

The Relationship of System-Level Quality Improvement With Quality of Depression Care

Andrea Charbonneau, MD, MSc; Victoria Parker, DBA; Mark Meterko, PhD;
Amy K. Rosen, PhD; Boris Kader, PhD; Richard R. Owen, MD; Arlene S. Ash, PhD;
Jeffrey Whittle, MD, MPH; and Dan R. Berlowitz, MD, MPH

Objective: To explore the relationship of systemwide continuous quality improvement (CQI) with depression care quality in the Veterans Health Administration (VHA).

Study Design: Observational study using data from 2 VHA studies.

Patients and Methods: The Depression Care Quality Study (DCQS) was a retrospective cohort study of depression care quality in the northeastern United States involving 12 678 patients cared for at 14 VHA facilities; it used guideline-based process measures (ie, dosage and duration adequacy). The VHA CQI survey was a cross-sectional survey of systemwide CQI among a representative sample of VHA hospitals; it assessed CQI and organizational culture (OC) at 116 VHA hospitals nationwide and provided data on the 14 study facilities. We used analysis of variance to identify differences in the adequacy of depression care among these facilities. Pearson's correlation was used to identify the relationship of CQI and OC with facility-level depression care adequacy.

Results: Mean depression care adequacy differed among the 14 DCQS facilities ($P < .0001$). Overall dosage adequacy was 90% (range: 87%-92%). Overall duration adequacy was 45% (range: 39%-64%). There was no correlation between CQI and either dosage adequacy ($r = .004$, $P = .98$) or duration adequacy ($r = -.17$, $P = .55$). Similarly, there was no correlation between OC and either dosage adequacy ($r = -.35$, $P = .22$) or duration adequacy ($r = -.12$, $P = .68$).

Conclusion: Although CQI may help bridge the healthcare quality gap, it may not be associated with higher disease-specific quality of care.

(*Am J Manag Care.* 2004;10(part 2):846-851)

Continuous quality improvement (CQI) has been defined as a "philosophy of continual improvement of the processes associated with providing a good or service that meets or exceeds customer expectations."¹ Given the positive CQI experience of industries other than healthcare, CQI has been embraced as a key mechanism of enhancing quality of care while preserving cost-effectiveness. In a recent review, Shortell et al¹ described how CQI has been implemented in varied healthcare settings and its diverse effects on healthcare processes and outcomes. These authors concluded that CQI is a potentially promising mechanism for bridging the existing gap in healthcare quality.

Key components necessary for successful implementation of CQI have been described for the private sector²

and the Veterans Health Administration (VHA).³ These include a dynamic, innovative quality leadership with representatives from top management, boards of trustees, and clinical arenas. An organizational culture (OC) that emphasizes teamwork and consensual decision making (ie, group culture), and innovation and risk taking (ie, developmental culture) has been linked to greater CQI implementation³ and to improved clinical outcomes.⁴ Furthermore, ongoing participation in quality improvement activities is a necessary element of a strong CQI leader and program. However, quality improvement activities alone are not sufficient to sustain overall organizational CQI.² Thus, quality improvement and quality assurance (QI/QA) principles have remained conceptually distinct from CQI.¹

There is considerable uncertainty in the literature regarding the link between CQI and specific healthcare quality domains. Understanding the relationship of systemwide CQI with disease-specific quality of care is important for targeting local CQI and QI/QA activities. Depression serves as a good disease role model for study because it is prevalent, disabling, and costly; lacks quality care; and has well-studied, guideline-based benchmarks that can serve as criteria for measuring the quality of care.

Depression is a prevalent medical condition with far-reaching and potentially serious consequences, especially if undertreated, yet gaps in the quality of depression care continue to be identified.⁵ To our knowledge, only 1 small study of the effect of CQI on depression care quality has been reported, and those results were equiv-

From the Division of General and Geriatric Medicine, University of Kansas Medical Center, Kansas City, Kan (AC, JW); the Center for Health Quality, Outcomes, and Economic Research, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, Mass (VP, AKR, BK, DRB); the Management Decision and Research Center, Boston VAMC, Boston, Mass (MM); the Center for Mental Healthcare and Outcomes Research, Central Arkansas Veterans Healthcare System, Little Rock, Ark (RRO); and the Section of General Internal Medicine, Boston Medical Center, Boston, Mass (ASA).

This work was supported by grants CPI 99-134 and MNH 98-001 (Mental Health QUERI) from the Health Services Research and Development Service, Department of Veterans Affairs.

Address correspondence to: Andrea Charbonneau, MD, MSc, Division of General and Geriatric Medicine, University of Kansas Medical Center, 3901 Rainbow Blvd, Wescoe 5026, Kansas City, KS 66160. E-mail: acharbonneau@kumc.edu.

ocal.⁶ Nonetheless, the positive effect of QI/QA multidisciplinary teams who provided guideline-based depression care^{7,8} was recently described. These interventions have demonstrated remission of short- and long-term depression, and improved depressive symptoms, patient satisfaction, work-related performance, health-related quality of life, and cost-effectiveness.⁹⁻¹⁹ Although the design and implementation of these QI/QA depression interventions have been somewhat diverse, a commonality among all has been collaborative care delivered via multidisciplinary teams in a chronic-care model approach.²⁰ In light of this evidence, the US Preventive Services Task Force recently recommended routine screening for depressive disorders in primary care settings with established systems for diagnosis, treatment, and follow-up.^{21,22}

Central VHA leadership has strongly advocated for the development of local VHA CQI activities since the early 1990s. After the creation of 22 VHA healthcare administrative regions in 1995, local CQI activities remained diffuse.³ A large, nationally representative survey of 116 VHA hospitals was undertaken in 1998 to describe VHA CQI and to define predictors of successful CQI in the VHA. Details of this work have been described elsewhere (referred to herein as the VHA CQI survey).³

In this study, we examine facility-level differences in depression care quality and explore the relationship of CQI to depression care quality by using existing data from the VHA CQI survey and our past work evaluating depression care quality in the VHA.^{23,24} Although the optimal design for studying the effects of CQI is still under debate,²⁵ these data presented a unique opportunity to conduct an observational study examining this crucial question. Because the VHA is the nation's largest managed healthcare system, its experience can serve as a leading example for other healthcare systems.

METHODS

Study Design and Sample Definition

In previous work, we conducted a retrospective cohort study of 14 VHA hospitals in New England and upstate New York. For that study, we used VHA centralized data to define a depressed cohort, to assess depression care quality using guideline-based process measures, to identify predictors of depression care adequacy,²³ and to determine the predictive validity of these guideline-based depression process measures by examining their relationship with subsequent overall and psychiatric hospitalizations.²⁴ We now seek to assess the link between depression care quality as measured by

those guideline-based process measures and CQI as measured by the VHA CQI survey.³ The methodologies of these studies are described in brief below.

Depression Care Quality Study

We used VHA administrative and centralized pharmacy records to define a depressed cohort of 12 678 patients who received antidepressant treatment during a 3-month period in 1999. Subject eligibility criteria were as follows: (1) at least 1 *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* diagnosis code 296.2x or 296.3x (major depression single or recurrent episode, respectively) noted in a psychiatry, primary care, emergency, or social work clinical setting during October 1, 1997, through September 30, 1999 (fiscal year [FY] 1998 or FY 1999) at 1 of the 14 northeastern VHA hospitals, or at least 1 *ICD-9-CM* diagnosis code 311.xx (depression, not otherwise specified) noted in a primary care clinical setting during FY 1998 or FY 1999, exclusive of other depression diagnosis codes; (2) no comorbid schizophrenia and/or bipolar disorder; and (3) receipt of at least 1 antidepressant from a VHA pharmacy during the time that depression care was profiled (June 1, 1999, through August 31, 1999). These criteria produced the final sample for which the process of depression care can be linked to system-level CQI. The characteristics of the sample are reported elsewhere.^{23,24}

Guideline-based Depression Process Measures

We compared 2 dimensions of antidepressant therapy with clinical guideline benchmarks using the 1997 VHA Depression Guidelines,²⁶ a compilation of recommendations from the Agency for Healthcare Research and Quality and the American Psychiatric Association depression guidelines.^{7,8}

We exclusively used centralized data sources, predominantly automated pharmacy records, to profile depression care quality in this study. Centralized pharmacy records have been examined in various disease models and validated as acceptable sources for assessing medication regimens.²⁷⁻²⁹ The accuracy of centralized pharmacy records in predicting suboptimal dosing and premature discontinuation of medications was recently confirmed in a depressed cohort by comparing centralized pharmacy records with patient self-report of medication administration.³⁰

We describe the guideline-based depression process measures in full in another paper.²³ In short, *dosage adequacy* was achieved when the average daily dosage of antidepressant during the 3-month profiling period met the guideline-recommended minimum daily

dosage. This resulted in a dichotomous outcome variable (ie, guideline-recommended minimum daily dosage, yes/no). *Duration adequacy* for each patient concerned the overall length of therapy with any eligible antidepressant during the profiling period. Antidepressant eligibility criteria are described elsewhere.²³ Duration adequacy was defined as a dichotomy, with inadequate duration being >21% of the profiling period without antidepressants. This boundary (.21) translates into 3 weeks of the 3-month period (or 1 week per month), and is consistent with other definitions of continuous dosing in the literature.³¹

The VHA CQI Survey

From April through September 1998, a large, nationally representative number of hospital employees (9993/14 892; 67% response³), managers (2406/3400; 71% response), and hospital directors (130/155; 84% response) were sampled from 162 VHA hospitals for the VHA CQI survey. There were 116 hospitals with available data for final quantitative analysis. The 14 facilities examined in this paper had lower response rates for employees (532/1237, 43% response) and managers (152/258, 59% response) than overall. Developed in the VHA Office of Quality Management with assistance from a private-sector consultant, the VHA CQI survey is conceptually based on criteria from the Malcolm Baldrige Award, granted by the US Department of Commerce to companies excelling in quality assurance.³ Although each of these items has significant face validity, to our knowledge, neither their construct nor criterion validity has been formally tested.

The VHA CQI survey consisted of 42 items divided into 5 categories of CQI: role of managers, information and analysis, strategic quality planning, human resources development and management, and management of process quality. An example from the information and analysis category is the following item, to which respondents were asked to indicate their level of agreement on a 5-point Likert scale: "We try to use data about quality to prevent problems, not to just fix them after they have occurred." An overall CQI measure derived from these items was scored from 1 through 5, with higher scores indicating greater CQI implementation. In previous research at VA medical centers, the 5 subscales have demonstrated good internal consistency (Cronbach's alphas ranged from .89 to .92).³²

The survey had 20 items divided into 4 dimensions of OC: group, developmental, hierarchical, and rational. These were assessed exclusively by hospital employees, who were asked to distribute 100 points among various statements best characterizing their organization. An example of a group-culture item is the following: "The

hospital is a very personal place. It is a lot like an extended family; people seem to share a lot of themselves." The overall OC measure was derived by combining the average percentage of points respondents allocated to the group culture and developmental culture. Higher scores indicated greater innovation and teamwork (ie, a stronger group and developmental culture vs a hierarchical and rational culture).

Assessing the Link Between System-level CQI and Depression Care Quality

We calculated the mean dosage and duration adequacy for each facility using patient-level adequacy scores. We observed crude facility-level differences in depression care adequacy (Table) and were concerned that patient-level differences might explain this variation. Therefore, we recalculated the facility-level mean dosage and duration adequacy, adjusting for patient-level demographic and clinical predictors of depression care adequacy identified in our previous work.²³ Specifically, we calculated the expected facility-level mean dosage and duration adequacy using probabilities generated by logistic regression from our prior models. In these, patient age, race, marital status, comorbid illness, and type of clinical care were identified as significant predictors of depression care adequacy.²³

Using patient-level adequacy scores, we calculated the mean risk-adjusted depression care adequacy for each facility by using the observed minus expected (ie, O – E) rate of depression care adequacy. We established these rates separately for dosage and duration adequacy. In our previous work, we used generalized estimating equations to adjust for clustering at the hospital level. Because we did not identify a clustering effect, we reported the simpler models examining individual-level predictors of depression care adequacy.²³ Nevertheless, we observed crude facility-level differences in depression care adequacy in the present study. Therefore, we computed and used facility-specific risk-adjusted depression care adequacy rates in all analyses.

We used analysis of variance to establish the statistical significance of differences in the adequacy of depression care among the facilities. We ranked facilities in order of depression care quality by comparing crude and risk-adjusted duration adequacy scores (Table). We assessed facility rankings by using scores for duration adequacy because these demonstrated more facility-level variation than dosage adequacy.

We identified facility-level CQI implementation and OC using data from the VHA CQI survey described above. We used Pearson's correlation to determine the relationship between CQI implementation and depres-

Table. Crude and Risk-adjusted Adequacy of Facility-level Depression Care

Facility	No. of Cases (% of Overall Sample)	Crude Rank Order*	Dosage Was Adequate (% Yes)	Duration Was Adequate (% Yes)	Adjusted Rank Order*	Risk-adjusted Dosage Adequacy (O% - E%) [†]	Risk-adjusted Duration Adequacy (O% - E%) [†]
A	558 (4.4)	1	89	64	1	-1.0	17.8
B	723 (5.7)	2	91	53	2	.3	4.6
C	824 (6.5)	3	90	51	3	-0.1	3.8
D	355 (2.8)	4	92	49	6	-1.3	3.3
E	1902 (15.0)	5	92	49	4	-0.6	0.5
F	1192 (9.4)	6	88	49	5	1.0	0.2
G	1458 (11.5)	7	91	48	8	-4.8	-0.3
H	596 (4.7)	8	85	47	10	.8	-0.4
I	1078 (8.5)	9	87	47	9	-2.1	-0.4
J	773 (6.1)	10	90	47	7	1.0	-2.1
K	494 (3.9)	11	88	46	12	1.4	-4.2
L	1254 (9.9)	12	90	45	11	-2.5	-4.6
M	127 (1.0)	13	89	40	13	1.2	-7.9
N	1344 (10.6)	14	92	39	14	2.9	-8.6

*Facilities are ranked in order from highest to lowest performance by using duration adequacy scores.

[†]O - E signifies observed minus expected. This value was calculated by using probabilities generated by logistic regression from prior models. For example, a positive estimate demonstrates that observed depression care adequacy was higher than expected (ie, O - E).

sion care adequacy (identified separately for dosage and duration adequacy). Because OC is strongly linked to CQI implementation,³ we separately examined the association of OC with depression care adequacy using Pearson’s correlation. We performed all analyses with SAS v. 8.1 (SAS Institute Inc, Cary, NC).

RESULTS

We presented the results of our evaluation of depression care adequacy in full in a prior paper.²³ To summarize briefly, we identified 90% dosage adequacy and 45% duration adequacy in the overall sample (12 678 patients). Dosage adequacy ranged from 85% to 92% among the 14 facilities ($P < .0001$ for differences). Duration adequacy ranged from 39% to 64% among the 14 facilities ($P < .0001$ for differences). Risk-adjusted dosage adequacy (O% - E%) ranged from -4.8% to 2.9% among the 14 facilities ($P < .0001$ for differences among means). Risk-adjusted duration adequacy (O% - E%) ranged from (-8.6% to 17.8% among the 14 facilities ($P < .0001$ for differences among means). Facility rankings for depression care quality did not change substantially after risk adjustment. The largest shift in ranking

was demonstrated with facility 10, which shifted to facility 7 after risk adjustment. In the Table, we present crude and risk-adjusted rates of depression care adequacy and facility rankings.

The median CQI score for the 14 facilities was 3.42 (range of 3.10 to 3.64, scored from 1-5; higher scores indicate greater CQI implementation). This compares with the overall CQI mean \pm SD of 3.69 \pm .06 from the top quartile scores, and the overall CQI mean \pm SD of 3.30 \pm .10 from the bottom quartile scores of the 116 facilities surveyed in the VHA CQI. The median OC score for the 14 facilities was 31.78 (range of 26.44-49.91, scored from 0-100; higher scores indicate greater group/developmental culture). This compares with the overall OC mean \pm SD of 39.34 \pm 3.26 from top quartile scores, and the overall OC mean \pm SD of 32.41 \pm 3.51 from the bottom quartile scores of the 116 facilities surveyed in the VHA CQI.

We did not identify a correlation between CQI and either dosage adequacy ($r = .004, P = .98$) or duration adequacy ($r = -.17, P = .55$). Similarly, we did not find a correlation between OC and either dosage adequacy ($r = -.35, P = .22$) or duration adequacy ($r = -.12, P = .68$). We repeated these analyses using crude dosage and

duration adequacy scores. This did not change our results (data not shown).

DISCUSSION

In this study, we examined facility-level differences in depression care quality among 14 VHA facilities in the northeastern United States and explored the relationship of depression care quality with CQI and OC. To our knowledge, this is the first study to examine how systemwide CQI implementation relates to depression care quality in VHA healthcare settings. This subject is important, given that veteran patients have a disproportionately higher prevalence of mental illness.³³ Additionally, VHA healthcare settings provide an excellent laboratory for such analyses, as annual completion of depression screening has been a systemwide quality measure for several years. Furthermore, the VHA has worked to narrow the depression care quality gap by dissemination of VHA-specific depression guidelines,²⁶ and provider education and performance feedback.

In this study, we have identified depression care that falls short of guideline recommendations, especially regarding duration of therapy. We also found variations in depression care among the 14 facilities concerning adequacy of both dosage and duration of antidepressant treatment, although the clinical significance of these differences is less clear. More clinically detailed information would allow us to assess the comprehensiveness of our process measures in capturing the quality of depression care. Nevertheless, our exclusive use of administrative and centralized pharmacy records mirrors the practical reality of quality profiling in a naturalistic clinical setting.

We did not identify a significant association of systemwide CQI or OC with depression care quality among these 14 VHA facilities. Our small sample size rendered the analyses underpowered for these purposes. In this study, we had 25% power to detect a true population relationship of strength $r = .35$, commonly accepted as a moderate effect size.³⁴ Future studies need to examine these relationships among at least 50 facilities to have 80% power for detecting a moderate effect size. Nevertheless, despite our underpowered examination of these important relationships, our study has wider implications for the VHA and all managed care settings.

Herein, we highlight the significance of furthering an understanding of the link between CQI, OC, and disease-specific healthcare quality. In the VHA and other managed care settings, precious resources are directed towards CQI efforts. Understanding how these efforts impact patient-level care is critical. The VHA is well poised to study these important relationships, given its vertical integration and large number of healthcare sites.

Other large managed care organizations with established CQI efforts and more than 50 healthcare sites (to render at least 80% power to detect a moderate effect size) should consider conducting a similar study to broaden our understanding of the relationship between CQI, OC, and disease-specific healthcare quality.

Despite our underpowered analyses, our results are similar to others that assessed the relationship between CQI and disease-specific healthcare quality. Berlowitz et al³² recently found no association between CQI and risk-adjusted pressure ulcer development in 35 VHA nursing homes (beta coefficient \pm SE [*P* value]: (1.73 \pm 1.30 [.19] from a regression model examining risk-adjusted pressure ulcer development as a function of CQI). Our average facility-level CQI implementation and OC were similar to, though lower than, those of the 35 VHA nursing homes. Our median CQI was 3.42 (range: 3.10-3.64), compared with theirs of 3.55 (range 2.98:4.08). Our median OC was 31.78 (range: 26.44-49.91), compared with theirs of 44.90 (range: 29.3-64.9).

To our knowledge, the clinical and policy significance of quantitative differences in OC and CQI scores has not been elucidated to date. (In other words, how many points on the CQI scale represent a meaningful, real-life change in CQI?) Nevertheless, qualitative interviews among 5 "top-performing" and 5 "low-performing" facilities from the VHA CQI survey highlighted important differences between these 2 groups.³ For example, several notable characteristics of OC among those with CQI scores in the highest quartile were as follows: high morale, low staff turnover, open communications, long tenure of leading quality assurance champions, collegiality of top management, and team-oriented staff.³ On the other hand, several characteristics of OC among those with CQI scores in the lowest quartile were high turnover at top management levels, lack of ownership, poor relationship between management and staff, and resistance to change.³

This study has limitations. Restricting depression care profiling to the northeastern United States limits generalizability to the national veteran population. These sites were chosen for ease of data collection. The national VHA pharmacy database was organized just before the study, and access to these data was unclear initially. Nevertheless, these hospitals represent diverse geographic settings and reflect the experience of the national veteran patient population.³⁵ One should maintain caution, however, in applying these findings to non-VHA healthcare settings. Our predominately male sample (92%) might not generalize to the national population undergoing depression care, the majority of whom are women.³⁶ Furthermore, our small sample size (ie, 12 678 subjects from 14 facilities) rendered our analyses underpowered in testing our study hypotheses. Moreover, restricted

resources limited our access to other important organizational characteristics that might affect depression care quality, such as teaching status, facility size, and on-site mental health care. Finally, we are unsure why the response rates for the VHA CQI survey were lower for our 14 facilities than for the 116 VHA hospitals nationwide. Additionally, we were unable to isolate the hospital director responses for our 14 facilities and thus cannot compare these to the hospital director responses overall. The survey methodologists who conducted the 1998 VHA CQI survey were unavailable for comment. Despite these limitations, this study adds to a growing literature assessing the relationship of system-level CQI and OC with disease-specific healthcare quality. To our knowledge, it is the first study to detail the association of system-level CQI and OC with mental healthcare quality.

Based on our experience, we would make a number of recommendations regarding future research in this area. First, given the difficulty of observational studies examining the link between CQI implementation and disease-specific healthcare quality, we suggest measuring CQI implementation among those most directly concerned with the disease-specific healthcare delivery. Our study examined CQI implementation on a broad, systemwide scale. This approach contrasts with the studies described above,^{4,32} which measured CQI implementation among those more closely involved with the disease-specific quality domain under study (ie, nursing home staff³²).

Future work should specifically explore the relationship between broad-based CQI efforts and mental healthcare quality. Although parity for mental healthcare access, coverage, and quality is now a legislative mandate, it is not yet a reality compared with the care rendered for general medical conditions.³⁷ Finally, future research might focus on the process by which CQI impacts disease-specific healthcare quality. This approach will require longitudinal study and will allow richer hypothesis testing than the dose-response evaluation that our cross-sectional analyses provided.

REFERENCES

- Shortell SM, Bennett CL, Byck GR. Assessing the impact of continuous quality improvement on clinical practice: what it will take to accelerate progress [see comment]. *Milbank Q*. 1998;76(4):593-624.
- Weiner BJ, Shortell SM, Alexander J. Promoting clinical involvement in hospital quality improvement efforts: the effects of top management, board, and physician leadership. *Health Serv Res*. 1997;32:491-510.
- Parker VA, Wubbenhorst WH, Young GJ, Desai KR, Charns MP. Implementing quality improvement in hospitals: the role of leadership and culture. *Am J Med Qual*. January-February 1999;14:64-69.
- Shortell SM, Jones RH, Rademaker AW, et al. Assessing the impact of total quality management and organizational culture on multiple outcomes of care for coronary artery bypass graft surgery patients. *Med Care*. 2000;38:207-217.
- Young AS, Klap R, Sherbourne CD, Wells KB. The quality of care for depressive and anxiety disorders in the United States. *Arch Gen Psychiatry*. 2001;58:55-61.
- Goldberg HI, Wagner EH, Fihn SD, et al. A randomized controlled trial of CQI teams and academic detailing: can they alter compliance with guidelines? *Jt Comm J Qual Improv*. 1998;24:130-142.
- Depression Guidelines Panel. *Clinical Practice Guideline Number 5: Depression in Primary Care, 2: Treatment of Major Depression*. Rockville, MD: Agency for Health Care Policy and Research; 1993. Publication 93-0551.
- American Psychiatric Association. Practice guideline for major depressive disorder in adults. *Am J Psychiatry*. 1993;150(suppl):S1-S26.
- Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA*. 2000;283:212-220.
- Katon W, Russo J, Von Korff M, et al. Long-term effects of a collaborative care intervention in persistently depressed primary care patients. *J Gen Intern Med*. 2002;17:741-748.
- Rost K, Nutting P, Smith J, Werner J, Duan N. Improving depression outcomes in community primary care practice: a randomized trial of the quEST intervention. Quality Enhancement by Strategic Teaming. *J Gen Intern Med*. 2001;16:143-149.
- Meredith LS, Orlando M, Humphrey N, Camp P, Sherbourne CD. Are better ratings of the patient-provider relationship associated with higher quality care for depression? *Med Care*. 2001;39:349-360.
- Schoenbaum M, Unutzer J, McCaffrey D, Duan N, Sherbourne C, Wells KB. The effects of primary care depression treatment on patients' clinical status and employment. *Health Serv Res*. 2002;37:1145-1158.
- Schoenbaum M, Unutzer J, Sherbourne C, et al. Cost-effectiveness of practice-initiated quality improvement for depression: results of a randomized controlled trial. *JAMA*. 2001;286:1325-1330.
- Katon W, Von Korff M, Lin E, et al. Collaborative management to achieve treatment guidelines. Impact on depression in primary care. *JAMA*. 1995;273:1026-1031.
- Hedrick SC, Chaney EF, Felker B, et al. Effectiveness of collaborative care depression treatment in Veterans' Affairs primary care. *J Gen Intern Med*. 2003;18:9-16.
- Unutzer J, Katon W, Callahan CM, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial [see comment]. *JAMA*. 2002;288:2836-2845.
- Unutzer J, Rubenstein L, Katon WJ, et al. Two-year effects of quality improvement programs on medication management for depression. *Arch Gen Psychiatry*. 2001;58:935-942.
- Katzelnick DJ, Simon GE, Pearson SD, et al. Randomized trial of a depression management program in high utilizers of medical care. *Arch Fam Med*. 2000;9:345-351.
- Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? *Eff Clin Pract*. August-September 1998;1:2-4.
- US Preventive Services Task Force. Screening for depression: recommendations and rationale. *Ann Intern Med*. 2002;136:760-776.
- Pignone MP, Gaynes BN, Rushton JL, et al. Screening for depression in adults: a summary of the evidence for the US Preventive Services Task Force. *Ann Intern Med*. 2002;136:765-776.
- Charbonneau A, Rosen AK, Ash AS, et al. Measuring the quality of depression care in a large integrated health system. *Med Care*. 2003;41:669-680.
- Charbonneau A, Rosen AK, Owen RR, et al. Monitoring depression care: in search of an accurate quality indicator. *Med Care*. 2004;42:522-531.
- Samsa G, Matchar D. Can continuous quality improvement be assessed using randomized trials? [See comment.] *Health Serv Res*. 2000;35:687-700.
- Department of Veterans Affairs. *Clinical Practice Guidelines for Major Depressive Disorder*. Washington, DC: Department of Veterans Affairs; 1997.
- Steiner JF, Koepsell TD, Fihn SD, Inui TS. A general method of compliance assessment using centralized pharmacy records. Description and validation. *Med Care*. 1988;26:814-823.
- Steiner JF, Prochazka AV. The assessment of refill compliance using pharmacy records: methods, validity, and applications. *J Clin Epidemiol*. 1997;50:105-116.
- Choo PW, Rand CS, Inui TS, et al. Validation of patient reports, automated pharmacy records, and pill counts with electronic monitoring of adherence to anti-hypertensive therapy. *Med Care*. 1999;37:846-857.
- Saunders K, Simon G, Bush T, Grothaus L. Assessing the feasibility of using computerized pharmacy refill data to monitor antidepressant treatment on a population basis: a comparison of automated and self-report data. *J Clin Epidemiol*. 1998;51:883-890.
- Kerr EA, McGlynn EA, Van Vorst KA, Wickstrom SL. Measuring antidepressant prescribing practice in a health care system using administrative data: implications for quality measurement and improvement. *Jt Comm J Qual Improv*. 2000;26:203-216.
- Berlowitz DR, Young GJ, Hickey EC, et al. Quality improvement implementation in the nursing home. *Health Serv Res*. 2003;38(1 pt 1):65-83.
- Hankin CS, Spiro A 3rd, Miller DR, Kazis L. Mental disorders and mental health treatment among US Department of Veterans Affairs outpatients: the Veterans Health Study. *Am J Psychiatry*. 1999;156:1924-1930.
- Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 1988.
- Perlin J, Kazis L, Skinner K, et al. *Health Status and Outcomes of Veterans: 1999 Health Survey of Veteran Enrollees*. Executive Report. Washington, DC: Department of Veterans Affairs; 2000.
- Goldman LS, Nielsen NH, Champion HC. Awareness, diagnosis, and treatment of depression. *J Gen Intern Med*. 1999;14:569-580.
- Frank RG, Goldman HH, McGuire TG. Will parity in coverage result in better mental health care? *N Engl J Med*. 2001;345:1701-1704.