Annals of Internal Medicine | ARTICLE

Case Management for Depression by Health Care Assistants in Small Primary Care Practices

A Cluster Randomized Trial

Jochen Gensichen, MD, MPH, MA; Michael von Korff, ScD; Monika Peitz, PhD; Christiane Muth, MD, MPH; Martin Beyer, MSc; Corina Güthlin, PhD; Marion Torge, MD; Juliana J. Petersen, MD, MPH; Thomas Rosemann, MD, PhD; Jochem König, ScD; and Ferdinand M. Gerlach, MD, MPH, for PRoMPT (PRimary care Monitoring for depressive Patients Trial)

Background: Case management by health care assistants in small primary care practices provides unclear benefit for improving depression symptoms.

Objective: To determine whether case management provided by health care assistants in small primary care practices is more effective than usual care in improving depression symptoms and process of care for patients with major depression.

Design: Cluster randomized, controlled trial. A central automated system generated the randomization scheme, which was stratified by urban and rural practices; allocation sequence was concealed until groups were assigned.

Setting: 74 small primary care practices in Germany from April 2005 to September 2007.

Patients: 626 patients age 18 to 80 years with major depression.

Intervention: Structured telephone interview to monitor depression symptoms and support for adherence to medication, with feedback to the family physician.

Measurements: Depression symptoms at 12 months, as measured by the Patient Health Questionnaire-9 (PHQ-9); secondary outcomes were patient assessment of chronic illness care, adherence to medication, and quality of life.

Results: A total of 310 patients were randomly assigned to case management and 316 to usual care. At 12 months, 249 intervention recipients and 278 control patients were assessed; 555 patients were included in a modified intention-to-treat-analysis (267 intervention recipients vs. 288 control patients). Compared with control patients, intervention recipients had lower mean PHQ-9 values in depression symptoms (−1.41 [95% CI, −2.49 to −0.33]; $P = 0.014$), more favorable assessments of care (3.41 vs. 3.11; $P = 0.011$), and increased treatment adherence (2.70 vs. 2.53; $P = 0.042$). Quality-of-life scores did not differ between groups.

Limitation: Patients, health care assistants, family physicians, and researchers were not blinded to group assignment, and 12-month follow-up of patients was incomplete.

Conclusion: Case management provided by primary care practice–based health care assistants may reduce depression symptoms and improve process of care for patients with major depression more than usual care.

Primary Funding Source: German Ministry of Education and Research.

Ann Intern Med. 2009;151:369-378. www.annals.org

For author affiliations, see end of text.

International Standard Randomized Controlled Trial Number: ISRCTN66386086.

Depression causes a substantial disease burden (1) and is responsible for annual health care costs of about $83.1 billion in the United States (2, 3). Most patients with depression are treated in primary care (4–6). Collaborative care can improve depression outcomes by providing decision support and clinical information for family physicians, as well as self-management support and follow-up for patients. However, evidence regarding collaborative depression care stems mostly from academic or managed care settings in the United States (7–11). In these trials, family physicians generally relied on mental health case managers and decision support from mental health specialists (12). Case management is a patient-centered element of collaborative care that may be effective in primary care (13). It comprises systematic tracking of patients, support for continuing the treatment, and taking action in the case of nonadherence or lack of improvement (14). Collaborative depression care has generally yielded positive results in diverse primary care settings (7). The IMPACT (Improving Mood—Promoting Access to Collaborative Treatment) trial found that care managers who were supervised by psychiatrists and who provided education and support for medication adherence reduced depression symptoms in older patients (15). Dietrich and colleagues (16) found that telephone support, provided by a trained, centrally based mental health care manager who was supervised by a psychiatrist, improved depression symptoms. Dobscha and colleagues (17) evaluated a primary care decision-support team for depression (comprising a psychiatrist and a mental health nurse) and found improved care processes but no differences in depression symptoms (17), possibly because of less intensive follow-up of patients.

See also:

Print
Editors' Notes 370
Editorial comment. 425

Web-Only
Appendix
Appendix Tables
Conversion of graphics into slides

Context

Few studies have evaluated whether health care assistants can improve care for depressed patients.

Contribution

Patients who were randomly assigned to receive telephone case management by health care assistants reported slightly greater improvements in depression symptoms, better adherence to antidepressant therapies, and more favorable assessments of the quality of their care than did patients randomly assigned to receive usual care.

Implication

Telephone case management facilitated by health care assistants may be a feasible mechanism for small primary care practices to improve care of their patients with depression.

—The Editors

Small, isolated primary care settings often have limited resources (18). In the United States, 26% of primary care practices are solo practices or 2-person partnerships, in which extensive collaborative models would be difficult to implement, and 22% are located in rural areas with limited access to mental health specialists (19). Health care assistants are established professionals in primary care settings. They have less training than U.S. physician assistants or nurse practitioners, who provide first-contact care, and need not be college graduates (20). In Germany, health care assistants have 3 years of on-the-job training. They are mainly responsible for administrative tasks in general practice but provide basic clinical procedures (21). Health care assistants are a potentially important resource for enhancing patient care in primary care settings (20).

Our aim was to evaluate whether case management by a practice-based health care assistant can reduce depression symptoms and improve the process of care for patients with major depression in small primary care practices.

METHODS

We designed a pragmatic, cluster randomized, controlled trial that used practice as the unit of randomization to avoid contamination (22). The institutional review board of Goethe University Frankfurt am Main, Frankfurt am Main, Germany, approved the study protocol on 25 April 2005 (23). We used written consent procedures for family physicians and patients. We recruited practices between February 2005 and May 2005 and patients between May 2005 and July 2006. We carried out the intervention between June 2005 and August 2007. We completed the last follow-up for study patients in September 2007.

Setting and Participants

After calculating the sample size (24), we informed all 1600 family physicians registered with the medical association of the state of Hesse, Germany (mandatory registration), about the trial and invited them by mail to participate in information meetings. As the registration list presents only names and addresses, we checked inclusion criteria only for those who participated in the meetings. We stopped recruiting when 74 practices had enrolled, even though more practices were interested in participating. Inclusion criteria for the practices were acceptance of all major health plans (90% of patients are covered by this type of insurance) (25); provision of a primary care service, according to the definition of Starfield and colleagues (26); and management by a family physician. Patients were screened on special dates and referred to the trial by the family physician in the primary care practice. Inclusion criteria for patients were diagnosis of major depression with indication for any antidepressive treatment, age 18 to 80 years, access to a private telephone, ability to give informed consent, and ability to communicate in German. The diagnosis of major depression was based on a score of more than 9 points and a categorical diagnosis in the Patient Health Questionnaire-9 (PHQ-9) (27), and was confirmed by the family physician by using the checklists in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, and International Classification of Diseases, Tenth Edition. New patients were double-screened with the same procedure within 2 weeks. Exclusion criteria were confirmed pregnancy, severe alcohol or illicit drug consumption, or acute suicidal ideation assessed by the family physician.

Randomization and Interventions

The data safety and monitoring board stratified the practices according to the size of the city and performed computer-based randomization. Patient random assignment status was nested within the practice status. The data safety and monitoring board was responsible for allocation concealment by keeping the randomization results in a secure database. Because of the practice staff training required for the behavioral intervention, patients, health care assistants, family physicians, and researchers were not blinded to assignment once the trial was started.

We designed our case management intervention in accordance with the Chronic Care Model (28, 29), which emphasizes proactive support for the patient by the entire practice team. We trained 1 health care assistant from each practice assigned to the intervention group in 2 workshops (an 11-hour and a 6-hour workshop). This interactive training included information on depression, communication skills, telephone monitoring, and behavioral activation for the patient (30–32). The health care assistants contacted their patients by telephone twice a week in the first month and than once a month for the following 11 months. They monitored depression symptoms and adherence to medication by using the Depression Monitoring

List (33). Health care assistants also encouraged patients to follow self-management activities, such as medication adherence and activation for pleasant or social activities. The assistants provided this information to the family physician in a structured report that stratified the urgency of the contact by a robot scheme. Family physicians in both the intervention and control groups received training on evidence-based depression treatment guidelines (34). During the trial, other forms of disease or case management programs were uncommon in Germany (35). No study practice carried out case management for any other diseases.

Outcomes and Follow-up

Self-rating questionnaires were handed out to the patients at baseline and at 6 and 12 months after baseline. Patients filled in the questionnaires at home and sent them back to the practice. We collected the questionnaires in the practices and collected data from patient records (number of family physician and specialist contacts, hospitalization, and prescribed medication). Research staff carried out data input and management (36). Serious adverse events were reported to the data safety and monitoring board.

Clinical Outcomes

The primary outcome was depression symptoms, which we assessed by using the primary care–validated PHQ-9 (37). Each item is scored from 0 (not at all) to 3 (nearly every day), for a total score that ranges from 0 to 27 (high scores indicate more severe depression). We assessed response (50% improvement in PHQ-9 score) and remission status (PHQ-9 score <5) as secondary outcomes. We also report data for the following secondary outcomes: quality of life, patient assessment of chronic illness care, and medication adherence. We measured health-related quality of life by using the Medical Outcomes Study Short Form 36 (SF-36) (38, 39) and the EuroQol-5D (40). The SF-36 allows the calculation of scores for physical health and mental health (range, 0 to 100; higher scores indicate better status). The EuroQol-5D is a generic instrument that measures health-related quality of life with a visual analogue scale (range, 0 to 100; higher ratings indicate higher quality of life). We determined the number of physical comorbid conditions by counting the documented diagnoses from different diagnostic groups listed in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, and International Classification of Diseases, Tenth Edition, excluding all psychiatric diagnoses in the patient record. We assessed severity of chronic physical diseases by using the Chronic Disease Score, on the basis of prescription data from the patient record (41).

Process-of-Care Outcomes

We assessed the number of family physician and mental health specialist contacts, as well as prescriptions for antidepressant medications, by using data from patient records. We evaluated patient medication adherence by using a modified Morisky patient self-report scale (42), in which patients are scored from 0 to 3 on the basis of their answers to the following 3 questions (coded higher values indicate higher adherence): Did you ever forget to take your medicine during the last 2 weeks? During the last 2 weeks, did you sometimes stop taking your medicine when you felt better? During the last 2 weeks, did you stop taking your medicine when you felt worse?

We assessed patient satisfaction with primary care services by using the European Task Force on Patient Evaluations of General Practice Care instrument. This instrument has 23 items, and responses are given on a 5-point scale from 1 (excellent) to 5 (poor) (43). We used the Patient Assessment of Chronic Illness Care (PACIC) scale to assess patient perception of the depression management support provided by the primary care practice team (44, 45). This scale comprises 20 items, each of which can be scored on a 5-point scale from 1 (almost never) to 5 (almost always). The PACIC scale comprises 5 subscales that represent the key components of the Chronic Care Model: patient activation, delivery system, goal setting, problem-solving, and follow-up. The PACIC scale can also be calculated as an overall score.

Statistical Analysis

We based our prospective sample size calculation on cluster randomization (23). We analyzed PHQ-9 scores by fitting a linear mixed-effects model to data from patients with a PHQ-9 value at baseline and at least 1 further follow-up (modified intention-to-treat analysis), as others have done (46, 47).

As missing scale values were present at 6 and 12 months, we used a 2-level model with random effects for follow-up assessments at both the practice and patient level. We allowed random practice-level effects and random patient-level effects to be correlated, which resulted in a 6-parameter covariance structure. We entered treatment group indicator and baseline PHQ-9 scores as fixed effects. We imputed missing item values for PHQ-9 and PACIC scores separately by treatment group and data assessment by using the normal theory maximum likelihood method together with the expectation–maximization algorithm. Imputation was based on SAS procedure MI (SAS System for Windows, version 9.1, SAS Institute, Cary, North Carolina). We managed missing item-level values for the SF-36 as prescribed in the SF-36 manual.

We analyzed secondary outcomes in a similar manner. We performed the mixed-model analysis by using the SAS PROC MIXED procedure and fit generalized estimating equation models by using the SAS GENMOD procedure. We performed basic analyses by using Statistical Package for Social Sciences for Windows, version 15 (SPSS, Chicago, Illinois).

To assess bias due to nonignorable missing data (48), we performed a Bayesian sensitivity analysis (49) under the

missing-at-random assumption for all available data on the 626 included patients, with model-based estimates of response and remission proportions. For a sensitivity analysis under nonignorable missing data scenarios, we further supplemented the model with 2 logistic models for missingness at both follow-up assessments. We allowed the probability of missingness to depend on all previous PHQ-9 scores, with separate parameters for each treatment group; practice; and the current PHQ-9 score, regardless of whether it was observed. The parameters for the current score were fixed (5 parameter sets, assuming an odds ratio of 0.5, 1, or 2, for a difference of 2 SDs). For all other parameters, we specified noninformative priors. We used a Markov Chain Monte Carlo simulation on WinBUGS software (Imperial College and Medical Research Council, Cambridge, United Kingdom) for estimation (50) (**Appendix**, available at www.annals.org).

All *P* values were 2-sided and reported as being statistically significant on the basis of a significance level of 0.05. We performed no interim analyses.

Role of the Funding Source

The German Ministry of Education and Research provided funding for this trial. The funding source did not have any role in the study design; data collection, analysis, or interpretation; or the decision to submit the findings for publication.

RESULTS

We checked 93 practices to recruit the required sample of practices. We excluded 19 practices before patient enrollment because they did not meet inclusion criteria (8 practices) or could not participate because of time constraints (11 practices). One practice, with 9 patients, dropped out of the intervention group during the trial; no practice assigned to usual care crossed over to case management or vice versa (**Figure**). All 3051 patients were referred to the study for screening by their family physician. Of these, 1434 patients were not clinically depressed at screening (PHQ-9 score ≤9), whereas 1617 (946 control patients and 671 intervention recipients) had positive depression screening results. A diagnosis of major depression was confirmed for 820 patients (428 control patients and 392 intervention recipients) by PHQ-9 score, the additional categorical criteria in the PHQ-9, and a structured clinical interview by the family physician. We included 626 of these patients (73.3%; 310 intervention recipients and 316 control patients) in the study; the remaining 194 patients (23.7%; 82 intervention recipients and 112 control patients) did not participate. We enrolled a mean of 8.1 patients (SD, 2.7) from each control group practice and 8.9 patients (SD, 2.6) from each intervention group practice. We collected follow-up data from 84.8% of the patients at 6 months and 84.2% at 12 months. We based our modified intention-to-treat analysis on 555 patients. The enrolled patients differed slightly in depression severity from the nonparticipants at baseline (mean PHQ-9 score, 17.3 in study sample vs. 17.0 in nonparticipants). Enrolled patients were also slightly older than nonparticipants (50.7 years vs. 46.2 years).

Demographic and Clinical Characteristics

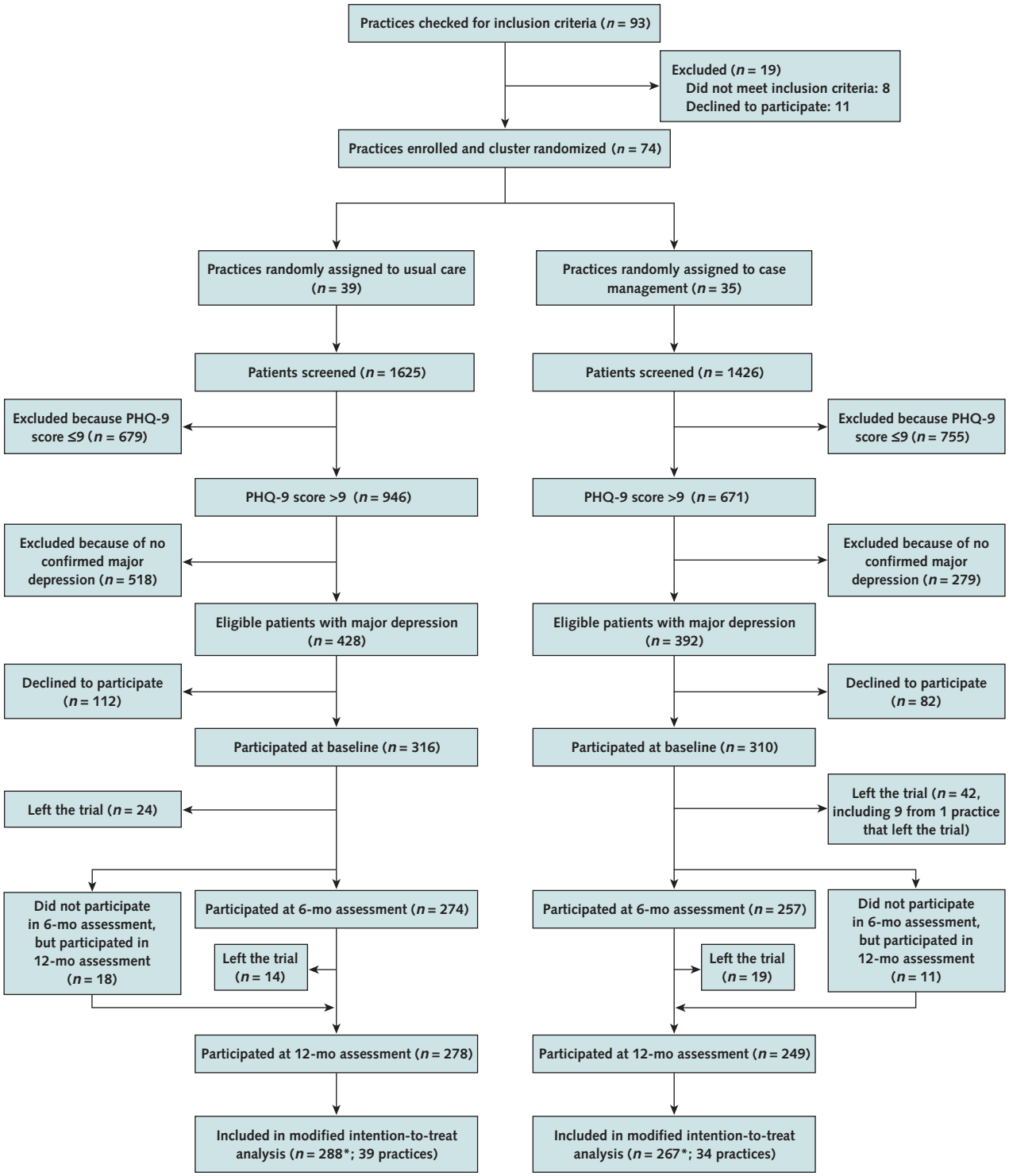
Most practices (100% in the intervention group vs. 92.3% in the control group) were self-owned by 1 or 2 physicians, and one third (32.4% vs. 46.2%) were located in a rural area. The practices had been run by family physicians for a mean of 13.4 years (SD, 9.3) in the intervention group and 10.7 years (SD, 7.7) in the control group. Health care assistants in the intervention group reported a mean of 17.3 years (SD, 11.2) of working experience, compared with a mean of 18.8 years (SD, 10.3) for those in the control group (**Table 1**). Sociodemographic and clinical characteristics ascertained at baseline seemed similar for both groups (**Table 2**). Overall, 76.4% of the patients were women and 45.9% were employed. Most patients (88.8% in the intervention group vs. 88.5% in the control group) had additional physical comorbid conditions, with a mean physical Chronic Disease Score of 1.29 (SD, 2.05) in the intervention group and 1.20 (SD, 1.98) in the control group. The percentage of patients receiving depression treatment (depression diagnosis known by the family physician before starting the trial) was 75.9%. The mean PHQ-9 depression score was 17.43 (SD, 3.60) in the intervention group and 17.17 (SD, 3.51) in the control group. The study intervention was realized as recommended in the study protocol. On average, patients in the intervention group were contacted by telephone 14 times during the study period. The mean duration of these interviews was 12 minutes.

Clinical Outcomes

Our intervention may have improved depression symptoms. Intervention recipients had significantly lower mean PHQ-9 scores than control patients after 6 months (11.9 vs. 13.2) and 12 months (10.7 vs. 12.1) (**Table 3**). After 12 months, intervention recipients also showed a lower mean value in the primary outcome (−1.41 [95% CI, −2.49 to −0.33]; $P = 0.014$) than control patients. At 12 months, intervention recipients had a higher depression treatment response rate than control patients (41.2% vs. 27.3%; $P = 0.003$; mean difference, 13.9 percentage points [CI, 4.8 to 22.9 percentage points]). At 12 months, the trend in the remission rate favored the intervention group (15.7% vs. 10.7%; $P = 0.057$; mean difference, 5.0 percentage points [CI, −0.3 to 10.4 percentage points]).

Our Bayesian sensitivity analysis revealed stable estimates under unfavorable scenarios. Under the missing-at-random scenario, mean PHQ-9 score difference at 12 months was −1.41 (95% credible interval [CrI], −2.48 to −0.35), the difference in proportions of response was 9.0 percentage points (95% CrI, 2.2 to 15.8 percentage points), and the difference in proportions of remission was

Figure. **Study flow diagram.**



PHQ-9 = Patient Health Questionnaire-9.
* Modified intention-to-treat analysis included patients with the primary outcome (PHQ-9 score) available at baseline and at least 1 follow-up assessment (6-mo or 12-mo assessment).

Table 1. **Practice, Family Physician, and Health Care Assistant Characteristics**

Characteristic*	Intervention Group (*n* = 34)	Control Group (*n* = 39)
Practices		
Maximum of 2 clinicians, *n (%)*	34 (100.0)	36 (92.3)
Location, *n (%)*		
Rural	11 (32.4)	18 (46.2)
Urban†	23 (67.6)	21 (53.8)
Panel size‡		
Mean patients per 3 months (SD), *n*	1065 (433)	1051 (435)
Age of patients within practice, %		
<18 y	13.5	12.6
18–65 y	51.3	56.3
>65 y	36.6	31.1
Family physicians		
Women, *n (%)*	16 (47.1)	17 (43.6)
Mean years at this practice (SD)	13.4 (9.3)	10.7 (7.7)
Health care assistants		
Mean years of job experience (SD)	17.3 (11.2)	18.8 (10.3)

* As contributing to the modified intention-to-treat analysis.
† Refers to a town with >50 000 inhabitants.
‡ In Germany, panel size is given as number of patient registrations in a practice in 3 months.

5.4 percentage points (CrI, 1.2 to 9.7 percentage points). Under the worst-case scenario, the respective results were −0.95 percentage points (CrI, −2.02 to 0.12 percentage points), 6.0 percentage points (CrI, −0.8 to 12.8 percentage points), and 3.7 percentage points (CrI, −0.5 to 7.9 percentage points).

The EuroQol-5D and SF-36 scores did not significantly differ between groups.

Process-of-Care Outcomes

Process-of-care outcomes differed at 12 months, with increased patient adherence to antidepressant medication (2.70 vs. 2.53; $P = 0.042$; mean difference, 0.17 [CI, 0.01 to 0.34]) and a more favorable patient assessment of the quality of chronic illness care (3.41 vs. 3.11; $P = 0.011$; mean difference, 0.31 [CI, 0.07 to 0.54]) among intervention recipients than among control patients (**Table 4**). Intervention recipients also gave higher ratings than control patients in delivery system (3.60 vs. 3.38; $P = 0.028$; mean difference, 0.22 [CI, 0.02 to 0.42]), goal setting (3.21 vs. 2.74; $P = 0.002$; mean difference, 0.48 [CI, 0.18 to 0.77]), and problem solving (3.82 vs. 3.56; $P = 0.034$; mean difference, 0.27 [CI, 0.02 to 0.51]). The groups did not differ in number of family physician or mental health specialist contacts or in use of prescribed antidepressant medications. European Task Force on Patient Evaluations of General Practice Care instrument ratings also did not differ between the intervention and control groups (data not shown).

DISCUSSION

Our results suggest that case management provided by primary care practice–based health care assistants may reduce depression symptoms and improve the process of depression care for patients with major depression. Most previously published evidence (8, 10) stems from academic or

Table 2. **Patient Characteristics at Baseline**

Characteristic	Intervention Group (*n* = 267)	Control Group (*n* = 288)	*P* Value
Sociodemographic characteristics*			
Mean age (SD), *y*	51.70 (14.05)	50.53 (14.32)	0.48†
Female, *n (%)*	200 (74.9)	224 (77.8)	0.51‡
Married, *n (%)*	153 (57.3)	150 (52.1)	0.26‡
Employed, *n (%)*	115 (43.1)	140 (48.6)	0.180‡
Clinical characteristics			
Received diagnosis of physical comorbid condition, *n (%)*	237 (88.8)	256 (88.5)	0.93‡
Mean physical comorbid conditions (SD), *n*	3.07 (2.24)	3.00 (2.27)	0.84†
Mean Chronic Disease Score (SD)§	1.29 (2.05)	1.20 (1.98)	0.56†
Receiving maintenance treatment, *n (%)*	205 (76.8)	216 (75.0)	0.72‡
Depression			
Mean PHQ-9 score (SD)	17.43 (3.60)	17.17 (3.51)	0.57†
Quality of life			
Mean SF-36 physical health score (SD)	40.34 (10.92)	40.88 (11.46)	0.64†
Mean SF-36 mental health score (SD)	28.35 (9.65)	27.56 (10.74)	0.59†
Mean EQ-5D score (SD)	45.82 (17.39)	46.19 (20.32)	0.80†

EQ-5D = EuroQol-5D; PHQ-9 = Patient Health Questionnaire-9; SF-36 = Medical Outcomes Study Short Form 36.
* Estimates are shown for the 555 patients included in the main analyses.
† Estimates and tests are based on a linear mixed model adjusted for cluster effects.
‡ Test is based on a logistic generalized estimating equation model adjusted for intracluster correlation. We adjusted proportions for cluster effects by using a linear mixed model.
§ Excluding psychiatric diagnoses.

Table 3. **Clinical Outcomes After 12 Months**

Outcome	Intervention Group (n = 267)	Control Group (n = 288)	*P* Value	Mean Difference (95% CI)
Depression				
Mean PHQ-9 score (SD)	10.72 (5.43)	12.13 (5.60)	0.014*	−1.41 (−2.49 to −0.33)
Response, *n/n (%)*†	100/242 (41.2)	74/272 (27.3)	0.003‡	13.9 (4.8 to 22.9)
Remission, *n/n (%)*§	38/242 (15.7)	29/272 (10.7)	0.057‡	5.0 (−0.3 to 10.4)
Quality of life				
Mean SF-36 physical health score (SD)‖	41.49 (11.40)	43.23 (12.09)	0.170¶	−1.77 (−4.29 to 0.75)
Mean SF-36 mental health score (SD)‖	35.58 (12.39)	33.24 (12.57)	0.051¶	2.45 (−0.01 to 4.90)
Mean EQ-5D score (SD)	55.30 (20.55)	53.86 (21.76)	0.52¶	1.33 (−2.71 to 5.37)

EQ-5D = EuroQol-5D; PHQ-9 = Patient Health Questionnaire-9; SF-36 = Medical Outcomes Study Short Form 36.
* Based on a 2-level linear mixed model for respective outcome (T1 and T2) adjusted for intracluster correlation and baseline depression.
† Defined as improvement of PHQ-9 score by 50%.
‡ Test is based on a logistic generalized estimating equation model adjusted for intracluster correlation. We adjusted proportions for cluster effects by using a linear effects model.
§ Defined as a PHQ-9 score <5.
‖ Based on an intervention group of 201 patients and a control group of 224 patients.
¶ Based on a linear mixed model adjusted for intracluster correlation but not for baseline depression.

highly structured (HMO) settings, whereas our trial highlights the benefits of a simple, depression case management intervention in a nonacademic, nonstructured, small primary care practice setting and thus closes a research gap (7). Evidence on the importance of a background in mental health for health professionals is heterogeneous. Recent reviews (51, 52) indicate that case management has positive effects when delivered by highly qualified nurses and psychologists. Our trial points to the potential role that less highly trained practice staff may play in improving depression care. Because health care resources are limited, the need for cost-effective health care interventions in primary care settings is increasing.

The effect size (Cohn d = 0.26) was small but similar to other case management trials on depression (53). We did not expect stronger effects because the intervention was in addition to usual care. The number of family physician and mental health specialist contacts and the prescribed antidepressive medication did not differ between groups. In addition to the effects on depression symptoms, our trial supports positive effects on process-of-care outcomes (PACIC and patient adherence to medication regimens). Because those outcomes are associated with actions taken by the health care assistants, our intervention's mechanism of action may be related to the supportive activity of the health care assistant. This case management intervention

Table 4. **Process-of-Care Outcomes After 12 Months**

Outcome	Intervention Group (n = 246)	Control Group (n = 274)	*P* Value	Mean Difference (95% CI)
Visits and medication				
Mean visits to primary care physician within last 6 months (SD), *n**	6.15 (2.44)	5.79 (2.66)	0.63†	1.06 (−0.84 to 1.34)
Mean visits to psychiatric specialist within last 6 months (SD), *n*	2.1 (4.4)	1.8 (3.2)	0.86†	0.26 (−0.48 to 1.00)
Antidepressant medications, *n (%)*	142 (57.9)	158 (57.6)	0.99‡	0.1 (−9.8 to 10.0)
Adherence				
Mean modified Morisky score (SD)§	2.70 (0.63)	2.53 (0.83)	0.042†	0.17 (0.01 to 0.34)
Assessment of chronic illness care				
Mean PACIC overall score (SD)‖	3.41 (0.80)	3.11 (0.76)	0.011†	0.31 (0.07 to 0.54)
Mean patient activation subscore (SD)	3.77 (1.04)	3.54 (1.06)	0.060	0.24 (−0.01 to 0.49)
Mean delivery system subscore (SD)	3.60 (0.79)	3.38 (0.78)	0.028	0.22 (0.02 to 0.42)
Mean goal setting subscore (SD)	3.21 (0.99)	2.74 (0.97)	0.002	0.48 (0.18 to 0.77)
Mean problem solving subscore (SD)	3.82 (0.89)	3.56 (0.95)	0.034	0.27 (0.02 to 0.51)
Mean follow-up subscore (SD)	2.96 (0.11)	2.68 (0.96)	0.063	0.28 (−0.02 to 0.58)

PACIC = Patient Assessment of Chronic Illness Care.
* We set the ratio of intervention to control to 0.5 for patients without any visit. Based on an intervention group of 240 patients and a control group of 268 patients.
† Estimates and tests are based on a linear mixed model adjusted for cluster effects.
‡ Test is based on a logistic generalized estimating equation model adjusted for intracluster correlation. We adjusted proportions for cluster effects by using a linear mixed model.
§ Analyses included only the 142 intervention patients and 158 control patients who were receiving medication.
‖ Based on an intervention group of 199 patients and a control group of 236 patients.

may have enhanced patient activation and adherence through improved practice team support. The practice teams' familiarity with their patients and long-time continuity of the patient–provider relationships, which are typical for small primary care practices, may have played a role in achieving the positive results we observed (54).

We acknowledge a potential patient selection bias that resulted from the physicians' selection of patients for the trial. We did not enroll all eligible patients with depression in the practices, and enrolled patients were slightly more depressed than nonenrolled patients. In addition, the extended patient recruitment period of 12 months meant that we had to randomly assign practices and train health care assistants before the last patient was enrolled—as other studies have done (17). However, because patient characteristics did not differ among the enrolled groups at baseline, we do not believe the enrollment was skewed. Study patients were similar to other primary care patients with depression in terms of major sociodemographic characteristics, symptom severity, and physical comorbid conditions (31, 55). More than half of enrolled patients were unemployed, which suggests a lower socioeconomic status, and most patients had 1 or more physical diseases. Both characteristics have been described as poor prognostic factors for depression, which may have influenced the depression treatment response rates (56–60). The measured mean depression symptoms score of 17 (moderately severe) on the PHQ-9 scale was similar to that of the primary care validation study (61). We enrolled more patients who were already receiving depression treatment than did other trials (15, 17), which may have reduced observed intervention effects.

Because our study protocol did not permit providing financial incentives for patients in this trial, a notable and unbalanced percentage of loss to follow-up occurred at 12 months (19.7% in the intervention group and 12.0% in the control group); this is a recurrent problem of trials in primary care. Dietrich and colleagues (16) reported that 80.7% of the patients in their control group completed the follow-up assessment, compared with 79.9% in their intervention group. Follow-up data from Dobscha and colleagues (17) (84% of participants at 3 months, 84% at 6 months, and 85% at 12 months) show the general difficulties of loss to follow-up in multicenter primary care research. Our sensitivity analyses show that the effects of the intervention on the primary outcome remain statistically significant and stable under unfavorable assumptions about nonparticipation at follow-up assessments. To check whether our results on response and remission rates were sensitive to missing value mechanisms, we performed a Bayesian analysis on PHQ-9 score data from all assessments and predicted response and remission rates. We reproduced the results of the primary analysis.

Although the random assignment to intervention and control groups yielded homogeneous results, our findings are not generalizable or applicable to all patients with depression in primary care practice because patients may have other types of depressive disorders, such as minor depression, or be less responsive to this type of intervention. Patients who voluntarily participated in such an intervention might be more similar to those we included. We also may have overestimated the effects because we could not use blinded assessment of the primary outcome—most family physicians would not allow independent contact calls to their mental health patients by research staff (62).

In terms of panel size, age, and the clinical experience of the staff, enrolled practices were representative of German primary care practices (63) and similar to practices enrolled in other case management trials (64). Because of the similarities between German practices and U.S. primary care practices outside HMO settings (such as a high percentage of solo or 2-person partnerships), our findings should be relevant to both countries.

Our findings suggest that health care assistants without specialized training in mental health care can be effective in primary care, as recently proposed (20). Assistants can support self-management strategies by monitoring symptoms regularly and providing behavioral activation to patients with depression. We designed our intervention to manage patients with depression without making excessive demands on the limited resources of the small primary care practices who treat most such patients (65, 66).

Our findings have supported innovations in the German health care delivery system. The government and major health insurance companies have reformed the reimbursement system. In 2009, health insurance companies began funding clinical work provided by practice-based health care assistants, such as systematic telephone monitoring (67). Our trial intervention has been adopted for national health care assistant training programs (68). These financial incentives may lead to a broader implementation of the intervention.

Our study suggests that case management provided by health care assistants in small primary care practices may reduce depression symptoms and improve the process of care for patients with major depression relative to usual care. Involving primary care practice–based health care assistants in patient care for depression may improve depression outcomes in primary care settings.

From Friedrich Schiller University, Jena, Germany; Goethe University Frankfurt am Main, Frankfurt am Main, Germany; Group Health Center for Health Studies, Seattle, Washington; University of Zürich, Zürich, Switzerland; and University Mainz, Mainz, Germany.

Acknowledgment: PRoMPT has received the German Research Award for Primary Care—Dr. Lothar–Beyer Award for 2008. The authors thank the jury and all contributing primary care practice teams.

Grant Support: By the German Ministry of Education and Research (grant 01GK0302).

Potential Financial Conflicts of Interest: None disclosed.

Reproducible Research Statement: *Study protocol:* Available from Dr. Gensichen (e-mail, jochen.gensichen@med.uni-jena.de). *Statistical code:* Available from Dr. König (e-mail, koenig@imbei.uni-mainz.de). *Data set:* Not available.

Requests for Single Reprints: Jochen Gensichen, MD, MPH, MA, Institute for General Practice, Friedrich Schiller University/University Hospital Jena, Bachstrasse 18, D-07743 Jena, Germany; e-mail, jochen.gensichen@med.uni-jena.de.

Current author addresses and author contributions are available at www.annals.org.

References

1. **Ustün TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ.** Global burden of depressive disorders in the year 2000. Br J Psychiatry. 2004;184:386-92. [PMID: 15123501]
2. **Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, et al.** The economic burden of depression in the United States: how did it change between 1990 and 2000? J Clin Psychiatry. 2003;64:1465-75. [PMID: 14728109]
3. **Stewart WF, Ricci JA, Chee E, Hahn SR, Morganstein D.** Cost of lost productive work time among US workers with depression. JAMA. 2003;289:3135-44. [PMID: 12813119]
4. **Katon W, Schulberg H.** Epidemiology of depression in primary care. Gen Hosp Psychiatry. 1992;14:237-47. [PMID: 1505745]
5. **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;62:629-40. [PMID: 15939840]
6. **Fernández A, Haro JM, Martinez-Alonso M, Demyttenaere K, Brugha TS, Autonell J, et al.** Treatment adequacy for anxiety and depressive disorders in six European countries. Br J Psychiatry. 2007;190:172-3. [PMID: 17267936]
7. **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;166:2314-21. [PMID: 17130383]
8. **Williams JW Jr, Gerrity M, Holsinger T, Dobscha S, Gaynes B, Dietrich A.** Systematic review of multifaceted interventions to improve depression care. Gen Hosp Psychiatry. 2007;29:91-116. [PMID: 17336659]
9. **Gilchrist G, Gunn J.** Observational studies of depression in primary care: what do we know? BMC Fam Pract. 2007;8:28. [PMID: 17493280]
10. **Kates N, Mach M.** Chronic disease management for depression in primary care: a summary of the current literature and implications for practice. Can J Psychiatry. 2007;52:77-85. [PMID: 17375862]
11. **Christensen H, Griffiths KM, Gulliver A, Clack D, Kljakovic M, Wells L.** Models in the delivery of depression care: a systematic review of randomised and controlled intervention trials. BMC Fam Pract. 2008;9:25. [PMID: 18454878]
12. **Casalino LP, Devers KJ, Lake TK, Reed M, Stoddard JJ.** Benefits of and barriers to large medical group practice in the United States. Arch Intern Med. 2003;163:1958-64. [PMID: 12963570]
13. **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;36:7-14. [PMID: 16356292]
14. **Von Korff M, Goldberg D.** Improving outcomes in depression [Editorial]. BMJ. 2001;323:948-9. [PMID: 11679372]
15. **Unützer J, Katon W, Callahan CM, Williams JW Jr, Hunkeler E, Harpole L, et al; IMPACT Investigators.** Improving Mood-Promoting Access to Collaborative Treatment. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA. 2002;288:2836-45. [PMID: 12472325]
16. **Dietrich AJ, Oxman TE, Williams JW Jr, Schulberg HC, Bruce ML, Lee PW, et al.** Re-engineering systems for the treatment of depression in primary care: cluster randomised controlled trial. BMJ. 2004;329:602. [PMID: 15345600]
17. **Dobscha SK, Corson K, Hickam DH, Perrin NA, Kraemer DF, Gerrity MS.** Depression decision support in primary care: a cluster randomized trial. Ann Intern Med. 2006;145:477-87. [PMID: 17015865]
18. **Bodenheimer T, Berenson RA, Rudolf P.** The primary care-specialty income gap: why it matters. Ann Intern Med. 2007;146:301-6. [PMID: 17310054]
19. **American Academy of Family Physicians.** Practice profile of family physicians by family medicine residency completion. Leawood, KS: American Academy of Family Physicians; 2009. Accessed at www.aafp.org/online/en/home/aboutus/specialty/facts/4.html on 21 July 2009.
20. **Bosley S, Dale J.** Healthcare assistants in general practice: practical and conceptual issues of skill-mix change. Br J Gen Pract. 2008;58:118-24. [PMID: 18307856]
21. **German Medical Association.** [Health care assistant] (German). Berlin, Germany: German Med Assoc; 2009. Accessed at www.baek.de/page.asp?his=1.99.3478.3860 on 21 July 2009.
22. **Richards DA, Lovell K, Gilbody S, Gask L, Torgerson D, Barkham M, et al.** Collaborative care for depression in UK primary care: a randomized controlled trial. Psychol Med. 2008;38:279-87. [PMID: 17803837]
23. **Gensichen J, Torge M, Peitz M, Wendt-Hermainski H, Beyer M, Rosemann T, et al.** Case management for the treatment of patients with major depression in general practices—rationale, design and conduct of a cluster randomized controlled trial—PRoMPT (PRimary care Monitoring for depressive Patient's Trial) [ISRCTN66386086]—study protocol. BMC Public Health. 2005;5:101. [PMID: 16207375]
24. **Campbell M, Grimshaw J, Steen N.** Sample size calculations for cluster randomised trials. Changing Professional Practice in Europe Group (EU BIOMED II Concerted Action). J Health Serv Res Policy. 2000;5:12-6. [PMID: 10787581]
25. **German Ministry of Health.** [Statistical pocket book health] (German). Berlin, Germany: German Ministry of Health; 2005.
26. **Starfield B, Shi L, Macinko J.** Contribution of primary care to health systems and health. Milbank Q. 2005;83:457-502. [PMID: 16202000]
27. **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;282:1737-44. [PMID: 10568646]
28. **Wagner EH, Austin BT, Von Korff M.** Improving outcomes in chronic illness. Manag Care Q. 1996;4:12-25. [PMID: 10157259]
29. **Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A.** Improving chronic illness care: translating evidence into action. Health Aff (Millwood). 2001;20:64-78. [PMID: 11816692]
30. **Katon W, Rutter C, Ludman EJ, Von Korff M, Lin E, Simon G, et al.** A randomized trial of relapse prevention of depression in primary care. Arch Gen Psychiatry. 2001;58:241-7. [PMID: 11231831]
31. **Anderson B.** Collaborative care and motivational interviewing: improving depression outcomes through patient empowerment interventions. Am J Manag Care. 2007;13:S103-6. [PMID: 18041866]
32. **Bilsker D, Goldner EM, Jones W.** Health service patterns indicate potential benefit of supported self-management for depression in primary care. Can J Psychiatry. 2007;52:86-95. [PMID: 17375863]
33. **Gensichen J, Peitz M, Torge M, Mosig-Frey J, Wendt-Hermainski H, Rosemann T, et al.** [The "Depression Monitoring list" (DeMoL) with integrated PHQ-D-Rationale and design of a tool for the case management for depression in primary care]. Z Arztl Fortbild Qualitatssich. 2006;100:375-82. [PMID: 16955623]
34. **Van Marwijk HWJ, Grundmeiyer HGLM, Bijl D, Gelderen M, van Haan M, Weel-Baumgarten EM, et al.** [The Netherlands General Physician Guidelines Depression]. Huisarts Wet. 2003;46:614-23.
35. **Busse R.** Disease management programs in Germany's statutory health insurance system. Health Aff (Millwood). 2004;23:56-67. [PMID: 15160803]
36. **Wendt-Hermainski H, Peitz M, Torge M, Beyer M, Gerlach FM, Gensichen J.** [Considerations of data management in clinical trials - good clinical practice in general practice research - the PRoMPT-project]. Z Allgemeinmed. 2005;81:565-70.
37. **Löwe B, Spitzer RL, Gräfe K, Kroenke K, Quenter A, Zipfel S, et al.** Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. J Affect Disord. 2004;78:131-40. [PMID: 14706723]
38. **Ware JE, Snow KK, Kosinski M, Gandek B.** SF-36 Health Survey Manual and Interpretation Guide. Boston: New England Medical Center, The Health Institute; 1993.

39. **Bullinger M, Kirchberger I, Ware J.** Der deutsche SF-36 Health Survey. [The German SF-36 Health Survey—Translation and psychometric testing of a generic instrument for the assessment of health-related quality of life] (German). Zeitschrift für Gesundheitswissenschaften. 1995;3:21-36.
40. **Greiner W, Claes C, Busschbach JJ, Graf von der Schulenburg JM.** Validating the EQ-5D with time trade off for the German population. Eur J Health Econ. 2004;6:124-30.
41. **Von Korff M, Wagner EH, Saunders K.** A chronic disease score from automated pharmacy data. J Clin Epidemiol. 1992;45:197-203. [PMID: 1573438]
42. **Morisky DE, Green LW, Levine DM.** Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care. 1986;24:67-74. [PMID: 3945130]
43. **Grol R, Wensing M, Mainz J, Jung HP, Ferreira P, Hearnshaw H, et al; European Task Force on Patient Evaluations of General Practice Care (EUROPEP).** Patients in Europe evaluate general practice care: an international comparison. Br J Gen Pract. 2000;50:882-7. [PMID: 11141874]
44. **Glasgow RE, Wagner EH, Schaefer J, Mahoney LD, Reid RJ, Greene SM.** Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). Med Care. 2005;43:436-44. [PMID: 15838407]
45. **Rosemann T, Laux G, Droesemeyer S, Gensichen J, Szecsenyi J.** Evaluation of a culturally adapted German version of the Patient Assessment of Chronic Illness Care (PACIC 5A) questionnaire in a sample of osteoarthritis patients. J Eval Clin Pract. 2007;13:806-13. [PMID: 17824876]
46. **Thiam S, LeFevre AM, Hane F, Ndiaye A, Ba F, Fielding KL, et al.** Effectiveness of a strategy to improve adherence to tuberculosis treatment in a resource-poor setting: a cluster randomized controlled trial. JAMA. 2007;297:380-6. [PMID: 17244834]
47. **Loeb M, Carusone SC, Goeree R, Walter SD, Brazil K, Krueger P, et al.** Effect of a clinical pathway to reduce hospitalizations in nursing home residents with pneumonia: a randomized controlled trial. JAMA. 2006;295:2503-10. [PMID: 16757722]
48. **Daniels MJ, Hogan JW.** Missing Data in Longitudinal Studies. New York: Chapman & Hall; 2008.
49. **Carpenter J, Pocock S, Lamm CJ.** Coping with missing data in clinical trials: a model-based approach applied to asthma trials. Stat Med. 2002;21:1043-66. [PMID: 11933033]
50. **Lunn DJ, Thomas A, Best N, Spiegelhalter D.** WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. Stat and Comput. 2000;10:325-37.
51. **Christensen H, Griffiths KM, Gulliver A, Clack D, Kljakovic M, Wells L.** Models in the delivery of depression care: a systematic review of randomised and controlled intervention trials. BMC Fam Pract. 2008;9:25. [PMID: 18454878]
52. **Bower P, Rowland N, Hardy R.** The clinical effectiveness of counselling in primary care: a systematic review and meta-analysis. Psychol Med. 2003;33:203-15. [PMID: 12622300]
53. **Wang PS, Simon GE, Avorn J, Azocar F, Ludman EJ, McCulloch J, et al.** Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial. JAMA. 2007;298:1401-11. [PMID: 17895456]
54. **Gensichen J, Jaeger C, Peitz M, Torge M, Güthlin C, Petersen J, et al.** Health care assistants in primary care depression management: role perception, burdening factors and disease conception. Ann Fam Med. [Forthcoming].
55. **Bramesfeld A, Grobe TG, Schwartz FW.** Who is diagnosed as suffering from depression in the German statutory health care system? An analysis of health insurance data. Eur J Epidemiol. 2007;22:397-403. [PMID: 17484028]
56. **Weich S, Churchill R, Lewis G, Mann A.** Do socio-economic risk factors predict the incidence and maintenance of psychiatric disorder in primary care? Psychol Med. 1997;27:73-80. [PMID: 9122310]
57. **Cole MG, Bellavance F, Mansour A.** Prognosis of depression in elderly community and primary care populations: a systematic review and meta-analysis. Am J Psychiatry. 1999;156:1182-9. [PMID: 10450258]
58. **Fortin M, Bravo G, Hudon C, Vanasse A, Lapointe L.** Prevalence of multimorbidity among adults seen in family practice. Ann Fam Med. 2005;3:223-8. [PMID: 15928225]
59. **Fulton MM, Allen ER.** Polypharmacy in the elderly: a literature review. J Am Acad Nurse Pract. 2005;17:123-32. [PMID: 15819637]
60. **Wells KB, Golding JM, Burnam MA.** Psychiatric disorder in a sample of the general population with and without chronic medical conditions. Am J Psychiatry. 1988;145:976-81. [PMID: 2969199]
61. **Löwe B, Spitzer RL, Zipfel S, Herzog W.** [PRIME MD Patient Health Questionnaire (PHQ)-German version]. 2nd ed. Karlsruhe, Germany: Pfizer; 2002.
62. **Linden M.** Psychiatric disorders in primary care. In: Henn F, Satorius N, Helmchen H, Lauter H, eds. Contemporary Psychiatry. vol 1. Berlin, Germany: Springer; 2001.
63. **Wittchen HU, Kessler RC, Beesdo K, Krause P, Höfler M, Hoyer J.** Generalized anxiety and depression in primary care: prevalence, recognition, and management. J Clin Psychiatry. 2002;63 Suppl 8:24-34. [PMID: 12044105]
64. **Rosemann T, Körner T, Wensing M, Gensichen J, Muth C, Joos S, et al.** Rationale, design and conduct of a comprehensive evaluation of a primary care based intervention to improve the quality of life of osteoarthritis patients. The PraxArt-project: a cluster randomized controlled trial [ISRCTN87252339]. BMC Public Health. 2005;5:77. [PMID: 16029495]
65. **Thielke S, Vannoy S, Unützer J.** Integrating mental health and primary care. Prim Care. 2007;34:571-92, vii. [PMID: 17868760]
66. **Katon WJ, Seelig M.** Population-based care of depression: team care approaches to improving outcomes. J Occup Environ Med. 2008;50:459-67. [PMID: 18404019]
67. Sozialgesetzbuch [SGB] Fünftes Buch [V]—Gesetzliche Krankenversicherung [Social Security Code V] 20 December 1988, BGBl. I S. 2477, § 73b Hausarztzentrierte Versorgung [General practice-centered care].
68. **German Medical Association.** Curriculum Health Care Assistants: "Patient support and co-ordination". Berlin, Germany: German Med Assoc; 2007.

Current Author Addresses: Dr. Gensichen: Institute for General Practice, Friedrich Schiller University/Hospital Jena University, Bachstrasse 18, D-07740 Jena, Germany.
Dr. von Korff: Group Health Center for Health Studies, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101-1448.
Drs. Peitz, Muth, Güthlin, Torge, Petersen, and Gerlach and Mr. Beyer: Institute for General Practice, Goethe University Hospital, Theodor-Stern-Kai 7, D-60590 Frankfurt am Main, Germany.
Dr. Rosemann: Institute for General Practice, University of Zürich, Rämistrasse 100, CH-8091 Zürich, Switzerland.
Dr. König: Institute of Medical Biostatistics, Epidemiology and Informatics, University Mainz, Obere Zahlbacherstrasse 69, D 55131 Mainz, Germany.

Author Contributions: Conception and design: Gensichen J, von Korff M, Muth C, Peitz M, Beyer M, Güthlin C, Torge M, Rosemann T, Gerlach FM.
Data collection: Gensichen J, Muth C, Peitz M, Torge M, Beyer M.
Analysis and interpretation: Gensichen J, von Korff M, Muth C, Peitz M, Beyer M, Güthlin C, Torge M, Petersen JJ, Rosemann T, Gerlach FM, König J.
Statistical analysis: König J, Gensichen J.
Drafting of the article: Gensichen J, Güthlin C, König J.
Critical revision of the article for important intellectual content: Gensichen J, von Korff M, Muth C, Peitz M, Beyer M, Güthlin C, Torge M, Petersen JJ, Rosemann T, König J, Gerlach FM.
Final approval of the article: Gensichen J, von Korff M, Muth C, Peitz M, Beyer M, Güthlin C, Torge M, Petersen JJ, Rosemann T, König J, Gerlach FM.
Obtaining of funding: Gensichen J, Gerlach FM.

Appendix: Bayesian Sensitivity Analysis

Reasons for Sensitivity Analysis

Nonparticipation and loss to follow-up on the primary outcome can cause bias in treatment effect estimates, if missing data depend on variables that are not observed or considered in the analysis. A primary analysis method is unbiased if the probability of missingness depends on observed variables only (the missing-at-random assumption [49]). For our analysis of PHQ-9 scores, missingness may depend on treatment group, practice, baseline PHQ-9 score, and (for the 12-month assessment) the score and missingness indicator from the 6-month assessment. Our primary analysis should therefore prove to be fairly robust against missingness.

However, our analysis of clinical response and remission is prone to bias because we did not use the 6-month assessment, and the dependency of missing values at 12 months on observations at 6 months would be a nonignorable missing data mechanism (48).

We performed 2 types of sensitivity analysis. The first was to investigate whether the analysis of clinical response and remission end points is stable, when analyzed by a method that allows missingness to depend on PHQ-9 scores at baseline and 6 months. The second was to investigate the extent to which our results are influenced if missingness depends on the unobserved PHQ-9 score (the missing-not-at-random assumption).

Methods

For the first sensitivity analysis, we extended the model of our primary analysis to a Bayesian model by specifying a normal distribution for baseline PHQ-9 score and noninformative priors. We assessed the proportions of response and remission in the same model by expressing them as functions of the continuous model parameters under the normal distribution assumption (which would not have been feasible without Bayesian analysis).

Our second group of sensitivity analyses was similar to those described by Carpenter and colleagues (49). We further supplemented the Bayesian model by 2 logistic models for missingness at both follow-up assessments, denoted as selection models (48, 49). We allowed the probability of missingness to depend on all previous PHQ-9 scores, with separate parameters for each treatment group; practice; and the current PHQ-9 score, regardless of whether it was observed. The parameters for the current score were fixed (5 parameter sets, assuming an odds ratio of 0.5, 1, or 2, for a difference of 2 SDs). In doing so, we followed Carpenter and colleagues (49), who recommend, "The way in which the non-random mechanism operates should be simple yet scientifically plausible." For all other parameters, we specified noninformative priors.

We used a Markov Chain Monte Carlo simulation on WinBUGS software (50) for estimation, on the basis of available data from 626 patients. Each simulation used a burn-in of 5000 and 10 000 additional runs. We checked convergence by inspecting trajectories and autocorrelation diagrams.

Results and Discussion

Our Bayesian analysis under the assumption of an ignorable missing data mechanism reproduces our main mixed-model analysis (**Appendix Table 1**; compare with **Table 3**). Results concerning proportions of response and remission are similar and point in the same direction as our main analysis.

Appendix Table 2 presents the sensitivity analyses under the assumption of nonignorable missing data mechanisms. Treatment effects were stable whenever we assumed the same missing data mechanism in both treatment groups. The 12-month mean difference in PHQ-9 score was reduced from −1.41 to −0.95 (CrI, −2.02 to 0.12) under the most unfavorable scenario.

Appendix Table 1. Bayesian Analysis*

Depression Outcome	Intervention Group (n = 310)	Control Group (n = 316)	*P* Value†	Mean Difference (95% CrI)
Mean PHQ-9 score (SD)	10.74 (5.87)	12.13 (5.84)	0.027	−1.402 (−2.46 to 0.34)
Response, %‡	35.4	26.6	0.008	8.9 (2.1 to 16.6)
Remission, %§	16.5	11.2	0.009	5.3 (1.2 to 9.6)

CrI = credible interval; PHQ-9 = Patient Health Questionnaire-9.
* Based on the same linear mixed model as our primary analysis and all observed PHQ-9 score data at the baseline, 6-month, and 12-month assessments; 626 patients at 74 practices. We estimated response and remission rates on the basis of the Gaussian distribution model, fitted for the PHQ-9 scale to all available data.
† Defined as twice the posterior probability of a positive effect, which is meant as an analogue of the 2-sided *P* value of classical inferential statistics.
‡ Defined as improvement of PHQ-9 score by 50%.
§ Defined as a PHQ-9 score <5.

Appendix Table 2. Sensitivity Analysis Concerning Non–Missing-at-Random Scenarios

Odds Ratio of Intervention to Control Patients*	Mean Score Difference (95% CrI)		Response Rate at 12 Months (95% CrI), %	Remission Rate at 12 Months (95% CrI), %
	6 Months	12 Months		
1/1	−1.36 (−2.36 to −0.37)	−1.41 (−2.48 to −0.35)	9.0 (2.2 to 15.8)	5.4 (1.2 to 9.7)
0.5/0.5	−1.36 (−2.32 to −0.42)	−1.45 (−2.50 to −0.41)	9.3 (2.6 to 16.0)	5.7 (1.4 to 10.1)
0.5/2	−1.72 (−2.67 to −0.75)	−1.79 (−2.84 to −0.72)	11.3 (4.5 to 17.9)	6.8 (2.6 to 11.1)
2/0.5	−0.92 (−1.90 to 0.06)	−0.95 (−2.02 to 0.12)	6.0 (−0.8 to 12.8)	3.7 (−0.5 to 7.9)
2/2	−1.32 (−2.25 to −0.38)	−1.33 (−2.38 to −0.27)	8.3 (1.7 to 14.9)	4.9 (1.0 to 9.0)

CrI = credible interval.
* Missingness odds ratio at 6 and 12 months, per 2 SDs of Patient Health Questionnaire-9 score.