



Published in final edited form as:

Ment Health Serv Res. 2005 June ; 7(2): 103–112.

Clinical Improvement Associated with Conformance to HEDIS-Based Depression Care

Kathryn Rost^{1,4}, L. Miriam Dickinson¹, John Fortney², John Westfall¹, and Richard C. Hermann³

¹Center for Studies in Family Medicine, Department of Family Medicine, University of Colorado Health Sciences Center.

²Health Services Research and Development, Center for Mental Health and Outcomes Research, Central Arkansas Veterans Healthcare System, North Little Rock, AR and Centers for Mental Healthcare Research, Department of Psychiatry, University of Arkansas for Medical Sciences, Little Rock, AR.

³Center for Quality Assessment and Improvement in Mental Health, Department of Psychiatry, Research Scientist, Health Institute, Tufts-New England Medical Center, Boston, MA.

Abstract

Background: Employers recently requested a valid metric of depression treatment quality. Such an indicator needs to measure the proportion of the population in need who receive high-quality care, and to predict clinical improvement. *Methods:* We constructed an administrative database indicator derived from HEDIS criteria for antidepressant medication management, and tested it in 230 employed patients in five health plans. *Results:* Indicator rates were 7.0% in the population in need. Conformance to indicator criteria in this population was associated with 23.0% improvement in depression severity over 1 year ($p = .02$). *Conclusions:* Administrative database indicators that predict clinical improvement are a very rare accomplishment. Existing depression indicators may need to be calculated for the population in need to provide a valid metric for employer purchasers.

Keywords

HEDIS; primary care; depression; outcomes; quality

In a move towards value-based purchasing, a recent forum of employers announced they were prepared to actively negotiate with their health plans to improve the quality of depression treatment provided their workforce 'if a valid metric of depression treatment could be identified' (Apgar, 2002). Among the most promising candidates for this metric are the National Committee on Quality Assurance's (NCQA) Health Plan Employer Data and Information Set (HEDIS) outpatient depression indicators. Derived from guidelines summarizing expert consensus on treatment research and clinical care (Depression Guideline Panel, 1993), HEDIS depression indicators are currently constructed from administrative databases in over 300 health plans to characterize the quality of care a health plan provides to the population *in treatment*, diagnosed patients who initiate antidepressant medication (Thompson, Bost, Ahmed, Ingalls, & Sennett, 1998).

⁴Correspondence should be directed to Dr Kathryn Rost, Center for Studies in Family Medicine, University of Colorado Health Sciences Center Department of Family Medicine, UCHSC at Fitzsimons, P.O. Box 6508 Mail Stop F496, Aurora, CO 80045-0508; e-mail: Kathryn.Rost@UCHSC.edu.

To be a valid metric of the quality of depression care from the perspective of a purchaser, an indicator needs to meet two criteria: (1) it should measure the quality of care a health plan provides to the population *in need*, and (2) it should predict improvement in clinical outcomes in this population. Little is known about whether indicators currently calculated for the population *in treatment* accurately characterize the population *in need* because (unlike other chronic conditions) most depressed patients are not in treatment (Rost, Smith, Guise, & Matthews, 1994; Rost et al., 2001). In addition, it is unclear whether HEDIS-based outpatient depression indicators predict improved outcomes in either the population in treatment or the population in need (Rost, Williams, Wherry, & Smith, 1995; Simon et al., 1995; Melfi et al., 1998; Katon et al., 2000; Fortney, Rost, Zhang, & Pyne, 2001; Bull et al., 2002; Schoenbaum et al., 2002); however, multiple studies question whether administrative databases contain too much measurement error to derive valid quality metrics for any condition (Hunt et al., 2000; Cotter, Smith, Rossiter, Pugh, & Bramble, 1999; Bloom, Harris, Thompson, Ahmed, & Thompson, 2000; Kobak et al., 2002; Jones et al., 2000). Conducted as a secondary analysis in a cooperative study database (Rost et al., 2001), the first objective of the study was to calculate HEDIS-based outpatient depression indicator rates for the population in need. The second objective of this study was to investigate whether this indicator significantly predicted improved health in this population.

METHODS

Recruitment and Data Collection

Our previously described methods (Rost et al., 2001; Wells, 1999), approved by the Human Research Advisory Committees at the University of California Los Angeles, RAND, and the VA Greater Los Angeles, are summarized here. Research teams from Partners in Care (PIC) and the Mental Health Awareness Project (MHAP) administered a depression screener to a consecutive sample of adult patients insured by five health plans who provided administrative data for participants recruited at an index primary care visit from 27 staff/group and network primary care practices. Patients who screened positive on the depression screener at the index visit were eligible for the parent study if they reported: (1) they intended to receive care in the clinic on an ongoing basis; (2) they had telephone access; (3) they were not pregnant/breastfeeding, cognitively impaired or seriously physically ill, and (4) spoke English or Spanish. In addition, one of the parent studies excluded patients who: (5) failed to meet criteria for past-year major depression; (6) screened positive for bereavement, mania or alcohol problems; or (7) did not speak English. Because HEDIS depression indicators are currently calculated for patients *beginning* treatment, the study team excluded patients who entered the parent study *in treatment* (see Fig. 1). Because this manuscript is part of a larger series of studies examining employer health benefit purchasing (Rost, Smith, & Fortney, 2000), the study team also excluded patients who reported no full or part time employment. Note that we included 32 depressed patients who did not meet criteria for current major depression because: (1) HEDIS-eligible patients include diagnosed patients with minor depression and (2) quality improvement initiatives improve outcomes comparably across the spectrum of depression severity, including minor depression (Wells et al., 2000; Rost, Elliott, & Dickinson, 2002). After completing baseline, patients from the parent study completed structured interviews at 6 and 12 months, with response rates of 97.4 and 88.7% in this cohort (see Fig. 1).

Operational Definition of Major Constructs

Acute Outpatient Depression Indicator—The indicator was calculated from complete medical, mental health specialty care, and pharmacy data for each participant for 1 year following the index visit. Parallel to HEDIS 3.0 criteria, patients were included in the indicator's numerator if: (a) a prescription for an antidepressant medication was noted from up to 30 days before to 14 days after index episode start date; (b) the prescription was filled a sufficient

number of times for patients to be able to take the medication for 84 out of 114 days following the first prescription; and (c) three non-emergency room visits with any mental health diagnosis (ICD9 code between 290 and 319 inclusive) to a primary care or mental health provider (if all three visits were to mental health providers, at least one of them had to be to a prescribing provider) were noted during the 12 weeks following the index episode start date. The index episode start date was defined as the date of the first primary care visit with a depression diagnosis, or the index visit for depressed patients who were not diagnosed. We elected to test an indicator combining criteria (a) through (c) above because the current HEDIS indicator requires all three criteria be met, and multicollinearity prohibited us from drawing meaningful conclusions about which criteria made stronger contributions to improved outcomes.

All patients meeting the study's eligibility criteria were included in the indicator's denominator as members of the population in need. Thus, all subjects received an administrative database diagnosis of depression (ICD 9 codes of 296.2, 296.3, 298.0, 300.4, 309.1 and 311) during the 3 months following the index visit and/or reported at screening 2 or more weeks of feeling depressed or losing interest in the past year with 1 or more weeks in the past month (Rost et al., 2001). We selected 3 months as the cutoff for diagnosis to be able to assess whether patients met numerator criteria (b).

Outcome—Depression severity was measured at each wave by the 23-item modified Center for Epidemiological Studies—Depression (mCES-D) (Rost, Nutting, Smith, Werner, & Duan, 2001) 100-point scale examining DSM-IV depressive symptoms ($\alpha = .80$), with higher scores indicating greater severity. Depression severity scores were log-transformed before analysis to achieve a more normally distributed dependent variable.

Covariates—Multivariate models presented in this manuscript used baseline measures of gender, minority status, education, depression and/or dysthymia diagnoses derived from structured interviews, role functioning, household income, physical comorbidities, and health plan. We selected these covariates from more than twenty sociodemographic, occupational, and clinical characteristics by using stepwise regression methods to identify all variables that entered the depression severity model at $p < .20$.

Data Analysis

To address the first objective, we calculated indicator rates in the population in need as defined above. To address the second objective, we investigated the relationship of the indicator to depression severity change using multilevel longitudinal models (Gibbons & Hedeker, 1994; Singer, 1998; Raudenbush & Bryk, 2002). Because both projects randomized patients to an effectiveness intervention, the dependent variable across all waves was modeled with a random intercept as a linear combination of indicator, time, intervention, indicator \times time, intervention \times time, and the covariates listed above, to control for possible intervention effects that were not mediated by the indicator (Kraemer, Wilson, Fairburn, & Agras, 2002). We tested the relationship between the indicator and depression severity change by examining the parameter estimate and associated statistical significance of the indicator \times time term, retransforming parameter estimates using smearing retransformation (Duan, 1983). We characterized the strength of the relationship between the indicator and depression severity change by calculating the area between the curves, reporting the percent improvement in depression severity over 1 year attributable to indicator–concordant care. To rule out the possibility that the indicator's impact on severity reflected concurrent psychotherapy, we conducted an analysis in the subgroup of patients whose administrative database records indicated they had no psychotherapy visits in the 6 months following the index visit. We were prohibited from using instrument variables to correct for selection bias by the small sample size; however, previous research suggests that non-instrumented estimates of indicator–outcome relationships are

conservative (Schoenbaum et al., 2002; Fortney et al., 2001). Power analyses indicated that the sample size provided us greater than 80% power to detect a .8 effect of the indicator on severity change using a two-sided test with alpha set at .05.

RESULTS

Participants

On average, the 230 patients eligible for this analysis were at baseline 41.6-years old ($SD = 11.2$), 66.5% female, 39.1% minority, 48.7% currently married, and 74.4% with some college education. Occupationally, 43.5% of participants were employed at baseline as professionals/administrators, 22.2% as managers/salespeople, and 34.3% as clerical/service workers. The majority of subjects met criteria for recent major depression and/or dysthymia on structured interviews with 74.8% meeting criteria for 1-year major depression, 11.3% meeting criteria for 1-year dysthymia, and 13.9% meeting criteria for substantial depressive symptoms. Patients reported an average of 1.2 ($SD = 1.4$) physical comorbidities.

Among the 230 patients comprising the population in need, 7.0% were positive on the indicator with 18.3% starting antidepressant medications (criteria a), 13.0% continuing antidepressant medications that were started (criteria b), and 11.7% receiving follow-up visits (criteria c). As Table 1 and Figure 2 show, conformance with the indicator was associated with an average annual improvement in depression severity of 23.0% (Indicator \times Time $t = -2.37$, $p = .019$, $df = 639$). By 1 year, patients who were indicator positive had realized a 51.8% reduction in depressive symptoms (from 45.4 to 21.9), while patients who were indicator negative realized a 16.5% reduction in depressive symptoms (from 45.4 to 37.9). Thus, patients who were indicator positive realized 3.1 times the improvement in depressive symptoms that their indicator negative counterparts realized. Conformance with the indicator was associated with a 32.1% improvement in the 203 patients in need who received no psychotherapy (Indicator \times Time $t = -2.29$, $p = .023$, $df = 563$).

DISCUSSION

To be a valid metric of depression treatment for a purchaser, an indicator should meet two criteria. First, the indicator should characterize the quality of care provided to the population in need. Second, the indicator should predict improved clinical outcomes in this population. In terms of the first criterion, the HEDIS indicator currently in use potentially overestimates the rate of high quality medication management in the population in need by as much as four-fold (NCQA, 2003). In terms of the second criterion, the HEDIS indicator currently in use predicts a 23% improvement in clinical outcomes in the broader population in need. This degree of improvement is clinically as well as statistically significant, associated with six times the average effect size of antidepressant medication compared to placebo in patients with major depression (Walsh, Seidman, Sysko, & Gould, 2002).

While providing evidence that depression indicators developed in administrative databases 'work,' these findings suggest that the indicator may need to be calculated in both narrowly and broadly defined populations to be relevant for the range of stakeholders whose decisions quality indicators are meant to influence (Hermann, 2002; Hermann & Palmer, 2002). Applied narrowly to the population in treatment, the indicator provides adherence data that health plans have used to guide adherence improvement initiatives (Roberts, Cockerham, & Waugh, 2002; Hoffman et al., 2003). Applied broadly to the population in need, the indicator provides data on diagnosis and medication initiation/adherence, for broader quality improvement initiatives. From the employer perspective, an indicator characterizing the population in need is more meaningful because companies absorb considerable costs when depressed workers are not diagnosed and effectively treated (Druss et al., 2001; Kessler & Frank, 1997; Kessler et

al., 1999; Greenberg, Kessler, Nells, Finkelstein, & Berndt, 1996; Russell, Patterson, & Baker, 1998; Martin, Blum, Beach, & Roman, 1996; Burton, Conti, Chen, Schultz, & Edington, 1999; Ettner, Frank, & Kessler, 1997). Employer purchasers need to evaluate the quality of the *entire* process, rather than the quality of *any single part* of the process to judge how depression treatment quality is affecting these costs.

Calculating depression indicators for the population in treatment can be done in pre-existing administrative databases. Calculating depression indicators for the population in need is more cumbersome because it requires a population-based survey. Economies of effort may be obtained by adding a highly sensitive and specific 2-item depression screener (Kroenke & Spitzer, 2002) to telephone surveys that NCQA may adopt as part of its accreditation process. The potential benefit of calculating depression indicators for the population in need is that the resulting indicator could motivate employer purchasers to support cost-effective quality improvement initiatives which modestly increase health care expenditures (Rost, Pyne, Dickinson, & LoSasso. The cost-effectiveness of ongoing management for primary care major depression, 2005) (Schoenbaum et al., 2001; Lave, Frank, Schulberg, & Kamlet, 1998; Simon et al., 2001).

The internal validity of the findings we report is strengthened by our careful construction of a 'real world' indicator using 'real world' databases. However, we recognize that future research needs to address challenges to the internal validity of these findings from selection bias and causal inference issues. In terms of selection bias (Newhouse & McClellan, 1998; Sturm, Unutzer, & Katon, 1999), our inability to use instrumented models in all likelihood downwardly biases the indicator–outcome relationships we report by an estimated 20% or greater (Fortney et al., 2001; Schoenbaum et al., 2002). In terms of inference, we recognize the non-experimental design prevents us from definitively concluding that high-quality depression care improves clinical outcomes; however, experimental studies showing that interventions that improve antidepressant medication management do in fact significantly reduce depressive symptoms (Wells et al., 2000; Rost, Nutting, Smith, Elliott, & Dickinson, 2002; Katon et al., 1995; Katon et al., 1996; Schulberg et al., 1996; Katzelnick et al., 2000; Katon et al., 1999; Simon, VonKorff, Rutter, & Wagner, 2000; Hunkeler et al., 2000), strengthen the causal link between quality and outcomes this study and other investigators observe (see Table 2).

The external validity of the study is strengthened by its ability to employ longitudinal models to test a quality indicator derived from administrative databases in a highly representative population of depressed primary care patients; however, we recognize limitations to external validity. Because we lacked pre-index visit administrative data that HEDIS 3.0 currently relies on, we had to use patient report to identify patients beginning a new treatment episode, allowing us to test the indicator in a comparable but not completely equivalent population. Because the parent study recruited a consecutively screened primary care sample, we could not include the estimated 6% of depressed patients who get their care exclusively from specialty care settings without seeing a primary care provider for any reason during their episode (Rost, Zhang, Fortney, Smith, & Smith, 1998). While it is possible the population we analyzed had a less severe form of depression than the population in which HEDIS indicators are currently calculated, we do not have sufficient power to determine whether the indicator is differentially effective in more or less severe populations. Future investigators are encouraged to examine this important question. While our methods can be criticized for proposing too broad a definition of the population in need, we note that quality improvement interventions targeting this very population achieve comparable improvements in clinical outcomes across the severity spectrum including those with significant depressive symptoms only (Wells et al., 2000; Rost, Elliott, & Dickinson, 2002). While there are clear advantages in demonstrating the value of quality indicators to employer purchasers, we recognize that the indicator–outcome relationships we report do not generalize to the non-employed patients insured through

Medicare and/or Medicaid, some of whom may have more severe disorder (Marcotte & Wilcox-Gok, 2001).

Although it is reassuring that we can demonstrate the expected indicator–outcome relationships, the fact that we have administrative data from only five health plans limits the generalizability of our results. We encourage future investigators to explore these relationships within and across a considerably larger and more diverse group of health plans and patients using analytic strategies capable of precisely differentiating the relative contribution of acute and continuation indicators to a broad range of outcomes. Such databases can also make an important contribution to determining whether a broader application of the indicator is more susceptible to case mix differences across health plans than the narrower application.

Given the enormous challenges in creating ‘real world’ metrics, the definition of an administrative database indicator that predicts clinical improvement is an almost unique accomplishment in health care (Krumholz et al., 1998; Ahmed, Elbasha, Thompson, Harris, & Sneller, 2002), particularly in mental health (Hermann et al., 2000; Buchanan, Kreyenbuhl, Zito, & Lehman, 2002). This study contributes to the quality improvement literature by being the first study to our knowledge to demonstrate that a HEDIS-derived indicator predicts patient-reported clinical improvement (Melfi et al., 1998; Krumholz et al., 1998; Ahmed et al., 2002). The study also raises an important question about whether HEDIS-derived depression indicators need to be calculated in both broadly and narrowly defined populations to be relevant to employers as well as the health plans they sponsor.

ACKNOWLEDGMENTS

We greatly appreciate the expert programming and encouragement that Carl Elliott contributed to this project. We also appreciate the help that Martha Mancewicz, Bernadette Benjamin, and the VA Center for the Study of Healthcare Provider Behavior provided with database construction. We gratefully acknowledge the efforts of Stephanie Mitchell, Mary Abdun-Nur, Maureen Carney, Carole Oken, Sarah Scholle, and our colleagues in the Quality Improvement for Depression (QID) Cooperative Study. We are also grateful to the managed care plans, clinicians, and patients who participated in this study from the Alamo Mental Health Group (San Antonio TX), Allina Medical Group (Twin Cities MN), Magellen/GreenSpring Behavioral Health (MD), Humana Health Care Plans (San Antonio TX), Kaiser Permanente (Northern CA), MedPartners (Los Angeles CA), PacifiCare of Texas (San Antonio TX), Patuxent Medical Group (MD), San Luis Valley Mental Health/Colorado Health Networks (San Luis Valley CO), Valley-Wide Health Services (San Luis Valley CO), and the Veterans Administration Greater Los Angeles Healthcare System (Los Angeles CA). This work was funded by the Agency for Healthcare Research and Quality (HS08349), by the National Institute of Mental Health (MH54444, MH54443, MH01170, MH63651, the Research Center on Managed Care for Psychiatric Disorders MH54623), and the John D. and Catherine T. MacArthur Foundation (96-42901A-HE).

REFERENCES

- Ahmed F, Elbasha EE, Thompson BL, Harris JR, Sneller VP. Cost-benefit analysis of a new HEDIS performance measure for pneumococcal vaccination. *Medical Decision Making* 2002;22:S58–S66. [PubMed: 12369232]
- Apgar, KR. *Linking pharmacy and mental health benefits: Collaborating to improve depression care.* Washington, DC: 2002.
- Bloom SA, Harris JR, Thompson BL, Ahmed F, Thompson J. Tracking clinical preventive service use: A comparison of the health plan employer data and information set with the behavioral risk factor surveillance system. *Medical Care* 2000;38:187–194. [PubMed: 10659692]
- Buchanan RW, Kreyenbuhl J, Zito JM, Lehman A. The schizophrenia PORT pharmacological treatment recommendations: Conformance and implications for symptoms and functional outcome. *Schizophrenia Bulletin* 2002;28:63–73. [PubMed: 12047023]
- Bull SA, Hu XH, Hunkeler EM, Lee JY, Ming EE, Fireman B. Discontinuation of use and switching of antidepressants: Influence of patient-physician communication. *JAMA* 2002;288:1403–1409. [PubMed: 12234237]

- Burton WN, Conti DJ, Chen C, Schultz AB, Edington DW. The role of health risk factors and disease on worker productivity. *Journal of Occupational and Environmental Medicine* 1999;41:863–877. [PubMed: 10529942]
- Cotter JJ, Smith WR, Rossiter LF, Pugh CB, Bramble JD. Combining state administrative databases and provider records to assess the quality of care for children enrolled in Medicaid. *American Journal of Medical Quality* 1999;14:98–104. [PubMed: 10446671]
- Depression Guideline Panel. *Depression in Primary Care: Vol. 2. Treatment of Major Depression. Clinical Practice Guideline, Number 5.* US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; Rockville, MD: 1993. AHCPR Publication No. 93-0551
- Druss BG, Marcus SC, Olfson M, Tanielian T, Elinson L, Pincus HA. Comparing the national economic burden of five chronic conditions. *Health Affairs* 2001;20:233–241. [PubMed: 11816664]
- Duan N. Smearing estimate: A nonparametric retransformation method. *Journal of the American Statistical Association* 1983;78:605–610.
- Ettner SL, Frank RG, Kessler RC. The impact of psychiatric disorders on labor market outcomes. *Industrial and Labor Relations Review* 1997;51:64–81.
- Fortney J, Rost K, Zhang M, Pyne J. The relationship between quality and outcomes in routine depression care. *Psychiatric Services* 2001;52:56–62. [PubMed: 11141529]
- Gibbons RD, Hedeker D. Application of random-effects probit regression models. *Journal of Consulting and Clinical Psychology* 1994;62:285–296. [PubMed: 8201066]
- Greenberg, PE.; Kessler, RC.; Nells, TL.; Finkelstein, SN.; Berndt, ER. Depression in the workplace: An economic perspective. In: Feighner, JP.; Boyer, WF., editors. *Selective Serotonin Re-Uptake Inhibitors: Advances in Basic Research and Clinical Practice.* 2nd. John Wiley and Sons, Ltd; Chichester, West Sussex England: 1996.
- Hermann, R. Linking outcome measurement with process measurement for quality improvement. In: IsHak, WBT.; Sederer, L., editors. *Outcome Measurement in Psychiatry: A Critical Review.* Washington, DC: 2002.
- Hermann RC, Leff S, Palmer H, Yang D, Teller T, Provost S, Jakubiak C, Chan J. Quality measures for mental health care: Research from a national inventory. *Medical Care Research and Review* 2000;57 (Suppl 2):136–154. [PubMed: 11105510]
- Hermann RC, Palmer RH. Common ground: A framework for selecting core quality measures for mental health and substance abuse care. *Psychiatric Services* 2002;53:281–287. [PubMed: 11875220]
- Hoffman L, Enders J, Luo J, Segal R, Pippins J, Kimberlin C. Impact of an antidepressant management program on medication adherence. *American Journal of Managed Care* 2003;9:70–80. [PubMed: 12549816]
- Hunkeler EM, Meresman JF, Hargreaves WA, Fireman B, Berman WH, Kirsch AJ, Groebe J, Hurt SW, Braden P, Getzell M, Feigenbaum PA, Peng T, Salzer M. Efficacy of nurse telehealth care and peer support in augmenting treatment of depression in primary care. *Archives of Family Medicine* 2000;9:700–708. [PubMed: 10927707]
- Hunt JP, Cherr GS, Hunter C, Wright MJ, Wang YZ, Steeb G, Buechter KJ, Meyer AA, Baker CC. Accuracy of administrative data in trauma: Splenic injuries as an example. *Journal of Trauma, Injury, Infection and Critical Care* 2000;49:679–686.
- Jones D, Hendricks A, Comstock C, Rosen A, Chang BH, Rothendler J, Hankin C, Prashker M. Eye examinations for VA patients with diabetes: Standardizing performance measures. *International Journal for Quality in Health Care* 2000;12:97–104.
- Katon W, Robinson P, Von Korff M, Lin E, Bush T, Ludman E, Simon G, Walker E. A multifaceted intervention to improve treatment of depression in primary care. *Archives of General Psychiatry* 1996;53:924–932. [PubMed: 8857869]
- Katon W, Rutter CM, Lin E, Simon G, Von Korff M, Bush T, Walker E, Ludman E. Are there detectable differences in quality of care or outcome of depression across primary care providers? *Medical Care* 2000;38:552–561. [PubMed: 10843308]
- Katon W, Von Korff M, Lin E, Simon G, Walker E, Unutzer J, Bush T, Russo J, Ludman E. Stepped collaborative care for primary care patients with persistent symptoms of depression: A randomized trial. *Archives of General Psychiatry* 1999;56:1109–1115. [PubMed: 10591288]

- Katon W, Von Korff M, Lin E, Walker E, Simon GE, Bush T, Robinson P, Russo J. Collaborative management to achieve treatment guidelines: Impact on depression in primary care. *JAMA* 1995;273:1026–1031. [PubMed: 7897786]
- Katzelnick DJ, Simon GE, Pearson SD, Manning WG, Helstad CP, Henk HJ, Cole SM, Lin EHB, Taylor LH, Kobak KA. Randomized trial of a depression management program in high utilizers of medical care. *Archives of Family Medicine* 2000;9:345–351. [PubMed: 10776363]
- Kessler RC, Barber C, Birnbaum HG, Frank RG, Greenberg PE, Rose RM, Simon GE, Wang P. Depression in the workplace: Effects on short-term disability. *Health Affairs* 1999;18:163–171. [PubMed: 10495604]
- Kessler RC, Frank RG. The impact of psychiatric disorders on work loss days. *Psychological Medicine* 1997;27:861–873. [PubMed: 9234464]
- Kobak KA, Taylor L, Katzelnick DJ, Olson N, Clagnaz P, Henk HJ. Antidepressant medication management and Health Plan Employer Data Information Set (HEDIS) criteria: Reasons for nonadherence. *Journal of Clinical Psychiatry* 2002;63:727–732. [PubMed: 12197454]
- Kraemer HC, Wilson T, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry* 2002;59:877–883. [PubMed: 12365874]
- Kroenke K, Spitzer RL, Williams JBW. The PHQ: Validity of a brief depression severity measure. *Journal of General Internal Medicine* 2001;16:606–613. [PubMed: 11556941]
- Krumholz HM, Radford MJ, Wang Y, Chen J, Heiat A, Marciniak TA. National use and effectiveness of beta-blockers for the treatment of elderly patients after acute myocardial infarction: National Cooperative Cardiovascular Project. *JAMA* 1998;280:623–629. [PubMed: 9718054]
- Lave JR, Frank RG, Schulberg HC, Kamlet MS. Cost-effectiveness of treatments for major depression in primary care practice. *Archives of General Psychiatry* 1998;55:645–651. [PubMed: 9672056]
- Marcotte DE, Wilcox-Gok V. Estimating the employment and earnings costs of mental illness: Recent developments in the United States. *Social Science and Medicine* 2001;53:21–27. [PubMed: 11386306]
- Martin JK, Blum TC, Beach SRH, Roman PM. Subclinical depression and performance at work. *Social Psychiatry and Psychiatric Epidemiology* 1996;31:3–9. [PubMed: 8821918]
- Melfi CA, Chawla AJ, Croghan TW, Hanna MP, Kennedy S, Sredl K. The effects of adherence to antidepressant treatment guidelines on relapse and recurrence of depression. *Archives of General Psychiatry* 1998;55:1128–1132. [PubMed: 9862557]
- NCQA Antidepressant Medication Management: The State of Health Care Quality 2002. 2003 Accessed May 2003. [On-line]. Available: http://www.ncqa.org/sohc2002/sohc_2002_amm.html
- Newhouse JP, McClellan M. Econometrics in outcomes research: The use of instrumental variables. *Annual Review of Public Health* 1998;19:17–34.
- Raudenbush, SW.; Bryk, AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd. Sage; Thousand Oaks, CA: 2002.
- Roberts K, Cockerham TR, Waugh WJ. An innovative approach to managing depression: Focus on HEDIS standards. *Journal for Healthcare Quality* 2002;24:11–17. [PubMed: 12432857]
- 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. *Journal of General Internal Medicine* 2001;16:143–149. [PubMed: 11318908]
- Rost K, Nutting P, Smith JL, Elliott CE, Dickinson M. Managing depression as a chronic disease: A randomised trial of ongoing treatment in primary care. *BMJ* 2002;325:934–939. [PubMed: 12399343]
- Rost K, Smith GR, Guise B, Matthews DB. The deliberate misdiagnosis of major depression in primary care. *Archives of Family Medicine* 1994;3:333–337. [PubMed: 8012621]
- Rost K, Smith J, Fortney J. Large employers' selection criteria in purchasing behavioral health benefits. *Journal of Behavioral Health Services and Research* 2000;27:334–338. [PubMed: 10932446]
- Rost K, Williams C, Wherry J, Smith GR Jr. The process and outcomes of care for major depression in rural family practice settings. *Journal of Rural Health* 1995;11:114–121. [PubMed: 10143272]
- Rost KM, Duan N, Rubenstein LV, Ford DE, Sherbourne CD, Meredith LS, Wells KB. The Quality Improvement for Depression collaboration: General analytic strategies for a coordinated study of

quality improvement in depression care. *General Hospital Psychiatry* 2001;23:239–253. [PubMed: 11600165]

- Rost KM, Smith JL, Dickinson LM. The effect of improving primary care depression management on employee absenteeism and productivity: A randomized trial. *Medical Care* 2004;42:1202–1210. [PubMed: 15550800]
- Rost KM, Pyne JM, Dickinson LM, LoSasso T. Cost-effectiveness of enhancing primary care depression management on an ongoing basis. *Annals of Family Medicine* 2005;3:15–22. [PubMed: 15671186]
- Rost KM, Zhang M, Fortney J, Smith J, Smith GR Jr. Rural-urban differences in depression treatment and suicidality. *Medical Care* 1998;36:1098–1107. [PubMed: 9674626]
- Russell JM, Patterson J, Baker AM. Depression in the workplace: Epidemiology, economics and effects of treatment. *Disease Management and Health Outcomes* 1998;4:135–142.
- 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 Services Research* 2002;37:1145–1158. [PubMed: 12479490]
- Schoenbaum M, Unutzer J, Sherbourne C, Duan N, Rubenstein LV, Miranda J, Meredith LS, Carney MF, Wells K. Cost-effectiveness of practice-initiated quality improvement for depression: Results of a randomized, controlled trial. *JAMA* 2001;286:1325–1330. [PubMed: 11560537]
- Schulberg HC, Block MR, Madonia MJ, Scott CP, Rodriguez E, Imber SD, Perel J, Lave J, Houck PR, Coulehan JL. Treating major depression in primary care practice: Eight-month clinical outcomes. *Archives of General Psychiatry* 1996;53:913–919. [PubMed: 8857868]
- Simon GE, Lin EHB, Katon W, Saunders K, VonKorff M, Walker E, Bush T, Robinson P. Outcomes of “inadequate” antidepressant treatment. *Journal of General Internal Medicine* 1995;10:663–670. [PubMed: 8770718]
- Simon GE, Manning WG, Katelnick DJ, Pearson SD, Henk HJ, Helstad CP. Cost-effectiveness of systematic depression treatment for high utilizers of general medical care. *Arch General Psychiatry* 2001;58:181–187.
- Simon GE, VonKorff M, Rutter C, Wagner E. Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. *BMJ* 2000;320:550–554. [PubMed: 10688563]
- Singer JD. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics* 1998;23:323–355.
- Sturm R, Unutzer J, Katon W. Effectiveness research and implications for study design: Sample size and statistical power. *General Hospital Psychiatry* 1999;21:274–283. [PubMed: 10514951]
- Thompson JW, Bost J, Ahmed F, Ingalls CE, Sennett C. The NCQA's quality compass: Evaluating managed care in the United States: A brief look at the NCQA's comparison of health plan performance. *Health Affairs* 1998;17:152–158. [PubMed: 9455025]
- Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: Variable, substantial, and growing. *JAMA* 2002;287:1840–1847. [PubMed: 11939870]
- Wells KB. The design of partners in care: Evaluating the cost-effectiveness of improving care for depression in primary care. *Social Psychiatry and Psychiatric Epidemiology* 1999;34:20–29. [PubMed: 10073117]
- Wells KB, Sherbourne CD, Schoenbaum M, Duan N, Meredith LS, Unutzer J, Miranda J, Carney MF, Rubenstein LV. Impact of disseminating quality improvement programs for depression in managed primary care: A randomized controlled trial. *JAMA* 2000;283:212–220. [PubMed: 10634337]



Fig. 1.
Patient selection.

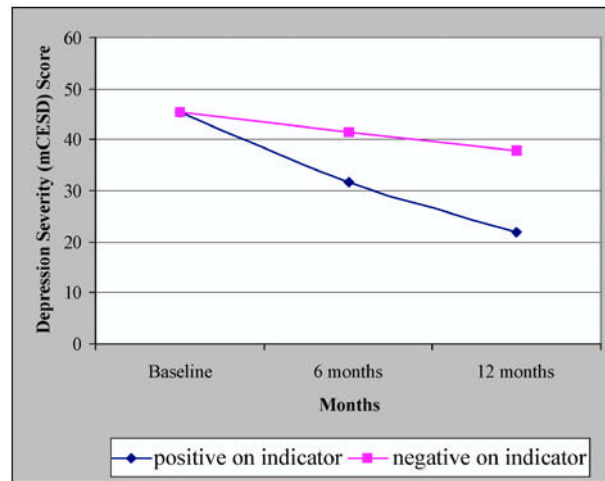


Fig. 2. Relationship of outpatient depression indicator to depression severity change ($n = 230$).

Table 1
Relationship of Outpatient Depression Indicator* To Depression Severity Over 1 Year

Depression severity	All patients (N = 230)			All patients with no psychotherapy (N = 203)		
	Est.	SE	P	Est.	SE	P
Intercept	3.899	0.124	<.0001	3.923	0.130	<.0001
Indicator	0.336	0.102	0.001	0.484	0.191	0.012
Intervention	0.013	0.076	0.865	-0.025	0.080	0.759
Time	-0.113	0.041	0.007	-0.107	0.042	0.012
Indicator × Time	-0.255	0.107	0.019	-0.330	0.144	0.023
Intervention × Time	-0.069	0.053	0.194	-0.080	0.055	0.0150

* Controlling for covariates noted in text.

** Non-significant interaction terms (Indicator × Intervention × Time) are not included.

Table 2
Outcome Relationships in Previous Studies of Depressed Patients

Study	Population	Indicator	Source	Time frame	Outcome	Source	Time frame	Correction for selection bias	Indicator–outcome relationship
Williams, Liem, & Smith, 1995	38 primary-care patients with current major depression identified by screening	75% or greater compliance with therapeutic dose of AD for 8+ weeks	Patient report	6 month window following index visit	2 or fewer depressive symptoms	Patient report	6 months following index visit	Covariates	Significant but strength not estimated
Monahan et al., 2005	88 primary-care patients identified by AD	Average daily dose diagnosed within 10% of recommended dose	Administrative database report	6 months following referral to study	No current major depression at 1, 4, or 7 months AND 50% or more reduction in depression severity	Patient report	7 months following index visit	Covariates	Significant improvement with adequate doses compared to inadequate doses (84% versus 64%)
Fiore et al., 2008	4052 patients beginning new treatment episode	Filled AD prescriptions 4+ times	Administrative database	6 months following initial prescription	New AD treatment episode OR hospital admission for depression OR electroconvulsive shock therapy OR emergency room visit for mental health OR attempted suicide	Administrative database	18 months following treatment period	Covariates	Significantly reduces risk of relapse/remission by 56%
Monahan et al., 2010	1,599 primary care patients identified by initial AD prescription for new treatment episode	AD prescription refill	Administrative database	6 weeks following initial prescription	Persistent depressive symptoms	Patient report	6–8 weeks following initial prescriptions	Covariates	Significantly reduces probability of persistent symptoms ($r = -.18$)
Stoney et al., 2001	106 primary-care patients with current major and minor depression identified	75% or greater compliance with therapeutic dose of AD for 8+ weeks OR 8+ specialty care visits over 12 week period	Patient report Administrative database report	6 month window not keyed to index visit	Depression severity MCS PCS	patient report	6 months following baseline interview	Covariates IV sensitivity analysis	Significant 10.1–12.2 point improvement in depression severity in patients with major depression

Men's Health Services Research Center manuscript available in PMC from 2006 January 25

Study	Population	Indicator	Source	Time frame	Outcome	Source	Time frame	Correction for selection bias	Indicator-outcome relationship
1 et al., 2002	by screening 401 primary-care patients identified by AD prescription for new episode major depression	≥3 depression related follow-up visits with physician	Administrative database report	Not specified	Continuation of AD medication ^a	Patient report	75–105 days following AD initiation	Covariates	Significantly more patients who continued ADs had 3+ follow-up visits (37% versus 18%)
oenbaum et al., 2002	938 primary care patients with current major and minor depression identified by screening	Any AD use at therapeutic level OR 4+ specialty counseling visits	Patient report	6 months following index visit	Probable depression MCS	Patient report	6 months following index visit	IV analysis	Significantly increases remission (76% versus 30%) and improves MCS by 15.4 points

Note. IV = Instrument variable, MCS = Mental Component Scale of SF-12 or SF-36, AD = Antidepressant medication, PCS = Physical Component Scale of SF-36.

^aRelated to improvement in depressive symptoms at 3 months.