

Telephone Screening, Outreach, and Care Management for Depressed Workers and Impact on Clinical and Work Productivity Outcomes

A Randomized Controlled Trial

Philip S. Wang, MD, DrPH

Gregory E. Simon, MD, MPH

Jerry Avorn, MD

Francisca Azocar, PhD

Evette J. Ludman, PhD

Joyce McCulloch, MS

Maria Z. Petukhova, PhD

Ronald C. Kessler, PhD

DEPRESSION IMPOSES ENORMOUS societal burdens,¹⁻³ with annual US economic costs of tens of billions of dollars due largely to productivity losses.^{4,5} Indeed, comparative cost-of-illness studies show that depression is among the most costly of all health problems to employers.⁶⁻¹⁰ Despite evidence that guideline-concordant treatment can be effective,¹¹⁻²¹ many depressed workers are untreated or inadequately treated.²²⁻²⁶ Effectiveness trials have demonstrated that organized depression screening and enhanced-care programs can significantly improve treatment and clinical outcomes.²⁷⁻⁴¹ Based on the magnitude of depression-related lost productivity, one might expect employer-purchasers (ie, those who purchase corporate health benefits) to invest in enhanced depression screening-treatment programs. However, widespread uptake has not occurred^{42,43} due in part to employer-purchasers being unsure of the return-on-investment of such programs.^{44,45} Few controlled trials have evaluated effects of

For editorial comment see p 1451.

Context Although guideline-concordant depression treatment is clearly effective, treatment often falls short of evidence-based recommendations. Organized depression care programs significantly improve treatment quality, but employer purchasers have been slow to adopt these programs based on lack of evidence for cost-effectiveness from their perspective.

Objective To evaluate the effects of a depression outreach-treatment program on workplace outcomes, a concern to employers.

Design, Setting, and Participants A randomized controlled trial involving 604 employees covered by a managed behavioral health plan were identified in a 2-stage screening process as having significant depression. Patient treatment allocation was concealed and assessment of depression severity and work performance at months 6 and 12 was blinded. Employees with lifetime bipolar disorder, substance disorder, recent mental health specialty care, or suicidality were excluded.

Intervention A telephonic outreach and care management program encouraged workers to enter outpatient treatment (psychotherapy and/or antidepressant medication), monitored treatment quality continuity, and attempted to improve treatment by giving recommendations to providers. Participants reluctant to enter treatment were offered a structured telephone cognitive behavioral psychotherapy.

Main Outcome Measures Depression severity (Quick Inventory of Depressive Symptomatology, QIDS) and work performance (World Health Organization Health and Productivity Questionnaire [HPQ], a validated self-report instrument assessing job retention, time missed from work, work performance, and critical workplace incidents).

Results Combining data across 6- and 12-month assessments, the intervention group had significantly lower QIDS self-report scores (relative odds of recovery, 1.4; 95% confidence interval, 1.1-2.0; $P=.009$), significantly higher job retention (relative odds, 1.7; 95% confidence interval, 1.1-3.3; $P=.02$), and significantly more hours worked among the intervention ($\beta=2.0$; $P=.02$; equivalent to an annualized effect of 2 weeks of work) than the usual care groups that were employed.

Conclusions A systematic program to identify depression and promote effective treatment significantly improves not only clinical outcomes but also workplace outcomes. The financial value of the latter to employers in terms of recovered hiring, training, and salary costs suggests that many employers would experience a positive return on investment from outreach and enhanced treatment of depressed workers.

Trial Registration clinicaltrials.gov Identifier: NCT00057590

JAMA. 2007;298(12):1401-1411

www.jama.com

such programs on work outcomes and those few focused on primary care samples rather than on the workplace samples that would be the focus of em-

Author Affiliations are listed at the end of this article. **Corresponding Author:** Philip Wang, MD, DrPH, Division of Services and Intervention Research, National Institute of Mental Health, 6001 Executive Blvd, Room 7141, MSC 9629, Bethesda, MD 20892-9629 (wangphi@mail.nih.gov).

ployer-based screening, outreach, and disease management efforts.^{36,41} Without such information, employer-purchasers are likely to remain hesitant to invest in enhanced depression care. In this report, we present the results of a randomized controlled trial designed explicitly to address this issue. The trial examined the impact of depression screening, vigorous outreach, and care management of depressed workers employed by a number of large national firms. The primary outcomes included not only depression symptom relief but also job retention, decreased sickness absence, and increased work productivity.

METHODS

Sample

Participants included 604 depressed workers 18 years and older enrolled in United Behavioral Health (UBH), a large managed behavioral health care company. Participating employers differed in whether UBH was the only behavioral health plan available, coverage criteria, the extent to which workers paid insurance premiums from salaries, and benefit design (co-payment size, visit caps, etc).

Recruitment

Recruitment occurred between January 2004 and February 2005 using a 2-phase procedure. Phase 1 included a health risk appraisal (HRA) survey conducted in 16 large companies from diverse sectors (airline, insurance, banking, public utility, state government, manufacturing) and containing broad distributions of occupations. An informed consent script was followed by a chronic-conditions checklist,⁴⁶ screen for psychological distress (K-6),^{47,48} questions about occupation and work performance, and sociodemographics.

Employees whose screen results were positive for possible depression (K-6 score ≥ 9) were invited by an introductory letter and telephone call from a survey interviewer to participate in a second-phase telephone interview that assessed depression more specifically using the Quick Inventory of Depres-

sion Symptoms Self-Report (QIDS-SR) version.⁴⁹⁻⁵¹ An initial informed consent script described the study and its voluntary and confidential nature; respondents were told they might be invited to participate in an innovative treatment program but there were no requirements to accept this or any specific treatments. Respondents with at least moderate depression severity (QIDS-SR score ≥ 8) were eligible for randomization. Exclusion criteria included: positive responses to Composite International Diagnostic Interview short form⁵² screening questions for a history of mania or substance dependence (such persons were informed and told to follow-up with a clinician), suicidal ideation or attempts in the prior week (those respondents were immediately connected by telephone to a UBH crisis counselor), and treatment by a mental health specialist in the prior year.

Phase 1 individual-level HRA invitations were sent to 113 843 workers (mainly by e-mail) in 12 of the 16 companies. Of those invited, 35 169 completed at least 1 HRA question, 2358 (7.7%) of respondents had screen results positive for depression, and 1422 consented to baseline eligibility assessments and had UBH coverage (all company employees were screened, not just those with UBH coverage).

In the 4 remaining companies (approximately 150 000 workers) where individual-level contact information was not available, mass e-mail invitations were sent to the entire workforce and individuals provided contact information only in the initial HRA; 11 715 responded by completing at least one HRA question, 942 (10.9%) of respondents had screen results positive for depression, of whom 331 also had UBH coverage and consented to baseline eligibility assessments. A separate enrollment procedure used the UBH member wellness survey of behavioral health symptoms and functional impairments mailed to members who received an authorization for outpatient treatment.^{53,54} Of 114 635 mailed surveys, 30 402 were completed and re-

turned. Of those who returned the surveys, 6225 individuals agreed to participate in eligibility assessments.

These recruitment and consent procedures, including use of tape-recorded oral informed consent, were approved by Harvard Medical School's Human Subjects Research Committee.

Randomization

Dispositions of eligible consenting phase 1 participants for whom phase 2 eligibility assessments were attempted are shown in the FIGURE. Of 604 eligible, 304 were randomized without blocking or stratification to the intervention and 300 to usual care. Randomization was carried out by the survey research firm conducting eligibility assessments with a computerized procedure that classified respondents for eligibility and used a random number generator to assign participants to intervention or usual care. Those assigned to usual care were informed that their responses indicated possible depression and advised to consult with a clinician; they could receive any normally available insurance benefit or service (eg, psychotherapy or pharmacotherapy), just not the additional telephone care management components provided to those in the intervention group.

Telephone Outreach, Care Management, and Psychotherapy

The structured telephone intervention program (provided without charge to participants who were randomly assigned to receive the intervention) systematically assessed needs for treatment, facilitated entry into in-person treatment (both psychotherapy and antidepressant medication), monitored and supported treatment adherence, and (for those declining in-person treatment) provided a structured psychotherapy intervention by telephone.⁵⁵ Specific treatments were provided according to both clinical need and participant's willingness to accept treatment.

Care managers were licensed master's degree-level mental health clini-

cians employed by UBH. Additional training for this study included 12 hours of didactic instruction, role-play, and observed care manager contacts. Care managers also received approximately 60 minutes of supervision each week (from G.E.S., E.J.L., P.S.W., or F.A.) throughout the study period and had approximately 50 to 70 individuals in their case-loads when functioning at capacity.

Initial telephone contacts included a structured assessment of depressive symptoms (Patient Health Questionnaire 9),⁵⁶ prior treatment, complicating factors (including substance abuse), and motivation for treatment. For all participants with significant depressive symptoms, care managers recommended in-person psychotherapy as well as medication evaluation and provided treatment authorization and referral information. For participants declining in-person treatment, care managers provided a brief motivational intervention⁵⁷

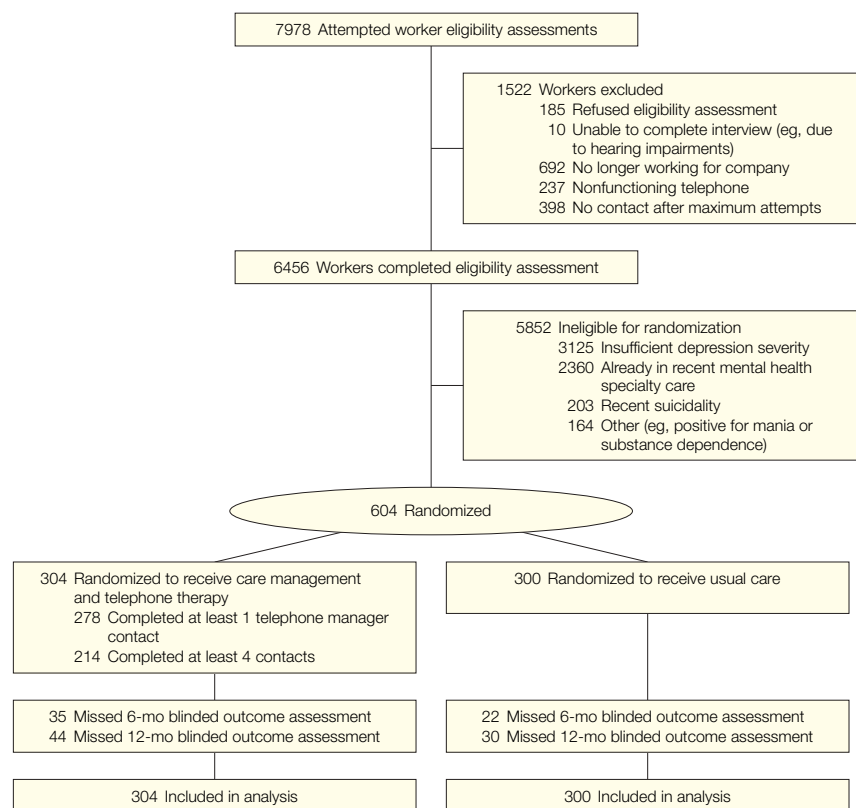
and asked permission for continued telephone contact. Following initial contacts, all intervention participants were mailed a psychoeducational workbook⁵⁸ emphasizing behavioral activation, identifying and challenging negative thoughts, and developing long-term self-care plans.

For participants entering in-person treatment, subsequent care manager contacts included structured assessments of depressive symptoms, treatment adherence, and barriers to continuing treatment. As needed, care managers provided feedback and algorithm-based recommendations to treating providers. A UBH psychiatrist was also available for consultation to clinicians if needed. For participants receiving one mode of treatment, adding a second mode was recommended if significant depressive symptoms persisted after 2 months. For participants declining in-person treatment, care

managers maintained regular telephone contacts. Intervals between contacts were determined by clinical need, ranging from weekly (for severe depressive symptoms) to bimonthly (for those considered stable or with mild or minimal symptoms).

Intervention participants declining in-person treatment and experiencing significant depressive symptoms after 2 months were offered a structured 8-session cognitive behavioral psychotherapy program.^{38,55,59} Approximately weekly sessions lasted 30 to 40 minutes and followed the workbook described above. Sessions included assessment of motivation for treatment and motivational enhancement exercises⁵⁷; focus on increasing pleasant and rewarding activities⁶⁰; identifying, challenging, and distancing from negative thoughts^{61,62}; and creating a personal self-care plan covering medication use, self-monitoring, and self management

Figure. Study Flow Diagram



HRA indicates health risk appraisal; UBH, United Behavioral Health (a managed behavioral health care company).

skills.⁶³ Booster sessions were scheduled every 4 to 8 weeks to monitor and support progress.

All care management activities were organized and supported by an electronic decision support system. Care manager training and procedural materials are available upon request

Blinded Outcome Assessments

Blinded outcome assessments were performed at baseline and at 6 and 12 months by trained survey interviewers at the research firm conducting telephone interviews. Participants were advised not to offer information regarding their intervention status to preserve blinding.

QIDS-SR

The QIDS-SR is a fully structured assessment^{50,51} that correlates significantly ($r=0.72$) with the 17-item clinician-administered Hamilton Rating Scale (HRSD)⁶⁴ and has good sensitivity to change.⁴⁹ QIDS-SR partial item-missing data were imputed using mean imputations (ie, by assigning the mean score obtained from respondents in the same treatment arm with valid scores). Item-missing data occurred in only a small proportion of cases and were unrelated to treatment status.

World Health Organization Health and Productivity Questionnaire

The World Health Organization Health and Productivity Questionnaire (HPQ) is a fully structured instrument^{65,66} assessing 4 broad dimensions of work functioning: (1) work hours (absenteeism); (2) job performance, with priming and decomposition questions followed by an anchored global 0-to-10 rating scale that is the measure used to define performance; (3) job turnover (fired, quitting, changing jobs, disability leave); and (4) critical workplace incidents, assessed in open-ended questions about job-related accidents or injuries, other major negative events, and major positive events (eg, receiving a promotion). Validation studies have documented significant associations ($r=0.61-0.87$) of HPQ work hours

assessments with payroll records⁶⁶ and job performance assessments with supervisor ratings ($r=0.52$)⁶⁵ and other administrative records (area under the curve, 0.58-0.72).⁶⁶ The HPQ partial item-missing data occurred in only a small proportion of cases, were unrelated to treatment status, and were imputed using mean imputations.

Care Process Measures

The care manager record system recorded contacts for recruitment monitoring and telephone psychotherapy. Telephone psychotherapy data were missing for the first 30% of participants and were imputed using multiple imputation.⁶⁷ Other mental health services were assessed by self-report in the telephone interviews.

Data Analysis

All comparisons were made in intent-to-treat analyses (based on original intervention assignment, regardless of treatments received). The propensity score method⁶⁸ was used to adjust for imperfect randomization in baseline characteristics. A logistic regression equation distinguishing intervention vs usual care groups based on baseline depression severity, absenteeism, job performance, and sociodemographics was used to generate predicted probabilities of intervention assignment. These predicted probabilities were used to weight the data without case-level matching so intervention and usual care groups had comparable distributions of characteristics (results available on request.)

Multiple imputation was used to adjust for some participants not completing either 6-month (35 intervention and 22 usual care) or 12-month (44 intervention and 30 usual care) interviews. Nonrespondents differed from respondents both on sociodemographic variables (nonrespondents were older, less well educated, and more likely to be male) and baseline scores on outcomes (nonrespondents had somewhat higher baseline depression severity and lower work performance). A 2-part process was used to adjust for

these differences: (1) estimates were generated for missing values using regression equations with all available data in the baseline and completed follow-up interviews as predictors; and (2) significance tests were adjusted for imputed values being estimated rather than observed using multiple imputation simulation methods.

Intervention effects on depression severity were estimated using multiple imputation multiple linear regression with simulated standard errors. QIDS-SR scores at 6 and 12 months were regressed on a dichotomous predictor for randomization status. Dichotomous measures of symptom improvement ($\geq 50\%$ reduction in QIDS-SR scores) and complete remission (QIDS-SR scores of ≤ 5) were also examined using multiple imputation multiple logistic regression.

Comparable multiple imputation regression analyses were used to estimate intervention effects on work outcomes. The primary outcome was a composite measure of the number of effective hours worked in the prior 7 days, for which participants no longer working contributed no hours and numbers of hours worked by employed respondents were weighted by job performance (eg, 30 hours worked with a rating of on-the-job performance of 8 out of a possible 10 points was assigned a score of 24 effective hours). Intervention effects on components of this composite (ie, probability of no longer working, number of hours worked among the employed, job performance during hours worked among the employed, and critical workplace incidents) were also examined. The multiple imputation multiple regression analyses estimated intervention effects at 6 and 12 months. We also estimated intervention effects constrained to be equal across these 2 time intervals to increase statistical power. This approach was implemented by pooling data across the 2 wave-pairs and including a dummy variable to distinguish the baseline-to-6-month panel from 6-to-12-month panel.

Table 1. Characteristics of Participants Assigned to Intervention and Usual Care

	Unweighted				Weighted ^a			
	Mean (SD)		t Value	P Value	Mean (SD)		t Value	P Value
	Intervention (n = 304)	Usual Care (n = 300)			Intervention (n = 304)	Usual Care (n = 300)		
Sociodemographics								
Age, mean, y	40.7 (10.5)	42.4 (10.8)	2.0	.05	41.3 (10.5)	41.3 (10.6)	0.0	>.99
Women, No. (%)	70.7 (20.7)	77.0 (17.7)	1.8	.08	74.4 (19.0)	74.4 (19.0)	0.0	>.99
College graduates, %	38.0 (23.6)	43.8 (24.6)	1.4	.15	40.7 (24.1)	40.7 (24.1)	0.0	>.99
Baseline values of outcomes								
Depression symptom severity, QIDS-SR score ^b	13.3 (3.3)	13.8 (3.6)	2.0	.04	13.5 (3.3)	13.5 (3.5)	0.0	>.99
Effective weekly work, h ^c	29.8 (13.5)	31.8 (14.4)	1.7	.08	30.4 (13.3)	31.2 (14.2)	0.7	.48
Actual weekly work, h	41.6 (14.4)	43.5 (14.4)	1.6	.10	42.5 (13.7)	42.5 (14.6)	0.0	>.99
On-the-job work performance	0.7 (0.2)	0.7 (0.2)	0.9	.38	0.7 (0.2)	0.7 (0.2)	1.3	.20

Abbreviation: QIDS-SR, Quick Inventory of Depressive Symptomatology self-report.

^aBased on a propensity score weight designed to reduce aggregate baseline differences between the 2 subsamples.

^bQIDS-SR score, with higher scores indicating greater depression severity.

^cThe product of relative hours and performance at work.

All linear regression coefficients (β) and odds ratios (ORs) from logistic regression equations were adjusted for outcomes at the prior assessment; expected hours of work at the prior assessment; and respondent age, sex, and education. In post hoc subgroup analyses, intervention effects were examined in subsamples defined by initial depression severity (moderate and severe cases corresponding to QIDS-SR score ≥ 11) and recruitment procedure (HRA vs UBH member wellness survey).

Power calculations indicated a sample of 300 per group would be needed to detect a difference of 0.2 SD of work productivity (power = 0.8, 2-sided $P = .05$) with 20% loss to follow-up. The minimum detectable effect size of a 0.2 SD was based on previous evidence that low to moderate intensity depression interventions yield effects on clinical outcomes of an SD of approximately 0.33^{35,39,40} and that clinical improvements are correlated with an SD of approximately 0.60 with improvements in work performance.¹³ Analyses were performed using SAS statistical software version 9.1.3 (SAS Institute Inc, Cary, NC).

RESULTS

Sample

Propensity score weights adjusted for baseline differences after randomiza-

tion (TABLE 1), including the somewhat younger age, lower proportions of women and college graduates, and lower baseline depression severity in the intervention vs usual care group.

When compared with workers meeting study criteria in the nationally representative National Comorbidity Survey Replication,⁶⁹ trial participants were somewhat older, more likely to be women, more educated, somewhat less depressed, and worked somewhat more hours, possibly reflecting the fact that the trial sample was confined to full-time workers in large national firms (results available upon request).

Effects on Depression Outcomes

QIDS-SR scores were significantly lower in the intervention than in the usual care group by 6 months, with an effect size ($\beta = -1.0$) of approximately one-third of an SD (TABLE 2). This advantage was retained at 12 months ($\beta = -1.1$). The proportion of participants whose symptoms improved substantially (50% QIDS-SR improvement) was also significantly higher among the intervention than the usual care group but not until the 12-month assessment (30.9% vs 21.6%; OR, 1.7; 95% CI, 1.1-2.5). The proportion of participants experiencing recovery (QIDS-SR score, ≤ 5) was also significantly higher in the intervention than the usual care group but not until 12

months (26.2% vs 17.7%; OR, 1.7; 95% CI, 1.1-2.4).

Post hoc subgroup analysis failed to find significant differences in intervention effects among participants with mild (QIDS-SR score, ≤ 10) vs moderate or severe (QIDS-SR score, ≥ 11) baseline depression and among participants recruited through HRAs vs the UBH member wellness survey (results available upon request).

Effects on Work Performance Outcomes

Scores on the summary effective hours worked measure were significantly higher in the intervention than usual care group at 6 ($\beta = 3.0$) and 12 ($\beta = 3.3$) months (TABLE 3). The 2.0 coefficient for hours worked among the employed means that workers in the intervention group worked an average of 2 more hours per week than workers in the usual care group, which is equivalent to an annualized effect of more than 2 weeks of work (ie, 2 hours per week \times 48 work weeks in the year). This overall effect was due to significant improvements in job retention (92.6% vs 88.0% by 12-month; OR, 1.7; 95% CI, 1.0-3.3) and hours worked among employed respondents ($\beta = 2.0$). Job retention was defined from the employee perspective (ie, the employee's continuing to hold a job) rather than from the employer perspective (ie, the

Table 2. The impact of the Intervention on Depression Outcomes at 6 and 12 Months After Randomization

	Mean (SD) ^a		Estimated Coefficient (95% Confidence Interval) ^b	t Value	P Value
	Intervention (n = 304)	Usual Care (n = 300)			
Symptom severity, mean QIDS-SR score ^c					
6 mo	10.2 (4.8)	11.2 (4.9)	-1.0 (-1.8 to 0.2) ^d	2.6	.01
12 mo	8.9 (4.8)	10.0 (4.7)	-1.1 (-1.8 to 0.3) ^d	2.8	.005
Pooled ^e			-1.0 (-1.7 to 0.4) ^d	3.2	.001
Substantial improvement, % ^f					
6 mo	21.7 (17.0)	17.4 (14.4)	1.2 (0.8 to 2.0)	1.3	.20
12 mo	30.9 (21.4)	21.6 (16.9)	1.7 (1.1 to 2.5) ^d	2.6	.01
Pooled ^e			1.4 (1.1 to 2.0) ^d	2.5	.01
Recovery, % ^g					
6 mo	18.2 (14.9)	12.6 (11.0)	1.7 (1.0 to 2.5)	1.9	.05
12 mo	26.2 (19.3)	17.7 (14.6)	1.7 (1.1 to 2.4) ^d	2.6	.01
Pooled ^e			1.4 (1.1 to 2.0) ^d	2.6	.009

^aThe means are reported without adjustment for control variables. The estimates of regression coefficients, in comparison, are based on analyses that include the control variables. As a result, the mean differences generally do not equal the linear regression coefficients and the odds ratios that can be calculated from the means generally do not equal the odds ratios obtained from the logistic regression equations.
^bLinear regression coefficient for reduction in symptom severity; odds ratio for substantial improvement, and recovery.
^cContinuous Quick Inventory of Depressive Symptomatology self-report (QIDS-SR) scores, with higher scores indicating greater depression severity.
^dSignificant at the .05 level, 2-sided test.
^eThe test for the pooled data used both 6- and 12-month outcomes and constrained the 2 coefficients (odds ratios or linear regression coefficient) to be equal.
^fThe percentage substantially improved was defined as a 50% or higher reduction in QIDS-SR score compared with baseline.
^gThe percentage recovered was defined as a QIDS-SR score of 5 or lower.

employee's continuing to work for the same employer). Subsequent analysis found that, among participants employed at follow-up, the proportion working for the same employer was unrelated to intervention. However, as overall job retention was significantly higher in the intervention group, the unconditional proportion of baseline respondents working for the same employer was higher in the intervention group than the usual care group (results available on request).

Further analysis showed that extreme values did not explain the significant intervention effect on hours worked among employed respondents, for this effect remained significant when the outcome variable was transformed using a square root transformation ($P = .049$). Decomposition showed, furthermore, that the overall effect on hours worked was due to an effect on the number of hours worked controlling for expected work hours ($\beta = 1.7, P = .04$) rather than on number of expected work hours ($\beta = 0.6, P = .18$). The

intervention effect on job performance, in comparison, was not significant ($\beta = 0.2, P = .11$), although it was consistently positive. No significant effects were found, finally, on critical workplace incidents or on taking a job with another employer (results available upon request). Subgroup analysis failed to find significant differences in intervention effects between mild vs moderate or severe cases or those recruited through HRAs vs the UBH member wellness survey (results available on request).

Effects on Use of Mental Health Care

Intervention group participants were significantly more likely than those in usual care to receive any mental health specialty treatment (OR, 1.6; 95% CI, 1.1-2.3), but somewhat less likely to obtain any depression treatment in primary care (OR, 0.7; 95% CI, 0.5-1.0) or nonmedical settings (OR, 0.6; 95% CI, 0.4-1.1; TABLE 4). The mean number of treatment contacts across all set-

tings (including care manager contacts) was nearly twice as large in the intervention vs the usual care group (12.7 vs 6.5; $t = 5.7; P < .001$). Of those randomized to receive the intervention, 50% completed initial care management contact by 8 days, 75% by 22 days, 90% by 114 days, and 9% were never reached.

COMMENT

The results suggest that enhanced depression care of workers has benefits not only on clinical outcomes but also on workplace outcomes. Although direct comparison to earlier studies is difficult because our trial is the first conducted exclusively among employed people, it is noteworthy that our effect size on clinical improvement (approximately one-third of an SD on the QIDS-SR distribution) is similar to earlier primary care trials using low to moderate intensity interventions.^{35,39,40} Our finding of similar effects among less severe and more serious cases is consistent with earlier primary care trials^{41,70} and suggests the intervention has benefit to a wide spectrum of depressed workers.

The significant 2.6-hour improvement per week in overall work functioning among intervention participants is due to a combination of increased job retention and increased hours worked among the employed. Earlier analyses of either primary care depressed patients⁴¹ or working subsamples of such patients⁷¹ have found generally comparable effects on retention and absenteeism. Although we did not find significant effects on work performance among the employed by 1 year, 1 primary care trial that followed up patients for 2 years did.³⁶ The apparent effects on absenteeism and performance among the employed in our trial may also have been downwardly biased if the intervention led to retention of employees with more absences or worse performances. Unfortunately, direct comparisons with these earlier studies are difficult due to differences in intervention intensity, follow-up, and stratum definitions.

Formal evaluation of our intervention's return on investment to employers is not currently possible because the latter requires information not yet available on duration of improvements, disability payments, overall health care expenditures, and hiring and training costs. However, the \$1800 annualized value of higher mean hours worked among intervention participants retaining their jobs (assuming the median annual salary in the US civilian labor force) by itself far exceeds the \$100 to \$400 outreach and care management costs associated with low- to moderate-intensity interventions of the sort we implemented^{38,39}; these saving might also exceed or closely approximate the costs of approximately 10 additional mental health specialty visits made by intervention participants over the course of a year.

These last observations suggests that outreach and enhanced care for depressed workers might be better conceptualized as an opportunity to invest in improving the productive capacity of workforces (referred to by employers as "human capital investments") than as workplace costs.^{72,73} That the intervention also had positive impacts on job retention and the costs of hiring and training new workers are typically high¹ reinforce this in-

Table 3. Effect of Intervention on Work Performance at 6 and 12 Months After Randomization

	Mean (SD) ^a		Estimated Coefficient (95% Confidence Interval) ^b	t Value	P Value
	Intervention (n = 304)	Usual Care (n = 300)			
Effective weekly hours worked ^c					
6 mo	30.1 (14.5)	27.1 (15.5)	3.0 (0.4 to 5.6) ^d	2.2	.03
12 mo	29.5 (14.5)	26.0 (15.8)	3.3 (0.9 to 5.8) ^d	2.7	.008
Pooled ^e			2.6 (1.0 to 4.3) ^d	3.2	.002
Job retention, %					
6 mo	96.1 (3.7)	90.1 (8.9)	2.5 (1.2 to 5.0) ^d	2.7	.007
12 mo	92.6 (6.8)	88.0 (10.6)	1.7 (1.0 to 3.3)	1.8	.07
Pooled ^e			1.7 (1.1 to 3.3) ^d	2.3	.02
Actual weekly hours worked among the employed					
6 mo	42.0 (15.4)	40.1 (15.6)	1.8 (-0.8 to 4.4)	1.3	.18
12 mo	42.3 (13.4)	39.5 (13.7)	2.1 (-0.4 to 4.5)	1.7	.09
Pooled ^e			2.0 (0.3 to 3.7) ^d	2.3	.02
On-the-job performance among the employed ^b					
6 mo	0.8 (0.2)	0.7 (0.2)	0.2 (-0.2 to 0.5)	0.9	.35
12 mo	0.8 (0.2)	0.7 (0.2)	0.2 (-0.2 to 0.6)	0.8	.40
Pooled ^e			0.2 (-0.0 to 0.1)	1.6	.11
Critical workplace incidents ^f					
6 mo	-0.2 (0.6)	-0.2 (0.6)	0.00 (-0.1 to 0.1)	0.1	.93
12 mo	-0.2 (0.6)	-0.2 (0.6)	0.05 (-0.0 to 0.2)	1.1	.29
Pooled ^e			0.02 (-0.0 to 0.1)	0.7	.51

^aThe means are reported without adjustment for control variables. The estimates of regression coefficients, in comparison, are based on analyses that include the control variables. As a result, the mean differences generally do not equal the linear regression coefficients and the odds ratios that can be calculated from the means generally do not equal the odds-ratios obtained from the logistic regression equations.
^bOdds ratio for job retention; linear regression coefficient for other outcomes.
^cRelative effective hours work is a product of job retention, relative hours among the employed, and on-the-job performance among the employed. The value is set to 0 for respondents who no longer work.
^dSignificant difference between intervention and usual care at the .05 level, 2-sided test.
^eThe test for the pooled data used both 6 and 12 mo outcomes and constrained the 2 coefficients (odds ratios or linear regression coefficient) to be equal.
^fCritical workplace incidents include accidents or injuries, other important work failures, and important work successes (reverse coded).

Table 4. Service Use of Participants Assigned to Intervention and Usual Care at 6 and 12 Months After Randomization

	Mean (SD) ^a		Estimated Coefficient (95% Confidence Interval) ^b	t Value	P Value
	Intervention (n = 304)	Usual Care (n = 300)			
Any contacts, %					
Case manager calls					
6 mo	90.5 (8.6)	0.0			
12 mo	71.0 (20.6)	0.0			
Pooled	91.4 (7.9)	0.0			
Case manager telephonic psychotherapy					
6 mo	33.6 (22.3)	0.0			
12 mo	30.7 (21.3)	0.0			
Pooled	33.6 (22.3)	0.0			
Visits to PCP for major depression					
6 mo	17.6 (14.5)	24.1 (18.3)	0.7 (0.4 to 1.0)	1.9	.05
12 mo	14.9 (12.7)	21.0 (16.6)	0.7 (0.4 to 1.1)	1.6	.10
Pooled	25.1 (18.8)	32.8 (22.0)	0.7 (0.5 to 1.0)	1.9	.06
Visits to MH specialist for major depression					
6 mo	34.8 (22.7)	27.3 (19.8)	1.4 (1.0 to 2.0) ^c	2.0	.05
12 mo	25.0 (18.8)	19.0 (15.4)	1.5 (1.0 to 2.3)	1.9	.06
Pooled	43.6 (24.6)	32.9 (22.1)	1.6 (1.1 to 2.3) ^c	2.5	.01

(continued)

interpretation.⁷⁴ However, it is important to recognize that these workplace benefits would not be realized by all employers because hiring and training costs and the extents to which employees are paid—piecemeal, hourly, or by salary—vary.

The intervention had modest effects on self-reported use of treatments, consistent with earlier trials of low-intensity interventions.^{38,39} However, intervention participants received twice as many contacts as usual care participants when care manager

contacts were included (12.7 vs 6.5, $t=5.7$, $P<.001$) and were 70% more likely to receive any mental health specialty treatment. Although it is difficult to identify active components from a trial with a single intervention group, the finding that intervention partici-

Table 4. Service Use of Participants Assigned to Intervention and Usual Care at 6 and 12 Months After Randomization (cont)

	Mean (SD) ^a		Estimated Coefficient (95% Confidence Interval) ^b	t Value	P Value
	Intervention (n = 304)	Usual Care (n = 300)			
Any contacts, %					
Visits to any other provider for major depression					
6 mo	6.8 (6.3)	11.8 (10.4)	0.5 (0.3 to 1.0) ^c	2.0	.04
12 mo	5.7 (5.4)	9.1 (8.3)	0.6 (0.3 to 1.2)	1.4	.18
Pooled	11.1 (9.9)	16.0 (13.4)	0.6 (0.4 to 1.1)	1.7	.08
Any treatment contacts					
6 mo	93.9 (5.7)	50.5 (25.0)	16.4 (9.5 to 28.3) ^a	10.0	<.001
12 mo	82.2 (14.6)	41.7 (24.3)	7.0 (4.6 to 10.6) ^a	9.3	<.001
Pooled	96.0 (3.8)	59.9 (24.0)	17.2 (9.0 to 32.8) ^a	8.7	<.001
Number of contacts					
Case manager calls					
6 mo	4.0 (2.6)	0.0			
12 mo	2.2 (2.0)	0.0			
Pooled	6.1 (4.2)	0.0			
Case manager telephonic psychotherapy					
6 mo	1.2 (1.9)	0.0			
12 mo	0.8 (1.5)	0.0			
Pooled	2.0 (3.3)	0.0			
Visits to PCP for major depression					
6 mo	0.4 (1.1)	0.6 (1.5)	-0.2 (-0.4 to 0.0)	1.8	.08
12 mo	0.3 (0.8)	0.5 (1.4)	-0.2 (-0.4 to -0.0)	2.2	.02
Pooled	0.6 (1.5)	1.0 (2.5)	-0.4 (-0.8 to -0.1)	2.6	.01
Visits to MH specialist for major depression ^d					
6 mo	2.8 (6.1)	2.1 (5.5)	0.7 (-0.2 to 1.6)	1.5	.14
12 mo	1.9 (4.7)	1.7 (5.3)	0.3 (-0.6 to 1.1)	0.6	.53
Pooled	4.7 (9.6)	3.9 (9.1)	0.9 (-0.6 to 2.4)	1.2	.25
Visits to any other clinician for major depression ^e					
6 mo	0.8 (5.8)	0.8 (5.4)	-0.0 (-0.9 to 0.9)	0.0	.97
12 mo	0.4 (2.7)	0.7 (4.0)	-0.2 (-0.8 to 0.3)	0.9	.37
Pooled	1.2 (6.7)	1.5 (8.9)	-0.3 (-1.5 to 1.0)	0.5	.66
Any treatment contacts ^f					
6 mo	8.0 (9.5)	3.5 (8.6)	4.5 (3.0 to 5.9) ^c	6.1	<.001
12 mo	4.7 (6.1)	2.9 (7.1)	2.0 (0.9 to 3.0) ^c	3.7	<.001
Pooled	12.7 (13.6)	6.5 (13.7)	6.3 (4.1 to 8.4) ^c	5.7	<.001
Taking antidepressant medication, %					
6 mo	30.4 (21.2)	35.1 (22.8)	1.0 (0.7 to 1.4)	0.0	.98
12 mo	30.5 (21.2)	34.1 (22.5)	0.8 (0.6 to 1.2)	1.1	.30
Pooled	41.1 (24.2)	41.5 (24.3)	1.0 (0.7 to 1.5)	0.0	.99

Abbreviations: MD, major depression; MH, mental health; PCP, primary care provider.

^aThe means are reported without adjustment for control variables. The estimates of regression coefficients, in comparison, are based on analyses that include the control variables. As a result, the mean differences generally do not equal the linear regression coefficients and the odds ratios that can be calculated from the means generally do not equal the odds ratios obtained from the logistic regression equations.

^bOdds ratio for the outcomes in the "Any contacts" and "Taking antidepressant medication" sections of the Table; linear regression coefficient for the outcomes in the "Number of contacts" section.

^cSignificant difference between intervention and usual care at the .05 level, 2-sided test.

^dMH specialists include psychiatrists, psychologists, therapists, and MH counselors.

^eOther care providers include spiritual advisors, other human services professionals, and complementary-alternative medical clinicians.

^fTotal contacts include contacts with care managers and visits to mental health specialists, physicians, and other clinicians.

pants received more mental health specialty treatment is noteworthy in light of other data suggesting that mental health specialty care is more likely to meet evidence-based recommendations than treatments in other sectors.²⁵ A recent meta-regression of 28 depression collaborative care trials is also instructive in that it identified 3 “active ingredients”: systematic screening to identify patients (vs other means such as clinician referral), use of mental health professionals as care managers, and regularly planned care manager supervision.²⁷ All 3 elements were included in our intervention. The telephone cognitive behavioral therapy may also have had similar beneficial effects as an earlier trial that found a higher proportion of patients whose care management included telephone cognitive behavioral therapy had experienced depression improvement than those who had received care management alone.³⁸

Several potential technical limitations are noteworthy. First, the QIDS-SR might have misclassified cases, although clinical reappraisal studies show it has high concordance with blinded clinician assessments.⁴⁹⁻⁵¹ Second, the HPQ might have been systematically biased,⁷⁵ although significant associations between HPQ scores and independent administrative or archival records of absenteeism and work performance have been documented across a broad range of occupations.^{65,66} Third, workers who participated in our initial screening phase may have had a different prevalence, severity, or impairment associated with their depression than nonparticipants. Although we had no way to evaluate this possibility, a prior study found initial participants in HRA screenings were comparable—in terms of depression severity, work impairments, and associations between the two—with initial nonrespondents who participated only after more intensive recruitment efforts.⁷⁶ Fourth, the generalizability of our findings is unclear because trial participants had less severe depressions and a different sociodemographic profile than a nationally representative sample

of depressed workers. Although we found no differences in intervention effects across levels of depression severity or method of recruitment, such differences could exist across other subgroups (eg, white-collar vs blue-collar workers) and have relevance to employers whose workforces vary in these characteristics.

Potential conceptual limitations also need to be considered. Simple human capital metrics such as absenteeism and job performance might overestimate true costs to employers, as would happen if unperformed work during absences is made up by coworkers or the absent worker upon return.⁷⁷ However, it is also possible that the burdens of depression to employers are underestimated herein because other costs, such as for hiring temporary workers, paying coworkers overtime, and adverse effects on coworkers’ productivity were not considered.⁷⁸ Likewise, we did not assess intervention effects on outcomes such as suffering, marital stability, caregiver burden, and employee contributions outside the workplace that could have value from a societal perspective and might lead to long-term improvements in productive capacity.⁷⁹ An exclusive focus on work outcomes might devalue benefits of intervening among groups not in the workforce (eg, the elderly) or in low-wage occupations, emphasizing that health care resource allocation decisions need to consider a societal as well as employer perspective.⁸⁰ Within the constraints of these limitations, this study suggests that enhanced care for depressed workers can have benefits for employers that go beyond improved health and diminished suffering in their workforces and extend to increased work productivity. Further study is needed to determine whether intervention costs are offset by these workplace benefits and the variation in this offset across different employment settings.^{72,73,81} Toward this end, it is noteworthy that increased depression treatment among intervention participants was largely provided through telephone contacts with care managers and

not provided through more expensive in-person visits with traditional clinicians. We also did not consider whether the intervention offsets any greater general medical utilization associated with depression, as has been observed in earlier primary care trials.^{32,34,35,82} Likewise, availability of Web-based, e-mail, and interactive voice-recognition technologies should ensure that the costs of screening and recruiting depressed workers into interventions are low.⁸³ These features may be critically important to potential purchasers, who are not just sensitive to interventions’ returns on investment but also to their absolute costs and impacts on per member per month charges.⁴⁴ Attention to these issues in future research is needed to ensure that successful programs of outreach and enhanced depression treatment are widely disseminated.⁴⁵

Author Affiliations: Division of Services and Intervention Research, National Institute of Mental Health, Rockville, Maryland (Dr Wang); Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women’s Hospital (Drs Wang and Avorn) and Department of Health Care Policy, Harvard Medical School (Drs Wang, Petukhova and Kessler), Boston, Massachusetts; Center for Health Studies, Group Health Cooperative, Seattle, Washington (Dr Simon and Ludman); and United Behavioral Health, San Francisco, California (Dr Azocar).

Author Contributions: Drs Wang and Kessler had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Wang, Simon, Avorn, Azocar, Ludman, McCulloch, Kessler.

Acquisition of data: Wang, Simon, Azocar, Ludman, McCulloch, Kessler.

Analysis and interpretation of data: Wang, Simon, Azocar, Ludman, McCulloch, Petukhova, Kessler.

Drafting of the manuscript: Wang, Simon, Avorn, Azocar, Ludman, McCulloch, Petukhova, Kessler.

Critical revision of the manuscript for important intellectual content: Wang, Simon, Avorn, Azocar, Ludman, McCulloch, Petukhova, Kessler.

Statistical analysis: Wang, Simon, Petukhova, Kessler.

Obtained funding: Wang.

Administrative, technical, or material support: Wang, Simon, Azocar, Ludman, McCulloch.

Study supervision: Wang, Simon, Ludman, Azocar.

Financial Disclosures: Dr Kessler reports consulting for AstraZeneca, Bristol-Myers Squibb, Eli Lilly & Co, GlaxoSmithKline, Pfizer, and Wyeth and reports receiving research support for his epidemiological studies from Bristol-Myers Squibb, Eli Lilly, Ortho-McNeil, Pfizer, and the Pfizer Foundation. Dr Azocar and Ms McCulloch are both employees of United Behavioral Health and report holding stocks and options from United Health Group. No other authors reported financial conflicts.

Funding/Support: The research reported herein was supported by grant R01 MH61941 from the National Institute of Mental Health (Dr Wang), grant 048123 from the Robert Wood Johnson Foundation (Dr Wang), and an unrestricted educational grant that

supported earlier pilot work of this study from the John D. and Catherine T. MacArthur Foundation.

Role of the Sponsor: None of these organizations had any role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

REFERENCES

- Broadhead WE, Blazer DG, George LK, Tse CK. Depression, disability days, and days lost from work in a prospective epidemiologic survey. *JAMA*. 1990;264(19):2524-2528.
- Murray CJL, Lopez AD. *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries and Risk Factors in 1990 and Projected to 2020*. Cambridge, MA: Harvard University Press; 1996.
- Wells KB, Stewart A, Hays RD, et al. The functioning and well-being of depressed patients: results from the Medical Outcomes Study. *JAMA*. 1989;262(7):914-919.
- Greenberg PE, Kessler RC, Nells TL, Finkelstein SN, Berndt ER. Depression in the workplace: an economic perspective. In: Feighner JP, Boyer WF, eds. *Selective Serotonin Re-Uptake Inhibitors: Advances in Basic Research and Clinical Practice*. 2nd ed. New York, NY: John Wiley & Sons; 1996:327-363.
- Kessler RC, Frank RG. The impact of psychiatric disorders on work loss days. *Psychol Med*. 1997;27(4):861-873.
- Burton WN, Conti DJ, Chen CY, Schultz AB, Edington DW. The role of health risk factors and disease on worker productivity. *J Occup Environ Med*. 1999;41(10):863-877.
- Druss BG, Marcus SC, Olfson M, Tanielian T, Elinson L, Pincus HA. Comparing the national economic burden of five chronic conditions. *Health Aff (Millwood)*. 2001;20(6):233-241.
- Kessler RC, Mickelson KD, Barber CB, Wang PS. The association between chronic medical conditions and work impairment. In: Rossi AS, ed. *Caring and Doing for Others: Social Responsibility in the Domains of Family, Work, and Community*. Chicago, IL: University of Chicago Press; 2001:403-426.
- Stewart WF, Ricci JA, Chee E, Hahn SR, Morganstein D. Cost of lost productive work time among US workers with depression. *JAMA*. 2003;289(23):3135-3144.
- Wang PS, Beck A, Berglund P, et al. Chronic medical conditions and work performance in the health and work performance questionnaire calibration surveys. *J Occup Environ Med*. 2003;45(12):1303-1311.
- Agosti V, Stewart JW, Quitkin FM. Life satisfaction and psychosocial functioning in chronic depression: effect of acute treatment with antidepressants. *J Affect Disord*. 1991;23(1):35-41.
- Barge-Schaapveld DQ, Nicolson NA, van der Hoop RG, De Vries MW. Changes in daily life experience associated with clinical improvement in depression. *J Affect Disord*. 1995;34(2):139-154.
- Berndt ER, Finkelstein SN, Greenberg PE, et al. Workplace performance effects from chronic depression and its treatment. *J Health Econ*. 1998;17(5):511-535.
- Hays RD, Wells KB, Sherbourne CD, Rogers W, Spritzer K. Functioning and well-being outcomes of patients with depression compared with chronic general medical illnesses. *Arch Gen Psychiatry*. 1995;52(1):11-19.
- Kocsis JH, Frances AJ, Voss C, Mason BJ, Mann JJ, Sweeney J. Imipramine and social-vocational adjustment in chronic depression. *Am J Psychiatry*. 1988;145(8):997-999.
- Mauskopf JA, Simeon GP, Miles MA, Westlund RE, Davidson JR. Functional status in depressed patients: the relationship to disease severity and disease resolution. *J Clin Psychiatry*. 1996;57(12):588-592.
- Mintz J, Mintz LI, Arruda MJ, Hwang SS. Treatments of depression and the functional capacity to work. *Arch Gen Psychiatry*. 1992;49(10):761-768.
- Mynors-Wallis LM, Gath DH, Lloyd-Thomas AR, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *BMJ*. 1995;310(6977):441-445.
- Ormel J, Von Korff M, Van den Brink W, Katon W, Brilman E, Oldehinkel T. Depression, anxiety, and social disability show synchrony of change in primary care patients. *Am J Public Health*. 1993;83(3):385-390.
- Simon GE, Katon W, Rutter C, et al. Impact of improved depression treatment in primary care on daily functioning and disability. *Psychol Med*. 1998;28(3):693-701.
- Von Korff M, Ormel J, Katon W, Lin EH. Disability and depression among high utilizers of health care: a longitudinal analysis. *Arch Gen Psychiatry*. 1992;49(2):91-100.
- Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003;289(23):3095-3105.
- Wang PS, Berglund P, Kessler RC. Recent care of common mental disorders in the United States: prevalence and conformance with evidence-based recommendations. *J Gen Intern Med*. 2000;15(5):284-292.
- Wang PS, Demler O, Kessler RC. Adequacy of treatment for serious mental illness in the United States. *Am J Public Health*. 2002;92(1):92-98.
- Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):629-640.
- Young AS, Klap R, Sherbourne CD, Wells KB. The quality of care for depressive and anxiety disorders in the United States. *Arch Gen Psychiatry*. 2001;58(1):55-61.
- Bower P, Gilbody S, Richards D, Fletcher J, Sutton A. Collaborative care for depression in primary care: making sense of a complex intervention: systematic review and meta-regression. *Br J Psychiatry*. 2006;189:484-493.
- Gensichen J, Beyer M, Muth C, Gerlach FM, Von Korff M, Ormel J. Case management to improve major depression in primary health care: a systematic review. *Psychol Med*. 2006;36(1):7-14.
- Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. *Arch Intern Med*. 2006;166(21):2314-2321.
- Gilbody S, Bower P, Whitty P. Costs and consequences of enhanced primary care for depression: systematic review of randomised economic evaluations. *Br J Psychiatry*. 2006;189:297-308.
- Gilbody S, Whitty P, Grimshaw J, Thomas R. Educational and organizational interventions to improve the management of depression in primary care: a systematic review. *JAMA*. 2003;289(23):3145-3151.
- Katon W, Robinson P, Von Korff M, et al. A multifaceted intervention to improve treatment of depression in primary care. *Arch Gen Psychiatry*. 1996;53(10):924-932.
- Katon W, Von Korff M, Lin E, et al. Stepped collaborative care for primary care patients with persistent symptoms of depression: a randomized trial. *Arch Gen Psychiatry*. 1999;56(12):1109-1115.
- Katon W, Von Korff M, Lin E, et al. Collaborative management to achieve treatment guidelines: impact on depression in primary care. *JAMA*. 1995;273(13):1026-1031.
- Katzelnick DJ, Simon GE, Pearson SD, et al. Randomized trial of a depression management program in high utilizers of medical care. *Arch Fam Med*. 2000;9(4):345-351.
- Rost K, Smith JL, Dickinson M. The effect of improving primary care depression management on employee absenteeism and productivity: a randomized trial. *Med Care*. 2004;42(12):1202-1210.
- Schulberg HC, Block MR, Madonia MJ, et al. Treating major depression in primary care practice: eight-month clinical outcomes. *Arch Gen Psychiatry*. 1996;53(10):913-919.
- Simon GE, Ludman EJ, Tutty S, Operskalski B, Von Korff M. Telephone psychotherapy and telephone care management for primary care patients starting antidepressant treatment: a randomized controlled trial. *JAMA*. 2004;292(8):935-942.
- 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(7234):550-554.
- Unutzer J, Katon W, Callahan CM, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA*. 2002;288(22):2836-2845.
- Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA*. 2000;283(2):212-220.
- Frank RG, Huskamp HA, Pincus HA. Aligning incentives in the treatment of depression in primary care with evidence-based practice. *Psychiatr Serv*. 2003;54(5):682-687.
- Pincus HA, Hough L, Houtsinger JK, Rollman BL, Frank RG. Emerging models of depression care: multi-level ("6 P") strategies. *Int J Methods Psychiatr Res*. 2003;12(1):54-63.
- Schoenbaum M, Kelleher K, Lave JR, Green S, Keyser D, Pincus H. Exploratory evidence on the market for effective depression care in Pittsburgh. *Psychiatr Serv*. 2004;55(4):392-395.
- Wang PS, Simon G, Kessler RC. The economic burden of depression and the cost-effectiveness of treatment. *Int J Methods Psychiatr Res*. 2003;12(1):22-33.
- Evaluation of National Health Interview Survey diagnostic reporting. *Vital Health Stat 2*. 1994;(120):1-116.
- Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med*. 2002;32(6):959-976.
- Kessler RC, Barker PR, Colpe LJ, et al. Screening for serious mental illness in the general population. *Arch Gen Psychiatry*. 2003;60(2):184-189.
- Rush AJ, Carmody T, Reimnitz P-E. The Inventory of Depressive Symptomatology (IDS): clinician (IDS-C) and self-report (IDS-SR) ratings of depressive symptoms. *Int J Methods Psychiatr Res*. 2000;9(2):45-59.
- Rush AJ, Giles DE, Schlessler MA, Fulton CL, Weisenburger J, Burns C. The Inventory for Depressive Symptomatology (IDS): preliminary findings. *Psychiatry Res*. 1986;18(1):65-87.
- Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med*. 1996;26(3):477-486.
- Kessler RC, Andrews G, Mroczek D, Üstün TB, Wittchen H-U. The World Health Organization Composite International Diagnostic Interview Short Form (CIDI-SF). *Int J Methods Psychiatr Res*. 1998;7(4):171-185.
- Azocar F, Cuffel BJ, McCulloch J, McCabe J, Tani S, Brody B. Monitoring patient improvement and treatment outcomes in managed behavioral health. *J Healthc Qual*. 2007;29(2):4-12.

54. Brodey BB, Cuffel B, McCulloch J, et al. The acceptability and effectiveness of patient-reported assessments and feedback in a managed behavioral healthcare setting. *Am J Manag Care*. 2005;11(12):774-780.
55. Ludman EJ, Simon G, Tutty S, Von Korff M. A randomized trial of telephone psychotherapy and pharmacotherapy support for depression: continuation and durability of effects. *J Consult Clin Psychol*. 2007;75(2):257-266.
56. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*. 1999;282(18):1737-1744.
57. Rollnick S, Mason P, Butler C. *Health Behavior Change: A Guide for Practitioners*. London, England: Churchill Livingstone; 1999.
58. Simon GE, Ludman EJ, Tutty S. *Creating a Balance: A Step-by-Step Approach to Managing Stress and Lifting Your Mood*. Victoria, BC: Trafford Press; 2006.
59. Tutty S, Simon G, Ludman E. Telephone counseling as an adjunct to antidepressant treatment in the primary care system: a pilot study. *Eff Clin Pract*. 2000;3(4):170-178.
60. Lewinsohn P, Munoz R, Youngren M, Zeiss A. *Control Your Depression*. New York, NY: Prentice Hall; 1986.
61. Beck A, Rush A, Shaw B, Emery G. *Cognitive Therapy of Depression*. New York, NY: Guilford Press; 1987.
62. Hays S, Strosahl K, Wilson K. *Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change*. New York, NY: Guilford Press; 1999.
63. Ludman E, Von Korff M, Katon W, et al. The design, implementation, and acceptance of a primary care-based intervention to prevent depression relapse. *Int J Psychiatry Med*. 2000;30(3):229-245.
64. Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol*. 1967;6(4):278-296.
65. Kessler RC, Ames M, Hymel PA, et al. Using the WHO Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med*. 2004;46(6)(suppl 6):S23-S27.
66. Kessler RC, Barber C, Beck A, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med*. 2003;45(2):156-174.
67. Rubin DB. Multiple Imputation after 18+ years. *J Am Stat Assoc*. 1996;91(434):473-489.
68. Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55.
69. Kessler RC, Merikangas KR. The National Comorbidity Survey Replication (NCS-R): background and aims. *Int J Methods Psychiatr Res*. 2004;13(2):60-68.
70. Bush T, Rutter C, Simon G, et al. Who benefits from more structured depression treatment? *Int J Psychiatry Med*. 2004;34(3):247-258.
71. Rollman BL, Belnap BH, Mazumdar S, et al. A randomized trial to improve the quality of treatment for panic and generalized anxiety disorders in primary care. *Arch Gen Psychiatry*. 2005;62(12):1332-1341.
72. Goetzel RZ, Ozminkowski RJ, Sederer LI, Mark TL. The business case for quality mental health services: why employers should care about the mental health and well-being of their employees. *J Occup Environ Med*. 2002;44(4):320-330.
73. Leatherman S, Berwick D, Iles D, et al. The business case for quality: case studies and an analysis. *Health Aff (Millwood)*. 2003;22(2):17-30.
74. Wang PS, Patrick A, Avorn J, et al. The costs and benefits of enhanced depression care to employers. *Arch Gen Psychiatry*. 2006;63(12):1345-1353.
75. Morgado A, Raoux N, Smith M, Allilaire JF, Widlocher D. Subjective bias in reports of poor work adjustment in depressed patients. *Acta Psychiatr Scand*. 1989;80(6):541-547.
76. Wang PS, Beck AL, McKenas DK, et al. Effects of efforts to increase response rates on a workplace chronic condition screening survey. *Med Care*. 2002;40(9):752-760.
77. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ*. 1995;14(2):171-189.
78. Nicholson S, Pauly MV, Polsky D, Sharda C, Szrek H, Berger ML. Measuring the effects of work loss on productivity with team production. *Health Econ*. 2006;15(2):111-123.
79. Simon GE, Katon WJ, VonKorff M, et al. Cost-effectiveness of a collaborative care program for primary care patients with persistent depression. *Am J Psychiatry*. 2001;158(10):1638-1644.
80. Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-effectiveness in Health and Medicine*. New York, NY: Oxford University Press; 1996.
81. Sturm R. Economic grand rounds: the myth of medical cost offset. *Psychiatr Serv*. 2001;52(6):738-740.
82. Von Korff M, Katon W, Bush T, et al. Treatment costs, cost offset, and cost-effectiveness of collaborative management of depression. *Psychosom Med*. 1998;60(2):143-149.
83. Ryder R. Implementation strategies and applications for health risk appraisals. In: Hyner GC, Peterson KW, Travis JW, Dewey JE, Foerster JJ, Frammer EM, eds. *SPM Handbook of Health Assessment Tools*. Pittsburgh, PA: Society for Prospective Medicine and the Institute for Health and Productivity Management; 1999:179-184.