Title

Depression Quality of Care: Measuring Quality over Time Using VA Electronic Medical Record Data.

Journal

Journal of General Internal Medicine, 31 Suppl 1(Suppl 1)

Authors

Farmer, Melissa
Rubenstein, Lisa
Sherbourne, Cathy
et al.

Publication Date

2016-04-01

DOI

10.1007/s11606-015-3563-4

Peer reviewed

Depression Quality of Care: Measuring Quality over Time Using VA Electronic Medical Record Data

Melissa M. Farmer, PhD[1], Lisa V. Rubenstein, MD, MSPH[1,2,3], Cathy D. Sherbourne, PhD[2], Alexis Huynh, PhD, MPH[1], Karen Chu, MS[1], Christine A. Lam, MD, MBA[1], Jacqueline J. Fickel, PhD[1], Martin L. Lee, PhD[1,3], Maureen E. Metzger, PhD, MPH[4], Lilia Verchinina, PhD[4], Edward P. Post, MD, PhD[4,5], and Edmund F. Chaney, PhD[6]

[1]VA HSR&D Center for the Study of Healthcare Innovation, Implementation & Policy, VA Greater Los Angeles Healthcare System, Sepulveda, CA, USA; [2]RAND, Santa Monica, CA, USA; [3]David Geffen School of Medicine and Fielding School of Public Health, University of California, Los Angeles, CA, USA; [4]VA Ann Arbor, HSR&D Center for Clinical Management Research, Ann Arbor, MI, USA; [5]University of Michigan Medical School, Ann Arbor, MI, USA; [6]University of Washington School of Medicine, Seattle, WA, USA.

BACKGROUND: The Veterans Health Administration (VA) has invested substantially in evidence-based mental health care. Yet no electronic performance measures for assessing the level at which the population of Veterans with depression receive appropriate care have proven robust enough to support rigorous evaluation of the VA's depression initiatives.

OBJECTIVE: Our objectives were to develop prototype longitudinal electronic population-based measures of depression care quality, validate the measures using expert panel judgment by VA and non-VA experts, and examine detection, follow-up and treatment rates over a decade (2000–2010). We describe our development methodology and the challenges to creating measures that capture the longitudinal course of clinical care from detection to treatment.

DESIGN AND PARTICIPANTS: Data come from the National Patient Care Database and Pharmacy Benefits Management Database for primary care patients from 1999 to 2011, from nine Veteran Integrated Service Networks.

MEASURES: We developed four population-based quality metrics for depression care that incorporate a 6-month look back and 1-year follow-up: detection of a new episode of depression, 84 and 180 day follow-up, and minimum appropriate treatment 1-year post detection. Expert panel techniques were used to evaluate the measure development methodology and results. Key challenges to creating valid longitudinal measures are discussed.

KEY RESULTS: Over the decade, the rates for detection of new episodes of depression remained stable at 7–8 %. Follow-up at 84 and 180 days were 37 % and 45 % in 2000 and increased to 56 % and 63 % by 2010. Minimum appropriate treatment remained relatively stable over the decade (82–84 %).

CONCLUSIONS: The development of valid longitudinal, population-based quality measures for depression care is a complex process with numerous challenges. If the full spectrum of care from detection to follow-up and treatment is not captured, performance measures could actually mask the clinical areas in need of quality improvement efforts.

Published online March 7, 2016

KEY WORDS: performance measurement; depression; Veterans; measurement; quality assessment.
J Gen Intern Med 31(Suppl 1):S36–45
DOI: 10.1007/s11606-015-3563-4

INTRODUCTION

Depression is the leading cause of reversible disability among primary care patients, the major precursor to suicide, contributes to the development and severity of chronic illnesses such as heart disease and diabetes, and increases cost among affected patients with comorbid diseases.[1–5] Veterans experience a high burden of depression, with approximately 12 % of Veterans attending the Veterans Health Administration (VA) primary care practices having symptoms of major depression.[6]

The VA has invested substantially in evidence-based mental health care. A major initiative focused on implementing collaborative care management (CCM) for depression care in primary care was incorporated into the VA Uniform Mental Health Services Handbook (2007)[7] mandating primary care/mental health integration nationally. It specifically recognized collocated mental health specialists in primary care and CCM as requirements for all primary care sites with at least 5000 patients. CCM clinical goals are to ensure that primary care patients identified with depression are assessed, treated, followed frequently, and receive self-management support as indicated in national depression guidelines.[8] Management of depression in primary care using CCM for improving treatment can significantly reduce depressive symptoms,[9–11] lower risk of depression reoccurrence,[12] prevent prolonged disability,[13] job loss,[14–16] negative life events,[17] and reduce suicide rates.[18]

Yet achieving these favorable outcomes in routine clinical practice is challenging, and a barrier has been the lack of valid, reliable measures to evaluate improvement efforts.[19–21] For many conditions, performance metrics have been a major force for improved care.[22] VA has required yearly screening

for depression over the last decade and has used screening rates to monitor performance (as with HbA1c). In 2008, VA introduced additional national performance measures to assess depression follow-up, only to withdraw them 2 years later. These additional measures encountered substantial resistance from primary care practice sites due to difficulties in interpreting site-level results—the level at which CCM improvements must be implemented. Therefore, our goal was to develop and validate prototype electronic measures suitable for evaluating VA's CCM initiative from 2000 to 2010 at the primary care practice site level. To this end, we developed site-level measures that used only electronic data, followed patients longitudinally through detection and treatment, and reflected care for the full primary care population at a given primary care location.[23, 24] We based the measures on depression guidelines and on those used in prior CCM evaluations.[14] We then reviewed the measures and our development methodology with an expert panel of VA and non-VA experts.

Measure development was guided by the Donabedian model of quality that links healthcare structure, process and outcomes.[25] Prior work has shown a link with depression process measures derived from administrative data and hospitalization outcomes.[26] In terms of reducing or eliminating depression symptoms, though, the lack of timed and electronically documented symptom assessment data means that neither administrative review nor electronic measures can measure depression symptom outcomes. As done in prior studies,[27, 28] our work focuses on actionable guideline-based depression care processes;[29] guidelines in turn reflect evidence linking these processes to outcomes.[30] The objectives are to 1) describe the development process and challenges to developing prototype measures that capture the longitudinal course of clinical care from detection through treatment; 2) assess measure validity and limitations through an expert panel; and 3) examine the proportion of Veterans in primary care that met those depression quality measures (detection, follow-up and minimal appropriate treatment) at the VA from 2000 to 2010.

METHODS

Design

We used a systematic approach to identify a population, determine inclusion and exclusion criteria for the cohorts, develop quality measures, and map how each patient was accounted for in each measure. Per the domains identified by Hermann,[31] we expected our measures to be meaningful, feasible, and actionable for quality improvement. For meaningfulness, we relied on literature, guidelines and an expert panel. For feasibility, we assessed whether we could program measures that accurately accounted for the full target population at each branch node of our measurement algorithm using counting trees. For actionability, we used benchmarks from the literature, interpretability by our expert panel, and applicability to assessing CCM.

Expert Panel

In March 2015, we convened a 1-day modified Delphi expert panel[17, 32] to review our development methods and the results of applying our measures. Panelists included VA and non-VA experts in quality and performance measurement, depression, primary care/mental health integration, and program evaluation. Prior to the meeting, panelists received a detailed report on measure development and results. They completed an online survey (available from the authors) to evaluate sampling decisions and definitions, and on the importance and feasibility of electronic depression quality measurement. Survey results were presented to the panelists at the meeting. Summary notes were taken on a flip chart, and panelists were asked to vote on the summaries in real time. Two investigators took detailed notes.

Data

We used existing VA electronic medical record data from the National Patient Care Database and prescription data from Pharmacy Benefits Management Database to pull cohorts of all primary care patients during the federal fiscal years (FY) 2000–2010 from nine Veteran Integrated Service Networks (VISNs). Our larger project focused on the VA implementation of CCM over the decade, so we included four VISNs that implemented CCM early and five additional VISNs from across the United States to represent diverse levels of involvement with mental health primary care models and geographical diversity. The Greater Los Angeles Human Subjects Institutional Review Board approved this study.

Identification of the Population

Identifying the appropriate patient population on which to apply quality measures is a critical component of measure development. In our case, this challenge included identifying patients seen in primary care, determining the index visit, establishing continuity of care, and validating an algorithm to exclude patients with recent prior diagnosis or depression treatment. Issues related to timing of visits, timing of exclusion criteria, and cut-points for prescription medications and refills had to be considered. Figure 1 illustrates the development process.

Primary Care Cohort and Index Visit. The cohorts of patients for each measure included all patients seen in primary care for each FY, 2000–2010. The baseline visit (defined "index visit") was a patient's first primary care visit after the start of a given FY based on primary care visit encounter identifiers (for the VA: "clinic stop codes").

Continuously Seen Cohort and Home Site. The patient must have been seen at their primary care site at least once within the 12 months prior to the index visit (T_0-12) and again in the 12 months after the index visit (T_0+12). This definition of

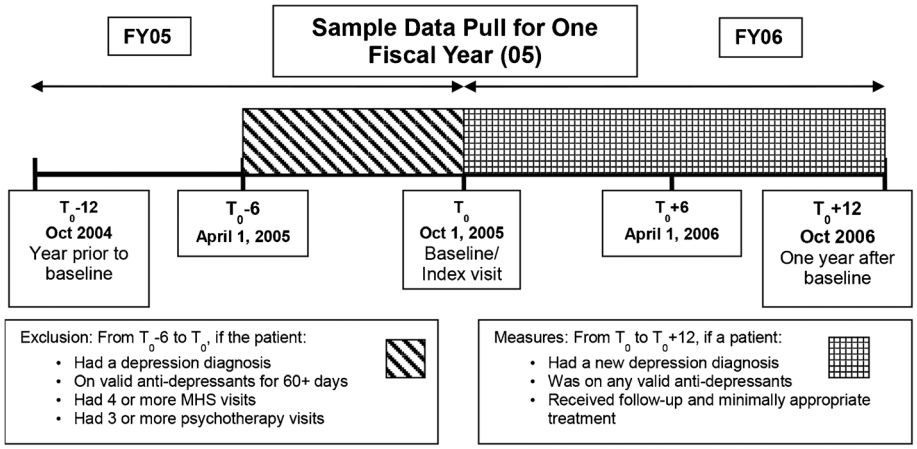

Figure 1. Identification of a new episode of depression for patients without depression diagnosis or minimal treatment in 6 months prior to index visit: example of how a patient was identified in federal fiscal year (FY) 2005.

"continuously seen" allows for sufficient time for follow-up and helps avoid truncated data. To assign each patient a home site, we used an algorithm similar to those used in a variety of primary care studies.[33, 34] The algorithm stipulates that a patient's "home site" is the site with the most primary care visits for that patient over the 2-year period. For ties, we used the site with the most recent visit, or when sites differed in complexity of services (e.g., a large medical center versus a smaller community based outpatient clinic), we chose the smaller, less complex site.

Exclusion for Depression Diagnosis or Treatment in Prior 6 Months. To limit the measures to patients with a new episode of depression, we excluded patients with a depression diagnosis (based on ICD-9 codes for depression, shown in Appendix 1 available online), or who had received minimally appropriate treatment in the 6 months prior to the index visit (T_0-6). Minimally appropriate treatment was defined as ≥60 days of depression prescriptions (list of antidepressant drugs shown in Appendix 2 available online), ≥4 mental health visits (VA clinic stop codes shown in Appendix 3 available online), or ≥3 psychotherapy visits (Current Procedural Terminology (CPT) codes shown in Appendix 4 available online).

Measures

Based on depression care literature, prior quality measures from the VA and the National Committee for Quality Assurance (NCQA), and depression guidelines, we developed four population-based quality metrics for depression care to follow patients electronically over time (see measures in Table 1). Each reported measure uses as a denominator only the subset of patients to whom the measure is applicable. For example, detection is the proportion of patients who had a newly detected episode of depression (numerator) over the eligible population of primary care patients without recent depression diagnosis or minimally appropriate treatment (denominator). The proportion of patients with follow-up and minimally appropriate treatment (numerators) uses the denominator of all patients with a newly detected episode of depression. The ICD-9 codes, medications, stop codes and CPT codes that we used are shown in Appendices 1, 2, 3 and 4 available online.

Measure 1: *Detection of new episode of depression*: Detection of a new episode of depression was defined as a clinic visit with an ICD-9 code for depression or any antidepressant prescription in the 12 month period after the index visit.

Measures 2 & 3: *Follow-up of patient with new episode of depression*: Following NCQA measures, we evaluated follow-up for a new depression diagnosis within 84 days and 180 days. Appropriate follow-up was defined as ≥3 MH visits, or ≥3 psychotherapy visits, or ≥3 primary care visits with a depression ICD-9 diagnosis within 84 or 180 days of the newly detected episode.

Measure 4: *Minimally appropriate treatment for patients with a new episode of depression*: Minimally appropriate treatment was defined as having ≥60 days of antidepressants, or ≥4 MHS visits, or ≥3 psychotherapy visits within 12 months post detection. For prescriptions, we used the cut point of ≥60 days of prescription medications to indicate at least one medication refill. We excluded prescriptions with non-depression indication/keywords written on the dosing instructions, and prescriptions with a subtherapeutic dose (see Appendix 2 available online for details).

Accounting for All Patients in Each Measure

We used a counting hierarchy (counting trees) to document the number of patients retained and excluded at *all* steps of measure construction. The counting hierarchy ensures that every patient in the original primary care population cohort is accounted for across each branch in the logic, such that at each node, the sum of patients meeting and not meeting each criterion equals the full initial population.

Table 1. Quality of Depression Care Measures

Constructed measure	Description of measure	Time frame
Detection of a new episode of depression	Number of patients detected with a new episode of depression (numerator) over the number of primary care patients who in the 6 months prior to the index visit were undiagnosed and untreated for depression (denominator)	12 months from index visit
Follow-up of a patient with a new episode of the depression detected within 84 days of detection	Number of patients who received follow-up with three or more MHS visits, three or more psychotherapy visits, or three or more primary care visits with a depression ICD-9 diagnosis within 84 days of detection of new depression (numerator) over the number of patients with a new episode of depression detected (denominator)	84 days from detection
Follow-up of a patient with a new episode of the depression detected within 180 days of detection	Number of patients who received follow-up with three or more MHS visits, three or more psychotherapy visits, or three or more primary care visits with a depression ICD-9 diagnosis within 180 days of detection of new depression (numerator) over the number of patients with a new episode of depression detected (denominator)	180 days from detection
Minimally appropriate treatment	Number of patients who complete 60 or more days of antidepressants, four or more MHS visits, or three or more psychotherapy visits within 12 months of the visit where the new episode of depression was detected (numerator) over the number of patients with a new episode of depression detected (denominator)	12 months from detection

Depression diagnosis based on ICD-9 codes are shown in Appendix 1, Prescriptions used as antidepressants are shown in Appendix 2; Mental health visits based on 500 level clinic stop codes are shown in Appendix 3; Psychotherapy visits based on CPT-codes are shown in Appendix 4. Appendices 1, 2, 3 and 4 are available online

RESULTS

Overview. Measure development focused on using counting trees to continuously verify application of measures to the primary care population, and measure application focused on using the measures to assess depression care from FY2000 to FY2010 in nine VISNs. The expert panel reviewed and critiqued measure development and application.

Counting Trees. Counting trees show that programming accurately accounted for all patients in the population at every branch node. Figure 2 is the counting tree for minimally appropriate treatment for FY2005. The tree begins on the far left with the cohort of all patients who had an index visit in primary care during FY2005 from the nine VISNs (n=2,011,849). The next branch shows which patients were continuously seen in their primary care site over the 2-year period (n=1,574,532). Twenty-two percent (n=437,317) had not been seen continuously and therefore were not eligible for the measure. The next branches exclude patients who had a prior diagnosis of depression (n=53,221) or had completed minimal treatment for depression in the prior 6 months (n=201,814). Of the remaining 1,319,497 patients, 94,130 patients had a new episode of depression detected in FY2005 (7 % of the eligible population). Among those patients with a new episode of depression detected, 82 % (n=77,533) completed minimally appropriate treatment within 12 months of detection.

Counting trees assessed meaningfulness (through face validity of the branches), but can also be used for quality improvement (to assess what happens to those not meeting the measure). For example, in FY2005, the bottom branch identified that 22 % of the patients had not been continuously seen twice in 2 years, 16,909 had an ICD-9 diagnosis for depression in 6 months prior to the index visit, and 13,008 patients had not received minimally appropriate treatment in the prior 6 months. Among the 13,008 patients, 86 % (n=11,167) had an additional episode of depression detected, and 12 % did not receive minimally appropriate treatment within 12 months of detection.

Measures Over Time in Nine VISNs. Over the decade, there was a substantial increase in the number of patients seen in primary care in the nine VISNs, from 1.19 million in FY2000 to over 2.26 million in FY2010 (Fig. 3). The cohort without a depression diagnosis or active treatment in the prior 6 months ranged from 790,000 to 1.35 million and the rates for detection of new episodes of depression remained stable at 7–8 % over the years (not shown). Follow-up at 84 and 180 days was 37 % and 45 % in FY2000, and increased to 56 % and 63 % by FY2010 (Fig. 4). Minimally appropriate treatment remained relatively stable at 84 % in FY2000, dropped to 82 % in FY2005, and was 83 % in FY2010.

Expert Panel Evaluation of the Measures as Developed and Applied. The expert panel included 14 panelists, four of whom were also on the project team. Analysis of the pre-meeting survey (86 % response rate) data showed a high level of agreement on the appropriateness of most measurement development and cohort construction methods, including the decision to use antidepressants prescribed in primary care as a signal for depression detection, even without an accompanying ICD-9 diagnosis. Panelists also validated the cohort definitions and counting tree based methods. For the definition of new depression, panelists suggested the name "new episode of depression" to indicate that the patient may have been previously diagnosed, but had not been diagnosed or treated for the previous 6 months. For the treatment measure, panelists judged the threshold for treatment completion we used, although based on prior

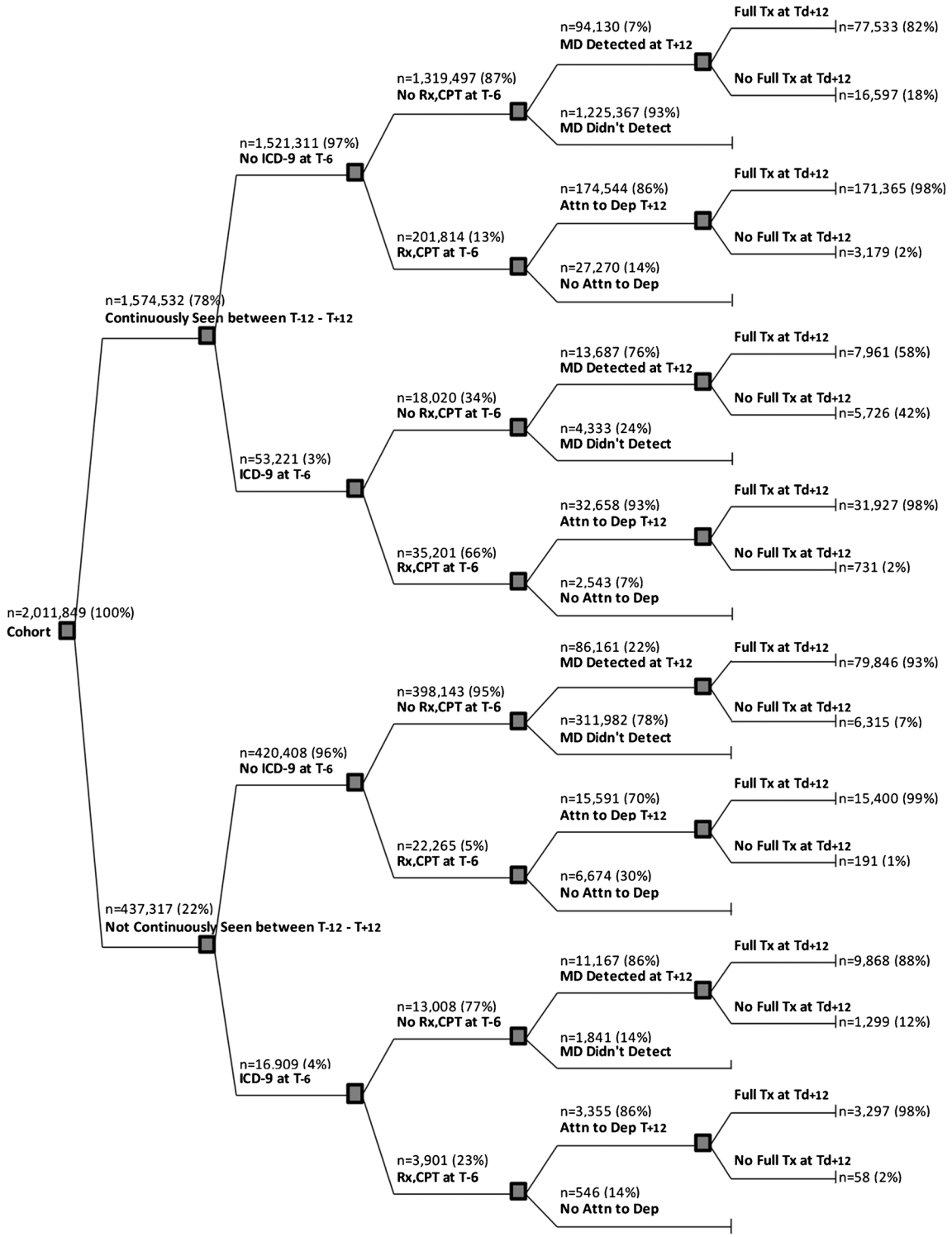

Figure 2. Counting tree for completion of minimally appropriate treatment 12 months after detection of a new episode of depression in federal fiscal year 2005. Abbreviations in figure: No ICD-9 at T_{-6} (vs. ICD-9 at T_{-6}) = Exclude patients who had a diagnosis of depression in the 6 months prior to index visit. No Rx, CPT at T_{-6} (vs. Rx, CPT at T_{-6}) = Exclude patients who completed a minimal course of treatment for depression in the prior 6 months. Attn to Dep (vs. No Attn to Dep) = There was attention (treatment) to depression found. MD Detected (vs. MD Didn't Detect) = A new episode of depression was detected. Full Tx at Td_{12} (vs. No Full Tx at Td_{12}) = Completed minimally appropriate treatment within 12 months of detection of a new episode of depression.

studies,[14] to be too low relative to optimal treatment, especially given the severity and complexity of depression among Veterans. Panelists suggested terming the measure "minimally appropriate treatment." They discussed future potential modifications, including requiring 90 days of antidepressants (expanded from the 60 day requirement) and/or 90 days of continuous antidepressants. They also discussed restricting the 60 days of medications to be within 90 days of detection (as

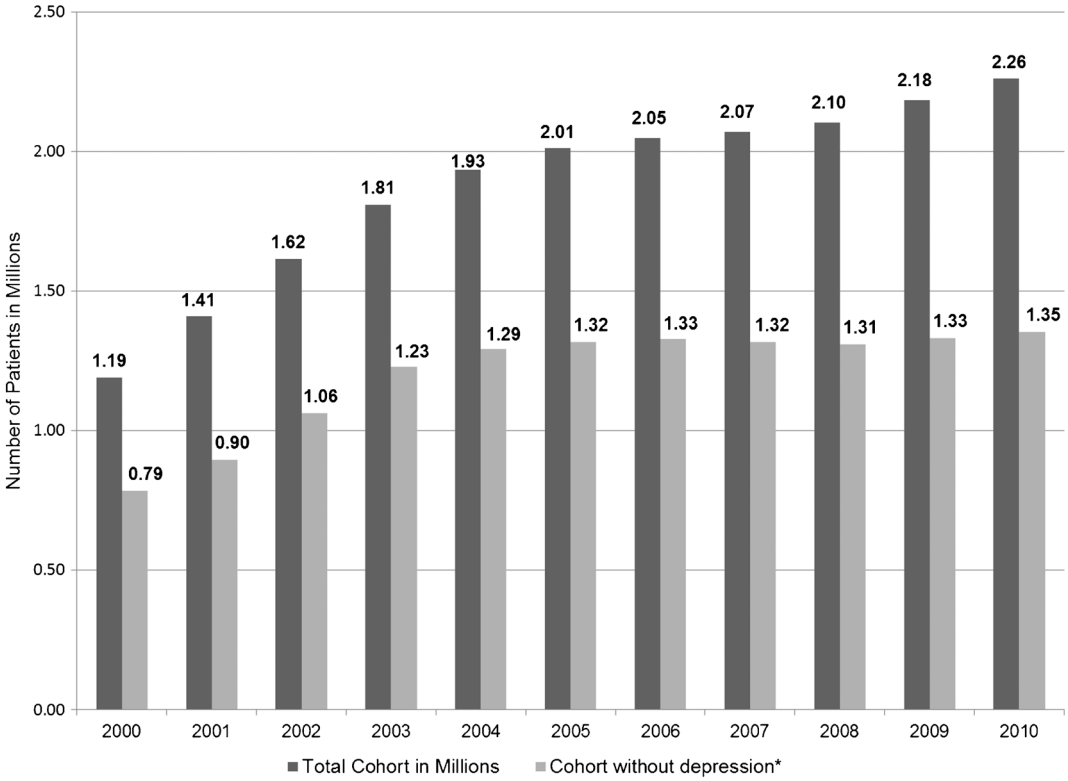

Figure 3 Cohorts of VA primary care patients from the nine Veteran integrated service networks in federal fiscal years 2000–2010 (in millions). *Patients without a depression diagnosis or actively treated in 6 months prior to index visit.

opposed to 12 months). To address continuity of treatment, panelists suggested that in the future, measures should require mental health visits (or psychotherapy visits) to be with the same provider. Panelists also endorsed the future goal of similar depression care measures for evaluating the population of patients screening positive for depression in primary care.

DISCUSSION

The VA saw rapid growth in primary care patients from FY2000 to FY2010, increasing by over one million patients. Despite rapid growth, our measures indicate that the detection of new episodes of depression (8 %) and minimally appropriate treatment rates (84 %) remained stable, suggesting VA was able to maintain a standard of care while treating significantly more patients each year.

While our measure of treatment completion mirrors standards used in prior clinical quality improvement trials,[14, 27] our expert panel judged that future iterations of this measure need to incorporate a higher minimal treatment threshold. This is feasible without changes to the basic measure approach we used. Future efforts to develop measures to identify excellent care should consider, however, the potential tradeoffs if electronic measures become too specific. Tradeoffs can include greater potential for error, lower measure reliability, and

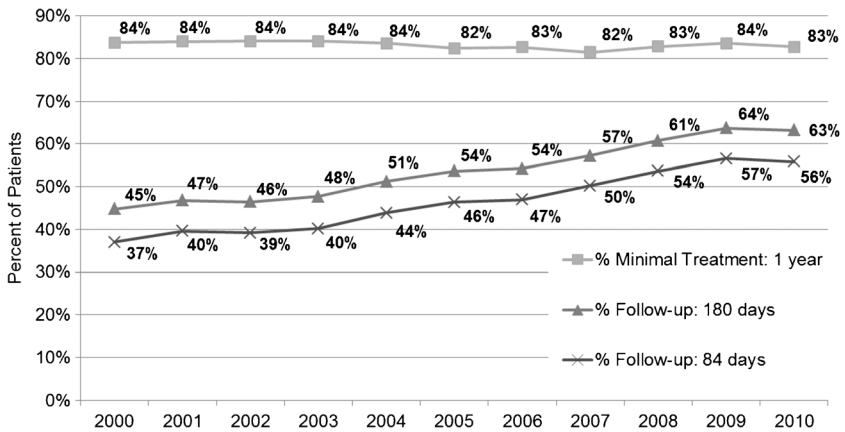

Figure 4. Among patients with a new episode of depression detected, percent of patients with follow-up and minimal treatment completion (FY2000–FY2010).

greater likelihood of using censored data that does not reflect the care for the full relevant population of interest. Furthermore, measures focused on enhancing access to basic care (e.g., performance of HbA1c or following up detected depression) often have large impacts on patient outcomes. Low threshold measures often identify patients receiving severely inadequate care and thus in most need of intervention.

Follow-up rates increased for both 84-day and 180-day follow-up (18–19 % increase), indicating improvements in timely treatment initiation and follow-up; yet showed lower adherence than did our treatment measure. Substantial evidence indicates that non-face-to-face modalities can effectively substitute for frequent visits but administrative codes for these were not available during our study period. Reliable coding of telephone, telehealth, and secure messaging encounters is now available in the VA, and future measures using our development methodology can feasibly incorporate these follow-up modalities.

The measures development methodology, including patient assignment, identifying new depression episodes, and evaluating treatment over time proved feasible and led to measures that were stable over time and across multiple primary care sites and regions. The counting trees verify and validate the algorithms underlying the measures and systematically reference the full primary care population to avoid errors due to loss of subjects. Additionally, counting trees can be used to make measures more actionable to local sites. The strict attention to the branching algorithm, the multiple data sources, and the timing of measures around an index visit creates programming challenges, yet is critical.

Using administrative data for measure development has limitations. The measures we developed advance electronic measurement by moving beyond detection based purely on ICD-9 codes to consider new antidepressant use.[20] Based on ICD-9 codes alone, the detection rate was 1 %, but adding use of antidepressants for depression resulted in detection of an additional 6–7 % of primary care patients. However, the 7–8 % detection rate was still considerably lower than published detection rates based on survey-based depression screening.[6] We preliminarily tested incorporating a PHQ-2 based detection approach into our measures (unpublished) using FY2010 data and found it to be feasible. In this approach, screening positive is used to identify the applicable population (denominator) and the cohort is then linked to the quality measures. However, because our project required stable measures spanning FY2000-FY2010, and standardized national data on screening (e.g., PHQ-2 and PHQ-9) was not available until FY2008, we could not incorporate these screening/symptom measures into the measurement algorithms. Finally, we developed an antidepressant algorithm to exclude prescriptions for non-depression indications; however, antidepressants used for other conditions (sleep, pain, migraines, etc.) may not have been entirely excluded. Future attention to promoting coding that indicates non-depression–related antidepressant use could improve measure accuracy.

We developed electronic population-based longitudinal depression quality measures that met reasonable standards as meaningful, feasible and actionable [31] for assessing VA depression care over a decade. Our current data shows that VA improved depression follow-up between FY2000 and FY2010, and that treatment rates compared favorably with non-VA benchmarks. Looking forward, our measure development methodology can feasibly be adapted to incorporate more stringent definitions of treatment, depression symptom screening data, and future enhancements in care. The methodology and techniques we used to address measurement challenges provide a basis for future performance measure development, especially for other chronic conditions where longitudinal care must be captured.

Acknowledegments: *Funding for this research was supported by a grant from VA Health Services Research & Development (Project #IIR 11-326; PI: Farmer). We would like to acknowledge the expert panelists for their contribution as well as Britney Chow, MPH and Ismelda Canelo, MPA for administrative support.*

The views expressed in the manuscript are solely those of the authors, and do not necessarily represent the views of the U.S. Department of Veterans Affairs or the United States government.

Corresponding Author: *Melissa M. Farmer, PhD; VA HSR&D Center for the Study of Healthcare Innovation, Implementation & Policy, VA Greater Los Angeles Healthcare System, 16111 Plummer Street (152), Sepulveda, CA 91343, USA (e-mail: Melissa.Farmer@va.gov).*

Compliance with Ethical Standards:

Conflict of Interest: *The authors declare that they do not have a conflict of interest.*

REFERENCES

1. **Katon WJ, Lin EH, Von Korff M, Ciechanowski P, Ludman EJ, Young B, et al.** Collaborative care for patients with depression and chronic illnesses. N Engl J Med. 2010;363(27):2611–20.
2. **Wells KB, Burnam MA, Rogers W, Hays R, Camp P.** The course of depression in adult outpatients. Results from the medical outcomes study. Arch Gen Psychiatry. 1992;49(10):788–94.
3. **Wells KB, Sherbourne CD.** Functioning and utility for current health of patients with depression or chronic medical conditions in managed, primary care practices. Arch Gen Psychiatry. 1999;56(10):897–904.
4. **Murray CJ, Lopez AD.** The global burden of disease: a comprehensive assessment of mortality and disability from disease, injuries, and risk factors in 1990 and projected to 2020. Boston, MA: The Harvard School of Public Health on behalf of the World Health Organization and the World Bank; 1996.
5. **Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE.** Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. Arch Gen Psychiatry. 2005;62(6):617–27.
6. **Yano EM, Chaney EF, Campbell DG, Klap R, Simon BF, Bonner LM, et al.** Yield of practice-based depression screening in VA primary care settings. J Gen Intern Med. 2012;27(3):331–8.
7. Uniform Mental Health Services; http://vaww1.va.gov/vhapublications/ViewPublication.asp?pub_ID=1762. Accessed November 18, 2015.
8. Panel DG. Depression in primary care, Vol II: treatment of major depression: clinical practice guide-line 5. In: AfHCPa, ed. Research. AHCPR Publication 93-0551: Rockville, MD; 1993.
9. **Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ.** Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. Arch Intern Med. 2006;166(21):2314–21.
10. **Coventry PA, Hudson JL, Kontopantelis E, Archer J, Richards DA, Gilbody S, et al.** Characteristics of effective collaborative care for treatment of depression: a systematic review and meta-regression of 74 randomised controlled trials. PLoS ONE. 2014;9(9):e108114.
11. **Archer J, Bower P, Gilbody S, Lovell K, Richards D, Gask L, et al.** Collaborative care for depression and anxiety problems. Cochrane Database Syst Rev. 2012;10:CD006525.

12. **Katon W, Unutzer J, Wells K, Jones L.** Collaborative depression care: history, evolution and ways to enhance dissemination and sustainability. Gen Hosp Psychiatry. 2010;32(5):456–64.
13. **Sherbourne CD, Edelen MO, Zhou A, Bird C, Duan N, Wells KB.** How a therapy-based quality improvement intervention for depression affected life events and psychological well-being over time: a 9-year longitudinal analysis. Med Care. 2008;46(1):78–84.
14. **Wells KB, Sherbourne C, Schoenbaum M, Duan N, Meredith L, Unutzer J, et al.** Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. JAMA. 2000;283(2):212–20.
15. **Wells K, Sherbourne C, Schoenbaum M, Ettner S, Duan N, Miranda J, et al.** Five-year impact of quality improvement for depression: results of a group-level randomized controlled trial. Arch Gen Psychiatry. 2004;61(4):378–86.
16. **Smith JL, Rost KM, Nutting PA, Libby AM, Elliott CE, Pyne JM.** Impact of primary care depression intervention on employment and workplace conflict outcomes: is value added? J Ment Health Policy Econ. 2002;5(1):43–9.
17. **Rubenstein LV, Chaney EF, Ober S, Felker B, Sherman SE, Lanto A, et al.** Using evidence-based quality improvement methods for translating depression collaborative care research into practice. Fam Syst Health. 2010;28(2):91–113.
18. **Unutzer J, Tang L, Oishi S, Katon W, Williams JW Jr, Hunkeler E, et al.** Reducing suicidal ideation in depressed older primary care patients. J Am Geriatr Soc. 2006;54(10):1550–6.
19. **Essock SM, Olfson M, Hogan MF.** Current practices for measuring mental health outcomes in the USA: international overview of routine outcome measures in mental health. Int Rev Psychiatry. 2015;27(4):296–305.
20. **Solberg LI, Engebretson KI, Sperl-Hillen JM, Hroscikoski MC, O'Connor PJ.** Are claims data accurate enough to identify patients for performance measures or quality improvement? The case of diabetes, heart disease, and depression. Am J Med Qual. 2006;21(4):238–45.
21. **Hermann RC, Provost S.** Best practices: interpreting measurement data for quality improvement: standards, means, norms, and benchmarks. Psychiatr Serv. 2003;54(5):655–7.
22. **Asch SM, McGlynn EA, Hogan MM, Hayward RA, Shekelle P, Rubenstein L, et al.** Comparison of quality of care for patients in the Veterans Health Administration and patients in a national sample. Ann Intern Med. 2004;141(12):938–45.
23. **Bynum JP, Bernal-Delgado E, Gottlieb D, Fisher E.** Assigning ambulatory patients and their physicians to hospitals: a method for obtaining population-based provider performance measurements. Health Serv Res. 2007;42(1 Pt 1):45–62.
24. **Ross JS, Siu AL.** The importance of population-based performance measures. Health Serv Res. 2007;42(1 Pt 1):1–6.
25. **Donabedian A.** The quality of care. How can it be assessed? JAMA. 1988;260(12):1743–8.
26. **Charbonneau A, Rosen AK, Owen RR, Spiro A 3rd, Ash AS, Miller DR, et al.** Monitoring depression care: in search of an accurate quality indicator. Med Care. 2004;42(6):522–31.
27. **Hepner KA, Rowe M, Rost K, Hickey SC, Sherbourne CD, Ford DE, et al.** The effect of adherence to practice guidelines on depression outcomes. Ann Intern Med. 2007;147(5):320–9.
28. **Kahn KL, Tisnado DM, Adams JL, Liu H, Chen WP, Hu FA, et al.** Does ambulatory process of care predict health-related quality of life outcomes for patients with chronic disease? Health Serv Res. 2007;42(1 Pt 1):63–83.
29. **Kerr EA, Krein SL, Vijan S, Hofer TP, Hayward RA.** Avoiding pitfalls in chronic disease quality measurement: a case for the next generation of technical quality measures. Am J Manag Care. 2001;7(11):1033–43.
30. Major Depression Working Group. VA/DoD clinical practice guideline for management of major depressive disorder. In: Do D, ed. Washington DC: Department of Veterans Affairs; 2009.
31. **Hermann RC, Palmer RH.** Common ground: a framework for selecting core quality measures for mental health and substance abuse care. Psychiatr Serv. 2002;53(3):281–7.
32. **Rubenstein LV, Fink A, Yano EM, Simon B, Chernof B, Robbins AS.** Increasing the impact of quality improvement on health: an expert panel method for setting institutional priorities. Jt Comm J Qual Improv. 1995;21(8):420–32.
33. **Farmer M.** Identifying successful approaches to smoking cessation treatment for veterans, final report for HSR&D project IIR 04-380. Sepulveda, CA: VHA HSR&D; 2010.
34. **Yano E.** VA's quality transformation: lessons for evidence-based management, final report for HSR&D project IIR 06-087. Sepulveda, CA: VHA HSR&D; 2010.

APPENDICES

Table 2. ICD-9 Codes Used for Depression Diagnosis

ICD-9	Description	Used for:	
		*Exclude	Follow-up
293.83	Organic affective syndrome	X	X
296.2x	Major depressive disorder single episode	X	X
296.21	Major depressive affective disorder single episode mild degree	X	X
296.22	Major depressive affective disorder single episode moderate degree	X	X
296.23	Major depressive affective disorder single episode severe degree without psychotic behavior	X	X
296.24	Major depressive affective disorder single episode severe degree specified with psychotic behavior	X	X
296.25	Major depressive affective disorder single episode in partial or unspecified remission	X	X
296.26	Major depressive affective disorder single episode in full remission	X	X
296.3x	Major depressive disorder recurrent episode unspecified degree	X	X
296.31	Major depressive affective disorder recurrent episode mild degree	X	X
296.32	Major depressive affective disorder recurrent episode moderate degree	X	X
296.33	Major depressive affective disorder recurrent episode severe degree without psychotic behavior	X	X
296.34	Major depressive affective disorder recurrent episode severe degree specified with psychotic behavior	X	X
296.35	Major depressive affective disorder recurrent episode in partial or unspecified remission	X	X
296.36	Major depressive affective disorder recurrent episode in full remission	X	X
296.5x	Bipolar I disorder, most recent episode (or current) depressed unspecified degree	X	X
296.51	Bipolar I disorder, most recent episode (or current) depressed mild degree	X	X
296.52	Bipolar I disorder, most recent episode (or current) depressed moderate degree	X	X
296.53	Bipolar I disorder, most recent episode (or current) depressed severe degree without psychotic behavior	X	X
296.54	Bipolar I disorder, most recent episode (or current) depressed severe degree specified with psychotic behavior	X	X
296.55	Bipolar I disorder, most recent episode (or current) episode in partial or unspecified remission	X	X
296.56	Bipolar I disorder, most recent episode (or current) depressed in remission	X	X
296.90	Unspecified episodic mood disorder		X
296.99	Other specified affective psychoses	X	
298.0x	Depressive type psychosis	X	X
300.4x	Neurotic depression	X	X
309.0x	Brief depressive reaction as adjustment reaction		X
309.1x	Prolonged depressive reaction as adjustment reaction	X	X
311x	Depressive disorder not elsewhere classified (NEC)	X	X

**Exclusion for prior depression diagnosis*

Table 2. List of Antidepressant Drugs Used in Measures*

Amitriptyline	Nefazodone
Desipramine	Sertraline
Doxepin	Trazodone
Imipramine	Mirtazapine
Nortriptyline	Amoxapine
Citalopram	Clomipramine
Fluoxetine	Escitalopram
Paroxetine	Fluvoxamine
Phenelzine	Isocarboxazid
Tranylcypromine	Maprotiline
Venlafaxine	Protriptyline
Buproprion	Trimipramine
Duloxetine	Selegiline

**We excluded prescriptions with a non-depression indication/keywords written on the dosing instructions including 'neuropathy', 'pain', 'peripheral', 'sleep', 'insomnia', 'pruritis', 'itching', 'migraine', 'foot', 'head', 'ache', and 'itch'. We also excluded prescriptions with a smoking indication written in the dosing instructions ('smoke', 'tobacco', 'smoking'). We also excluded prescriptions with a low dose including if Amitriptyline HCL is less than 75 MG TAB, if Trazodone is less than 300MG, and if Mirtazapine is less than 45MG*

Table 4. VA Clinic Stop Codes for Mental Health Visits

Clinic stop code	Description
143	Sleep study
165	Bereavement counseling
292	Observation psychiatry
501	Homeless mentally ill outreach
502	Mental health clinic – Individual
503	Mental health residential care – Individual
504	IPCC medical center visit
505	Day treatment – Individual
506	Day hospital – Individual
507	Department of Housing & Urban Development/VA Supportive Housing (HUD/VASH)-Group
508	Health care for homeless Veterans/Homeless chronically mentally ill (HCHV/HCMI) – Group
509	Psychiatry – Individual
510	Psychology – Individual
511	Grant and per diem – Individual
512	Mental health consultation
513	Substance use disorder – Individual
514	Substance abuse – Home visit
515	Compensated work therapy (CWT) / TR-HCMI
516	Post traumatic stress disorder (PTSD) – Group
517	Compensated work therapy (CWT) / Substance abuse
518	Compensated work therapy (CWT) / TR-Substance abuse
519	Substance use disorder / PTSD teams
520	Long-term enhancement – Individual
521	Long-term enhancement – Group
522	Department of Housing and Urban Development (HUD) -VA Shared Housing (VASH)
523	Opioid substitution
524	Active duty sexual trauma
525	Women's stress disorder treatment teams
526	Telephone / Special psychiatry
527	Mental health telephone
528	Telephone homeless chronically mentally ill (HCMI)
529	Health care for homeless Veterans (HCHV) / HCMI
530	Telephone HUD/VASH
531	Mental health primary care – Individual
532	Psychosocial rehabilitation – Individual
533	Mental health intervention biomedical care – Individual
534	Mental health integrated care
535	Mental health vocational assistance – Individual
536	Telephone mental health vocational assistance
537	Telephone psychosocial rehabilitation
538	Psychological testing
539	Mental health integration care – Group
540	Post-traumatic stress disorder (PTSD) clinical team (PCT) Post-traumatic stress – Individual
541	Post-traumatic stress disorder (PTSD)
542	Telephone PTSD
543	Telephone/alcohol dependence
544	Telephone/drug dependence
545	Telephone/substance use disorder
546	Telephone mental health intensive care management (MHICM)
547	Intensive substance abuse treatment – Group
548	Intensive substance abuse disorder – Individual
550	Mental health clinic – Group
551	IPCC community clinic/day program visit
552	Mental health intensive case management (MHICM)
553	Day treatment – Group
554	Day hospital – Group
555	Drug dependence – Group
556	Alcohol treatment – Group
557	Psychiatry – MD Group
558	Psychology – Group
559	Psychosocial rehabilitation – Group
560	Substance Abuse – Group
561	PCT- Post traumatic stress – Group
562	PTSD – Individual
563	Mental health primary care team – Group
564	Mental health team case management
566	Mental health risk-factor reduction education group
567	Mental health intensive case management (MHICM) – Group
568	Mental health compensated work therapy/Supported employment (CWT/SE) face-to-face
569	Mental health CWT/Supported employment (CWT/SE) non-face-to-face CBO non-count
570	MH CWT/Transitional work experience (TWE) Non-face-to-face CBO Non-count
571	RVOEC (Returning Veterans outreach, education and care) Individual
572	RVOEC (Returning Veterans outreach, education and care) Group
573	Mental health incentive therapy group face-to-face
574	MH CWT/Transitional work experience (TWE) face-to-face
575	Mental health vocational assistance – Group
576	Psycho-geriatric clinic – Individual
577	Psycho-geriatric clinic – Group
578	Psycho-geriatric day program
579	Telephone psycho-geriatrics
580	PTSD day hospital
581	PTSD day treatment
582	Psychosocial rehabilitation recovery center (PRRC) – Individual
583	Psychosocial rehabilitation recovery center (PRRC) – Group
584	Telephone psychosocial rehabilitation recover center (PRRC)
588	Residential rehabilitation treatment program (RRTP) aftercare – Individual
589	Non-active duty sexual trauma
590	Community outreach homeless Veterans by staff other than HCHV and RRTP
591	Incarcerated Veterans re-entry
592	Veterans justice outreach
593	Residential rehabilitation treatment program (RRTP) outreach services
594	RRTP aftercare – Community
595	RRTP aftercare – VA
596	RRTP admission screening services
597	Telephone/RRTP
598	RRTP pre-admission – Individual
599	RRTP pre-admission – Group

(continued on next page)

Table 5. Current Procedural Terminology (CPT) Codes for Psychotherapy

CPT codes	Description
90804	Outpatient – individual psychotherapy 20–30 min
90805	Outpatient – individual psychotherapy 20–30 min with medical management
90806	Outpatient – individual psychotherapy 45–50 min
90807	Outpatient – individual psychotherapy 45–50 min with medical management
90808	Outpatient – individual psychotherapy 75–80 min
90809	Outpatient – individual psychotherapy 75–80 min with medical management
90810	Outpatient – interactive psychotherapy 20–30 min
90811	Outpatient – interactive psychotherapy 20–30 min with medical management
90812	Outpatient – interactive psychotherapy 45–50 min
90813	Outpatient – interactive psychotherapy 45–50 min with medical management
90814	Outpatient – interactive psychotherapy 75–80 min
90815	Outpatient – interactive psychotherapy 75–80 min with medical management
90816	Inpatient – individual psychotherapy 20–30 min
90817	Inpatient – individual psychotherapy 20–30 min with medical management
90818	Inpatient – individual psychotherapy 45–50 min
90819	Inpatient – individual psychotherapy 45–50 min with medical management
90821	Inpatient – individual psychotherapy 75–80 min
90822	Inpatient – individual psychotherapy 75–80 min with medical management
90823	Inpatient – interactive psychotherapy 20–30 min
90824	Inpatient – interactive psychotherapy 20–30 min with medical management
90826	Inpatient – interactive psychotherapy 45–50 min
90827	Inpatient – interactive psychotherapy 45–50 min with medical management
90828	Inpatient – interactive psychotherapy 75–80 min
90829	Inpatient – interactive psychotherapy 75–80 min with medical management
90845	Psychoanalysis
90847	Family psychotherapy (conjoint psychotherapy) (with patient present)
90849	Multiple-family group psychotherapy
90853	Group psychotherapy (other than of a multiple-family group)
90857	Interactive group psychotherapy
90875	Individual psychological therapy incorporating biofeedback, 30 min
90876	Individual psychological therapy incorporating biofeedback, 45 min
90801	Psychiatric diagnostic evaluation
90802	Interactive psychiatric diagnostic evaluation
90870	Electroconvulsive therapy, single seizure
90871	Electroconvulsive therapy, multiple seizure