# NATIONAL QUALITY FORUM

**Measure Evaluation 4.1**
**January 2010**

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 sub-criteria and most of the footnotes from the evaluation criteria are provided in Word comments and will appear if your cursor is over the highlighted area (or in the margin if your Word program is set to show revisions in balloons). Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized): Complete all <mark>yellow highlighted</mark> areas of the form. Evaluate the extent to which each sub-criterion 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* sub-criteria (*yellow* highlighted areas).

**Steering Committee**: Complete all <mark>pink</mark> highlighted areas of the form. Review the workgroup/TAP assessment of the sub-criterion, 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 sub-criteria as indicated)

(for NQF staff use) **NQF Review #**: OT3-011-10        **NQF Project**: Patient Outcomes Measures: Child Health and Mental Health (Phase III)

## MEASURE DESCRIPTIVE INFORMATION

**De.1 Measure Title:** Depression Remission at Twelve Months

**De.2 Brief description of measure***:* Adult patients age 18 and older with major depression or dysthymia and an initial PHQ-9 score > 9 who demonstrate remission at twelve months defined as a PHQ-9 score less than 5. This measure applies to both patients with newly diagnosed and existing depression whose current PHQ-9 score indicates a need for treatment.
The Patient Health Questionnaire (PHQ-9) tool is a widely accepted, standardized tool [Copyright © 2005 Pfizer, Inc. All rights reserved] that is completed by the patient, ideally at each visit, and utilized by the provider to monitor treatment progress.
This measure additionally promotes ongoing contact between the patient and provider as patients who do not have a follow-up PHQ-9 score at twelve months (+/- 30 days) are also included in the denominator.

**1.1-2 Type of Measure:** outcome
**De.3 If included in a composite or paired with another measure, please identify composite or paired measure**
This measure is related to two other measures that are included in this phase III submission for mental health measures. The other measures are 1) Depression Remission at Six Months (outcome) and 2) Depression Utilization of the PHQ-9 Tool (process)

**De.4 National Priority Partners Priority Area:** patient and family engagement
**De.5 IOM Quality Domain**: effectiveness
**De.6 Consumer Care Need:** Living With Illness

## 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 (<u>measure steward agreement</u>) is signed.<br>*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.*<br>**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<br>**A.2 Indicate if** Proprietary Measure (*as defined in measure steward agreement*):<br>**A.3** Measure Steward Agreement: agreement signed and submitted<br>**A.4** Measure Steward Agreement attached: NQF data steward agreement_signed 2009.pdf | A<br>A<br>Y☐<br>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<br>Y☐<br>N☐ |
| C. The intended use of the measure includes <u>both</u> public reporting <u>and</u> quality improvement.<br>►**Purpose:** public reporting, quality improvement Accountability | C<br>Y☐<br>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.<br>**D.1**Testing: Yes, fully developed and tested<br>**D.2** Have NQF-endorsed measures been reviewed to identify if there are similar or related measures? Yes | D<br>D<br>Y☐<br>N☐ |
| (for NQF staff use**)** Have all conditions for consideration been met?<br>**Staff Notes to Steward** (*if submission returned*): | Met<br>Y☐<br>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.<br>*Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.* (<u>evaluation criteria</u>)<br>**1a. High Impact** | <u>Eval Rating</u> |
| (for NQF staff use) <u>Specific NPP goal</u>: | |
| **1a.1 Demonstrated High Impact Aspect of Healthcare:** affects large numbers, severity of illness, patient/societal consequences of poor quality<br>**1a.2**<br><br>**1a.3 Summary of Evidence of High Impact:** The Centers for Disease Control and Prevention states that nationally 15.7% of people report being told by a health care professional that they had depression at some point in their lifetime. Persons with a current diagnosis of depression and a lifetime diagnosis of depression or anxiety were significantly more likely than persons without these conditions to have cardiovascular disease, diabetes, asthma and obesity and to be a current smoker, to be physically inactive and to drink heavily. According to National Institute of Mental Health (NIMH), 6.7 percent of the U.S. population ages 18 and older (14.8 million people) in any given year have a diagnosis of a major depressive disorder. Major depression is the leading cause of disability in the U.S. for ages 15 - 44. Additionally, dysthymia accounts for | 1a<br>C☐<br>P☐<br>M☐<br>N☐ |

an additional 3.3 million Americans. In Minnesota, the rates for current depression are 6 -7.9% and the percent of Minnesotans who have a lifetime diagnosis of depression is between 13 and 15%.
Suicide rates for Minnesotans are 10.4 per 100,000 or 1.3 suicides per day, with the highest rates among the following groups: males (4 times greater than females), ages 30 to 49 years, and non-hispanic whites.

**1a.4** Citations for Evidence of High Impact: Centers for Disease Control and Prevention: Anxiety and Depression Effective Treatments Exist: People with depression and anxiety should seek help as early as possible to reduce health effects and improve quality of life. March 2009. Based on 2006 Behavior Risk Factor Surveillance System www.cdc.gov/Features/dsBRFSSDepressionAnxiety/
Suicide Prevention Resource Center: Minnesota Suicide Fact Sheet; Suicides 1999 - 2005 www.sprc.org/
National Institute of Mental Health: The Numbers Count: Mental Disorders in America August 2009
www.nimh.nih.gov/health/publications/the-numbers-count-mental-disorders-in-america

## 1b. Opportunity for Improvement

**1b.1** Benefits (improvements in quality) envisioned by use of this measure: • Improve the outcomes of symptom control and functioning for patients with major depression and dysthymia. Major depression is a treatable cause of pain, suffering, disability and death, yet primary care providers detect major depression in only 1/3 to 1/2 of their patients with major depression (Schonfeld, 1997 [C]Williams Jr, 2002 [R])
• Improve the frequency of assessment of the response to treatment
• Improve the communication between the primary care and behavioral health providers, have a common tool to document response.
• Use of a standardized tool (PHQ-9) to measure outcomes over time
Source: ICSI Guideline for Major Depression in Adults in Primary Care 12th edition May 2009
www.icsi.org/guidelines_and_more/gl_os_prot/behavioral_health/depression_5/depression__major__in_adults_in_primary_care_4.html

**1b.2** Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
Data collection and submission for this measure started in September 2008, currently there is enough data to capture preliminary twelve month remission rates for patients with an index contact date between 1-1-2008 and 8-31-2008, n = 17,085 patients representing 123 clinics.  For these patients, the twelve month remission results are as follows:
Average :         4.6%  (790/17,085)
Range:            0.0% to 29.7%
StDev:            3.8
8 of the 92 clinics included in the data set of patients with an index contact date between 1/1/2008 and 8/31/2008 and a denominator that is greater than equal to 30 had confidence intervals fully above the mean, demonstrating meaningful difference and variability within the data.
Low remission rates at twelve months are in partially related to the medical group's ability to obtain a PHQ-9 score at twelve months +/- 30 days; currently only 20.3% of the eligible denominator patients have a 12 month PHQ-9.  Groups are diligently working on process improvements in workflow to impact ongoing contact with patients.
We plan to publicly report twelve month remission scores in June of 2010 when enough data has accumulated for patients with an index contact date 1-1-2008 to 12-31-2008, as a run-out of thirteen months of data is needed to capture the twelve month remission rates.  At this point we will have over 25,000 patients for twelve month remission rate calculation.

**1b.3** Citations for data on performance gap:
Publicly reported data with clinic level rates will available on the MN HealthScores website www.mnhealthscores.org in June 2010.  Additionally, rates for twelve month remission will be published in a future edition of the Health Care Quality Report posted on our corporate website at www.mncm.org.

**1b.4** Summary of Data on disparities by population group:
Major depressive disorder is a common disorder, widely distributed in the population, and usually associated with substantial symptom severity and role impairment. While the recent increase in treatment is encouraging, inadequate treatment is a serious concern. Emphasis on screening and expansion of treatment needs to be accompanied by a parallel emphasis on treatment quality.
Risk factors for major depression include family or personal history of major depression or substance abuse,

1b
C ☐
P ☐
M ☐
N ☐

recent loss, chronic medical illness, stressful life events that include loss, domestic abuse/ violence, traumatic events and major life changes. Although depression can affect any one in their lifetime, adults in the age ranges of 49 to 54 have the highest rates of depression. Other major risk factors include being female, being African-American and living in poverty. Women, regardless of nationality, race, ethnicity or socioeconomic level have twice the rate of depression than men.

Depression in the elderly is widespread, often undiagnosed and usually untreated. It is a common misperception that it is a part of normal aging. Losses, social isolation and chronic medical problems that older patients experience can contribute to depression. The rate of depression in adults older than 65 years of age ranges from 7% to 36% in medical outpatient clinics and increases to 40% in the hospitalized elderly. Comorbidities are more common in the elderly. The highest rates of depression are found in those with strokes (30% to 60%), coronary artery disease (up to 44%), cancer (up to 40%), Parkinson's disease (40%), and Alzheimer's disease (20% to 40%). The recurrence rate is also extremely high at 40%

**1b.5** Citations for data on Disparities:
ICSI Guideline for Major Depression in Adults in Primary Care 12th edition May 2009
www.icsi.org/guidelines_and_more/gl_os_prot/behavioral_health/depression_5/depression__major__in_adults_in_primary_care_4.html
Depression Risk Factors- New York Times Jan 22, 2009
The Epidemiology of Major Depressive Disorder Results From the National Comorbidity Survey Replication (NCS-R) Ronald C. Kessler, et al JAMA. 2003;289:3095-3105.

**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*): Improvement in the symptoms of depression and an ongoing assessment of the current treatment plan is crucial to the reduction of symptoms and psychosocial well being of patients with major depression. Most people treated for initial depression need to be on medication at least six to twelve months after adequate response to symptoms, patients with recurrent depression need to be treated for three years or more and response with psychotherapy can take eight to twelve weeks of regular and frequent therapy to show improvement. Remission is defined as a PHQ-9 score of less than five at twelve months. The Patient Health Questionnaire (PHQ-9) tool is a widely accepted, standardized tool [Copyright © 2005 Pfizer, Inc. All rights reserved] that is completed by the patient, ideally at each visit, and utilized by the provider to monitor treatment progress. This tool was selected for measuring outcomes for this population because it is 1) validated with a sensitivity of .080 and a specificity of 0.92 with substantial heterogeneity I2 = 82%, 2) widely accepted and utilized in our state, 3) available for clinical use, 4) translated into many languages and 5) easy for the patient to complete and the provider to score. This nine question tool contains the following questions which are scored on a scale of 0 to 27 based on the scale of Not at All (0), Several Days (1), More Than Half the Days (2), or Nearly Every Day (3) for responses to the questions over the last 2 weeks.
• Little interest or pleasure in doing things
• Feeling down, depressed, or hopeless
• Feeling tired or having little energy
• Poor appetite or overeating
• Feeling bad about yourself — or that you are a failure or have let yourself or your family down
• Trouble concentrating on things, such as reading the newspaper or watching television
• Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual
• Thoughts that you would be better off dead or of hurting yourself in some way

**1c.2-3.** Type of Evidence:  cohort study, evidence based guideline, expert opinion, meta-analysis, observational study, randomized controlled trial, other (specify) Consensus Statement

**1c.4** Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*):
ICSI Guideline Major Depression in Adults in Primary Care 12th Editions- May 2009
Major depression is a treatable cause of pain, suffering, disability and death, yet primary care providers detect major depression in only 1/3 to 1/2 of their patients with major depression (Schonfeld, 1997 [C]; Williams Jr, 2002 [R]). Additionally, more than 80% of patients with depression have a medical comorbidity (Klinkman, 2003 [R]). Usual care for depression in the primary care setting has resulted in only about half of

1c
C ☐
P ☐
M ☐
N ☐

depressed adults getting treated (Kessler, 2005 [C]) and only 20%-40% showing substantial improvement over 12 months (Katon, 1999 [A]; Unutzer, 2002 [A]).

Scope and Target Population:

To assist primary care in developing systems that support effective assessment, diagnosis and ongoing management of new or existing diagnosis of major depression in adults age 18 and over and assist patients to achieve remission of symptoms, reduce relapse and return to previous level of functioning.

Key Objectives of Treatment:

- achieve remission of symptoms in the acute treatment phase for major depression
- reduce relapse and reduction of symptoms
- return patient to previous level of occupational and psychosocial function

Priority Aims:

The aims and measures in this guideline are based upon evidence supporting impact of system elements, process elements, promoting actual symptom and functional patient improvement and outcomes, and are aligned with MN Community Measurement and the DIAMOND Initiative where there is overlap.

1. Increase the accuracy of diagnosis of major depression. (Annotations #1, 2, 3)

2. Improve the frequency of assessment of response to treatment in patients with major depression. (Annotation#12)

3. Improve the outcomes of treatment for major depression. (Annotations #11, 12)

4. Improve the frequency of assessment of patients with major depression for the presence of substance abuse. (Annotations #7, 8)

5. Increase the assessment for major depression of primary care patients presenting with additional high risk conditions such as diabetes, cardiovascular disease, post-stroke, chronic pain and all perinatal women. (Annotation #9, 13)

6. Improve communication between the primary care physician and the mental health care provider (if patient is co managed). (Annotations #8, 11, 14)

7. Decrease the number of completed suicides in patients managed for their depression in primary care. (Annotation #5)

Source: ICSI Guideline for Major Depression can be obtained at: www.icsi.org/guidelines_and_more/gl_os_prot/behavioral_health/depression_5/depression__major__in_adults_in_primary_care_4.html

**1c.5 R**ating of strength/quality of evidence (*also provide narrative description of the rating and by whom*):
Institute Clinical Systems Improvement, ICSI Evidence Grading System:
www.icsi.org/guidelines_and_more/evidence_grading_system_6/

**1c.6 Method for r**ating evidence: ICSI Evidence Grading System
A. Primary Reports of New Data Collection:
Class A: Randomized, controlled trial
Class B: Cohort study
Class C: Non-randomized trial with concurrent or historical controls, Case-control study, Study of sensitivity and specificity of a diagnostic test, Population-based descriptive study
Class D: Cross-sectional study, Case series, Case report
B. Reports that Synthesize or Reflect Upon Collections of Primary Reports:
Class M: Meta-analysis, Systematic review, Decision analysis, Cost-effectiveness analysis
Class R: Consensus statement, consensus report narrative review
Class X: Medical opinion
Citations are listed in the guideline utilizing the format of (Author, YYYY [report class]).
ICSI Evidence Grading System: www.icsi.org/guidelines_and_more/evidence_grading_system_6/

**1c.7 Summary of Controversy/Contradictory Evidence:** There is some controversy in the community in regards to the use of the PHQ-9 tool. Though widely accepted and implemented by primary care providers, behavioral specialists (psychiatrists, therapists) were more reluctant to use the PHQ-9 tool to assess the patient's status and to be measured in general. There has been an ongoing collaborative for behavioral providers the last 2 years to gain understanding and share experiences. There has been significant momentum and acceptance achieved and we currently have both primary care and behavioral health providers submitting data for this measure.

**1c.8 Citations for Evidence (***other than guidelines***):** Screening for Depression in Medical Settings with the

Patient Health Questionnaire (PHQ) A Diagnostic Meta Analysis. Gilbody, Simon et al Journal of General Internal Medicine Sept 2007
The PHQ-9: Validity of a Brief Depression Severity Measure Kurt Kroenke, MD, Robert Spitzer, MD J Gen Intern Med 2001 16:606-613
Collaborative Care for Depression: a Cumulative Meta-analysis and Review of Longer-term Outcomes. Gilbody, Simon et al Archives Internal Medicine Dec 2006

**1c.9** Quote the Specific guideline recommendation (*including guideline number and/or page number*):
ICSI Guideline Major Depression in Adults in Primary Care; page # referenced for each section.
The U.S. Preventive Services Task Force (USPSTF) recommends routine depression screening for all adults but only in clinical practices that have systems in place to assure accurate diagnosis, effective treatment and follow-up (U.S. Preventive Services Task Force, 2002 [R]). The purpose of this guideline is to assist ICSI members to develop systems that support effective diagnosis and treatment of major depression. A reasonable way to evaluate whether a system is successfully functioning in its diagnosis, treatment and follow-up of major depression would be to consider the following:
1. Diagnosis: The clinic or medical group should have a mechanism to assure that they are routinely evaluating for and documenting the presence for two weeks of at least five vegetative signs and symptoms of major depression (and that one includes sadness or loss of interest or pleasure in usual activities) in order to substantiate that the patient meets the DSM-IV TR criteria for major depression.
2. The clinic or medical group should have a systematic way to provide and document:
a. Engagement Education: The patient and his/her family is actively engaged and participating in self-management, based on knowledge of the nature of the disease, risk/benefits of treatment options, and consideration of patient preferences.
b. Ongoing Contacts: A documented system to assure ongoing contacts with the patient during the first six to twelve months of care (scheduled follow-up appointments, phone calls and some way to react and/or reach out if the patient drops out of treatment) based on use of a standardized, objective tool used at each contact to document and track treatment response.
3. Outcomes: The system should have a way of reliably and consistently monitoring outcomes of individuals and systemwide to improve individual care and the effectiveness of the clinical practice overall. [Page 8]

Screening
If depression is suspected, asking the two-question screen about mood and anhedonia may be as effective as using longer questionnaires:
Over the past month, have you been bothered by:
• Little interest or pleasure in doing things?
• Feeling down, depressed or hopeless?
If the patient answers "yes" to either of the above questions, use a quantitative, standardized instrument to document depressive symptoms and track treatment response (Pignone, 2002 [M]). This is used to supplement but not replace the clinical interview.
Multiple, practical questionnaires with reasonable performance characteristics are available to help clinicians identify and diagnose patients with major depression. In case-finding studies, average questionnaire administration times ranged from less than one minute to five minutes. While the two-question screen is effective with a broad population in primary care, a recent meta-analysis (Gilbody, 2007 [M]) concluded that for high-risk patients, screening with a nine-item Patient Health Questionnaire (PHQ-9) is more valid. The PHQ-9 has been validated for measuring depression severity (Kroenke, 2001 [C]; Spitzer, 1999 [C]). It can be administered telephonically (Pinto-Meza, 2005) and read to the patient. The factor structure of the nine items is comparable when tested with African Americans, Chinese Americans, Latino and non-Hispanic white patient groups (Huang, 2006 [C]). Other language versions that are validated for use in primary care are Spanish (Wulsin, 2002 [C]) and Chinese (Yeung, 2008 [C]). A Thai-language version has also been validated; however, the sensitivity is low (53%). This version could therefore be a useful and reasonable tool to help confirm a suspected depression but less so to screen general populations (Lotrakul, 2008 [C]).
Elderly patients with mild cognitive impairment can reliably fill out the PHQ-9 (Löwe, 2004 [C]).
A recent study found the PHQ-9 useful in psychiatric practices, as well. PHQ-9 scores influenced clinical decision-making for 93% of more than 6,000 patient contacts (Duffy, 2008 [C]).[Page 11]
Coding
The assessment of major depressive disorders should include the DSM-IV TR numerical rating of the disorder with all five digits, thus including a severity rating. For example, 296.22 (Major depressive disorder, single

episode, moderate severity).(American Psychiatric Association, 2000a [Not Assignable])
Depressive Disorder Not Otherwise Specified (Depression NOS), with a diagnosis code of 311 is designed for patients who do not meet criteria for Major Depression Disorder, Dysthymic Disorder, Adjustment Disorder with Depressed Mood or Adjustment Disorder with Mixed Anxiety and Depressed Mood. This is not a homogenous group of patients where there is evidence for best practice. If the patient meets criteria for Major Depressive Disorder or Dysthymic Disorder, it is important to diagnose and code them as such in order to proceed with evidence-based treatment. [Page 14]
Collaborative Care Model
More than 37 randomized controlled trials have demonstrated the effectiveness of the Collaborative Care Model. The work group recommends three key references (Gilbody, 2006 [M]; Hunkeler, 2006 [A]; Katon,1999 [A]) in which primary care treatment of depression is provided by a team (depression care manager, primary physician, consulting psychiatrist, others). This model has demonstrated improvement in treatment adherence, patient quality of life, and depression outcomes. Preliminary evidence suggests the collaborative care model is also effective for depression during pregnancy and postpartum (Gjerdingen, 2008 [M]).
The design of a team-based collaborative care approach involves:
• primary care providers using evidence-based approaches to depression care and a standard tool for measuring severity, response to treatment plan and remission;
• a systematic way of tracking and reminding patients at appropriate intervals of visits with their primary care physician and monitoring of treatment adherence and effectiveness;
• a team member (care manager role) to utilize the tracking system and make frequent contacts with the patients to provide further education, self-management support, and monitor for response in order to aid in facilitating treatment changes and in relapse prevention; and
• communication between primary care team and psychiatry to consult frequently and regularly regarding patient under clinical supervision, as well as direct patient visits as needed.(Unutzer, 2002 [A]) [Page 26]
Depression Treatment and Follow-Up Intervals Based on Severity
Treatment Recommendation: Education, Pharmacotherapy or psychotherapy. Start treatment and follow up plan, regularly re-evaluate and revise treatment plan.
Follow-up Interval: All depressed patients initially need weekly follow-up (phone or in person) for engagement in treatment, determine if following treatment plan, address side effects, and check if following through on any referrals.
PHQ-9 Score 10 – 14 Moderate- If patient is responding, contacts can extend as far as monthly.
PHQ-9 Score 15-19 Moderately Severe- If patient is responding, contacts can extend to every 2-4 weeks.
PHQ-9 Score > /= 20 Severe - Until significant response is achieved, contacts should remain weekly. Referral to mental health specialist may be warranted by PCP or psychiatrist.
Adapted from Kroenke and Spitzer, Psychiatric Annals 32:9 / September 2002 [Page 37]

If the primary care provider is seeing some improvement, continue working with that patient to augment or increase medication dosage to reach remission. This can take up to three months. Don't give up on the patient whether treating in primary care or referring. Stay connected through consultation or collaboration and take the steps needed to get the patient to remission. This can take longer and can take several medication interventions or other steps. The STAR*D study has shown that primary care can be just as successful as specialty care (Trivedi, 2006a [A]). [Page 37]
Measures
Improve the outcomes of treatment for major depression. (Annotations #11, 12)
Possible measures of accomplishing this aim:
a. Percentage of patients who have had a response to treatment at six months (+/- 30 days) after initiating treatment, e.g., have had a PHQ-9 score decreased by 50% from initial score at six months (+/- 30 days).
b. Percentage of patients who have reached remission at six months (+/- 30 days) after initiating treatment, e.g., have any PHQ-9 score less than 5 at six months (+/- 30 days).
c. Percentage of patients who have had a response to treatment at twelve months (+/- 30 days) after initiating treatment, e.g., have had a PHQ-9 score decreased by 50% from initial score at twelve months (+/- 30 days).
d. Percentage of patients who have reached remission at twelve months (+/- 30 days) after initiating treatment, e.g., have had any PHQ-9 score less than 5 at twelve months (+/- 30 days). [Page 77]

**1c.10** Clinical Practice Guideline Citation: ICSI Institute for Clinical Systems Improvement Health Care Guideline for Major Depression in Adults in Primary Care. 12th Edition May 2009

http://www.icsi.org/guidelines_and_more/gl_os_prot/behavioral_health/depression_5/depression__major__in_adults_in_primary_care_4.html

**1c.11 National Guideline Clearinghouse or other URL:** Please note that the ICSI guideline referenced is also listed in the National Guideline Clearinghouse but needs to be updated to the May 2009 version. Major depression in adults in primary care. Institute for Clinical Systems Improvement - Private Nonprofit Organization. 1996 Jan (revised 2008 May). 84 pages. [NGC Update Pending] NGC:006525. www.guideline.gov/summary/summary.aspx?doc_id=12617&nbr=6525

**1c.12 Rating of strength of recommendation** (*also provide narrative description of the rating and by whom*):
ICSI Scientific Document Development Process located at:
www.icsi.org/guidelines_and_more/document_development_process/

**1c.13 Method for r**ating strength of recommendation (*If different from* USPSTF system, *also describe rating and how it relates to USPSTF*):
ICSI's Conclusion Grade definitions parallel with USPSTF ratings of High, Moderate & Low.
CONCLUSION GRADES
Key conclusions (as determined by the work group) are supported by a conclusion grading worksheet that summarizes the important studies pertaining to the conclusion.
Grade I: The evidence consists of results from studies of strong design for answering the question addressed. The results are both clinically important and consistent with minor exceptions at most. The results are free of any significant doubts about generalizability, bias, and flaws in research design. Studies with negative results have sufficiently large samples to have adequate statistical power.
Grade II: The evidence consists of results from studies of strong design for answering the question addressed, but there is some uncertainty attached to the conclusion because of inconsistencies among the results from the studies or because of minor doubts about generalizability, bias, research design flaws, or adequacy of sample size. Alternatively, the evidence consists solely of results from weaker designs for the question addressed, but the results have been confirmed in separate studies and are consistent with minor exceptions at most.
Grade III: The evidence consists of results from studies of strong design for answering the question addressed, but there is substantial uncertainty attached to the conclusion because of inconsistencies among the results from different studies or because of serious doubts about generalizability, bias, research design flaws, or adequacy of sample size. Alternatively, the evidence consists solely of results from a limited number of studies of weak design for answering the question addressed.
Grade Not Assignable: There is no evidence available that directly supports or refutes the conclusion.

**1c.14 Rationale for using this guideline over others:**
The Institute for Clinical Systems Improvement (ICSI) is a unique organization that is widely respected for its collaborative efforts with guideline development. ICSI's purpose is to help improve patient care in Minnesota through collaboration and innovations in evidence-based medicine. The collaborative is unique in that it brings medical organizations, health plans and business representatives into the decision-making process. Providers in MN are engaged and respect this process and the resulting guideline recommendations.

| | |
|---|---|
| **TAP/Workgroup:** What are the strengths and weaknesses in relation to the sub-criteria 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, <u>as specified</u>, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (<u>evaluation criteria</u>) | 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*):
Adults age 18 and older with a diagnosis of major depression or dysthymia and an initial PHQ-9 score greater than nine who achieve remission at twelve months as demonstrated by a twelve month (+/- 30 days) PHQ-9 score of less than five.

**2a.2 Numerator Time Window** (*The time period in which cases are eligible for inclusion in the numerator*):
PHQ-9 scores are collected for each patient from the time they meet the inclusion criteria of diagnosis ICD-9 codes and PHQ-9 score greater than nine (this is the index or anchor date) until thirteen months have elapsed. This allows for calculation of a remission rate +/- 30 days from the index date.

**2a.3 Numerator Details** (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):
Adults age 18 and older; no upper age limit
Have the diagnosis of major depression or dysthymia defined by any of the following ICD-9* codes:
296.2x Major depressive disorder, single episode
296.3x Major depressive disorder, recurrent episode
300.4 Dysthymic disorder
AND
PHQ-9 Score is greater than nine.

Of the patients meeting the above inclusion criteria, the numerator is defined as those patients with a twelve month (+/- 30 days) PHQ-9 score of less than five.
The numerator rate is calculated as follows:
# adult pts with major depression or dysthymia (296.2x, 296.3x or 300.4) with a PHQ-9 score < 5 at 12 months(+/- 30 days)/
# adult pts with major depression or dysthymia (296.2x, 296.3x or 300.4)with index contact PHQ-9 > 9

Patients who do not have a twelve month +/- 30 day PHQ-9 score obtained are included in the denominator for this measure.

* For primary care providers the diagnosis codes can be in any position (primary or secondary). For behavioral health providers the diagnosis codes need to be in the primary position. This is to more accurately define major depression and exclude patients who may have other more serious mental health diagnoses (e.g. schizophrenia, psychosis) with a secondary diagnosis of depression.

**2a.4 Denominator Statement** (*Brief, text description of the denominator - target population being measured*):
Adults age 18 and older with a diagnosis of major depression or dysthymia and an initial PHQ-9 score greater than nine.

**2a.5 Target population gender:** Female, Male
**2a.6 Target population age range:** Age 18 and older

**2a.7 Denominator Time Window** (*The time period in which cases are eligible for inclusion in the denominator*):
PHQ-9 scores are collected for each patient from the time they meet the inclusion criteria of diagnosis ICD-9 codes and PHQ-9 score greater than nine (this is the index or anchor date) until thirteen months have elapsed. This allows for calculation of a remission rate +/- 30 days from the index date.

**2a.8 Denominator Details** (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):
Adults age 18 and older; no upper age limit
Have the diagnosis of major depression or dysthymia defined by any of the following ICD-9* codes:

2a-specs
C ☐
P ☐
M ☐
N ☐

296.2x Major depressive disorder, single episode
296.3x Major depressive disorder, recurrent episode
300.4 Dysthymic disorder
AND
PHQ-9 Score is greater than nine.
* For primary care providers the diagnosis codes can be in any position (primary or secondary). For behavioral health providers the diagnosis codes need to be in the primary position. This is to more accurately define major depression and exclude patients who may have other more serious mental health diagnoses (e.g. schizophrenia, psychosis) with a secondary diagnosis of depression.
Patients who do not have a twelve month +/- 30 day PHQ-9 score obtained are included in the denominator for this measure.

**2a.9 Denominator Exclusions** (*Brief text description of exclusions from the target population*): Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who are initially diagnosed with major depression and after further treatment are determined to have bipolar or personal disorders are excluded.

**2a.10 Denominator Exclusion Details** (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*):
•Patients who die during the measurement time frame
•Patients who are a permanent nursing home resident during the measurement time frame
•Patients who are enrolled in hospice during the measurement time frame
•Bipolar Disorder (Principal Diagnosis; initially diagnosed as depression but upon further treatment & evaluation primary diagnosis changed to bipolar disorder). See bipolar disorder codes below.
•Personality Disorder (Principal Diagnosis; initially diagnosed as depression but upon further treatment & evaluation primary diagnosis changed to personality disorder). See personality disorder codes below.
For patients with bipolar or personality disorder:
Do not exclude patients who have these bipolar or personality codes just because the codes are present. If the patient has major depression codes and bipolar or personality codes, the patient needs to be included. Exclusions are only to be used if the patient is initially thought to have major depression or dysthymia and it is determined at a later date that the patient has bipolar or personality disorder. For example, a patient is diagnosed in April with major depression and a PHQ-9 score of 23, therefore meeting the inclusion criteria. Several visits/ contacts with PHQ-9s occur in April and May. In June the patient has a first manic episode and is determined to have bipolar disorder. At this point the patient can be excluded from the denominator.
Bipolar Disorder Codes:
296.00 Bipolar I Disorder, Single Manic Episode, Unspecified
296.01 Bipolar I Disorder, Single Manic Episode, Mild
296.02 Bipolar I Disorder, Single Manic Episode, Moderate
296.03 Bipolar I Disorder, Single Manic Episode, Severe Without Psychotic Features
296.04 Bipolar I Disorder, Single Manic Episode, Severe With Psychotic Features
296.05 Bipolar I Disorder, Single Manic Episode, In Partial Remission
296.06 Bipolar I Disorder, Single Manic Episode, In Full Remission
296.10 Manic disorder, recurrent episode; Unspecified
296.11 Manic disorder, recurrent episode; Mild
296.12 Manic disorder, recurrent episode; Moderate
296.13 Manic disorder, recurrent episode; Severe Without Psychotic Features
296.14 Manic disorder, recurrent episode; Severe With Psychotic Features
296.15 Manic disorder, recurrent episode; In Partial Remission
296.16 Manic disorder, recurrent episode; In Full Remission
296.40 Bipolar I Disorder, Most Recent Episode Manic, Unspecified
296.41 Bipolar I Disorder, Most Recent Episode Manic, Mild
296.42 Bipolar I Disorder, Most Recent Episode Manic, Moderate
296.43 Bipolar I Disorder, Most Recent Episode Manic, Severe Without Psychotic Features
296.44 Bipolar I Disorder, Most Recent Episode Manic, Severe With Psychotic Features
296.45 Bipolar I Disorder, Most Recent Episode Manic, In Partial Remission
296.46 Bipolar I Disorder, Most Recent Episode Manic, In Full Remission
296.50 Bipolar I Disorder, Most Recent Episode Depressed, Unspecified
296.51 Bipolar I Disorder, Most Recent Episode Depressed, Mild
296.52 Bipolar I Disorder, Most Recent Episode Depressed, Moderate

296.53 Bipolar I Disorder, Most Recent Episode Depressed, Severe Without Psychotic Features
296.54 Bipolar I Disorder, Most Recent Episode Depressed, Severe With Psychotic Features
296.55 Bipolar I Disorder, Most Recent Episode Depressed, In Partial Remission
296.56 Bipolar I Disorder, Most Recent Episode Depressed, In Full Remission
296.60 Bipolar I Disorder, Most Recent Episode Mixed, Unspecified
296.61 Bipolar I Disorder, Most Recent Episode Mixed, Mild
296.62 Bipolar I Disorder, Most Recent Episode Mixed, Moderate
296.63 Bipolar I Disorder, Most Recent Episode Mixed, Severe Without Psychotic Features
296.64 Bipolar I Disorder, Most Recent Episode Mixed, Severe With Psychotic Features
296.65 Bipolar I Disorder, Most Recent Episode Mixed, In Partial Remission
296.66 Bipolar I Disorder, Most Recent Episode Mixed, In Full Remission
296.7 Bipolar I Disorder, Most Recent Episode Unspecified
296.80 Bipolar Disorder NOS
296.89 Bipolar II Disorder
Personality Disorder Codes:
301.0 Paranoid personality disorder
301.1 Affective personality disorder
301.10 Affective personality disorder unspecified
301.11 Chronic hypomanic personality disorder
301.12 Chronic depressive personality disorder
301.13 Cyclothymic disorder
301.2 Schizoid personality disorder
301.20 Schizoid personality disorder unspecified
301.21 Introverted personality
301.22 Schizotypal personality disorder
301.3 Explosive personality disorder
301.4 Obsessive-compulsive personality disorder
301.5 Histrionic personality disorder
301.50 Histrionic personality disorder unspecified
301.51 Chronic factitious illness with physical symptoms
301.59 Other histrionic personality disorder
301.6 Dependent personality disorder
301.7 Antisocial personality disorder
301.8 Other personality disorders
301.81 Narcissistic personality disorder
301.82 Avoidant personality disorder
301.83 Borderline personality disorder
301.84 Passive-aggressive personality
301.89 Other personality disorders
301.9 Unspecified personality disorder

**2a.11 Stratification Details/Variables** (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*):
This measure is currently not stratified. We will be convening a workgroup in the spring of 2010 to determine if stratification by severity of depression is clinically meaningful for data stratification and reporting.

**2a.12-13 Risk Adjustment Type:** Other (specify) Currently under exploration.

**2a.14 Risk Adjustment Methodology/Variables** (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*):
We are currently assessing the best variables for risk adjustment in this population. In preparing for this we are starting to collect gender, zip code, race & ethnicity, country of origin and primary language. We will be convening a workgroup in the spring of 2010 determine the best variables for risk adjustment for this population.

**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*):

This measure is calculated by submitting a visit level file for the eligible patients, each record in the file represents a contact with the patient and PHQ-9 score associated with this contact. Data file is submitted to a HIPAA secure data portal. Programming within the data portal determines the starting point (index visit) and then calculates based on dates if a twelve month +/- 30 days PHQ-9 was obtained and the resulting score.

Calculation logic:

Is patient eligible for inclusion with diagnosis codes of either 296.2x, 296.3x or 300.4 and PHQ-9 > 9?

If yes, mark the visit as index (anchor) and include this patient in the denominator.

Does patient have a PHQ-9 score completed with a contact date that is +/- 30 days from the index date?

If yes, include this score to calculate rate. Programming logic includes the most recent score within the +/- 30 day window.

If no, patient is included in the denominator only. Not having a PHQ-9 score within the 60 day window is considered a numerator miss.

If the patient does have a twelve month +/- 30 day PHQ-9 score is it less than five?

If twelve month +/- 30 day PHQ-9 is less than five; is considered a numerator case for rate calculation.

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

This measure is currently being collected on the full population of eligible patients; therefore significance testing of a sample methodology is not indicated. Of the approximately 165 clinics participating in this measure since September 2009, 97% have an EMR system in place. It is possible for a clinic currently on a paper chart system to use a registry or spreadsheet to track and report this information. Using an EMR with the PHQ-9 summary score as a reportable field does make this data collection more feasible and efficient.

**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)*:

The measure and its denominator is not based on a sample.

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

paper medical record/flowsheet, electronic Health/Medical Record, registry data

**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.*):

An excel template with formatted columns for data fields is provided. 97% of the 165 clinics participating currently in this measure extract the information from their EMR. Registries can be used as a source of information to create the data file; however groups must insure that all of their eligible patients are included. All data is uploaded in electronic format (.csv file) to a HIPAA secure, encrypted and password protected data portal.

**2a.26-28** Data source/data collection instrument reference web page URL or attachment: Attachment Depression_Template_Formatted-634001084723450218.xls

**2a.29-31** Data dictionary/code table web page URL or attachment: Attachment Depression DDS Measurement and Field Specifications 12-30-2009-634001085453445546.pdf

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

Clinicians: Other Clinic Site Level Reporting

**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*)

Clinicians: Psychologist/LCSW, Clinicians: PA/NP/Advanced Practice Nurse, Clinicians: Physicians (MD/DO), Other, Behavorial Health: Mental Health Psychiatrist

| TESTING/ANALYSIS |
|---|

**2b. Reliability testing**

**2b.1** Data/sample *(description of data/sample and size)*: Data collection started in September of 2008 with historical collection dates of service 1/1/2008 going forward. To date, there are over 165 clinics

**2b**
C ☐
P ☐
M ☐

participating and over 43,700 patients with major depression and an initial PHQ-9 > 9 submitted for outcome rate calculation. This measure was designed for EMR-based full population reporting; there is no sampling of the population. All patients meeting the inclusion criteria are included for rate calculation. However, it is possible for clinics on paper charts systems or with a registry to participate, but it is more labor intensive. Data reliability is further insured by data validation processes for every medical group for a sample of 30 patients flowing NCQA's "8 and 30" auditing processes. Groups must pass at 90% for data to be included for reporting.

**2b.2 Analytic Method** *(type of reliability & rationale, method for testing)*:
All groups submitting data undergo a denominator certification process to insure that the patient population is being identified and submitted correctly. Data specifications define all fields for the data submission file and include instructions for correct extraction. The data is extracted directly from the EMR and then the outcome rates are calculated based on rules & programming logic. Audit of all medical groups submitting data has demonstrated that the information is reliably collected and submitted with all groups passing an NCQA "8 and 30" audit method at greater than 90%.
The measure itself is determined to have content validity based on expert panel and workgroups associated with ICSI Adult Major Depression Guideline and the DIAMOND "Depression Improvement Across Minnesota, Offering a New Direction" project. For more information please refer to www.icsi.org/health_care_redesign_/diamond_35953/

**2b.3 Testing Results** *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
Reliability of data collection methods has been demonstrated by ongoing audits of all groups submitting data against the patient's medical record. Issues related to sampling or sample size are not in issue for this measure.

N ☐

**2c. Validity testing**

**2c.1 Data/sample** *(description of data/sample and size)*: For patients with an index contact date between 1/1/2008 and 8/31/2008 (n = 17,085); the twelve month remission rate is 4.6% (790/17,085). This is a very similar to our currently reported six month remission rate of 4.2%, however demonstrates significant room for improvement. The unit of measure for analysis is the clinic site location. This data set represents 123 clinics that had sufficient history submitted to calculate a twelve month remission rate and a denominator of at least 30 patients.

**2c.2 Analytic Method** *(type of validity & rationale, method for testing)*:
The measure itself is determined to have content validity based on expert panel and workgroups associated with ICSI Adult Major Depression Guideline and the DIAMOND "Depression Improvement Across Minnesota, Offering a New Direction" project. Experts agreed on the use of a common tool (PHQ-9) and that a score of less than five represented remission or an absence of depression symptoms. For more information please refer to www.icsi.org/health_care_redesign_/diamond_35953/.

**2c.3 Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
Data collection and submission for this measure started in September 2008, currently there is enough data to capture preliminary twelve month remission rates for patients with an index contact date between 1-1-2008 and 8-31-2008, n = 17,085 patients representing 123 clinics. For these patients, the twelve month remission results are as follows:
Average :        4.6%  (790/17,085)
Range:           0.0% to 29.7%
StDev:           3.8

8 of the 92 clinics included in the data set of patients with an index contact date between 1/1/2008 and 8/31/2008 and a denominator that is greater than equal to 30 had confidence intervals fully above the mean, demonstrating meaningful difference and variability within the data.
Low remission rates at twelve months are in partially related to the medical group's ability to obtain a PHQ-9 score at twelve months +/- 30 days; currently only 20.3% of the eligible denominator patients have a 12 month PHQ-9. Groups are diligently working on process improvements in workflow to impact ongoing contact with patients.

**2c**
C ☐
P ☐
M ☐
N ☐

We plan to publically report twelve month remission scores in June of 2010 when enough data has accumulated for patients with an index contact date 1-1-2008 to 12-31-2008, as a run-out of thirteen months of data is needed to capture the twelve month remission rates. At this point we will have over 25,000 patients for twelve month remission rate calculation.

## 2d. Exclusions Justified

**2d.1 Summary of Evidence supporting exclusion(s):**
Population definitions and exclusions are in accordance with ICSI guidelines as the focus is on major depression & dysthymia. ICD-9 codes selected for inclusion were given careful consideration by the work group in order to select a heterogeneous population for measuring treatment outcomes. The assessment of major depressive disorders should include the DSM-IV TR numerical rating of the disorder with all five digits, thus including a severity rating. For example, 296.22 (Major depressive disorder, single episode, moderate severity). (American Psychiatric Association, 2000a [Not Assignable])
Depressive Disorder Not Otherwise Specified (Depression NOS), with a diagnosis code of 311 is designed for patients who do not meet criteria for Major Depression Disorder, Dysthymic Disorder, Adjustment Disorder with Depressed Mood or Adjustment Disorder with Mixed Anxiety and Depressed Mood. This is not a homogenous group of patients where there is evidence for best practice. If the patient meets criteria for Major Depressive Disorder or Dysthymic Disorder, it is important to diagnose and code them as such in order to proceed with evidence-based treatment. [Page 14, ICSI guidelines]
During the initial stages of the pilot, it was realized that patients who were initially diagnosed with major depression, but after further visits/ time was discovered to have bipolar or personality disorder. Because this is a visit level data submission, a method was needed to exclude these patients and this was developed.

**2d.2 Citations for Evidence:**
ICSI Institute for Clinical Systems Improvement Health Care Guideline for Major Depression in Adults in Primary Care. 12th Edition May 2009
http://www.icsi.org/guidelines_and_more/gl_os_prot/behavioral_health/depression_5/depression__major__in_adults_in_primary_care_4.html

**2d.3 Data/sample** (description of data/sample and size): Exclusions are utilized for only 0.7% (325/44,621) of the total population of patients submitted; the most frequent exclusion utilized were for bipolar and personality disorders.

**2d.4 Analytic Method** (type analysis & rationale):
Descriptive, simple analysis of exclusions was used to analyze the data submitted by medical groups.

**2d.5 Testing Results** (e.g., frequency, variability, sensitivity analyses):
Exclusions represent a very small portion of the eligible denominator patients and do not impact overall rates. In the current data set of over 44,000 patients, exclusions were utilized for 0.7% of the patients.

| | 2d |
|---|---|
| | C☐ |
| | P☐ |
| | M☐ |
| | N☐ |
| | NA☐ |

## 2e. Risk Adjustment for Outcomes/ Resource Use Measures

**2e.1 Data/sample** (description of data/sample and size): Measure is not currently risk-adjusted. We are currently assessing the best variables for risk adjustment in this population. In preparing for this we are starting to collect gender, zip code, race & ethnicity, country of origin and primary language. We will be convening a workgroup in the spring of 2010 determine the best variables for risk adjustment for this population.

**2e.2 Analytic Method** (type of risk adjustment, analysis, & rationale):
Measure is not currently risk-adjusted.

**2e.3 Testing Results** (risk model performance metrics):
Measure is not currently risk-adjusted.

**2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:** We are currently assessing the best variables for risk adjustment in this population. In preparing for this we are starting to collect gender, zip code, race & ethnicity, country of origin and primary language. We will be convening a workgroup in the spring of 2010 determine the best variables for risk adjustment for this population.

| | 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)*:  This is a fairly new measure for our community and the state of Minnesota. Measures for depression were developed in concert with ICSI during the development of the DIAMOND "Depression Improvement Across Minnesota, Offering a New Direction". DIAMOND, a new collaborative care management model, was launched in March of 2008. Data collection for the population based depression measures began in September 2008 with some historical data being captured back to dates of service 1-1-2008. Over 165 clinics have participated in the depression measures, one of which is the twelve month remission. Currently we have over 43,700 patients in our data set and this continues to grow as more clinics join the program and more patients are identified. The percentages of initial PHQ-9 scores are as follows: (a patient's PHQ-9 score needs to be 10 or above for inclusion) 43% moderate; PHQ-9 score 10 to 14, 34% moderately severe; PHQ-9 score 15 to 19, and 23% severe; PHQ-9 score 20 to 27. During our first cycle of reporting, groups had the option of reporting publicly or suppressing reporting but having the results available for their own.  Preliminary twelve month remission rate is 4.6%, based on over 17,000 patients, we plan to publicly report scores in June of 2010 at www.mnhealthscores.org.  These initial results demonstrate significant opportunity for improvement; one of the most difficult things impacting measurement, but is also reflective in the care and management of patients with depression is maintaining an ongoing connection. One of the reasons for the DIAMOND project's early success is the care coordination & case manager role with an expectation of frequent contact between patient and provider.

**2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*:
Outcome results will be displayed in June 2010 on the public website MN HealthScores at www.mnhealthscores.org and can be ranked in order of performance or by the name of the clinic and comparisons are made to the state average. This measure will also be included in the annual Health Care Quality Report located at: www.mncm.org/site/ which incorporates rankings in this report based on percent and confidence intervals.

**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)*:
 Data collection and submission for this measure started in September 2008, currently there is enough data to capture preliminary twelve month remission rates for patients with an index contact date between 1-1-2008 and 8-31-2008, n = 17,085 patients representing 123 clinics.  For these patients, the twelve month remission results are as follows:
Average :         4.6%  (790/17,085)
Range:            0.0% to 29.7%
StDev:            3.8

8 of the 92 clinics included in the data set of patients with an index contact date between 1/1/2008 and 8/31/2008 and a denominator that is greater than equal to 30 had confidence intervals fully above the mean, demonstrating meaningful difference and variability within the data.

**2f**
C ☐
P ☐
M ☐
N ☐

**2g. Comparability of Multiple Data Sources/Methods**

**2g.1 Data/sample** *(description of data/sample and size)*:  Multiple data sources are not used. The data source for this information is the patient's medical record. No other sources of information are applicable (e.g. is not a claim based measure as serial PHQ-9 scores are needed to calculate this measure). Information can be obtained either from a query of the electronic medical record or via registry or chart abstraction. The measure was designed for full population extraction from an EMR and 97% of the current participating clinics have an EMR in place.

**2g.2 Analytic Method** *(type of analysis & rationale)*:
NA

**2g.3 Testing Results** *(e.g., correlation statistics, comparison of rankings)*:
NA

**2g**
C ☐
P ☐
M ☐
N ☐
NA ☐

| | |
|---|---|
| **2h. Disparities in Care**<br><br>**2h.1 If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: Measure is currently not stratified. This measure is currently not stratified. We will be convening a workgroup in the spring of 2010 to determine if stratification by severity of depression is clinically meaningful for data stratification and reporting. Additionally, we will be assessing the best variables for risk adjustment in this population. In preparing for this we are starting to collect gender, zip code, race & ethnicity, country of origin and primary language.<br><br>**2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:**<br>Future direct data submissions will include fields for gender, race/ethnicity, country of origin and primary language and will allow further analysis of potential disparities. | **2h**<br>C ☐<br>P ☐<br>M ☐<br>N ☐<br>NA ☐ |
| **TAP/Workgroup: What are the strengths and weaknesses in relation to the** sub-criteria for *Scientific Acceptability of Measure Properties?* | 2 |
| Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met?<br>Rationale: | *2*<br>C ☐<br>P ☐<br>M ☐<br>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**<br><br>**3a.1 Current Use:** in use<br><br>**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)*:<br>The twelve month remission rates for depression are scheduled to be publicly reported by MN Community Measurement on their consumer website located on the MN HealthScores Website at www.mnhealthscores.org in June 2010. MN Community Measurement is a collaborative effort in our community among those who believe that you cannot improve what you don't measure. Our collaborative includes medical groups, clinics, physicians, hospitals, health plans, employers, consumer representatives and quality improvement organizations. These stakeholders support the notion that greater transparency in our health care system will lead to better health outcomes for the people of Minnesota. MN Community Measurement's mission to accelerate the improvement of health by publicly reporting health care information is having a positive effect on the health care provided in Minnesota. For more information please visit our corporate website at www.mncm.org.<br><br>**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)*:<br>The six month remission measure (related ) was the measure selected in our community for pay for performance, Bridges to Excellence and the Minnesota Department of Health Statewide Quality Reporting System for 2011. The twelve month remission, though currently not a part of an external pay for performance program, will be publicly reported in June 2010. Use of the data for quality improvement efforts is encouraged and results reporting within the data portal assist groups in understanding potential opportunity for several process and outcome measures for depression. There is a compare function built into the public reporting website so that consumers (or providers) can pick clinics to be compared; additionally medical groups have access to their own detailed patient level results with numerator calculation within our HIPAA secure data portal. Groups can use this information to better understand their depression population in terms of processes and outcomes. | **3a**<br>C ☐<br>P ☐<br>M ☐<br>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*):  Consumer: In November of 2009, Robert Wood Johnson Foundation in conjunction with the Aligning Forces for Quality AF4Q Communities, which includes Minnesota, conducted a study "Producing Online Performance Reports that People Understand" to identify best practices for performance reporting that consumers understand. Six focus groups were conducted that included a total of 50 individuals.
In June of 2007, a series of three consumer focus groups were interviewed (28 individuals) to provide feedback about our old website. A new, enhanced website was launched in 2009 and additional feedback was sought from a focus group (5 individuals)
Providers: August 2008 and August 2009 (102 respondents)
Direct Data Submission Users: July 2009 (96 respondents)

**3a.5 Methods** (*e.g., focus group, survey, QI project*):
Feedback was obtained from consumers via focus groups, while surveys were utilized for both provider and direct data submission user feedback.

**3a.6 Results** (*qualitative and/or quantitative results and conclusions*):
Consumer:
In November of 2009, Robert Wood Johnson Foundation in conjunction with the Aligning Forces for Quality AF4Q Communities, which includes Minnesota, conducted a study "Producing Online Performance Reports that People Understand" to identify best practices for performance reporting that consumers understand. Six focus groups were conducted that included a total of 50 individuals. The study compared six different AF4Q communities (Minnesota, Detroit, Seattle, Wisconsin, Memphis, Pennsylvania and Maine), Minnesota ranked # 1 for easiest and clearest graphics and the search function which was rated as one of the most important consumer "wants".
In June of 2007, a series of three consumer focus groups were interviewed (28 individuals) to provide feedback about our old website. Some interesting feedback was obtained about our composite measures: accept responsibility for their own health outcomes, health care quality is not uniform across sites, awareness of the website is low, value having the information available during open enrollment and that the website is fairly easy to use. A new, enhanced website was launched in 2009 and additional feedback was sought from a focus group (5 individuals) that reacted positively about the new search and compare capabilities.
Providers: August 2008- Physicians were involved in the data portal redesign of the results display in terms of what additional information would be useful to them in using the data for quality improvement efforts. Providers liked the enhancements, display of the breakdown of the individual components and ability to download their own group's specific patient level data for use in further analysis.
August 2009- Survey to medical groups with 102 respondents
* 65% feel that MNCM is selecting measures that drive the most important improvement in health care
* 59% MNCM is accelerating the improvement of care by publicly reporting information
* 67% have visited the new public website MNHealthScores and 74% the corporate website

**3b/3c. Relation to other NQF-endorsed measures**

**3b.1 NQF # and Title of similar or related measures:**
NQF # 0105 Title: New Episode of Depression: (a) Optimal Practitioner Contacts for Medication Management, (b) Effective Acute Phase Treatment, (c) Effective Continuation Phase Treatment and NQF # 0418 Title: Screening for Clinical Depression

(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?**
The twelve month remission measure is similar in terms of the target population, however the measures are different. The twelve month remission measure is an outcome determining the effectiveness of the treatment and plan of care and measures the improvement of depression symptoms. The existing NQF endorsed measures are related to medication adherence and screening of the general population, and are not related to the outcome of treatment.

**3b**
C ☐
P ☐
M ☐
N ☐
NA ☐

| | |
|---|---|
| **3c. Distinctive or Additive Value**<br>**3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:**<br>This measure is similar only in terms of target population; there is no similar measure that is capturing an outcome for patients with depression. The twelve month remission measure provides improved and additive value because it demonstrates, based on a standardized tool, an objective measurement of the improvement in a patient's depression symptoms.<br><br>**5.1 Competing Measures** 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:<br>There is no current NQF endorsed measure that demonstrates the effectiveness of treatment and measures remission (relief of symptoms). Patients who are able to maintain at PHQ-9 score of less than five are demonstrating improved mental health and functional status. | **3c**<br>C☐<br>P☐<br>M☐<br>N☐ |
| **TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for *Usability*?** | **3** |
| **Steering Committee: Overall, to what extent was the criterion, *Usability*, met?**<br>**Rationale:** | **3**<br>C☐<br>P☐<br>M☐<br>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 Rati ng** |
| **4a. Data Generated as a Byproduct of Care Processes**<br><br>**4a.1-2 How are the data elements that are needed to compute measure scores generated?**<br>data generated as byproduct of care processes during delivery, coding/abstraction performed by someone other than person obtaining original information, | **4a**<br>C☐<br>P☐<br>M☐<br>N☐ |
| **4b. Electronic Sources**<br><br>**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*)<br>Yes<br><br>**4b.2 If not, specify the near-term path to achieve electronic capture by most providers.** | **4b**<br>C☐<br>P☐<br>M☐<br>N☐ |
| **4c. Exclusions**<br><br>**4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?**<br>No<br><br>**4c.2 If yes, provide justification.** | **4c**<br>C☐<br>P☐<br>M☐<br>N☐<br>NA☐ |
| **4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences**<br><br>**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.**<br>MN Community Measurement has modeled the direct data submission to minimize inaccuracies, errors and unintended consequences. All groups participating sign a terms of use agreement that delineates the group's responsibilities for submission of data and consequences for not participating in good faith. Additionally all groups sign a Business Associate Agreement that outlines the use of the data. Denominator certification prior to any data collection insures that groups are following the specifications and correctly identifying their population and serves as a point of correction prior to the expenditure of resources for data collection. | **4d**<br>C☐<br>P☐<br>M☐<br>N☐ |

Groups provide documentation of cases that are excluded and this is reviewed by MNCM staff prior to approval of the data submission. Extensive audit processes also support the data's accuracy. After data submission, in person validation audits are conducted comparing the submission to the patient's medical record using NCQA's 8 and 30 rule for audit requiring a 90% accuracy rate. Groups are only allowed three patient records with error out of 30 reviewed in order to achieve 90%. Audits are conducted in the following instances: 1) a random sample of clinics with prior successful submission, 2) for all groups who are new to the submission process, 3) a group who has had a change in system or process (e.g. went from paper charts to EMR) since the last submission or 4) any group with a history of prior unsuccessful audit. It has been our experience that the post submission audits have identified both issues with data extraction programming from an EMR and abstraction errors when data is collected from the chart. Groups have been amenable to remedy plans, resubmission and re-audit. The only reason groups tended to fail an audit was by the inadvertent omission of some of the PHQ-9 scores that should have been submitted. This was remedied by an adjustment to the EMR extract process and the majority of groups were able to successfully resubmit the data and pass audit at > 90%.

**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:
Over the last three years during the direct data submission process for three measures (Optimal Diabetes Care, Optimal Vascular Care and Depression Measures) we have learned the following:
1. Data Submission- Providing data collection software for medical groups wishing to submit data was not always the best and most efficient way of collecting data. As electronic health records use becomes more pervasive in our state, providing templates of data file submissions proved to be more efficient.
2. Specifications- Detailed specifications with instructions on how to handle most situations (e.g. detailed instructions on blood pressure values) has been valuable to medical groups, increased data accuracy and resulted in 98% of groups submitting data successfully.
3. Audit- Audit methods have insured the accuracy of our data and we are able to successfully compare providers because everyone is pulling their data the same way and subject to the same rules.
4. Confidentiality- Patient confidentiality has been addressed by numerous mechanisms. MNCM only receives the patient level information needed to calculate the rates, determine eligibility for inclusion in the measure and support the administration of pay for performance programs. The PHI submitted is minimal and the data is protected by 1) password protection with password only available to the medical group submitting data, 2) file upload process is encrypted as data is transferred and 3) Data is stored on a separate secure server and meets all HIPAA protection rules.
6. Acceptance of Data- Vast improvement in terms of the timeliness of the data submitted by medical groups six weeks after the end of the measurement period as compared to prior method of health plan's samples and the results over a year old. Providers are more accepting of the results as compared to previous methods of pooling health plan samples.
7. Data Collection Burden- We have learned that for additional future measures we will need to stagger the data collection time frames and submission deadlines as to not burden the medical groups in terms of abstraction/ extraction (e.g. can't always have a measurement period Jan 1st to Dec 31st reported the second week of February, may need to consider July 1st to June 30th with data submission in August)
8. Health Plans: pay for performance and the inclusion of measures within contracts significantly impacts the number of groups participating in each measure.

**4e.2** Costs to implement the measure (*costs of data collection, fees associated with proprietary measures*):
Medical Groups: There are no fees charged to medical groups to submit their data to MNCM. There are no fees for the procurement and use of the PHQ-9 tool. Data collection costs (staff time to write an extract program from EMR) are absorbed by the medical groups submitting data. Clinics that have the PHQ-9 tool embedded in their EMR and the total score as a reportable field are more efficient than clinics that do not have these features.
Administrative (Costs to MNCM): Costs are associated with staffing. Currently, there is one full time project manager and one part time project coordinator dedicated to the direct data submission project and services for validation audits are contracted with abstractor during a 4 – 6 week period each year. Responsibilities include creation and annual update of the direct data submission guide, recommendations for data portal enhancements, communication to users, denominator certification, training of auditors for validation,

4e
C☐
P☐
M☐
N☐

availability for all questions & problems related to specs and submission, planning and performing some of the validation audits and approving data for publication.
It is estimated that the startup costs for the development of our data portal was approximately $25,000 for both the diabetes and ischemic vascular composite measures; additional costs of $10,000 were incurred to add multiple depression measures.

**4e.3** Evidence for costs:
MNCM contracts with portal vendor (historical) and budget.
Staff's experience with data collection at numerous clinic sites.

**4e.4** Business case documentation: Prior to implementing the direct data submission process for the depression measures, MN Community Measurement and it stakeholders knew there was great variability in the care and management that was being provided to patients and preliminary results for remission at twelve months demonstrated very low overall rates and significant room for improvement. Groups were not experienced in collecting data for depression, however over 400 primary care clinics were already used to collecting and submitting data for two composite measures for diabetes and vascular disease and demonstrated success with the DDS Direct Data Submission process.

| | |
|---|---|
| TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for *Feasibility*? | 4 |
| Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met?<br>Rationale: | 4<br>C ☐<br>P ☐<br>M ☐<br>N ☐ |

## RECOMMENDATION

| | |
|---|---|
| (for NQF staff use)  Check if measure is untested and only eligible for time-limited endorsement. | Time-limited<br>☐ |
| Steering Committee: Do you recommend for endorsement?<br>Comments: | Y ☐<br>N ☐<br>A ☐ |

## CONTACT INFORMATION

**Co.1** Measure Steward (Intellectual Property Owner)
**Co.1** Organization
MN Community Measurement  | 3433 Broadway Street NE , Suite # 455  | Minneapolis  | Minnesota | 55413

**Co.2** Point of Contact
Anne  | Snowden , MPH, CPHQ  | snowden@mncm.org  | 612-454-4811

Measure Developer If different from Measure Steward
**Co.3** Organization
MN Community Measurement  | 3433 Broadway Street NE , Suite # 455  | Minneapolis  | Minnesota | 55413

**Co.4** Point of Contact
Anne  | Snowden , MPH, CPHQ  | snowden@mncm.org  | 612-454-4811

**Co.5** Submitter If different from Measure Steward POC
Collette  | Pitzen , RN, BSN, CPHQ  | pitzen@mncm.org  | 612-454-4815-  |MN Community Measurement

**Co.6** Additional organizations that sponsored/participated in measure development
ICSI – Institute for Clinical Systems Improvement

## 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.

Nancy Jaeckels - Institute for Clinical Systems Improvement, Vice President Member Relations & Strategic Initiatives

Katrina Beckstrom - HealthPartners - Care Innovation & Measurement, Senior Quality Coordinator

Michael Trangle, MD - HealthPartners Clinics/ Regions Hospital, Associate Medical Director

Kenneth Joslyn, MD - Medica Health Plan, Medical Director Quality and Population Health

Jim Chase - MN Community Measurement, President

Diane Mayberry - MN Community Measurement, Chief Operating Officer

Anne Snowden - MN Community Measurement, Director of Performance Measurement & Reporting

Carrie Trygstad- MN Community Measurement, Project Manager

This group worked in concert with the ICSI DIAMOND project for measure development; the original charter of the workgroup includes:

o Develop population-wide, ambulatory care measures(s) of the quality of care for patients diagnosed with Depression, consistent with the work of the ICSI DIAMOND project.

• Process measures (ie: quality of coding, depression screening, completion of PHQ-9 survey, 3 month follow-up visit)

• Outcome measures (response and remission rates)

o Develop direct data collection, submission and reporting plan

• Physicians and non-physicians

• Primary care and Behavioral Health Care

**Ad.2** If adapted, provide name of original measure: The measure is not adapted.
**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: 2009-12
**Ad.8** What is your frequency for review/update of this measure? Annual
**Ad.9** When is the next scheduled review/update for this measure? 2010-12

**Ad.10** Copyright statement/disclaimers: © MN Community Measurement, 2009. All rights reserved
PHQ-9: Copyright © 2005 Pfizer, Inc. All rights reserved

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

Date of Submission (*MM/DD/YY*): 01/27/2010