Depression Decision Support in Primary Care
A Cluster Randomized Trial

Steven K. Dobscha, MD; Kathryn Corson, PhD; David H. Hickam, MD, MPH; Nancy A. Perrin, PhD; Dale F. Kraemer, PhD; and Martha S. Gerrity, MD, PhD

Background: Intensive collaborative interventions improve depression outcomes, but the benefit of less intensive interventions is not clear.

Objective: To determine whether decision support improves outcomes for patients with depression.

Design: Clinician-level, cluster randomized, controlled trial.

Setting: 5 primary care clinics of 1 Veterans Affairs medical center.

Participants: 41 primary care clinicians, and 375 patients with depression (Patient Health Questionnaire [PHQ-9] depression scores of 10 to 25 or Hopkins Symptom Checklist-20 [SCL-20] scores \( \geq 1.0 \)).

Measurements: The primary outcome was change in depression score (SCL-20) at 6 and 12 months. Secondary outcomes were health-related quality-of-life (36-item Short Form for Veterans [SF-36V] score), patient satisfaction, antidepressant use, and health care utilization.

Intervention: Clinicians received depression education and were randomly assigned to depression decision support or usual care. The depression decision support team, which consisted of a psychiatrist and nurse, provided 1 early patient educational contact and depression monitoring with feedback to clinicians over 12 months.

Results: Although SCL-20 depression scores improved in both groups, the intervention had no effect compared with usual care. The difference in slopes comparing intervention and control over 12 months was 0.20 (95% CI, –0.37 to 0.78; \( P = 0.49 \)), which was neither clinically nor statistically significant. Changes in SF-36V scores also did not differ between groups. At 12 months, intervention patients reported greater satisfaction (\( P = 0.002 \)) and were more likely to have had at least 1 mental health specialty appointment (41.1% vs. 27.2%; \( P = 0.025 \)), to have received any antidepressant (79.3% vs. 69.3%; \( P = 0.041 \)), and to have received antidepressants for 90 days or more (76.2% vs. 61.6%; \( P = 0.008 \)).

Limitations: Usual care clinicians received depression education and had on-site mental health support, which may have mitigated intervention effectiveness.

Conclusions: Decision support improved processes of care but not depression outcomes. More intensive care management or specialty treatment may be needed to improve depression outcomes.

Depression is a common problem worldwide. According to projections from the World Health Organization (WHO), depression will be the second leading cause of disability in the developed world by 2020 (1). One in 10 primary care patients meets criteria for major depression (2), yet underrecognition and undertreatment are common (3, 4). Untreated depression is associated with increased deaths, adverse medical outcomes, deficits in function and well-being, and increased use of health services (3, 5–9).

Multifaceted, collaborative interventions have been shown to improve depression-related outcomes in primary care (10–22). These interventions include decision support for clinicians, self-management support for patients, clinical information systems modifications, and care management. Care management typically consists of patient education and activation, symptom and treatment adherence monitoring, and self-management reinforcement (23).

Most collaborative depression interventions have relied on intensive involvement of care managers and specialists, usually shifting responsibility and workload toward mental health or research teams. Clinical systems may not be capable of sustaining this level of intensity. We therefore developed a multifaceted depression decision support intervention, which was designed to optimize primary care clinicians’ abilities to treat depression without adding substantial new resources. We aimed to determine the effect of depression decision support on clinical outcomes and processes of care among patients with depression in a Veterans Affairs (VA) primary care setting.

Methods
Design Overview
Our clinician-level, cluster randomized, controlled trial studied depression decision support versus usual care. We randomly assigned clinicians to either group and nested patients within clinician group assignment. We recruited participants between July 2002 and October 2003.

See also:
Print
Editors’ Notes ........................................ 478
Editorial comment .................................. 544
Summary for Patients ............................ I-10
Web-Only
CME quiz
Conversion of figures and tables into slides
from 5 primary care clinics of a VA medical center and followed patients for 12 months. The local institutional review board approved the study, and all patients and clinicians gave written informed consent.

Setting

Approximately 25,000 veterans were treated during the study period in the 5 primary care clinics (3 urban and 2 rural clinics). Mental health clinicians are available on site in all clinics to provide consultation and brief treatment. A separate, more traditional mental health clinic serves approximately 8000 patients with chronic mental illnesses.

Participants

Full- and part-time staff physicians, fellows, physician assistants, and nurse practitioners were eligible to participate, and 41 (95%) of 43 eligible clinicians agreed to participate. To decrease variability in baseline depression-related knowledge and skills, we invited all clinicians to participate in the MacArthur Foundation depression education program (24–27) before randomization. In two 4-hour sessions, the program addresses communication skills and knowledge related to recognizing and managing depression.

All patients of participating clinicians were eligible for the study. Research assistants reviewed medical records of patients with upcoming primary care appointments (within 4 to 6 weeks). They excluded patients who had received treatment from mental health specialists within the previous 6 months; who had received a diagnosis of psychotic disorder, dementia, or bipolar disorder; or who were considered to be terminally ill. They mailed study introduction letters to nonexcluded patients. The letters informed the patients that the research team would call them within 2 weeks unless they declined screening by notifying the study office.

Telephone screening measures included the Patient Health Questionnaire (PHQ-9) (28) and the Short Blessed Test screening for dementia (29). We invited patients with PHQ-9 scores of 10 to 25 (moderate to severe depression) (30) to attend in-person enrollment interviews. We referred patients with PHQ-9 scores greater than 25 (very severe depression) or active dangerous ideation for urgent care, and we excluded them from participation. Enrollment interviews were scheduled within 2 weeks of primary care visits and usually took place the same day. The primary inclusion criterion was a repeated PHQ-9 score of 10 to 25 or a Hopkins Symptom Checklist-20 (SCL-20) (31) score of 1.0 or greater at the enrollment interview.

Of the 5434 patients who were mailed study introduction letters, 3500 (64%) were reached by telephone and were offered screening (Figure 1), and 3103 (89%) of those patients completed telephone screening. Of these patients, 560 (18%) had PHQ-9 scores between 10 and 25 and were eligible for in-person enrollment interviews. Of the 402 patients who completed the enrollment interviews, 375 had repeated PHQ-9 scores between 10 and 25 or SCL-20 scores of 1.0 or greater and were enrolled in the study.

Randomization

After clinicians participated in the MacArthur Foundation depression education program, we randomly assigned clinicians to receive depression decision support or usual care. Patients were nested within clinician assignment group. A stratified technique used a random-number generator to produce equivalent distributions across clinician type (physician vs. physician’s assistant or nurse practitioner) and clinic site. Clinicians in one clinic with substantial variation in caseload size were ranked by caseload and underwent block randomization.

Intervention

The depression decision support team consisted of 1 psychiatrist who was assigned up to 4 hours per week and 1 nurse care manager who was assigned up to 8 hours per week. Within 1 to 2 weeks after enrollment, the depression decision support care manager attempted to call each intervention patient to provide education, explore barriers, emphasize adherence to treatment, and encourage communication with clinicians about depression (Table 1). Supplemental educational materials were mailed to all intervention patients. During the telephone call, patients were invited to attend a 2-hour group depression education program led by the care manager or a depression education class offered by the mental health team in 1 of the urban clinics. Aside from this single early telephone contact, only rare additional contact between the depression decision support care manager and patients was expected.

The depression decision support team met weekly and
reviewed PHQ-9 scores (collected by the research team at baseline and at 1, 3, 6, 9, and 12 months) and medication and appointment data from the medical records. The team reviewed each intervention patient record at least monthly. Using a database, the depression decision support care manager compiled symptom severity and adherence data, posttraumatic stress disorder and alcohol screening results (obtained from routine clinic screening), and treatment recommendations into a treatment progress report. The report was mailed to each intervention clinician for all of their enrolled patients quarterly. If primary care clinicians did not respond to initial PHQ-9 scores greater than 15 or when patients’ depression did not adequately improve (PHQ-9 score >10 at 3 months or PHQ-9 score >5 at 6, 9, or 12 months), the depression decision support team reviewed records again and then contacted clinicians or their nurses to discuss treatment strategies or to offer consultation. When the depression decision support team and primary care clinician agreed that psychiatric consultation might be helpful, the psychiatrist arranged a visit. When ongoing mental health specialty care was indicated, the depression decision support team facilitated a referral.

**Figure 1. Study flow chart.**

PHQ-9 = Patient Health Questionnaire.
Usual Care

Clinicians received notifications when their patients enrolled in the study and were provided baseline PHQ-9 scores (Table 1). Usual care clinicians had access to all initial and follow-up PHQ-9 scores (available in the medical record), but usual care clinicians did not receive notifications, reminders, or recommendations about scores from the depression decision support team. Usual care clinicians and their patients also had access to mental health services, including on-site mental health teams.

Outcomes and Measurements

Blinded research assistants collected baseline patient data in person and PHQ-9 scores (at 1, 3, 6, 9, and 12 months) and outcome data (at 6 and 12 months) by telephone. When patients could not be reached by telephone, research assistants mailed questionnaires to them. Baseline measures included demographic information, Medical Outcomes Study 36-item Short Form for Veterans (SF-36V) scores for health-related quality of life (32), SCL-20 score for depression severity (31), Alcohol Use Disorders Identification Test score (33, 34), PHQ scores for anxiety and panic disorders (35), Posttraumatic Stress Disorder Checklist score (36), Medical Outcomes Study pain effects score (37), and the dysthymia stem from the WHO Composite International Diagnostic Interview (38). The PHQ-9 from the enrollment interview was used to diagnose major depression (DSM-IV [Diagnostic and Statistical Manual of Mental Disorders, fourth edition] criteria method) (28). We measured general medical comorbid conditions by using RxRisk-V, a version of the Chronic Disease Score derived from VA pharmacy data (39). Clinician baseline measures addressed confidence in caring for patients with depression (40), job satisfaction (41), burnout (42), and satisfaction with mental health services.

The primary study outcome was mean depression severity score (SCL-20) at 6 and 12 months. Additional outcomes included SF-36V and PHQ-9 scores, 4 items addressing satisfaction with care (43–46), antidepressant prescription durations and dosages, and health care utilization. We rated satisfaction from poor to excellent by using 5-point Likert scales. For patients who were not taking antidepressants at study entry, we determined the numbers of initial depression assessments and follow-up contacts between patients and primary care clinicians that resulted in depression-related clinical actions. Two research assistants reviewed medical charts to rate depression-related actions, which included monitoring and adjusting antidepressant treatments, making mental health referrals, or overly documenting watchful waiting (47). An investigator arbitrated coding discrepancies. We obtained antidepressant and VA health care utilization data from the Veterans Integrated Service Network 20 (VISN 20) Data Warehouse (48), which contains reliable local and national VA data.

Statistical Analysis

On the basis of previous collaborative trials (10–12), we hypothesized a mean difference in change of SCL-20 score of 0.35 (SD, 0.5). For a cluster randomization model with one level of nesting and a 2-tailed α level of 0.05, 20 clinicians per group and 7 patients per clinician (280 patients total) would provide 80% power to detect a treatment effect size of 0.33 or greater (determined by using PASS 2002, NCSS, Kaysville, Utah).

We used t-tests and chi-square statistics to assess the relationships of baseline variables with SCL-20 scores at 6 or 12 months. Because we randomly assigned by clinician and nested patients within clinician assignment, we used multilevel modeling with HLM 6.0 (Scientific Software International, Lincolnwood, Illinois) for the analyses. For the outcomes, SCL-20 scores, SF-36V physical health component and mental health component scores, and time (baseline, 6 months, or 12 months) formed the first level of the model. The slope associated with the linear trend across time reflects the change in outcome from baseline to 12 months. We nested outcome scores across time within patients at the second level of the model, which included the patient-level covariates. The clinician formed the third level of the model, which included the independent variable of interest (intervention vs. usual care). We used the independent variable of intervention as a predictor of the
slope across time, which was our primary test of the effectiveness of the intervention. Therefore, if the intervention variable is a statistically significant predictor of slope across time, then the intervention and usual care groups differ in the degree of change over time. We also used 3-level models for PHQ-9 scores (baseline and 1, 3, 6, 9, and 12 months), in which we fitted both linear and quadratic models to the data. We used 2-level continuous and logistic models for satisfaction, process-of-care, and utilization variables at 12 months, with patients at the first level of the model and clinician and the independent variable of intervention versus usual care as the second level of the model. Full maximum likelihood estimation was used for all models.

Hierarchical linear modeling incorporates all patients with at least 1 time point in the first level of the model (49), and thus, we did not exclude patients from the analyses because of missing dependent variables at follow-up. We excluded patients with missing covariate data from the analyses if those covariates were included in the model \( (n = 13 \ [3\%]) \). After enrollment, 11% of enrollees deviated from group assignment or study protocol, and 5 clinicians (1 intervention clinician and 4 usual care clinicians) left the VA. All results are based on intention-to-treat analyses.

Unadjusted means and proportions are presented to describe the baseline characteristics. We reported adjusted means and proportions and corresponding \( P \) values for all other models and derived them from the multilevel models using the method described by Raudenbush and Bryk (49).

Role of the Funding Source

Our study was funded by the VA Health Services Research and Development Service (Project Mental Health Initiative [MHI 20-020]). The funding source had no role in the recruitment of participants; study intervention; collection, analysis, or interpretation of the data; or preparation or review of the manuscript.

### RESULTS

We collected follow-up data on 84% of enrollees at 3 months, 84% at 6 months, and 85% at 12 months. Seventy-six percent of enrollees completed both the 6- and 12-month follow-ups. Therefore, 76% of enrollees completed all 3 time points, 18% completed 2 time points, and 7% completed only 1 time point (baseline). Missingness did not vary by intervention group or baseline depression scores. Approximately two thirds of follow-up surveys were administered by telephone, and the remainder were completed by mail.

Baseline characteristics did not statistically significantly differ between the intervention and usual care clinicians (Table 2) and patients (Table 3). The patient sample is representative of local and national VA populations (51). Comorbid psychiatric and medical conditions were prevalent. The SF-36V scores were in the 10th percentile of functioning range according to national norms (52).

#### Clinical Outcomes

We tested dysthymia, pain, anxiety, and posttraumatic stress disorder for possible correlations with the main outcome because these variables have been identified as important predictors of depression (6, 53–55). We confirmed these relationships in bivariate analyses: Baseline dysthymia, pain effects, anxiety, and posttraumatic stress disorder diagnosis were statistically significantly correlated with SCL-20 scores. We therefore included these variables in analyses as patient-level covariates.

In the linear models, both intervention and usual care groups showed significant improvement in SCL-20 scores overall (slope, \(-0.382 \ [95\% \ CI, -0.488 to -0.276]; P < 0.001\) (Figure 2). On average, SCL-20 scores decreased by 0.382 at each time period, controlling for the covariates, for a total decrease in SCL-20 scores of 0.764 from baseline to 12 months (mean baseline SCL-20 score, 1.9). However, the slope of the change in SCL-20 depression scores over 12 months did not differ between the groups.

### Table 2. Baseline Characteristics of Participating Clinicians

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Usual Care ((n = 21))</th>
<th>Intervention ((n = 20))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD) age, y*</td>
<td>45.1 (8.6)</td>
<td>45.1 (6.6)</td>
</tr>
<tr>
<td>Women, n (%)</td>
<td>10 (48)</td>
<td>11 (55)</td>
</tr>
<tr>
<td>Nurse practitioners or physicians’ assistants, n (%)</td>
<td>7 (33)</td>
<td>6 (30)</td>
</tr>
<tr>
<td>Physicians, n (%)</td>
<td>14 (67)</td>
<td>14 (70)</td>
</tr>
<tr>
<td>Mean (SD) time since training, y</td>
<td>14.6 (8.6)</td>
<td>15.4 (8.5)</td>
</tr>
<tr>
<td>Mean (SD) number of patients in panel, n</td>
<td>483 (244)</td>
<td>493 (223)</td>
</tr>
<tr>
<td>Mean (SD) proportion of patient panel with antidepressant prescription, %</td>
<td>24.5 (6.2)</td>
<td>25.2 (6.7)</td>
</tr>
<tr>
<td>Mean (SD) confidence in providing depression care score (4 items)†</td>
<td>2.6 (0.59)</td>
<td>2.6 (0.59)</td>
</tr>
<tr>
<td>Mean (SD) burnout score (2 items)†</td>
<td>3.5 (1.2)</td>
<td>3.8 (1.0)</td>
</tr>
<tr>
<td>Mean (SD) job satisfaction score (5 items)†</td>
<td>3.9 (1.2)</td>
<td>4.2 (0.90)</td>
</tr>
<tr>
<td>Mean (SD) satisfaction with mental health services score (2 items)†</td>
<td>3.2 (0.71)</td>
<td>3.0 (0.60)</td>
</tr>
</tbody>
</table>

* \( P \geq 0.389 \) for all comparisons.
† Measured by using Likert scales: confidence, 1–4; burnout, 1–6 (reverse scoring—higher scores represent less burnout); job satisfaction, 1–6; and satisfaction with mental health services, 0–4. Greater scores represent higher satisfaction.
Depression Decision Support in Primary Care

Table 3. Baseline Characteristics of Patients*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Usual Care</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 186)</td>
<td>(n = 189)</td>
</tr>
<tr>
<td>Mean (SD) age, y†</td>
<td>56.3 (11.2)</td>
<td>57.3 (10.9)</td>
</tr>
<tr>
<td>Women/men, n/n</td>
<td>13/173</td>
<td>13/176</td>
</tr>
<tr>
<td>Race or ethnicity, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>84 (45.2)</td>
<td>93 (49.2)</td>
</tr>
<tr>
<td>Not white</td>
<td>8 (4.3)</td>
<td>2 (1.1)</td>
</tr>
<tr>
<td>Not recorded</td>
<td>94 (50.5)</td>
<td>94 (49.7)</td>
</tr>
<tr>
<td>Married, %</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>Mean (SD) education, y</td>
<td>13.5 (2.4)</td>
<td>13.4 (2.6)</td>
</tr>
<tr>
<td>Worked during past 12 months, %</td>
<td>35</td>
<td>38</td>
</tr>
<tr>
<td>Mean (SD) baseline SCL-20 score</td>
<td>1.9 (0.50)</td>
<td>1.9 (0.57)</td>
</tr>
<tr>
<td>Mean (SD) baseline PHQ-9 score†‡</td>
<td>14.1 (5.0)</td>
<td>13.6 (4.4)</td>
</tr>
<tr>
<td>Meets PHQ-9 major depression diagnosis criteria, %§</td>
<td>51</td>
<td>47</td>
</tr>
<tr>
<td>Dysthymia, %</td>
<td>48</td>
<td>46</td>
</tr>
<tr>
<td>Moderate or greater pain effects, %</td>
<td>71</td>
<td>65</td>
</tr>
<tr>
<td>Active alcohol use disorder, %</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>Panic disorder, %</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>Other anxiety disorder, %</td>
<td>29</td>
<td>21</td>
</tr>
<tr>
<td>Active posttraumatic stress disorder, %**</td>
<td>44</td>
<td>35</td>
</tr>
<tr>
<td>Mean (SD) medical comorbid condition categories (RxRisk-V††), n</td>
<td>4.1 (2.6)</td>
<td>4.0 (2.7)</td>
</tr>
<tr>
<td>Taking an antidepressant at study entry, %</td>
<td>43</td>
<td>39</td>
</tr>
<tr>
<td>Took antidepressant in the year before study entry, %</td>
<td>58</td>
<td>48</td>
</tr>
<tr>
<td>Mean (SD) SF-36V subscale and component score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical functioning</td>
<td>44.1 (26.2)</td>
<td>44.2 (27.6)</td>
</tr>
<tr>
<td>Role physical</td>
<td>36.3 (27.7)</td>
<td>39.4 (27.5)</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>33.8 (21.2)</td>
<td>35.0 (20.7)</td>
</tr>
<tr>
<td>General health</td>
<td>37.5 (20.3)</td>
<td>37.9 (17.9)</td>
</tr>
<tr>
<td>Vitality</td>
<td>24.7 (17.4)</td>
<td>23.8 (15.9)</td>
</tr>
<tr>
<td>Social functioning</td>
<td>37.9 (21.1)</td>
<td>37.9 (22.4)</td>
</tr>
<tr>
<td>Role emotional</td>
<td>43.7 (22.8)</td>
<td>48.6 (23.0)</td>
</tr>
<tr>
<td>Mental health</td>
<td>46.0 (15.4)</td>
<td>45.8 (15.9)</td>
</tr>
<tr>
<td>Physical component score</td>
<td>32.9 (10.2)</td>
<td>33.1 (10.6)</td>
</tr>
<tr>
<td>Mental component score</td>
<td>32.8 (8.0)</td>
<td>33.2 (8.2)</td>
</tr>
</tbody>
</table>

* PHQ-9 = Patient Health Questionnaire; SCL-20 = Hopkins Symptom Checklist-20; SF-36V = 36-item Short Form for Veterans.
† P ≥ 0.073 for all comparisons.
‡ A PHQ-9 score between 10 and 15 indicates moderate depression severity (50).
§ ≥5 symptoms occurring more than half the days in past 2 weeks (suicide ideation counts regardless of frequency); ≥1 depressed mood or anhedonia item must be endorsed (28).
∥ ≥3, on average, of the 6 Medical Outcomes Study pain effects items.
¶ Alcohol Use Disorder Identification Test (AUDIT-C) score ≥ 5 (34).
** Posttraumatic Stress Disorder Checklist score ≥ 50 (36).
†† Version of the Chronic Disease Score derived from Veterans Affairs pharmacy data (39).

(difference, 0.020 [CI, −0.037 to 0.780]; P = 0.49). Similarly, overall mental health component scores improved over time in both groups (slope, 1.105 [CI, 0.420 to 1.791]; P = 0.003), and physical health component scores decreased in both groups (slope, −1.100 [CI, −1.645 to −0.553]; P < 0.001). On average, mental health component scores improved by 1.105 and physical health component scores decreased by 1.100 at each time period, controlling for covariates, for a total increase in the mental health component score of 2.21 (baseline mean score, 33.0) and a total decrease in the physical health component score of 2.20 (baseline mean score, 33.0) over 12 months. However, the difference in slopes of the mental and physical health component scores over time did not differ between the intervention group (difference, −0.008 [CI, −0.916 to 0.908]; P = 0.98) and the usual care group (difference, 0.636 [CI, −1.401 to 0.128]; P = 0.111).

The PHQ-9 scores also improved from baseline (Figure 3). In the quadratic multilevel model, the intervention group had greater initial improvement from baseline than the usual care group (difference in slopes, 0.248 [CI, 0.032 to 0.465]; P = 0.019). The 2 groups did not significantly differ in the degree to which initial improvement in PHQ-9 scores reversed over time (difference in slopes, 0.080 [CI, −0.027 to 0.187]; P = 0.062).

Process of Care

Intervention patients reported greater satisfaction with care than usual care patients at 12 months (adjusted means, 3.58 vs. 3.16; P = 0.002). Over 12 months, intervention patients were more likely to receive antidepressants than usual care patients (79.3% vs. 69.3%; P = 0.041) and were more likely to receive them for 90 days or more (76.2% vs. 61.6%; P = 0.008) but not for 180 days or more (63.1% vs. 52.3%; P = 0.065) (Table 4). Intervention patients were also more likely to receive more than 1 antidepressant at therapeutic dosages (2) (23.9% vs. 13.6%; P = 0.028).
Table 4 includes process-of-care variables for patients who were not taking antidepressants at baseline. Intervention patients in this subgroup were more likely to be assessed for depression by their clinicians over 12 months (93.5% vs. 77.4%; P < 0.003). Within this same subgroup, intervention clinicians were also more likely to perform 1 or more follow-up depression-related clinical actions per patient than usual care clinicians (84.8% vs. 53.6%; P < 0.001).

Utilization

Compared with those in the usual care group, patients in the intervention group were more likely to attend at least 1 mental health specialty appointment (41.1% vs. 27.2%; P = 0.025) (Table 4). Utilization of primary care visits, psychiatric or medical–surgical inpatient services, or emergency care did not statistically significantly differ between the groups.

Intervention Implementation

We reviewed logs and medical charts for the 172 (91%) intervention patients who received care from their assigned clinician for the entire 12 months. The depression decision support care manager conducted educational telephone calls that lasted an average of 16.6 minutes (SD, 5.5) with 155 (90%) patients. For 118 (76%) of these patients, this telephone call was the only contact with the depression decision support team. The psychiatrist met with 13% of patients and called 2 patients who declined in-person psychiatric consultation. Eight percent of patients attended a depression group education session. The depression decision support team communicated with primary care clinicians or their nurses an average of 2.2 times (SD, 1.7 [range, 0 to 8]) per patient, primarily via e-mail. Per independent ratings of the care manager and the psychiatrist and ratings of depression decision support logs by an investigator, approximately half of clinicians performed clinical actions in response to recommendations from the depression decision support team most of the time. Including training time, the depression decision support team spent 3.5 hours (SD, 2.3) per patient, of which approximately 0.5 hour involved direct patient contact.

Discussion

The depression decision support intervention had a positive effect on the rates of clinicians recognizing and treating depression and on patient satisfaction but did not generate sustained improvements in depression severity or health-related quality of life compared with usual care. While PHQ-9 scores in the intervention group initially improved, this effect diminished over time. The findings suggest that the intervention influenced clinician behaviors and that differences in clinician behaviors may have been perceived by patients, as evidenced by improvements in satisfaction scores.

The high prevalence of comorbid conditions in the sample may have influenced our results. According to SF-36V scores, these patients had very low levels of function at baseline. Two thirds of patients reported substantial pain, most patients had several chronic medical conditions, and

![Figure 2. Unadjusted mean Hopkins Symptom Checklist-20 (SCL-20) depression severity scores over time.](http://annals.org/)

Mean ± SE scores for the intervention group were 1.89 ± 0.05 at baseline, 1.54 ± 0.05 at 6 months, and 1.63 ± 0.06 at 12 months. Mean ± SE scores for the usual care group were 1.92 ± 0.05 at baseline, 1.58 ± 0.06 at 6 months, and 1.62 ± 0.06 at 12 months.

![Figure 3. Unadjusted mean Patient Health Questionnaire-9 (PHQ-9) depression scores over time.](http://annals.org/)

\[ P = 0.019 \] for initial improvement from baseline (quadratic model), comparing intervention and usual care groups.
40% of patients had active posttraumatic stress disorder. Moreover, dysthymia, pain effects, anxiety, and posttraumatic stress disorder were associated with worse depression outcomes. Previous reports suggest that although patients with substantial medical and psychiatric comorbid conditions often respond to depression treatment, they tend to respond to a lesser degree than patients without those conditions (53–61). Finally, 41% of patients were taking antidepressants at baseline, suggesting that some patients may have been refractory to antidepressant treatment.

We designed our study to help define the most essential elements of collaborative care (Table 1). While our data suggest that the intervention was implemented as intended, the intervention was not robust enough to influence clinical outcomes over time. Recent reviews of collaborative depression studies (17, 19, 23) suggest a possible association among the intensity and frequency of care management, specialty treatment, and response to depression treatment. In contrast to interventions involving more intensive contact between intervention team members and patients (10, 14), our intervention included only 1 early direct contact between the depression decision support team and each patient. In addition, all clinicians received baseline depression education and had access to on-site mental health consultation. These factors may have contributed to the limited effectiveness of depression decision support compared with usual care. We note that our PHQ-9 data echo Hedrick and colleagues’ findings of an early intervention effect that dissipated over time. Hedrick and colleagues’ study was also performed in a VA setting with a usual care condition that included on-site mental health support.

Several other limitations are worth noting. Sampling bias might have been present because of whom we were able to reach by telephone (for example, patients who were not employed may have been overrepresented). On the other hand, treatment settings.
In conclusion, while our results suggest that depression decision support can improve the process of care, the approach is insufficient to improve outcomes in a patient population with high rates of comorbid conditions, compared with usual care that includes on-site mental health support. Our study helps to define the boundaries of effective collaborative care for VA primary care settings and specifically indicates that more intensive and direct care management activities (for example, patient activation, self-management, and problem solving) or mental health specialty treatment over time may be critical elements of effective collaborative care for patients with complex illnesses.

From Portland Veterans Affairs Medical Center, Oregon Health & Science University, and Oregon State University, Portland, Oregon.

Acknowledgments: The authors thank Ginger Hanson, MS, for assistance with statistical analysis; Megan Crutchfield, BS, and Marsha W. Perkett, BA, for assistance with chart review; and the Portland Veterans Affairs Medical Center’s primary care clinicians and staff whose participation made this study possible.

Grant Support: By the VA Health Services Research & Development Service (Project Mental Health Initiative [MHI 20-020]).

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Steven K. Dobscha, MD, Portland Veterans Affairs Medical Center, PO Box 1034 (P3MHDC), Portland, OR 97207; e-mail, steven.dobscha@va.gov.

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

References


Current Author Addresses: Dr. Dobscha: Portland Veterans Affairs Medical Center, PO Box 1034 (P3MHDC), Portland, OR 97207. Drs. Corson, Hickam, and Gerrity: Portland Veterans Affairs Medical Center, 3710 SW U.S. Veterans Hospital Road, Portland, OR 97239. Dr. Perrin: Oregon Health & Science University, 3455 SW U.S. Veterans Hospital Road, Portland, OR 97239. Dr. Kraemer: Oregon State University, 840 SW Gaines, MC GH212, Portland, OR 97239-2985.