

Five-Year Impact of Quality Improvement for Depression

Results of a Group-Level Randomized Controlled Trial

Kenneth Wells, MD, MPH; Cathy Sherbourne, PhD; Michael Schoenbaum, PhD; Susan Ettner, PhD; Naihua Duan, PhD; Jeanne Miranda, PhD; Jürgen Unützer, MD, MPH; Lisa Rubenstein, MD, MSHS

Background: Quality improvement (QI) programs for depressed primary care patients can improve health outcomes for 6 to 28 months; effects for longer than 28 months are unknown.

Objective: To assess how QI for depression affects health outcomes, quality of care, and health outcome disparities at 57-month follow-up.

Design: A group-level randomized controlled trial.

Setting: Forty-six primary care practices in 6 managed care organizations.

Patients: Of 1356 primary care patients who screened positive for depression and enrolled in the trial, 991 (73%, including 451 Latinos and African Americans) completed 57-month telephone follow-up.

Interventions: Clinics were randomly assigned to usual care or to 1 of 2 QI programs supporting QI teams, provider training, nurse assessment, and patient education, plus resources to support medication management (QI-meds) or psychotherapy (QI-therapy) for 6 to 12 months.

Main Outcome Measures: Probable depressive dis-

order in the previous 6 months, mental health–related quality of life in the previous 30 days, primary care or mental health specialty visits, counseling or antidepressant medications in the previous 6 months, and unmet need, defined as depressed but not receiving appropriate care.

Results: Combined QI-meds and QI-therapy, relative to usual care, reduced the percentage of participants with probable disorder at 5 years by 6.6 percentage points ($P=.04$). QI-therapy improved health outcomes and reduced unmet need for appropriate care among Latinos and African Americans combined but provided few long-term benefits among whites, reducing outcome disparities related to usual care ($P=.04$ for QI-ethnicity interaction for probable depressive disorder).

Conclusions: Programs for QI for depressed primary care patients implemented by managed care practices can improve health outcomes 5 years after implementation and reduce health outcome disparities by markedly improving health outcomes and unmet need for appropriate care among Latinos and African Americans relative to whites; thus, equity was improved in the long run.

Arch Gen Psychiatry. 2004;61:378-386

DEPRESSIVE DISORDERS ARE a leading cause of disability worldwide.¹⁻³ Yet, many depressed patients, especially historically underserved minority groups, such as Latinos and African Americans, do not receive guideline-concordant care. Depressed primary care patients who are Latino or African American tend to have poorer health outcomes, without practice interventions to improve quality, than do whites.^{2,4-7} Quality improvement (QI) interventions for depression within primary care can improve quality of care and health outcomes over 6 to 28 months and yield cost-effectiveness ratios relative to usual care (UC) that are comparable to

commonly used medical therapies; however, not all interventions, particularly those focused only on providers, improve outcomes.⁸⁻¹⁵ Furthermore, interventions of longer duration may provide larger long-term benefits.¹⁶

Partners in Care (PIC) is a group-level randomized controlled trial of 2 QI interventions, compared with UC, for depressed primary care patients.¹⁷ Both interventions supported local teams for intervention management and provider education and nurses for patient education and assessment. One intervention trained nurses to support medication management by primary care providers (QI-meds), and the other trained local therapists in cognitive behavior therapy (CBT)

Author affiliations are listed at the end of this article.

and lowered the patient copay for use of those therapists (QI-therapy), but patients could have any or no treatment in either intervention. During 6 to 24 months of follow-up, relative to patients receiving UC, those in QI programs had a lower likelihood of probable depressive disorder, better quality of life, fewer “depression burden days,” and a higher rate of employment.¹⁷⁻¹⁹ In the first year, clinical outcome benefits were greatest for Latinos and African Americans combined, but employment gains were significant only among whites.²⁰ Health benefits continued into the second year under QI-therapy,¹⁹ and increased antidepressant medication use continued into the second year for QI-meds.²¹ The QI programs increased direct costs \$300 to \$400 for 24 months, but the interventions had 2-year cost-effectiveness ratios comparable to those of commonly used medical treatments.¹⁸

This study examines the effects of PIC interventions on health outcomes and quality of care 57 months after study enrollment. To our knowledge, no study has examined such long-term effects of short-term QI for any medical condition. Patients who participate in short-term QI could experience long-term benefits if they acquire new skills, such as learning to seek appropriate care when sick, that help them prevent or manage their depression. We hypothesized long-term health outcome benefits for the QI groups. Since improved health may reduce the need for services, we did not expect long-term QI effects on use of treatments or services overall; instead, we expected QI to reduce unmet need for appropriate care for depression. Nevertheless, practice and payer concerns about increased utilization of services in the long run may serve as a barrier to implementing effective QI programs, so we report QI effects on long-term probabilities of having different health care services as a context for our main results.

Based on our previous findings, we hypothesized that any average effects on health outcomes at 5 years would favor QI over UC and that there would be greater benefits of QI on outcomes in Latinos and African Americans combined than in whites.¹⁷⁻²¹ We did not expect any mean effects of QI on the use of services either for all participants or among Latinos and African Americans.²⁰ Low levels of utilization of, satisfaction with, and quality of care for mental disorders among Latino and African Americans have been the focus of national policy and research attention.^{6,20,22} Disparities in mental health care for other ethnic groups have been studied less often, and ethnic minorities other than Latinos and African Americans were represented in insufficient numbers in PIC for separate study.^{6,20}

METHODS

EXPERIMENTAL DESIGN AND IMPLEMENTATION

The data are from PIC, a group-level, randomized controlled trial of practice-implemented QI programs for depression.^{17-21,23} Six managed care organizations, 46 of 48 eligible primary care practices, and 181 of 183 eligible primary care clinicians participated. Within organizations, practices were matched into blocks of 3 clusters based on specialty mix, patient socioeco-

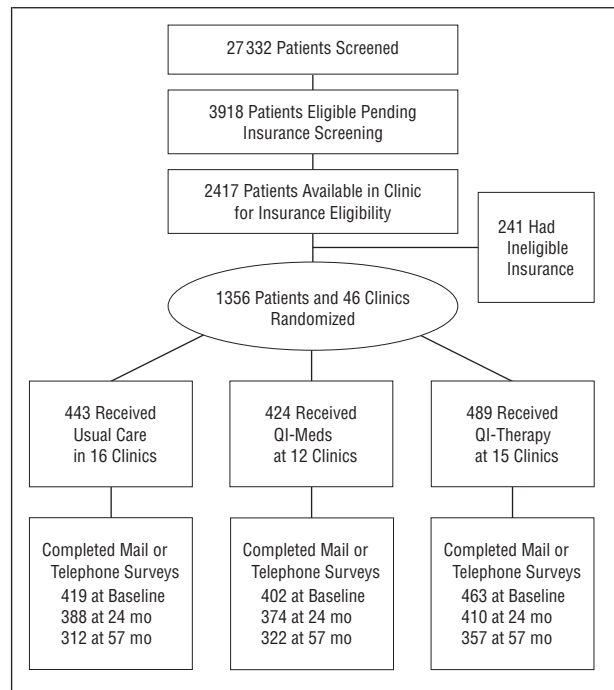


Figure 1. Partners in Care patient screening and enrollment. Forty-six clinics were randomized, but 43 had patients in this analysis. QI-meds indicates quality improvement plus medication management; QI-therapy, quality improvement plus psychotherapy.

nom and demographic factors, and having mental health specialists on-site. Practice clusters were randomized within blocks to UC (mailing of written practice guidelines to medical directors), QI-meds, or QI-therapy by using a random-number generator.

Study staff screened 27 332 consecutive patients in a 5- to 7-month period for each practice, between June 6, 1996, and February 28, 1997. Patients were eligible if they intended to use the practice for 12 months and if they had positive screening results for current depressive symptoms and probable major depressive or dysthymic disorder in the past year, using lead-in items from the World Health Organization’s 12-month *Composite International Diagnostic Interview (CIDI)*.²⁴ Patients were ineligible if they were younger than 18 years, if they were not fluent in English or Spanish, or if practice providers or therapists participating in the interventions were not covered by their insurance. The study was approved by the institutional review boards of RAND and the practices.

Of individuals completing the screening, 3918 were potentially eligible, but many left the clinic (eg, for a laboratory test) before their insurance status could be checked; 2417 patients were available for confirming insurance, and 241 (10%) had ineligible insurance. Of individuals reading informed consent, 1356 (70%) enrolled: 443 in UC, 424 in QI-meds, and 489 in QI-therapy (**Figure 1**). Enrollment did not differ by ethnicity ($\chi^2_3=3.5$; $P=.33$).

INTERVENTIONS

The interventions are published elsewhere²⁵; all intervention materials are posted at <http://www.rand.org/organization/health/pic.products/order.html>. We estimated each managed care organization’s participation costs based on expected enrollment and per case costs for health care and research time. We provided half that amount (\$35 000-\$70 000 per managed care organization).

The interventions provided practices with training and resources to initiate and monitor QI programs adapted to local goals and resources. Patients and clinicians retained choice of treatment and use of intervention materials; the randomization was to resources for improved care, not mandated treatment. The study provided training materials and staff training and offered limited implementation support.

For both interventions, local teams were trained in a 2-day workshop to educate primary care clinicians through lectures, detailing, or audit and feedback and to supervise staff and conduct team oversight. Practice nurses were trained to help with patient assessment, education, and activation for treatment. Practice teams were given patient education pamphlets, videotapes, tracking forms, clinician manuals, lecture slides, and pocket reminder cards. The materials described guideline-concordant care for depression; for example, they presented psychotherapy and antidepressant medication therapy as being equally effective for most patients, encouraged attention to patient preferences, and advised adjusting treatment plans to patient needs and preferences.^{26,27}

In QI-meds, nurse specialists were trained to support medication adherence through monthly visits or telephone contacts for 6 or 12 months, randomized at the patient level. In QI-therapy, practice therapists were trained to provide individual and group CBT,^{28,29} available at the primary care copay (approximately \$5-\$10) for 6 months after enrollment. All patients could have other therapy at their usual copay (approximately \$20-\$35). Supervision was provided by local experts, who were assisted by study experts. In all conditions, patients could have medications, therapy, both, or neither. Previously, for example, we showed that in the first and second 6 months of the study, 40% and 35% of QI-therapy patients, respectively, received some antidepressant medications and 38% and 34% received at least 4 psychotherapy sessions; 52% and 43% of QI-meds patients received some antidepressant medications and 30% and 29% received at least 4 psychotherapy sessions.^{21,30}

The study was advertised in each practice using posters featuring physicians and staff from that practice. The patient videotape featured minority physicians and patients. Study CBT was specifically developed for low-income minorities.^{28,29} The CBT training was provided in English and Spanish, and practices were asked to provide bilingual depression nurse specialists and therapists. Intervention staff received cultural sensitivity training, and physician training discussed issues in providing depression treatment to minorities.

DATA COLLECTION

Patients were asked to complete the screening instrument, a telephone *CIDI* for depression and economic status, and a mailed survey at baseline. We mailed follow-up surveys every 6 months for 24 months, with telephone follow-up, and we conducted a telephone survey at 24 months. At 57 months, we fielded a telephone follow-up. We would expect any survey mode effect (telephone vs self-administered via mail) to be equivalent

across intervention groups. Completion rates relative to all initial enrollees ($N=1356$) are 95% ($n=1286$) for the baseline telephone survey and 73% ($n=991$) for the 57-month survey, representing 86% of 1152 participants still living and enrolled.

MEASURES

Intervention Status

We used indicators for QI-meds and QI-therapy compared with UC.

Outcomes

We repeated the screening measure of probable major depressive disorder at follow-up, eliminating the dysthymic disorder item because it refers to the previous 2 years.¹⁷ The screening measure has a positive predictive value of 55% against the 12-month *CIDI*.⁹ Mental health-related quality of life was measured using the Mental Health Composite Score (MCS12), a global mental health scale scored from the Short-Form 12^{31,32}; scores were normalized to a population mean (SD) of 50 (10), with higher scores indicating better health.¹⁷⁻²⁰

Quality of Care

We used self-reports of 4 or more specialty visits as an indicator of counseling and of 2 months or more of antidepressant medication use during the 6 months before follow-up. At 6-month follow-up, having 1 or both types of treatments greatly improved health and employment outcomes.³³ At long-term follow-up, patients may no longer require treatments if their health has improved. To account for this, we use an indicator of "unmet need for appropriate care," defined as having probable major depressive disorder in the 6 months before follow-up but not receiving counseling or antidepressant medication, contrasted with persons without probable disorder or who received treatment.¹⁷

Services

As a context for our main results, we used survey items to develop counts of general medical visits for any purpose, general medical visits for emotional problems, specialty mental health visits, and overnight hospital stays for any purpose. We focus on dichotomous indicators of each type of use, but conclusions were similar using counts of visits and alternative transformations of counts.

Race or Ethnicity

For analyses of the main effect of intervention status on outcome, ethnicity is a covariate, and we contrast whites and all minorities grouped together by patient self-designation. For the analysis of QI effects on health disparities, we contrast Latinos and African Americans (combined) with non-Hispanic whites. Latinos and African Americans are 2 ethnic groups with strong national evidence of mental health care disparities.^{6,34,35} As a com-

bined group, Latinos and African Americans had a favorable clinical outcome of QI relative to whites at 1-year follow-up in PIC.²⁰ Combining these 2 groups affords greater statistical precision,^{36,37} but in sensitivity analyses, we compared QI effects for African Americans and Latinos relative to whites separately, with qualitatively similar conclusions as the combined comparisons. We base the main conclusions on the combined comparisons but provide estimates for African Americans and Latinos separately to facilitate interpretation.

Other ethnic groups in PIC included Asian Americans (n=17, 3 of whom self-designated themselves also as white), Native Americans (n=10), and whites/Native Americans (n=40). Qualitative descriptive findings for Native Americans and white/Native Americans combined suggest levels of treatment and outcomes similar to those for whites; for example, the weighted percentage with appropriate care was 48.2% at baseline and 45.2% at follow-up, and changes in mental health status from baseline to follow-up were similar across intervention groups (eg, for MCS12, increases of 7-9 scale points). Consistent with recent expert recommendations,^{6,38} we did not consider it appropriate to aggregate all ethnic minority groups into one category; instead, we focused on 2 ethnic minority groups with balanced representation across intervention groups, national data on mental health care disparities, and previous outcome data from PIC.^{20,34,38}

Covariates

Using the patient screening instrument, we measured age; sex; education (less than high school, completed high school, some college, completed college, or more); a count of having 0, 1, 2, or more than 2 of 19 chronic medical conditions; and the global mental and physical health scores on the Short-Form 12.^{31,32} We used data from the screening measure and the baseline CIDI^{17,24} to categorize patients as having 30-day depressive symptoms plus 1-year disorder (ie, "disorder") vs symptoms only. Using items modeled after the Health and Retirement Survey,³⁵ we developed a baseline household wealth variable. We used indicators for practice group.

DATA ANALYSIS

We conducted patient-level, intent-to-treat analyses, maintaining participants in the intervention group of the clinic in which they were screened for enrollment. For a group-level randomized trial with a moderate number of groups, control for baseline differences that may affect outcomes or that predict nonresponse is recommended.³⁶ To test the effects of QI, we estimated multivariate regression models with intervention status as the independent variable, controlling for ethnicity (white vs minority) and all the covariates listed in the previous subsection, which include factors predictive of nonresponse in this study (see the "Results" section). For dichotomous measures, we estimated logistic regression models. For MCS12, we conducted ordinary least squares regression on untransformed scores. To examine differences in QI effects for African Americans and Latinos combined vs whites, we conducted an additional set of analyses with

intervention status, ethnicity group, and intervention-ethnicity interactions as the main independent variables, excluding Asian Americans (n=17), Native Americans (n=10), and whites/Native Americans (n=40) (see the "Race or Ethnicity" subsection). We also conducted sensitivity analyses comparing whites with Latinos, and whites with African Americans, and stratifying the model interacting white and African Americans/Latinos combined by baseline disorder status (1-year disorder vs symptoms only). We adjusted for the clustering of patients within providers and clinics using a bias reduction method for standard error estimation developed by Bell and McCaffrey.³⁷

The significance of comparisons by intervention status and tests of interactions were based on regression coefficients. Consistent with recommendations for group-level randomized trials,³⁶ *df* are based on the number of practices (43 had patients in the follow-up). The presented results are standardized predictions generated from each regression model. We used the regression parameters and each individual's actual values for covariates other than intervention status to calculate predicted outcomes assuming the patient had been assigned to each intervention condition, in turn; we then calculated the mean prediction under each intervention condition. The analytic sample size for the main effect analysis is 991. For analysis of whites vs African Americans and Latinos, the total is 924. Specific models differ by a few individuals owing to missing data in the dependent variables. We used multiple imputations for missing items in independent variables; we averaged predictions from 5 randomly imputed data sets and adjusted standard errors for uncertainty due to imputation.^{39,40}

Nonresponse weighting is applied to the observed data, weighting each observation by the reciprocal of the probability of study enrollment and follow-up response, so as to facilitate generalization to the sample eligible for the study. For sensitivity analyses, we conducted unweighted, unadjusted analyses with no change in qualitative results or substantive conclusions.

We applied a 2-tailed significance level of $P \leq .05$, and we report actual *P* values. As a first study of such long-term outcomes, our analysis is somewhat exploratory; for most hypotheses, findings are consistent across measures, suggesting that formal correction for multiple statistical comparisons may be too conservative.

RESULTS

Table 1 indicates that the intervention groups had similar ethnic distributions at baseline among participants completing 57-month follow-up. However, baseline sex and global mental health (MCS12) differed by intervention status, with QI-therapy patients more likely to be female and to have worse global mental health; a similar health difference was reported for the baseline sample.^{17,18}

Additional factors related to nonresponse at the 57-month follow-up were ethnicity (African Americans were less likely to respond than whites, $t = 2.89$; $P = .006$), poorer physical and mental functioning, and education (high school graduates were less likely to respond than individuals with some college). Intervention status did not

Table 1. Baseline Characteristics of 991 Respondents Completing 57-Month Follow-up*

Characteristic	UC Group (n = 312)	QI-Meds Group (n = 322)	QI-Therapy Group (n = 357)	Difference Across Groups F _{2,41} (P Value)
Age, mean (SD), y	42.7 (14.1)	44.0 (14.2)	44.0 (14.1)	0.28 (.76)
Female, %	69.2	66.2	77.7	3.55 (.04)
Married, %	53.9	56.0	55.5	0.11 (.90)
Ethnicity, No. (%)				
White	191 (56.9)	210 (63.0)	200 (54.8)	0.27 (.76)
Latino	88 (30.2)	76 (24.4)	105 (30.8)	0.12 (.89)
African American	15 (6.6)	14 (6.4)	25 (7.0)	0.01 (.99)
Other	18 (6.3)	22 (6.2)	27 (7.3)	0.18 (.84)
MCS12 score, mean (SD)	36.9 (11.4)	35.6 (10.7)	34.6 (10.1)	2.84 (.07)
Chronic conditions, %				
0	20.6	21.9	22.8	0.20 (.82)
1	22.1	27.2	22.9	1.75 (.19)
2	22.3	18.8	20.8	0.31 (.73)
≥3	34.9	32.1	33.5	0.23 (.80)

Abbreviations: MCS12, Mental Health Composite Score; QI-meds, quality improvement plus medication management resources; QI-therapy, quality improvement plus psychotherapy resources; UC, usual care.

*Weighted data but actual number.

Table 2. Intervention Effects on 57-Month Outcomes and Unmet Need Overall (N = 991)*

	UC Group	QI-Meds Groups	QI-Therapy Group
Outcomes			
Probable disorder, % (95% CI)	43.6 (37.0-50.1)	37.9 (33.2-42.5)	36.2 (30.2-42.3)
<i>t</i> ₄₂ † (P value)	NA	1.80 (.08)	1.98 (.05)
MCS12 score, mean (95% CI)	42.6 (40.9-44.3)	43.9 (42.5-45.3)	44.3 (42.5-46.0)
<i>t</i> ₄₂ † (P value)	NA	1.27 (.21)	1.49 (.14)
Unmet Need			
Depressed and not in appropriate care, % (95% CI)	22.9 (17.9-27.8)	17.7 (12.4-22.9)	18.9 (13.6-24.3)
<i>t</i> ₄₂ † (P value)	NA	1.61 (.11)	1.22 (.23)

Abbreviations: CI, confidence interval; MCS12, Mental Health Composite Score; NA, not applicable; QI-meds, quality improvement plus medication management resources; QI-therapy, quality improvement plus psychotherapy resources; UC, usual care.

*Adjusted for baseline health status, sociodemographics, and randomization blocks; weighted analyses. Sample size varies by analysis between 989 and 991 owing to missing data.

†Comparing each intervention group with UC.

significantly predict nonresponse ($F_{2,41} = 1.76; P = .18$), and there was no statistically significant interaction between ethnicity (African American, Latino, other, white) and intervention status (combined QI vs UC) in follow-up response ($F_{3,40} = 1.52; P = .22$, unweighted analysis).

MEAN QI EFFECTS FOR ALL PARTICIPANTS

Relative to UC, QI-therapy lowered the adjusted rate of probable depressive disorder at follow-up by 7.4 percentage points (95% confidence interval [CI], 0.1 to 14.6 percentage points) and QI-meds lowered the rate by 5.7 percentage points (95% CI, -0.5 to 11.9 percentage points), with QI-therapy significant at $P = .05$ (**Table 2**). Pooling QI-meds and QI-therapy vs UC yields a 6.6–percentage point difference (95% CI, 0.4 to 12.8 percentage points, $t_{42} = 2.08; P = .04$). As seen in Table 2, the effect of each QI intervention on the MCS12 score is small and nonsignificant. The reductions in unmet need were also not statistically significant.

EFFECTS OF QI FOR AFRICAN AMERICANS AND LATINOS VS WHITES

Table 3 indicates that under UC, African Americans and Latinos combined have worse health outcomes than whites. Participation in QI-therapy lowered the adjusted rate of probable disorder among African Americans and Latinos but had little effect among whites; the interaction is significant at $P = .04$. Results are qualitatively similar for QI-meds but are not statistically significant; pooling the 2 QI groups vs UC yields a 16.1–percentage point reduction (95% CI, 5.2 to 27.0 percentage points, $t_{42} = 2.74; P = .009$) in probable disorder among African Americans and Latinos. Participation in QI-therapy improved global mental health (MCS12) among African Americans and Latinos by 4.5 scale points (95% CI, 0.5 to 8.5 points), a moderate to large effect.³² There is no effect of QI-therapy on MCS12 among whites; the interaction term is significant at $P = .02$.

Table 3. Intervention Effects on 57-Month Outcomes and Unmet Need Overall (N = 924)*

	UC Group	QI-Meds Groups	QI-Therapy Group
Outcomes			
Probable disorder, % (95% CI)			
Latino/African American combined	55.8 (46.7-64.9)	45.4 (34.4-56.5)	35.6 (23.1-48.0)
t_{42} (P value)	NA	1.55 (.13)	2.67 (.01)
White	36.0 (28.6-43.4)	31.9 (26.0-37.8)	34.4 (27.0-41.8)
t_{42} (P value)	NA	0.97 (.34)	0.33 (.74)
QI-ethnicity interaction, t_{42} (P value)	NA	0.67 (.50)	2.10 (.04)
Latino	57.1 (46.8-67.4)	48.3 (34.8-61.7)	42.5 (27.0-58.1)
African American	51.1 (27.2-75.1)	32.6 (18.7-46.5)	4.0 (0.0-12.0)
MCS12 score, mean (95% CI)			
Latino/African American combined	40.0 (37.2-42.8)	41.6 (39.5-43.8)	44.5 (41.6-47.5)
t_{42} † (P value)	NA	0.94 (.35)	2.23 (.03)
White	44.5 (42.9-46.1)	45.4 (43.5-47.3)	44.6 (42.9-46.3)
t_{42} (P value)	NA	0.77 (.45)	0.10 (.92)
QI-ethnicity interaction, t_{42} (P value)	NA	0.38 (.71)	2.38 (.02)
Latino	40.0 (37.1-42.9)	40.8 (38.0-43.6)	43.8 (40.2-47.4)
African American	39.8 (36.2-43.3)	45.3 (40.2-50.5)	47.7 (44.8-50.7)
Unmet Need			
Depressed and not in appropriate care, % (95% CI)			
Latino/African American combined	34.3 (27.2-41.5)	24.6 (16.5-32.7)	21.7 (13.1-30.3)
t_{42} † (P value)	NA	1.87 (.07)	2.27 (.03)
White	14.7 (8.3-21.1)	12.8 (5.3-20.4)	16.4 (10.6-22.1)
t_{42} (P value)	NA	0.38 (.70)	0.39 (.70)
QI-ethnicity interaction, t_{42} (P value)	NA	0.67 (.51)	1.83 (.07)
Latino	34.6 (26.7-42.5)	26.7 (15.7-35.8)	26.0 (15.6-36.5)
African American	33.7 (6.7-60.7)	18.9 (6.6-31.2)	3.1 (0-9.8)

Abbreviations: CI, confidence interval; MCS12, Mental Health Composite Score; NA, not applicable; QI-meds, quality improvement plus medication management resources; QI-therapy, quality improvement plus psychotherapy resources; UC, usual care.

*Adjusted for baseline health status, sociodemographics, and randomization blocks; weighted analyses. Analysis excludes individuals with "other" ethnicity. Sample size varies between 921 and 924 owing to missing data. Inference is based on white vs African American and Latino combined, but separate results for each ethnic group are provided for descriptive purposes.

The QI-therapy intervention reduced unmet need for appropriate care among African Americans and Latinos, an effect comparable to the pooled QI effect during the active intervention period.¹⁷ Use of QI-meds also reduced unmet need for appropriate care among African Americans and Latinos, but with weaker statistical significance ($P = .07$); pooling intervention groups yields a reduction of 11.4 percentage points (95% CI, 3.1 to 19.7 percentage points, $t_{42} = 2.65$; $P = .01$), relative to UC. No noteworthy QI effect on unmet need for appropriate care occurred among whites, but interaction terms are not statistically significant.

Table 3 illustrates that results were qualitatively similar for Latinos vs whites and for African Americans vs whites, but intervention effects were especially strong among African Americans. For example, the effect of QI-therapy on probable disorder is significant for this group relative to whites ($t = 3.07$; $P = .004$).

Sensitivity analyses stratified by baseline disorder status suggest qualitatively similar effects of pooled QI on reducing probable disorder for the combined minority groups, regardless of baseline disorder status. The trend for more favorable QI effects among minorities, relative to whites, holds among those with baseline disorder ($n = 565$, $t = 1.66$, $P = .10$) but not those with symptoms only ($n = 360$, $t = 0.15$, but $P = .88$). Similarly, the greater effect of the interventions on reducing unmet need among minorities, relative to whites, applies primarily among

those with baseline disorder ($t = 3.25$, $P = .002$), not symptoms only ($t = 0.76$, $P = .45$).

As expected, there were no noteworthy or statistically significant QI effects on the probability of having any type of visit or hospitalization or on counts of visits or hospital stays for the whole sample or for whites or African Americans and Latinos combined.

COMMENT

We found a modest mean reduction across the whole sample in the likelihood of having probable depressive disorder at follow-up for pooled QI interventions compared with UC, equivalent to a 3–percentage point reduction in the prevalence of depressive disorder.¹⁷ Even such a modest reduction in continued depression could have important public health implications given the high prevalence and effect on morbidity of depression. Our findings build on promising results from other studies^{15,16} showing substantial health outcome improvements from QI for depression in primary care at 24 to 28 months.

This result, however, was mainly due to a very large effect of QI, especially QI-therapy, on health outcomes among Latinos and African Americans at 5-year follow-up. Without intervention, 56% of minorities and 36% of whites had probable depressive disorder at follow-up, indicating a large outcome disparity. With QI-therapy, how-

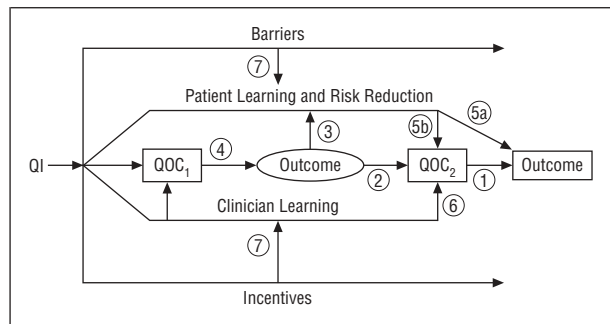


Figure 2. Seven alternative pathways for long-term quality improvement (QI) effects. QOC indicates quality of care. See the "Comment" section for a description of each pathway.

ever, outcomes were similar for whites and for the combined African American and Latino group. The effect of QI-therapy on the global mental health of African Americans and Latinos, almost 0.5 SD, is within a range considered clinically meaningful on the Short-Form 12.³² We also found that QI-therapy reduced unmet need for appropriate care among African Americans and Latinos at 57 months, which might lead to subsequent health outcome improvements, a subject for future study. Because the minority individuals participating in PIC preferred psychotherapy for treatment of depression at baseline,⁴¹ QI-therapy may have been a particularly culturally appropriate intervention. These findings are relevant to national debates over how to reduce disparities in the context of few evidence-based approaches.^{6,22}

Miranda et al²⁰ observed a similar differential QI effect during the first year of follow-up and proposed that this finding may have been due to the lower previous exposure of minorities to appropriate depression treatment. The observation of this effect at 5-year follow-up, however, suggests that the interventions may have also addressed fundamental causes of health care disparities. Factors cited in the literature as contributing to health care disparities²² include (1) cultural discordance of provider and patient,^{42,43} which the cultural sensitivity training in PIC may have attenuated somewhat; (2) provider uncertainty about diagnosis or treatment benefit, enhancing reliance on previous conceptions,⁴² which PIC education may have reduced; (3) language or other incompatibilities between patients and providers,^{44,45} which the use of bilingual intervention staff [ie, depression nurse specialists and therapists] reduced; (4) health care system problems in settings frequented by minorities, such as poor services availability^{46,47} and difficulty accessing specialty referrals,⁴⁸ which PIC resources may have reduced; (5) interpersonal communication and participation style factors,^{42,49} which PIC training in patient preference may have improved; (6) discrimination and racism,^{50,51} which PIC clinician training acknowledged but did not directly address; (7) overuse of services by whites,^{52,53} which PIC indirectly addressed through advocating tailoring treatment to need; (8) insurance coverage differences,⁵⁴⁻⁵⁶ which PIC partially addressed through reducing the CBT copay; (9) differences in having a regular provider,⁵⁶ which PIC did not address; and (10) differences in clinician diagnosis and treatment,⁵⁷⁻⁵⁹ which PIC standardized through clinician train-

ing. Causes of disparities could be mitigated by features of the underlying collaborative care approach, such as enhancing leadership commitment to support appropriate care.^{60,61} Support for this view is provided by recent studies showing the effectiveness of a stepped-care program for major depression among impoverished women in Chile⁶² and of group interpersonal psychotherapy for depression in rural Uganda.⁶³

The fundamental question posed by our findings is why there were long-term health benefits of QI at all.

Figure 2 illustrates 7 alternative pathways for long-term (time 2) gains after initial (time 1) benefits. The first pathway is a main QI effect on the use of treatments at long-term follow-up, that is, quality of care; we found no evidence of such main effects. The second pathway is a targeted long-term QI effect on reducing unmet need for appropriate care. Our findings support this explanation among African Americans and Latinos. The third pathway is an indirect effect through QI reducing risk factors for depression, such as the number of episodes. The PIC interventions reduced the likelihood of probable depressive disorder in the first year, particularly among African Americans and Latinos.²⁰ The fourth pathway is an early (first 2 years) QI effect on increasing maintenance treatment, which would reduce future recurrences. Unützer and colleagues²¹ found increased use of antidepressant medications through the second year under QI-meds, providing partial support. Pathways 5a and 5b (Figure 2) reflect indirect QI effects on outcomes through patient learning (eg, symptom self-management) as a result of education or experiences with care. We cannot test these pathways with our data. Pathway 6 is a long-term effect on clinician learning. Although QI increased clinician knowledge at 18 months,⁶⁴ most patients switched providers and most practices switched management by 2 years, reducing the plausibility of clinician-based explanations. The seventh pathway is a persistent effect of QI on barriers/incentives, such as reduced long-term financial barriers to care. We do not have the data to test this pathway, and we have no reason to think that QI would have had such persistent structural effects. Thus, PIC findings support several explanations for long-term QI effects, especially among African Americans and Latinos. Further exploration of such pathways in future studies may help design interventions that enhance long-term benefits.

Our findings are subject to limitations. Measures are based on self-report. We recruited patients from specific community-based managed care practices. Underserved minorities, especially Latinos, are more likely than whites to be uninsured³⁴; our findings apply mostly to insured patients, but one public sector site had uninsured patients. We had relatively few African Americans, who had lower response at long-term follow-up, but intervention groups were balanced in ethnic distribution, and there was no statistically significant interaction between ethnicity and intervention status in the probability of response. Our conclusions are largely based on pooled analyses of Latinos and African Americans; however, separate analyses show large effects of these QI programs in African Americans. This exploratory find-

ing should be reexamined in larger representative groups of African Americans. The effectiveness of QI should also be examined in ethnic groups we did not specifically study, such as Asian Americans, Pacific Islanders, and Native Americans.

We conclude that implementing QI programs for depressed primary care patients can lead to long-term improvements in health and reduced unmet need for appropriate care among Latinos and African Americans, reducing substantial health outcome disparities in UC for these groups relative to whites. The PIC therapy-resource intervention seemed particularly effective in this regard. In this study, gains for minority groups led to an overall modest health improvement across participants, without noticeable long-term increases in the use of services. In the long run, QI improved equity, a major goal of medicine,⁶⁵ after being effective and cost-effective in the short run.¹⁸

Submitted for publication May 19, 2003; final revision received September 4, 2003; accepted September 23, 2003.

From the Health Program, RAND, Santa Monica, Calif (Drs Wells, Sherbourne, Schoenbaum, and Rubenstein); the Neuropsychiatric Institute (Drs Wells, Duan, Miranda, and Unützer), the Departments of Psychiatry and Biobehavioral Services (Drs Wells, Duan, Miranda, and Unützer) and Biostatistics (Dr Duan), the Health Services Research Center (Drs Wells, Duan, Miranda, and Unützer), and the Center for Community Health (Dr Duan), University of California at Los Angeles; University of California at Los Angeles Department of Medicine, Division of General Internal Medicine and Health Services Research (Drs Ettner and Rubenstein); and the Department of Medicine, VA Greater Los Angeles Healthcare System, Sepulveda, Calif (Dr Rubenstein).

This work was funded by grants R10MH57992 and P50MH54623 from the National Institute of Mental Health, Bethesda, Md, and by grant R01-HS08349 from the Agency for Healthcare Research and Quality, Rockville, Md.

We thank Maureen Carney, MS, for coordination of follow-up; Barbara Levitan, BA, for oversight of the survey; Bernadette Benjamin, MS, for excellent programming support; the practice organizations participating in this study for providing access to their expertise and patients, implementing interventions, and providing in-kind resources: Allina Medical Group (Twin Cities, Minn), Patuxent Medical Group (Maryland), Humana Health Care Plans (San Antonio, Tex), MedPartners (Los Angeles), PacifiCare of Texas (San Antonio), and Valley-Wide Health Services (Colorado); their associated behavioral health organizations: Alamo Mental Health Group (San Antonio), San Luis Valley Mental Health/Colorado Health Networks, and Magellan/GreenSpring Behavioral Health (Maryland); and the clinicians and patients who contributed their time and efforts to this study.

Corresponding author: Kenneth Wells, MD, MPH, RAND, 1700 Main St, Santa Monica, CA 90407-2138 (e-mail: kwells@ucla.edu).

REFERENCES

- Hays RD, Wells KB, Sherbourne CD, Rogers W, Spritzer K. Functioning and well-being outcomes of patients with depression compared with chronic medical illnesses. *Arch Gen Psychiatry*. 1995;52:11-19.
- Wells KB, Sturm R, Sherbourne CD, Meredith LS. *Caring for Depression*. Cambridge, Mass: Harvard University Press; 1996.
- 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, Mass: Harvard School of Public Health, on behalf of the World Health Organization and the World Bank; 1996.
- Young AS, Klap R, Sherbourne DC, Wells KB. The quality of care for depressive and anxiety disorders in the United States. *Arch Gen Psychiatry*. 2001;58:55-61.
- Wang PS, Berglund P, Kessler RC. Recent care for common mental disorders in the United States: prevalence and conformance with evidence-based recommendations. *J Gen Intern Med*. 2000;15:284-292.
- US Department of Health and Human Services. *Mental Health: Culture, Race, and Ethnicity: A Supplement to "Mental Health: A Report of the Surgeon General."* Rockville, Md: US Dept of Health and Human Services, Public Health Services, Office of the Surgeon General; 2001.
- Thompson C, Kinmonth AL, Stevens L, Peveler RC, Stevens A, Ostler KJ, Pickering RM, Baker NG, Henson A, Preece J, Cooper D, Campbell MJ. Effects of a clinical-practice guideline and practice-based education on detection and outcome of depression in primary care: Hampshire Depression Project randomized controlled trial. *Lancet*. 2000;355:185-191.
- King M, Davidson O, Taylor F, Haines A, Sharp D, Turner R. Effectiveness of teaching general practitioners skills in brief cognitive behavioral therapy to treat patients with depression: randomized controlled trial. *BMJ*. 2002;324:947-950.
- 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.
- 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. *Arch Gen Psychiatry*. 1996;53:924-932.
- Simon GE, Manning WG, Katzelnick DJ, Pearson SD, Henk HJ, Helstad CS. Cost-effectiveness of systematic depression treatment for high utilizers of general medical care. *Arch Gen Psychiatry*. 2001;58:181-187.
- Hunkeler EM, Meresman J, 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. *Arch Fam Med*. 2000;9:700-708.
- 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.
- Von Korff M, Katon W, Bush T, Lin EH, Simon GE, Saunders K, Ludman E, Walker E, Unützer J. Treatment costs, cost offset, and cost-effectiveness of collaborative management of depression. *Psychosom Med*. 1998;60:143-149.
- Katon W, Russo J, Von Korff M, Lin E, Simon G, Bush T, Ludman E, Walker E. 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 JL, Elliott CE, Dickinson M. Managing depression as a chronic disease: a randomized trial of ongoing treatment in primary care. *BMJ*. 2002;325:934-939.
- Wells KB, Sherbourne C, Schoenbaum M, Duan N, Meredith L, Unützer J, Miranda J, Carney MF, Rubenstein LV. Impact of disseminating quality improvement for depression in managed primary care: a randomized controlled trial. *JAMA*. 2000;283:212-220.
- Schoenbaum M, Unützer J, Sherbourne C, Duan N, Rubenstein LV, Miranda J, Carney MF. Cost-effectiveness of practice-initiated quality improvement for depression: results of a randomized controlled trial. *JAMA*. 2001;286:1325-1330.
- Sherbourne CD, Wells KB, Duan N, Miranda J, Unützer J, Jaycox L, Schoenbaum M, Meredith LS, Rubenstein LV. Long-term effectiveness of disseminating quality improvement for depression in primary care. *Arch Gen Psychiatry*. 2001;58:696-703.
- Miranda J, Duan N, Sherbourne CD, Schoenbaum M, Lagomasino I, Jackson-Triche M, Wells KB. Improving care for minorities: can quality improvement interventions improve care and outcomes for depressed minorities? results of a randomized, controlled trial. *Health Serv Res*. 2003;38:613-630.
- Unützer J, Rubenstein L, Katon W, Tang L, Duan N, Lagomasino IT, Wells KB. Two-year effects of quality improvement programs on medication management for depression. *Arch Gen Psychiatry*. 2001;58:935-942.
- Smedley BD, Stith AY, Nelson AR, eds. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington, DC: National Academies Press; 2003.
- Wells KB. The design of Partners in Care: evaluating the cost-effectiveness of improving care for depression in primary care. *Soc Psychiatry Psychiatr Epidemiol*. 1999;34:20-29.

24. World Health Organization. *Composite International Diagnostic Interview (CIDI), Version 2.1*. Geneva, Switzerland: World Health Organization; 1995.
25. Rubenstein, LV, Jackson-Triche M, Unützer J, Miranda J, Minnium K, Pearson ML, Wells KB. Evidence-based care for depression in managed primary care practices. *Health Aff (Millwood)*. 1999;18:89-105.
26. Depression Guidelines Panel. *Depression in Primary Care, I: Detection and Diagnosis*. Rockville, Md: US Dept of Health and Human Services, US Public Health Service; 1993. Agency for Health Care Policy & Research publication 93-0550.
27. Depression Guidelines Panel. *Depression in Primary Care, II: Treatment of Major Depression*. Rockville, Md: US Dept of Health and Human Services, US Public Health Service; 1993. Agency for Health Care Policy & Research publication 93-0551.
28. Muñoz RJ, Miranda J. *Group Therapy for Cognitive Behavioral Treatment of Depression, San Francisco General Hospital Depression Clinic, 1986*. Santa Monica, Calif: RAND; 2000. Document MR01198/4.
29. Muñoz RF, Aguilar-Gaxiola S, Guzmán J. *Manual de Terapia de Grupo para el Tratamiento Cognitivo-conductual de Depresión, Hospital General de San Francisco, Clínica de Depresión, 1986*. Santa Monica, Calif: RAND; 2000. Document MR-1198/5.
30. Jaycox LH, Miranda J, Meredith LS, Duan N, Benjamin B, Wells K. Impact of primary care quality improvement intervention on use of psychotherapy for depression. *Ment Health Serv Res*. 2003;5:109-120.
31. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36), I: conceptual framework and item selection. *Med Care*. 1992;30:473-483.
32. Ware JE Jr, Kosinski M, Keller SD. *SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales*. Boston, Mass: Health Institute, New England Medical Center; 1995.
33. Schoenbaum M, Unützer J, McCaggrey D, Duan N, Sherbourne C, Wells KB. The effects of primary care depression treatment on patients' clinical status and employment. *Health Serv Res*. 2003;37:1145-1158.
34. Wells KB, Klap R, Sherbourne CD, Koike A. Ethnic disparities in care for alcohol, drug abuse, and mental health problems. *Am J Psychiatry*. 2001;158:2027-2032.
35. Smith JP. Racial and ethnic differences in wealth in the Health and Retirement Survey. *J Hum Resources*. 1995;30(suppl):S158-S183.
36. Murray DM. *Design and Analysis of Group-Randomized Trials*. Oxford, England: Oxford University Press; 1998.
37. Bell RM, McCaffrey DF. *Bias Reduction in Standard Errors for Linear Regression With Multi-Stage Samples*. Florham Park, NJ: AT&T Labs-Research; 2002. Publication TD-4S9H9T.
38. Committee on Communication for Behavior Change in the 21st Century: Improving the Health of Diverse Populations, Board on Neuroscience and Behavioral Health. *Speaking of Health: Assessing Health Communication Strategies for Diverse Populations*. Washington, DC: National Academies Press; 2002.
39. Little RJA. Pattern-mixture models for multivariate incomplete data. *J Am Stat Assoc*. 1993;88:125-134.
40. Schafer JL. *Analysis of Incomplete Multivariate Data*. London, England: Chapman & Hall; 1997.
41. Dwight-Johnson M, Sherbourne CD, Liao D, Wells KB. Treatment preferences among depressed primary care patients. *J Gen Intern Med*. 2000;15:527-534.
42. Cooper-Patrick L, Gallo JJ, Gonzales JJ, Vu HT, Powe NR, Nelson C, Ford DE. Race, gender and partnership in the patient-physician relationship. *JAMA*. 1999;282:583-589.
43. Saha S, Komaromy M, Koepsell TD, Bindman AB. Patient-physician racial concordance and the perceived quality and use of health care. *Arch Intern Med*. 1999;159:997-1004.
44. Baker DW, Stevens CD, Brook RH. Determinants of emergency department use: are race and ethnicity important? *Ann Emerg Med*. 1996;28:677-682.
45. David RA, Rhee M. The impact of language as a barrier to effective health care in an underserved urban Hispanic community. *Mt Sinai J Med*. 1998;65:393-397.
46. Leape LL, Hilborne LH, Bell R, Kamberg C, Brook RH. Underuse of cardiac procedures: do women, ethnic minorities, and the uninsured fail to receive needed revascularization? *Ann Intern Med*. 1999;130:183-192.
47. Kahn KL, Pearson ML, Harrison ER, Desmond KA, Rogers WH, Rubenstein LV, Brook RH, Keeler EB. Health care for black and poor hospitalized Medicare patients. *JAMA*. 1994;271:1169-1174.
48. Hargraves JL, Stoddard JJ, Trude S. Minority physicians' experiences obtaining referrals to specialists and hospital admissions. *Medscape General Medicine* [serial online]. August 9, 2001. Available at: http://www.medscape.com/viewarticle/408160_print. Accessed July 29, 2003.
49. Cooper LA, Roter DL. Patient-provider communication: the effect of race and ethnicity on process and outcomes of healthcare. In: Smedley BD, Stith AY, Nelson AR, eds. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington, DC: National Academies Press; 2003.
50. La Veist TA, Nickerson KJ, Bowie JV. Attitudes about racism, medical mistrust and satisfaction with care among African American and white cardiac patients. *Med Care Res Rev*. 2000;57:146-161.
51. Lillie-Blanton M, Marinez RM, Salganicoff A. Site of medical care: do racial and ethnic differences persist? *Yale J Health Policy Law Ethics*. 2001;1:1-17.
52. Hannan EL, van Ryn M, Burke J, Stone D, Kumar D, Arani D, Pierce W, Rafii S, Sanborn TA, Sharma S, Slater J, DeBuono BA. Access to coronary artery bypass surgery by race/ethnicity and gender among patients who are appropriate for surgery. *Med Care*. 1999;37:68-77.
53. Laourie M, Kravitz RL, French WJ, Yang I, Milliken JC, Hilborne L, Wachsner R, Brook RH. Underuse of coronary revascularization procedures: application of clinical method. *J Am Coll Cardiol*. 1997;29:891-897.
54. Lowe RA, Chaya S, Nasci K, Gavin LJ, Shaw K, Zwanger ML, Zeccardi JA, Dalsey WC, Abuhl SB, Feldman H, Berlin JA. Effect of ethnicity on denial of authorization for emergency department care by managed care gatekeepers. *Acad Emerg Med*. 2001;8:259-266.
55. Tai-Seale M, Freund D, LoSasso A. Racial disparities in service use among Medicaid beneficiaries after mandatory enrollment in managed care: a difference in differences approach. *Inquiry*. 2001;38:49-59.
56. Lillie-Blanton M, Marinez RM, Salganicoff A. Site of medical care: do racial and ethnic differences persist? *Yale J Health Policy Law Ethics*. 2001;1:1-17.
57. Abreu JM. Conscious and nonconscious African American stereotypes: impact on first impression and diagnostic rating by therapists. *J Consult Clin Psychol*. 1999;67:387-393.
58. Rathore SS, Lenert LA, Weinfurt KP, Tinoco A, Taleghani CK, Harless W, Schulman KA. The effects of patient sex and race on medical students' rating of quality of life. *Am J Med*. 2000;108:561-566.
59. van Ryn M, Burke J. The effect of patient race and socio-economic status on physicians' perceptions of patients. *Soc Sci Med*. 2000;50:813-828.
60. Von Korff M, Gruman J, Schaefer J, Curry SJ, Wagner EH. Collaborative management of chronic illness. *Ann Intern Med*. 1997;127:1097-1102.
61. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *Milbank Q*. 1996;74:511-544.
62. Araya R, Rojas G, Fritsch R, Gaete J, Shaw K, Simon G, Peters TJ. Treating depression in primary care in low-income women in Santiago, Chile: a randomized controlled trial. *Lancet*. 2003;361:995-1000.
63. Bolton P, Bass J, Neugebauer R, Verdell H, Clougherty KF, Wickramaratne P, Spellman L, Ndogoni L, Weissman M. Group interpersonal psychotherapy for depression in rural Uganda: a randomized trial. *JAMA*. 2003;289:3117-3124.
64. Meredith LS, Jackson-Triche M, Duan N, Rubenstein LV, Camp P, Wells KB. Quality improvement for depression enhances long-term treatment knowledge for primary care clinicians. *J Gen Intern Med*. 2000;15:868-877.
65. Committee on Quality of Health Care in America, Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press; 2001.