

Effectiveness of Disease Management Programs in Depression: A Systematic Review

Enkhe Badamgarav, M.D., M.P.H.

Scott R. Weingarten, M.D., M.P.H.

James M. Henning, M.S.

Kevin Knight, M.D., M.P.H.

Vic Hasselblad, Ph.D.

Anacleto Gano, Jr., M.P.H.

Joshua J. Ofman, M.D., M.S.H.S.

Objective: The authors systematically evaluated the published evidence to assess the effectiveness of disease management programs in depression.

Method: English-language articles on depression were identified through a MEDLINE search for the period from January 1987 to June 2001. Two reviewers evaluated 16,952 published titles, identified 24 depression disease management programs that met explicit inclusion criteria, and extracted data on study characteristics, interventions used, and outcome measures. Pooled effect sizes were calculated by using a random-effects model.

Results: Pooled results for disease management program effects on symptoms of depression showed statistically significant improvements (effect size=0.33, N=24). Programs also had statistically significant effects on patients' satisfaction with treatment (effect size=0.51, N=6), patients' compliance with the recommended treat-

ment regimen (effect size=0.36, N=7), and adequacy of prescribed treatment (effect size=0.44, N=11). One program with an explicit screening component showed significant improvement in the rate of detection of depression by primary care physicians (effect size=0.66); two other programs lacking a screening component showed small nonsignificant improvements in the detection rate (effect size=0.18). Disease management programs increased health care utilization (effect size=-0.10, N=8), treatment costs (effect size=-1.03, N=3), and hospitalization (effect size=-0.20, N=2).

Conclusions: Disease management appears to improve the detection and care of patients with depression. Further research is needed to assess the cost-effectiveness of disease management in depression, and consideration should be given to more widespread implementation of these programs.

(*Am J Psychiatry* 2003; 160:2080-2090)

Depression is a common medical condition, and although it is eminently treatable, it is associated with significant social and functional impairment as well as high direct and indirect health care costs (1, 2). Estimates of the prevalence of depression vary from 15% to 25% for lifetime prevalence and from 10% to 20% for 12-month prevalence (3-8). Depression has been reported to cause greater functional disability than diabetes, chronic lung disease, hypertension, or back pain (9). The total direct and indirect costs of depression have been estimated to be \$44 billion annually in the United States, with annual costs estimated at \$5,400 per patient (10, 11). Within the next 20 years, depression is predicted to become one of the leading causes of disability worldwide—second only to ischemic heart disease in terms of the total cost to society (12, 13). Although cost estimates depend heavily on the researcher's perspective, it is generally accepted that the indirect costs of depression outweigh direct costs by a ratio of as much as 7:1 (2, 9-11, 14-17).

Despite the enormous burden it places on the health care delivery system, depression remains a condition with suboptimal management in the primary care setting (18-20). Deficiencies in care relate to underdiagnosis, inadequate treatment, and lack of patient follow-up after

treatment is initiated (12, 21-25). Some researchers have suggested that these deficiencies mainly relate to the organization of care for depression, because the complex needs of patients and families struggling with chronic illnesses are unlikely to be met by a primary care system designed for acute care (25).

To improve delivery of care for patients with chronic illnesses, Wagner et al. (26) proposed an evidence-based model that incorporates the following main components: use of evidence-based practice guidelines, practice reorganization to meet the needs of chronically ill patients, patient education, and expert systems or multidisciplinary approaches to care. Similarly, disease management programs included in this review had multiple components such as the use of an evidence-based guideline akin to the Wagner model. We defined disease management as an intervention to manage or prevent a chronic condition by using a systematic approach to care (i.e., evidence-based practice guidelines) and potentially employing multiple treatment modalities (27). Several studies assessed the effect of single treatment modalities such as psychotherapy on outcomes of patients with depression. However, the effectiveness of multimodal disease management programs has rarely been demonstrated in rigorously designed eval-

uations, and few descriptions of programs with the goal of improving care for patients with depression have been published. Our study had two objectives: 1) to identify rigorously conducted studies of depression disease management programs and 2) to systematically evaluate the effectiveness of these programs.

Method

Literature Search and Review

We performed a systematic review of the published medical literature to identify studies evaluating the effectiveness of disease management programs in improving care or reducing costs for patients with a variety of common chronic conditions, including depression. In collaboration with a librarian with expertise in searching computerized bibliographic databases, we conducted a search of the MEDLINE, HealthSTAR, and Cochrane databases for English-language articles published between January 1987 and June 2001. The search used the following medical subject headings: patient care team, patient care planning, primary nursing care, case management, critical pathways, primary health care, continuity of patient care, guidelines, practice guidelines, disease management, comprehensive health care, ambulatory care, and the title words "disease state management" and "disease management." We performed an additional search using the terms "depression" and "randomized controlled trial." Hand searches of bibliographies from relevant articles and reviews as well as consultations with experts in the field yielded additional references.

Our working definition of disease management was as follows: an intervention designed to manage or prevent a chronic condition by using a systematic approach to care and potentially employing multiple treatment modalities (27). We defined a systematic approach to care (or guideline) as a set of systematically developed statements to assist practitioners' and patients' decisions about appropriate health care for a specific clinical circumstance (28). To determine if a program incorporated a systematic approach to care, we searched for keywords, including guidelines, protocols, algorithms, quality improvement programs, care plans, and standardized patient and provider education. We excluded programs exclusively evaluating single treatment modalities (e.g., psychotherapy or specific pharmacologic agents) or patients' compliance with medication regimens. Articles were rejected if they included only pediatric cases, or if they were reviews, case reports, editorials, letters, or abstracts of meeting presentations. Articles were rejected if they did not use acceptable experimental or quasi-experimental study designs as defined by the criteria for acceptable study design of the Cochrane Effective Practice and Organization of Care Group (29) or if they did not report sufficient information to allow for estimation of at least one measure of a program effect of interest and its variance. Failure to meet this criterion could be due to inadequate reporting of results or to the lack of an appropriate comparison group.

Based on these explicit inclusion and exclusion criteria for titles, abstracts, and articles, two reviewers trained in health services research and the principles of critical appraisal independently reviewed random samples of titles, abstracts, and articles. Interrater agreement was assessed by using the kappa statistic, and reviews were split between reviewers if a sufficient level of agreement was achieved (kappa >0.7). The findings from accepted articles reporting results for depression disease management programs were used to address study objectives.

Extraction of Data

By using a standardized abstraction form, data describing study design, population characteristics, sample size, intervention strategies, and processes and outcomes of care were collected from unmasked articles that met the inclusion criteria. Multiple published reports from a single study were treated as a single program evaluation. Studies with multiple intervention arms contributed more than one observation, as did those reporting results in different subgroups of patients.

When appropriate, we used changes from baseline values as opposed to follow-up values in our analyses. Several studies did not report variances for changes from baseline values. In those cases we assumed that the variance for the change was equal to the average of the variances of the baseline and follow-up distributions, if both were given, or, if both were not given, to the variance of the follow-up distribution.

Conceptual Model for Analysis

To guide our analysis, we developed a conceptual model of the processes and outcomes of care for patients with depression. We identified available assessments for various processes and outcomes of care by using the following framework: detection of depression → prescribed treatment → patients' compliance with recommended treatment → treatment outcomes. Outcomes potentially affected by both practitioners' adherence to guidelines and patients' compliance (e.g., the proportion of patients taking antidepressant drugs) were analyzed in a separate category.

Meta-Analysis

Effect sizes, defined as a difference between the means of the treatment and control arms divided by the pooled estimate of the standard deviation (continuous variables) or the log odds ratio multiplied by a constant variance term (binary variables) (30), were calculated for each study outcome to allow pooling of similar outcomes (31–35). Effect sizes were constructed such that positive numbers denote benefit.

We used the more conservative, random-effects, empirical Bayesian method of Hedges and Olkin to pool the estimated effects (35, 36). We pooled results for each category defined by the conceptual model and the additional category reflecting effects influenced by both providers' and patients' adherence. When more than one process or outcome result within a category were reported for the same group of subjects, the one associated with the smallest effect size was used. Results are reported as the pooled effect size with the 95% confidence interval (CI) for each group of process/outcome variables.

Publication Bias

We explored the evidence for publication bias with funnel plots (37) for the four areas with at least seven effect estimates (38). We plotted the effect size on the x-axis by the inverse of its standard error on the y-axis. A plot that is asymmetrical or that has a relative paucity of estimates favoring no treatment with large standard errors (the bottom left region of the graph) suggests publication bias.

Results

Literature Review

The initial disease management search strategy identified 16,952 references published between January 1987 and June 2001. We accepted 2,998 titles for further screening, and 581 abstracts met our explicit inclusion criteria. Eighty-five percent (N=491) of the accepted ab-

TABLE 1. Studies Estimating the Effect of Disease Management Programs on Symptoms of Depression

Study	Number of Subjects		Outcome Measure
	Intervention Group	Comparison Group	
Whooley et al., 2000 (44)	97	109	Geriatric Depression Scale score
Wells et al., 2000 (49)	913	443	Modified CES-D Scale score
Katon et al., 2001 (42)	194	192	20-item Depression Symptom Checklist score
Rost et al., 2001 (43)	130	59	Modified Center for Epidemiologic Studies Depression (CES-D) Scale score
Simon et al., 2000 (39)			HSCL score
Effect of providing feedback (patients' pharmacy and visit data) to doctors	221	196	
Effect of care management	196	196	
Hunkeler et al., 2000 (45)	159	97	Hamilton depression scale score
Katzelnick et al., 2000 (46)	218	189	Hamilton depression scale score
Tutty et al., 2000 (48)	26	94	HSCL score
Lin et al., 1999 (51)	63	53	HSCL score
Llewellyn-Jones et al., 1999 (52)	86	83	Geriatric Depression Scale score
Worrall et al., 1999 (53)	91	56	CES-D Scale score
Banerjee et al., 1996 (54)	33	36	Change in Montgomery-Åsberg Depression Rating Scale score
Thompson et al., 2000 (47)	2,730	3,308	Number of patients with Hamilton Depression Rating Scale score ≥ 8
Brown et al., 2000 (57)			Change in HSCL depression scale score
Effect of academic detailing ^d	464	464	
Effect of continuous quality improvement	488	440	
Goldberg et al., 1998 (56)			Change in HSCL depression scale score
Effect of academic detailing ^d	173	171	
Effect of academic detailing plus continuous quality improvement	210	171	
Katon et al., 1999 (50)	114	114	20-item Depression Symptom Checklist score or proportion of patents who recovered
Katon et al., 1996 (40)			20-item Depression Symptom Checklist score or proportion of patients with 50% improvement
Effect for patients with minor depression	46	42	
Effect for patients with major depression	31	34	
Katon et al., 1995 (41)			
Effect for patients with minor depression	59	67	
Effect for patients with major depression	49	42	
Callahan et al., 1994 (55)	75	75	Proportion of patients with Hamilton depression scale score ≥ 10
Pooled estimate (N=24)	6,866	6,354 ^e	

^a Not available.

^b Total N at follow-up=2,224.

^c Total N at follow-up=2,819.

^d Targeted one-on-one education for doctors about implementation of practice guideline for treatment of depression.

^e Studies with multiple treatment arms that were compared to a single comparison group were counted once to avoid double counting.

stracts failed to meet inclusion criteria when the article was reviewed. Bibliographic hand searches and expert consultation yielded an additional 57 articles for consideration, of which 16 were accepted. A total of 106 studies met our criteria for inclusion, and 18 of those 106 studies

examined treatment of depression. We also identified an additional study through the depression-specific search. Thus, both searches yielded 19 relevant references on depression disease management from January 1987 to June 2001 (39–57).

Results									
Baseline				Follow-Up				Effect Size	95% CI
Intervention Group		Comparison Group		Intervention Group		Comparison Group			
Mean	SE	Mean	SE	Mean	SE	Mean	SE		
8.2	2.1	8.4	2.4	1.8	0.4	2.2	0.4	-0.11	-0.38 to 0.17
75.4	1.6	75.7	2.1	54.5	1.9	61.4	2.6	0.26	0.08 to 0.43
Mean	SD	Mean	SD	Mean	SD	Mean	SD		
0.83	0.4	0.84	0.4	0.67	0.4	0.75	0.4	0.21	0.01 to 0.41
55.1	— ^a	52.7	— ^a	33.4	— ^a	39.2	— ^a	0.90	0.44 to 1.37
1.67	0.72	1.74	0.77	0.95	1.39	0.98	1.39	0.02	-0.17 to 0.21
1.66	0.76	1.74	0.77	0.83	0.57	0.98	0.57	0.26	0.06 to 0.46
17.4	6.6	16.6	6.2	8.1	6.4	10.4	6.4	0.22	-0.02 to 0.47
19.1	0.2	19.2	0.3	9.9	11.1	13.6	11.1	0.33	0.14 to 0.53
2.01	0.5	1.76	0.7	0.79	0.1	0.95	0.1	0.53	0.09 to 0.97
46.6	10.9	46.6	10.9	16.4	13.1	16.3	12.5	0.01	-0.36 to 0.37
13.38	3.0	13.18	3.3	11.81	4.7	12.57	4.1	0.25	-0.05 to 0.55
37.3	8.9	38.7	8.1	19.4	13.6	22.2	11.7	0.26	-0.08 to 0.59
27.5	6.2	25.1	6.3	9.2	6.8	13.5	6.5	1.03	0.52 to 1.53
N	%	N	%	N	%	N	%		
577	21	670	20	385 ^b	17	575 ^c	20	0.11	0.03 to 0.19
Mean	Mean Change		Mean	Mean Change					
1.48	1.44		1.56	1.57		1.58	1.43 to 1.73		
1.44	1.48		1.55	1.58		-0.22	-0.35 to -0.09		
1.45	1.48		1.49	1.58		0.51	0.29 to 0.72		
1.53	1.48		1.61	1.58		0.13	-0.08 to 0.33		
Mean	SD	Mean	SD	%		%			
1.9	0.5	1.9	0.5	44		31		0.31	0.01 to 0.61
1.77	0.49	1.62	0.54	66.7		52.8		0.34	-0.13 to 0.81
2.46	0.53	2.35	0.51	70.4		42.3		0.67	0.11 to 1.22
1.67	0.40	1.72	0.56	60.0		67.9		-0.18	-0.58 to 0.21
2.35	0.49	2.23	0.48	74.4		43.8		0.70	0.22 to 1.18
%		%		%		%			
22		22		13		12		0.06	-0.46 to 0.58
								0.33	0.16 to 0.49

Five (39–41, 56, 57) of the 19 selected studies had multiple treatment arms or subgroup analyses that were treated as separate observations, bringing the total number of disease management programs studied to 24.

Characteristics of Studies

Fifteen of the 19 selected studies were conducted in the United States, two were carried out in the United Kingdom, one was done in Australia, and one was done in Can-

ada. Seventeen studies used a randomized, controlled design, and eight of those 17 used a cluster randomization scheme (39, 43, 44, 46, 47, 49, 53, 56). One study (48) with a quasi-experimental design (controlled before and after trial) was included in the analysis. The report by Brown et al. (57) included results from a randomized, controlled trial and a quasi-experimental, controlled trial. The number of patients included in each study ranged from 65 to 6,055, and study duration ranged from 6 to 30 months. Selected studies included mostly patients with major depression or mostly patients with minor depression. The studies used the following types of disease management interventions: patient education programs (16 programs), provider feedback (12 programs), provider education programs (17 programs), multidisciplinary teams of providers (11 programs), provider reminders (six programs), and financial incentives for providers (one program). A qualitative overview of individual studies is available on request.

Estimates of program effect were available for 10 broad categories including 1) symptoms of depression (24 estimates), 2) providers' adherence to guidelines (16 estimates), 3) patients' adherence to recommended treatment regimens (seven estimates), 4) health services utilization such as primary care visits for depression (eight estimates), 5) patients' satisfaction with treatment (six estimates), 6) measures affected by both providers' and patients' adherence (eight estimates), 7) physical functioning and disability (seven estimates), 8) overall health status (six estimates), 9) health care costs (three estimates), and 10) hospitalization (two estimates).

Meta-Analysis of Outcomes of Care

Symptoms of depression. Twenty-four estimates (39–57) of program effects on depressive symptoms, based on data for a total of 13,220 patients, were available (Table 1). Specific instruments used to measure symptoms of depression included the 20-item Depression Symptom Checklist (40–42, 50), the HSCL (39, 48, 51, 56, 57), the Hamilton Depression Rating Scale (45–47, 55), the Modified Center for Epidemiologic Studies Depression Scale (CES-D Scale) (43, 49, 53), the Geriatric Depression Scale (44, 52), and the Montgomery-Åsberg Depression Rating Scale (54).

Twenty-one of the estimates favored the disease management group, with 13 (39–43, 46–50, 54, 56, 57) of 24 observations (54%) showing statistically significant improvements in symptoms of depression. However, three observations (41, 45, 57) favored usual care. The pooled estimate of effect showed a statistically significant decrease in symptoms of depression in disease management program participants, compared to patients who received usual care (effect size=0.33, 95% CI=0.16 to 0.49).

Physical functioning. Seven estimates of the effect of programs on physical functioning or disability (40, 41, 45, 49, 50, 57 [reference 57 provided two estimates]) based on observations of 3,852 patients were available (Table 2).

Measures of physical functioning/disability included the physical health summary scale score of the Medical Outcomes Study 12-item Short-Form Health Survey (45), the Medical Outcomes Study 36-item Short-Form Health Survey (57), the Sheehan Disability Scale score (50), and the proportion of patients unable to work (40, 41). Although six estimates favored disease management programs, only two estimates (50, 57) demonstrated a statistically significant improvement in physical functioning. The pooled estimate showed that the effect of disease management programs was close to that of usual care. This effect was not statistically significant (effect size=-0.05, 95% CI=-0.72 to 0.62).

Social and health status. Six estimates (40, 41, 49, 50, 57 [reference 57 provided two estimates]) based on data from 3,596 patients measured the effect of disease management on perceived health status (Table 2). Specific measures of perceived health status included the mental health summary scale score from the Medical Outcomes Study 12-item Short-Form Health Survey (49), the social function subscale score from the Medical Outcomes Study 36-item Short-Form Health Survey (40, 57), and the self-rated health questionnaire from the National Health Interview Survey (40, 41). Four (40, 41, 49, 57) of the six estimates showed statistically significant improvements in patients' health status associated with disease management. One study showed statistically significant change in health status favoring patients in usual care (57). The pooled estimate of effect showed a small but nonsignificant improvement in health status in disease management program participants (effect size=0.06, 95% CI=-0.51 to 0.62).

Patients' satisfaction with treatment. Six estimates (40 [two estimates], 41 [two estimates], 45, 50) involving 854 patients assessed patients' satisfaction with the treatment of depression (Table 2). Satisfaction with treatment was measured on a 5-point ordinal scale from "very dissatisfied" to "very satisfied." Of the six estimates, five (40, [two estimates], 41, 45, 50) (85%) showed that satisfaction levels were significantly greater among program participants, compared to nonparticipants. The pooled estimate of effect showed a statistically significant improvement in patients' satisfaction associated with disease management (effect size=0.51, 95% CI=0.33 to 0.68).

Health care utilization. Eight estimates (39 [two estimates], 42, 46, 48–50, 53) using data from 3,366 patients assessed the effect of disease management on health care utilization (number of outpatient visits) (Table 2). Two estimates (39, 53) indicated lower utilization among disease management program participants, and the remaining six observations showed fewer visits among usual-care patients. The pooled estimate indicated that disease management programs resulted in a small but statistically significant increase in primary care visits (effect size=-0.1, 95% CI=-0.18 to -0.02).

TABLE 2. Pooled Size of Effects of Disease Management Programs on Measures of Outcomes and Processes of Care in the Treatment of Depression^a

Measure	Number of Estimates	Total Number of Patients	Pooled Effect Size	95% CI
Outcomes of care				
Physical functioning	7	3,852	-0.05	-0.72 to 0.62
Health status	6	3,596	0.06	-0.51 to 0.62
Patients' satisfaction	6	854	0.51	0.33 to 0.68
Health care utilization	8	3,366	-0.10	-0.18 to -0.02
Hospitalization	2	443	-0.20	-0.35 to 0.04
Health care costs	3	1,148	-1.03	-2.62 to 0.54
Outcomes influenced by both providers' and patients' adherence	8	3,891	0.57	-0.11 to 1.26
Processes of care				
Detection of depression in a program with a screening component	1	175	0.66	0.22 to 1.1
Detection of depression in a program without a screening component	2	299	0.18	-0.11 to 0.18
Referral to specialized care	2	297	0.13	-0.32 to 0.57
Adequacy of prescribed antidepressant treatment	11	2,647	0.44	0.30 to 0.59
Patients' adherence	7	941	0.36	0.17 to 0.54

^a Based on a total of 24 estimates of the effects of disease management programs in 19 studies.

Hospitalization. Two estimates (46, 54) based on data from 443 patients assessed the effects of disease management on hospitalization, measured as either the mean number of hospitalizations or the number of patients admitted to a psychiatric unit (Table 2). Both estimates indicated higher levels of hospitalization among program participants. The pooled estimate showed a small, statistically nonsignificant, increase in hospitalization rates associated with depression in disease management program participants (effect size=-0.2, 95% CI=-0.35 to 0.04).

Health care costs. Three estimates (39 [two estimates], 46) based on data from 1,148 patients examined program effects on health care costs (Table 2). All three programs measured total health services cost associated with treatment of depression and indicated that program participants incurred higher costs. One estimate (46) achieved statistical significance. Consistent with the individual estimates, the pooled estimate of effect indicated higher costs among program participants than among nonparticipants, but the effect was not statistically significant (effect size=-1.03, 95% CI=-2.62 to 0.54).

Outcomes affected by both providers' and patients' adherence. Eight estimates (42, 43, 49, 52-54, 57 [two estimates]) based on data for 3,891 patients evaluated the effect of disease management on outcomes that are potentially influenced by both providers' adherence to guidelines and patients' adherence to treatment regimens (Table 2). Included in this category were measures of medications taken and adequacy of treatment received, regardless of the regimen prescribed. Of the eight estimates of program effects in this category, six (42, 43, 49, 53, 54, 57) showed statistically significant improvements. One study reported that more patients in usual care adhered to the prescribed treatment, compared to those in the disease management program (57). The pooled estimate favored the disease management programs. However, this effect was not statistically significant (effect size=0.57, 95% CI=-0.11 to 1.26).

Meta-Analysis of Processes of Care

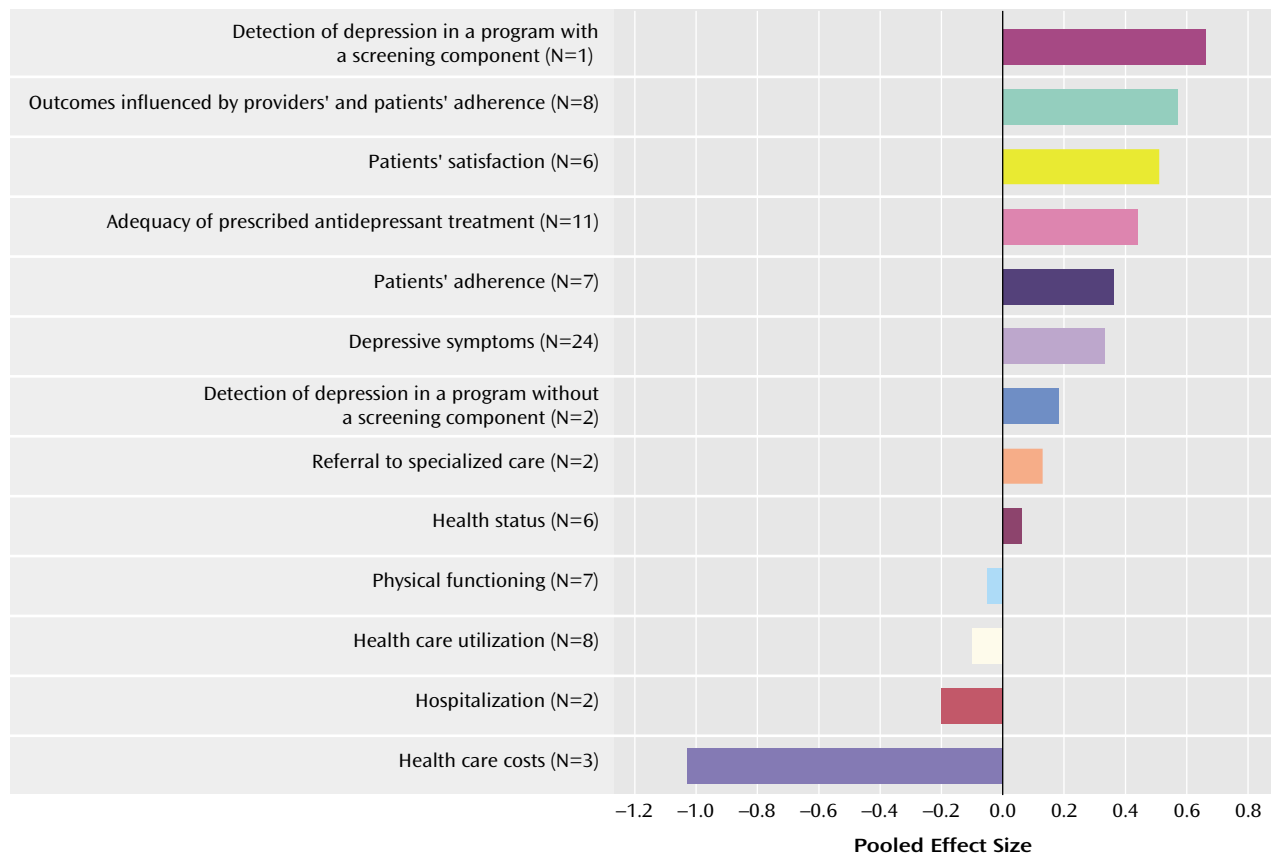
Detection of depression. Three observations (47, 53, 55) involving 474 patients evaluated the effect of disease management programs on detection of depression by primary care physicians (Table 2). One program with an explicit screening component (55) intended to increase the detection rate showed statistically significant improvement in recognition of depression (effect size=0.66, 95% CI=0.22 to 1.10) (Figure 1). The two other programs, which lacked a screening component, showed statistically nonsignificant improvements (effect size=0.18, 95% CI=-0.11 to 0.18).

Referral to specialized care. Two estimates (53, 55) based on data for 297 patients examined the effect of disease management on primary care physicians' rate of referral to psychiatrists (Table 2). Both studies assessed program effects on the proportion of patients referred to psychiatrists, but neither result was statistically significant. A pooled estimate showed that there was no significant program effect on the referral rate (effect size=0.13, 95% CI=-0.32 to 0.57).

Adequacy of prescribed treatment. Eleven estimates (40 [two estimates], 41 [two estimates], 42, 49 [two estimates] based on results for mutually exclusive subgroups of patients for this outcome), 51, 55, 56 [two estimates]) based on data from 2,647 patients assessed the effects of disease management programs on measures of treatment adequacy, including appropriateness of medication type, dose, and duration of treatment (Table 2). Of these, seven (40, 41 [two estimates], 49 [two estimates], 55, 56) estimates indicated that a greater proportion of the patients in disease management programs received adequate treatment, compared to usual-care patients. The pooled estimate of effect showed a significant beneficial program effect on treatment adequacy (effect size=0.44, 95% CI=0.30 to 0.59).

Patients' adherence with treatment regimens. Seven estimates (40 [two estimates], 48, 50, 51, 53, 55) involving

FIGURE 1. Rank Order of Pooled Size of Effects of Disease Management Programs on Measures of Outcomes and Processes of Care in the Treatment of Depression



941 patients evaluated the effect of disease management on patients' adherence with recommended treatment regimens (Table 2). All seven assessed compliance with prescribed antidepressant medication regimens. Of these, four (40 [two estimates], 50, 55) estimates indicated statistically significant improvement in patients' adherence to prescribed regimens. When all seven estimates were pooled, the resulting estimate indicated a significant improvement in patients' adherence (effect size=0.36, 95% CI=0.17 to 0.54).

Publication Bias

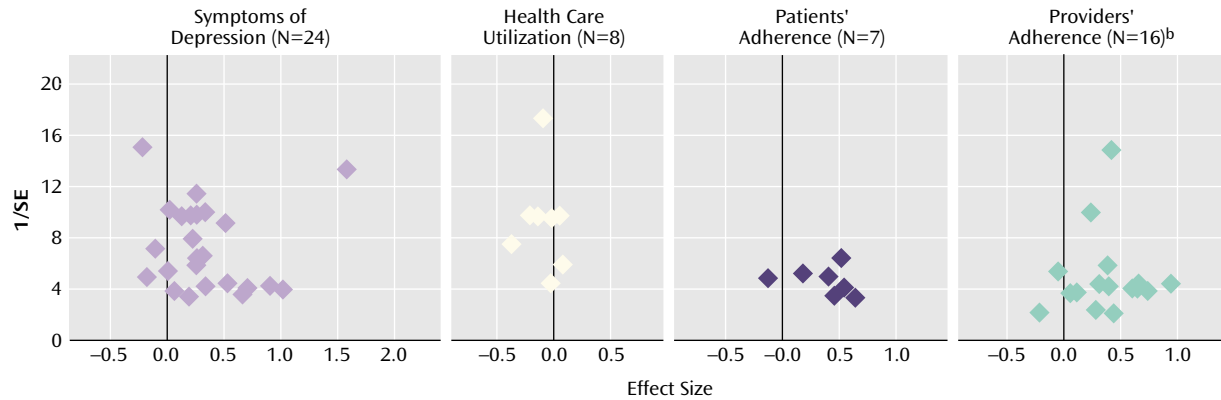
Data for four outcomes (depressive symptoms, health care utilization, providers' adherence, and patients' adherence) were sufficient to explore the possibility of publication bias by using the funnel plot method (Figure 2). In the absence of publication bias, the data points for studies with negative and positive findings are distributed symmetrically, creating a funnel. However, for each outcome we evaluated, there appeared to be asymmetry or missing values in the lower left area of the plots, suggesting that small studies with negative results were less likely to have been published than studies of similar size reporting positive results.

Discussion

We systematically evaluated and appraised published evaluations of depression disease management programs and found that such programs appear to result in some improvements in both processes and outcomes of care. Pooled results indicated statistically significant improvements in patients' symptoms of depression, physical functioning, health status, satisfaction of treatment, and adherence to treatment regimens, as well as in the rate of detection of depression, adequacy of treatment with antidepressants, and outcomes that are influenced by both providers' and patients' adherence. The largest effect was found in one program with an explicit screening component intended to increase the detection of depression in primary care patients (effect size=0.66).

This large effect is an important finding because 40%–50% of psychiatric disorders in primary care patients are undetected (24). Several studies that have assessed the effect of structured screening on rates of detection of depression have reported increases in the diagnosis of depression between 10% and 47%, suggesting that screening is effective for identifying depression among primary care patients (58). Systematic screening has been advocated as a means of improving detection, treatment, and outcomes

FIGURE 2. Funnel Plots for Analysis of Publication Bias in Studies Providing Estimates of the Effects of Disease Management Programs on Measures of Four Outcomes and Processes of Care in the Treatment of Depression^a



^a The plots display the relationship between effect sizes for four outcomes with at least seven effect estimates and the inverse of the standard error for each estimate. When the plot is not funnel shaped and asymmetry is present, it is suggestive of the presence of publication bias.

^b Based on estimates for detection of depression (N=3), referral to specialized care (N=2), and adequacy of prescribed treatment (N=11).

of depression. Moreover, screening implemented with interventions aimed at increasing recognition and management of depression has been reported to result in favorable outcomes. For example, in one study that implemented screening in combination with quality improvement activities to increase the percentage of patients receiving appropriate care according to national guidelines, the proportion of patients receiving appropriate treatment was higher in the disease management group, compared to usual care (49).

Even when depression is recognized, diagnosis does not necessarily result in appropriate treatment. In addition to their use in detecting depression among primary care patients, screening tools can also be used to establish thresholds for treatment initiation. Incorporating structured screening into disease management programs is likely to lead to increases in the detection and management of depression in the primary care setting.

Improved detection and treatment may increase costs to the extent that they result in more medications being prescribed and more visits to health care providers. Such increases are reflected in our pooled analysis. We found that disease management programs resulted in increases in measures of health care utilization, such as the number of primary care visits, the cost of treatment, and the number of hospitalizations for depression (Figure 1). Potential cost savings resulting from improved treatment, such as reductions in nonpsychiatric health care costs and improvements in productivity, were not evaluated in the studies we identified.

To assist our analysis, we developed a conceptual model to describe key aspects of depression care such as detection, prescription of appropriate treatment, and patients' compliance with treatment. These key domains represent potential targets for disease management interventions in addition to broader efforts in system reorganization. Treatment success, defined as improvement in symptoms

and reduction of impairments in functional status, depends on the providers' ability to recognize depression and prescribe appropriate treatment, as well as the patients' adherence to prescribed treatment.

Some measures, such as the quantity of medication taken by patients, are potentially affected both by providers' adherence to treatment guidelines and by patients' adherence to treatment recommendations, since medications must be prescribed in order to be taken. We grouped together outcomes influenced by both providers' and patients' adherence, as well as examined them separately. The pooled result showed a statistically significant positive effect of disease management programs on this group of measures. This improvement could result from program effects on providers' treatment patterns, on patients' compliance with prescribed treatments, or both. The pooled estimate of program effect on measures of providers' adherence (i.e., measures of adequacy of prescribed treatment) was slightly higher (effect size=0.44) than the pooled effect on measures of patients' compliance (effect size=0.36). These results suggest that disease management programs can improve both providers' adherence and patients' compliance. Similar findings were reported in a study by Weingarten et al. (59) that examined the effect of disease management interventions on the processes and outcomes of care among patients with chronic conditions.

Few studies evaluated the effects of multicomponent interventions on providers' satisfaction. One study assessed the effect of a new clinical system of care for elderly patients and found that providers in the intervention group reported being "very satisfied" with the management of patients participating in the study (60). However, studies measuring the relationship between providers' satisfaction and the rate of adherence to practice guidelines were not identified. Future research should focus on the development of appropriate measures to assess pro-

viders' satisfaction with disease management programs and explore whether an association exists between the rate of adherence to practice guidelines and providers' satisfaction.

While our study has several strengths, it has some limitations. Our definition of disease management, based on a published definition (27), was established a priori; however, disease management is a broad term, and its definition depends on the perspective employed. Our results may be subject to different interpretations, according to the operational definition of disease management.

The studies we identified may not be representative of all evaluations of depression disease management programs. Results from disease management programs implemented by health plans or disease management organizations may not be published for a variety of reasons, including negative findings, competitive or proprietary concerns, lack of expertise in publishing research studies, or a paucity of funding to support submission of programmatic results to peer-reviewed journals. Studies demonstrating statistically significant benefits, particularly small ones, may be more likely to be published than studies with nonsignificant or negative results. Examination of the funnel plots in this study suggested that such publication bias has, in fact, occurred to some extent.

In addition to using qualitative evaluation of disease management, we used effect sizes as a common metric to assess the magnitude of the effect of disease management across studies with different processes and outcomes of care. Some researchers have attempted to develop parameters to assist in interpreting effect sizes. We used the convention of Kazis et al. (61), in which an effect size less than 0.6 is considered a "small effect" and an effect size between 0.6 and 1.2 is characterized as a "moderate" effect. In our study, most effect sizes were less than 0.6, including those for program effects on depressive symptoms, patients' satisfaction, treatment adequacy, and patients' adherence with recommended treatment.

Finally, disease management is in an early stage of evolution. Although population-based approaches to care have existed for quite some time, the mechanisms for implementation (e.g., technologies, strategies to change physicians' and patients' behavior, and assessment methods) are still evolving. Thus, the effectiveness of disease management in depression may change over time, particularly when it is implemented in broader patient populations and in less controlled settings. Effectiveness may increase as implementation strategies are refined and improved.

As with other chronic and relapsing illnesses, depression is a condition for which results may depend on the therapeutic alliance that results when the patient and a team of providers participate in treatment. It has been suggested that depressed patients benefit most when interactions with their providers are recurrent and varied; the ultimate purpose and length of these visits are less important than their frequency. Disease management pro-

grams may improve outcomes of treatment by fostering a structured treatment environment in which such therapeutic alliances can be forged more readily.

Disease management programs can improve quality of care and outcomes for patients with depression, as reflected in improvements in measures of both processes and outcomes of care. However, such programs also increase treatment costs. Although investment of resources may be required to achieve improved outcomes for patients, potential cost savings could result in other areas, such as other health care utilization and employee productivity.

Received Oct. 2, 2002; revision received March 11, 2003; accepted April 10, 2003. From Zynx Health, a Cerner Company, Beverly Hills, Calif.; TAP Pharmaceutical Products Inc., Lake Forest, Ill.; Duke Clinical Research Institute, Duke University, Durham, N.C.; and the Department of Medicine, Cedars-Sinai Health System, Los Angeles. Address reprint requests to Dr. Badamgarav, Zynx Health, 9100 Wilshire Blvd., East Tower, Suite 655, Beverly Hills, CA 90212; ebadamgarav@cerner.com (e-mail).

Supported in part by a research grant from TAP Pharmaceutical Products Inc. Mr. Henning is an employee of TAP Pharmaceutical Products Inc. During the course of the research project, Dr. Ofman was a full-time employee of Zynx Health; currently he is with Amgen, Inc., Thousand Oaks, Calif.

References

1. Public Health Service Agency for Health Care Policy and Research: Depression in Primary Care: Treatment of Major Depression: AHCPR Publication 93-051. Rockville, Md, US Department of Health and Human Services, 1993
2. Wells KB: Caring for depression in primary care: defining and illustrating the policy context. *J Clin Psychiatry* 1997; 58(suppl 1):24-27
3. Angst J: Comorbidity of mood disorders: a longitudinal prospective study. *Br J Psychiatry* 1996; 30(suppl):31-37
4. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen H-U, Kendler KS: Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994; 51:8-19
5. Kessler RC, Nelson CB, McGonagle KA, Liu J, Swartz M, Blazer DG: Comorbidity of DSM-III-R major depressive disorder in the general population: results from the US National Comorbidity Survey. *Br J Psychiatry* 1996; 30(suppl):17-30
6. Kessler RC, Zhao S, Blazer DG, Swartz M: Prevalence, correlates, and course of minor depression and major depression in the National Comorbidity Survey. *J Affect Disord* 1997; 45:19-30
7. Sartorius N, Ustun TB, Lecrubier Y, Wittchen HU: Depression comorbid with anxiety: results from the WHO Study on Psychological Disorders in Primary Health Care. *Br J Psychiatry* 1996; 30(suppl):38-43
8. Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, Joyce PR, Karam EG, Lee CK, Lellouch J, Lepine JP, Newman SC, Rubio-Stipec M, Wells JE, Wickramaratne PJ, Wittchen H, Yeh EK: Cross-national epidemiology of major depression and bipolar disorder. *JAMA* 1996; 276:293-299
9. Panzarino PJ Jr: The costs of depression: direct and indirect; treatment versus nontreatment. *J Clin Psychiatry* 1998; 59(suppl 20):11-14
10. Druss BG, Rosenheck RA, Sledge WH: Health and disability costs of depressive illness in a major US corporation. *Am J Psychiatry* 2000; 157:1274-1278

11. Greenberg PE, Stiglin LE, Finkelstein SN, Berndt ER: The economic burden of depression in 1990. *J Clin Psychiatry* 1993; 54:405-418
12. Ballenger JC, Davidson JR, Lecrubier Y, Nutt DJ, Goldberg D, Magruder KM, Schulberg HC, Tylee A, Wittchen HU: Consensus statement on the primary care management of depression from the International Consensus Group on Depression and Anxiety. *J Clin Psychiatry* 1999; 60(suppl 7):54-61
13. Murray CJ, Lopez AD: Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997; 349:1498-1504
14. Conner TM, Crismon ML, Still DJ: A critical review of selected pharmacoeconomic analyses of antidepressant therapy. *Ann Pharmacother* 1999; 33:364-372
15. Kind P, Sorensen J: The costs of depression. *Int Clin Psychopharmacol* 1993; 7:191-195
16. Rice DP, Miller LS: The economic burden of affective disorders. *Br J Psychiatry* 1995; 27(suppl):34-42
17. Stoudemire A, Frank R, Hedemark N, Kamlet M, Blazer D: The economic burden of depression. *Gen Hosp Psychiatry* 1986; 8: 387-394
18. Keller MB, Klerman GL, Lavori PW, Fawcett JA, Coryell W, Endicott J: Treatment received by depressed patients. *JAMA* 1982; 248:1848-1855
19. Katon W, Von KM, Lin E, Bush T, Ormel J: Adequacy and duration of antidepressant treatment in primary care. *Med Care* 1992; 30:67-76
20. Simon GE, Von Korff M, Wagner EH, Barlow W: Patterns of antidepressant use in community practice. *Gen Hosp Psychiatry* 1993; 15:399-408
21. Ormel J, Oldehinkel T, Brilman E, vanden Brink W: Outcome of depression and anxiety in primary care: a three-wave 3 1/2-year study of psychopathology and disability. *Arch Gen Psychiatry* 1993; 50:759-766
22. Simon GE, Maier W, Ustun TB, Linden M, Boyer P: Research diagnosis of current depressive disorder: a comparison of methods using current symptoms and lifetime history. *J Psychiatr Res* 1995; 29:457-465
23. Wells KB, Hays RD, Burnam MA, Rogers W, Greenfield S, Ware JE Jr: Detection of depressive disorder for patients receiving prepaid or fee-for-service care: results from the Medical Outcomes Study. *JAMA* 1989; 262:3298-3302
24. Eisenberg L: Treating depression and anxiety in primary care: closing the gap between knowledge and practice. *N Engl J Med* 1992; 326:1080-1084
25. Von Korff M, Katon W, Unutzer J, Wells K, Wagner EH: Improving depression care barriers, solutions, and research needs. *J Fam Pract* 2001; 50:E1
26. Wagner EH, Austin BT, Von Korff M: Organizing care for patients with chronic illness. *Milbank Q* 1996; 74:511-544
27. Ellrod G, Cook DJ, Lee J, Cho M, Hunt D, Weingarten S: Evidence-based disease management. *JAMA* 1997; 278:1687-1692
28. Woolf SH: Practice guidelines: a new reality in medicine, I: recent developments. *Arch Intern Med* 1990; 150:1811-1818
29. Cochrane Effective Practice and Organisation of Care Group (EPOC). <http://www.epoc.uottawa.ca/>
30. Cohen J: *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, Lawrence Erlbaum Associates, 1988
31. Cohen J: *Statistical Power Analysis for the Behavioral Sciences*, revised ed. New York, Academic Press, 1977
32. Rosenthal R, Rubin D: Combining results from independent studies. *Psychol Bull* 1979; 85:185-193
33. Glass G: *Summarizing Effect Sizes: New Directions for Methodology of Social and Behavioral Science: Quantitative Assessment of Research Domains*. San Francisco, Jossey-Bass, 1980
34. Rosenthal R: *Meta-Analytic Procedures for Social Research*. Beverly Hills, Calif, Sage Publications, 1984
35. Hedges LV, Olkin I: *Statistical Methods for Meta-Analysis*. Orlando, Fla, Academic Press, 1985
36. DerSimonian R, Laird N: Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7:177-188
37. Egger M, Davey SG, Schneider M, Minder C: Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997; 315: 629-634
38. Light RJ, Pillemer DB: *Summing Up: The Science of Reviewing Research*. Cambridge, Mass, Harvard University Press, 1984
39. Simon GE, Von Korff M, Rutter C, Wagner E: Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. *Br Med J* 2000; 320:550-554
40. Katon W, Robinson P, Von Korff M, Lin E, Bush T, Ludman E, Simon G, Walker E: A multifaceted intervention to improve treatment of depression in primary care. *Arch Gen Psychiatry* 1996; 53:924-932
41. Katon W, Von Korff M, Lin E, Walker E, Simon GE, Bush T, Robinson P, Russo J: Collaborative management to achieve treatment guidelines: impact on depression in primary care. *JAMA* 1995; 273:1026-1031
42. Katon W, Rutter C, Ludman EJ, Von Korff M, Lin E, Simon G, Bush T, Walker E, Unutzer J: A randomized trial of relapse prevention of depression in primary care. *Arch Gen Psychiatry* 2001; 58:241-247
43. Rost K, Nutting P, Smith J, Werner J, Duan N: Improving depression outcomes in community primary care practices: a randomized trial of the QuEST intervention. *J Gen Intern Med* 2001; 16:143-149
44. Whooley MA, Stone B, Soghikian K: Randomized trial of case-finding for depression in elderly primary care patients. *J Gen Intern Med* 2000; 15:293-300
45. Hunkeler EM, Meresman JF, Hargreaves WA, Fireman B, Berman WH, Kirsch AJ, Groebe J, Hurt SW, Braden P, Getzell M, Feigenbaum PA, Peng T, Salzer M: Efficacy of nurse telehealth care and peer support in augmenting treatment of depression in primary care. *Arch Fam Med* 2000; 9:700-708
46. Katelnick DJ, Simon GE, Pearson SD, Manning WG, Helstad CP, Henk HJ, Cole SM, Lin EH, Taylor LH, Kobak KA: Randomized trial of a depression management program in high utilizers of medical care. *Arch Fam Med* 2000; 9:345-351
47. Thompson C, Kinmouth AL, Stevens L, Peveler RC, Stevens A, Ostler KJ, Pickering RM, Baker NG, Henson A, Preece J, Cooper D, Campbell MJ: Effects of a clinical-practice guideline and practice-based education on detection and outcome of depression in primary care: Hampshire Depression Project randomized controlled trial. *Lancet* 2000; 355:185-191
48. 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; 4:170-178
49. Wells KB, Sherbourne C, Schoenbaum M, Duan N, Meredith L, Unutzer J, Miranda J, Carney MF, Rubenstein LV: Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA* 2000; 283:212-220
50. Katon W, Von Korff M, Lin E, Simon G, Walker E, Unutzer J, Bush T, Russo J, Ludman E: Stepped collaborative care for primary care patients with persistent symptoms of depression: a randomized trial. *Arch Gen Psychiatry* 1999; 56:1109-1115
51. Lin EHB, Simon GE, Katon WJ, Russo JE, VonKorff M, Bush TM, Ludman EJ, Walker EA: Can enhanced acute-phase treatment of depression improve long-term outcomes? a report of randomized trials in primary care. *Am J Psychiatry* 1999; 156:643-645

DISEASE MANAGEMENT FOR DEPRESSION

52. Llewellyn-Jones RH, Baikie KA, Smithers H, Cohen J, Snowdon J, Tennant CC: Multifaceted shared care intervention for late life depression in residential care: randomized controlled trial. *Br Med J* 1999; 319:676–682
53. Worrall G, Angel J, Chaulk P, Clarke C, Robbins M: Effectiveness of an educational strategy to improve family physicians' detection and management of depression: a randomized controlled trial. *Can Med Assoc J* 1999; 161:37–40
54. Banerjee S, Shamash K, Macdonald AJ, Mann AH: Randomised controlled trial of effect of intervention by psychogeriatric team on depression in frail elderly people at home. *Br Med J* 1996; 313:1058–1061
55. Callahan CM, Hendrie HC, Dittus RS, Brater DC, Hui SL, Tierney WM: Improving treatment of late life depression in primary care: a randomized clinical trial. *J Am Geriatr Soc* 1994; 42: 839–846
56. Goldberg HI, Wagner EH, Fihn SD, Martin DP, Horowitz CR, Christensen DB, Cheadle AD, Diehr P, Simon G: A randomized controlled trial of CQI teams and academic detailing: can they alter compliance with guidelines? *Jt Comm J Qual Improv* 1998; 24:130–142
57. Brown JB, Shye D, McFarland BH, Nichols GA, Mullooly JP, Johnson RE: Controlled trials of CQI and academic detailing to implement a clinical practice guideline for depression. *Jt Comm J Qual Improv* 2000; 26:39–54
58. Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, Lohr KN: Screening for depression in adults: a summary of the evidence for the US Preventive Services Task Force. *Ann Intern Med* 2002; 136:765–776
59. Weingarten SR, Henning JM, Badamgarav E, Knight K, Hasselblad V, Gano A Jr, Ofman JJ: Interventions used in disease management programmes for patients with chronic illness—which ones work? Meta-analysis of published reports. *BMJ* 2002; 325: 925–928
60. Counsell SR, Holder CM, Liebenauer LL, Palmer RM, Fortinsky RH, Kresevic DM, Quinn LM, Allen AR, Covinsky KE, Landerfeld CS: Effects of a multicomponent intervention on functional outcomes and process of care in hospitalized older patients: a randomized controlled trial of Acute Care for Elders (ACE) in a community hospital. *J Am Geriatr Soc* 2000; 48:1572–1581
61. Kazis LE, Anderson JT, Meenan RF: Effect sizes for interpreting changes in health status. *Med Care* 1989; 27(suppl 3):S178–S189