☐

2d. Exclusions Justified

2d.1 **Summary of Evidence supporting exclusion(s):**
No Exceptions are allowed for this measure.

2d.2 **Citations for Evidence:**

2d
C☐
P☐
M☐
N☐
NA☐

2d.3 Data/sample *(description of data/sample and size)*: **2d.4 Analytic Method** *(type analysis & rationale)*: **2d.5 Testing Results** *(e.g., frequency, variability, sensitivity analyses)*:	
2e. Risk Adjustment for Outcomes/ Resource Use Measures **2e.1 Data/sample** *(description of data/sample and size)*: **2e.2 Analytic Method** *(type of risk adjustment, analysis, & rationale)*: **2e.3 Testing Results** *(risk model performance metrics)*: **2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:**	**2e** C☐ P☐ M☐ N☐ NA☐
2f. Identification of Meaningful Differences in Performance **2f.1 Data/sample from Testing or Current Use** *(description of data/sample and size)*: **2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: **2f.3 Provide Measure Scores from Testing or Current Use** *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*:	**2f** C☐ P☐ M☐ N☐
2g. Comparability of Multiple Data Sources/Methods **2g.1 Data/sample** *(description of data/sample and size)*: **2g.2 Analytic Method** *(type of analysis & rationale)*: **2g.3 Testing Results** *(e.g., correlation statistics, comparison of rankings)*:	**2g** C☐ P☐ M☐ N☐ NA☐
2h. Disparities in Care **2h.1 If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: The measure is not stratified by patient groups or cohorts that could potentially be affected by disparities in care, nor are we aware of any existing research identifying disparities in care that may be relevant to this measure. **2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** We are not aware of any relevant disparities that have been identified.	**2h** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	***2*** C☐ P☐ M☐

	N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	**Eval Rating**
3a. Meaningful, Understandable, and Useful Information **3a.1 Current Use:** In use **3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*)**:** This measure is in its adult form is currently utilized in the CMS PQRI Program **3a.3 If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):** **Testing of Interpretability** (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) **3a.4 Data/sample** *(description of data/sample and size)***:** **3a.5 Methods** *(e.g., focus group, survey, QI project)***:** **3a.6 Results** *(qualitative and/or quantitative results and conclusions)***:**	**3a** C☐ P☐ M☐ N☐
3b/3c. Relation to other NQF-endorsed measures **3b.1 NQF # and Title of similar or related measures:** 104: Major Depressive Disorder: Suicide Risk Assessment	
(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): **3b.2 Are the measure specifications harmonized? If not, why?** Yes	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value **3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** **5.1 If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:**	**3c** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐ P☐ M☐ N☐
4. FEASIBILITY	

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition)	**4a** C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) Yes 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.**	**4b** C☐ P☐ M☐ N☐
4c. Exclusions 4c.1 **Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?** No 4c.2 **If yes, provide justification.**	**4c** C☐ P☐ M☐ N☐ NA☐
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences 4d.1 **Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** We are not aware of any unintended consequences related to this measurement.	**4d** C☐ P☐ M☐ N☐
4e. Data Collection Strategy/Implementation 4e.1 **Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:** This pediatric MDD measure has a corresponding adult measure, which differs only in having an different age range. Therefore, implementation results for the adult measures are expected to be applicable to the pediatric measures. Through a partnership with the American Medical Association (AMA) and Healthcare Information and Management Systems Society (HIMSS), the Alliance of Chicago Community Health Centers developed the AHRQ-funded 3-year Enhancing Quality in Patient Care (EQUIP) project to augment its EHR implementation. This project implemented all 5 AMA-PCPI Adult MDD measures in the EHR. As part of the AHRQ-funded Effecting Change in Chronic Care: The Tipping Point project, 3 physicians implemented performance measures into existing electronic health record systems. One additional physician implemented a paper flow sheet documentation system where the flow sheet was placed in each chart at the time of the visit. This project found that the adult MDD measures were feasible to collect after the process changes were put into place. Additionally, the adult MDD version of this measure was utilized in the CMS PQRI program, in 2008, 2009, and 2010. The average performance rate for the 2008 PQRI program for the Suicide Risk Assessment measure was 81%, with n=5440. 4e.2 **Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** Costs to implement the measure have not been calculated. 4e.3 **Evidence for costs:**	**4e** C☐ P☐ M☐ N☐

4e.4 **Business case documentation:**	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility?*	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	**4** **C☐** **P☐** **M☐** **N☐**
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited ☐**
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y☐** **N☐** **A☐**

CONTACT INFORMATION

Co.1 **Measure Steward (Intellectual Property Owner)**
Co.1 **Organization**
American Medical Association, 515 N State St., Chicago, Illinois, 60654

Co.2 **Point of Contact**
Mark, Antman, DDS, MBA, mark.antman@ama-assn.org, 312-464-5056-

Measure Developer If different from Measure Steward
Co.3 **Organization**
American Medical Association, 515 N State St., Chicago, Illinois, 60654

Co.4 **Point of Contact**
Mark, Antman, DDS, MBA, mark.antman@ama-assn.org, 312-464-5056-

Co.5 **Submitter If different from Measure Steward POC**
Mark, Antman, DDS, MBA, mark.antman@ama-assn.org, 312-464-5056-, American Medical Association

Co.6 **Additional organizations that sponsored/participated in measure development**
American Psychiatric Association, American Academy of Child and Adolescent Psychiatry

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 **Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.**
Boris Birmaher, MD (child/adolescent psychiatry)
Mary Dobbins, MD, FAAP (pediatrics/psychiatry)
Scott Endsley, MD, MSc (family medicine)
William E. Golden, MD, FACP (internal medicine)
Margaret L. Keeler, MD, MS, FACEP (emergency medicine)
Louis J. Kraus, MD (child/adolescent psychiatry)
Laurent S. Lehmann, MD (psychiatry)
Karen Pierce, MD (child/adolescent psychiatry)
Reed E. Pyeritz, MD, PhD, FACP, FACMG (medical genetics)
Laura Richardson, MD, MPH (internal medicine/pediatrics)
Sam J.W. Romeo, MD, MBA (family medicine)
Carl A. Sirio, MD (critical care medicine)
Sharon Sweede, MD (family medicine)
Scott Williams, PsyD (The Joint Commission)

PCPI measures are developed through cross-specialty, multi-disciplinary work groups. All medical specialties and other health care professional disciplines participating in patient care for the clinical condition or topic under study must be equal contributors to the measure development process. In addition, the PCPI strives to include on its work groups individuals representing the perspectives of patients, consumers, private health plans, and employers. This broad-based approach to measure development ensures buy-in on the measures from all stakeholders and minimizes bias toward any individual specialty or stakeholder group. All work groups have at least two co-chairs who have relevant clinical and/or measure development expertise and who are responsible for ensuring that consensus is achieved and that all perspectives are voiced.

Ad.2 **If adapted, provide name of original measure:**
Ad.3-5 **If adapted, provide original specifications URL or attachment**

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 **Year the measure was first released:** 2008
Ad.7 **Month and Year of most recent revision:** 09, 2008
Ad.8 **What is your frequency for review/update of this measure?** Every 3 years or as new evidence becomes available that materially affects the measures
Ad.9 **When is the next scheduled review/update for this measure?** 09, 2011

Ad.10 **Copyright statement/disclaimers:** Physician Performance Measures (Measures) and related data specifications are developed by the American Medical Association (AMA) in collaboration with the Physician Consortium for Performance Improvement® (PCPI).

These performance Measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications.

The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measures require a license agreement between the user and the AMA, (on behalf of the PCPI). Neither the AMA, the PCPI nor its members shall be responsible for any use of the Measures.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. The AMA, NCQA, the PCPI and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT®) or other coding contained in the specifications.

CPT® contained in the Measures specifications is copyright 2007 American Medical Association. LOINC® copyright 2004 Regenstrief Institute, Inc. SNOMED CLINICAL TERMS (SNOMED CT®) copyright 2004 College of American Pathologists (CAP). All Rights Reserved. Use of SNOMED CT® is only authorized within the United States.

Ad.11 -13 **Additional Information web page URL or attachment:** Attachment NQF Aug 2010 Submission Letter.pdf

Date of Submission (*MM/DD/YY*)**:** 08/30/2010