Cost-effectiveness of a Primary Care Depression Intervention

Jeffrey M. Pyne, MD, Kathryn M. Rost, PhD, Mingliang Zhang, PhD, D. Keith Williams, PhD, Jeffrey Smith, BS, John Fortney, PhD

OBJECTIVE: To determine the incremental cost-effectiveness of a quality improvement depression intervention (enhanced care) in primary care settings relative to usual care.

DESIGN: Following stratification, we randomized 12 primary care practices to enhanced or usual care conditions and followed patients for 12 months.

SETTING: Primary care practices located in 10 states across the United States.

PATIENTS/PARTICIPANTS: Two hundred eleven patients beginning a new treatment episode for major depression.

INTERVENTIONS: Training the primary care team to assess, educate, and monitor depressed patients during the acute and continuation stages of their depression treatment episode over 1 year.

MEASUREMENTS AND MAIN RESULTS: Cost-effectiveness was measured by calculating incremental (enhanced minus usual care) costs and quality-adjusted life years (QALYs) derived from SF-36 data. The mean incremental cost-effectiveness ratio in the main analysis was $15,463 per QALY. The mean incremental cost-effectiveness ratios for the sensitivity analyses ranged from $11,341 (using geographic block variables to control for pre-intervention service utilization) to $19,976 (increasing the cost estimates by 50%) per QALY.

CONCLUSIONS: This quality improvement depression intervention was cost-effective relative to usual care compared to cost-effectiveness ratios for common primary care interventions and commonly cited cost-effectiveness ratio thresholds for intervention implementation.

KEY WORDS: cost-benefit analysis; depression; quality of life; primary health care.
J GEN INTERN MED 2003;18:432–441.

With major depression projected to constitute an increasing share of the global burden of disease,[1] there have been concerted efforts worldwide to develop more effective depression management strategies. In the United States, efforts to improve depression management have focused on the primary care setting,[2,3] where a wide diversity of "best practice models" to diagnose and manage depression have been tested.[4–15] Incremental cost-effectiveness analyses provide a useful framework for comparing a wide variety of best practice models.[16]

Interventions that integrate mental health professionals into the primary care setting have demonstrated remission rates similar to those in specialty care efficacy studies.[5,7] However, widespread dissemination of integrated interventions is unlikely, because the majority of primary care clinics do not employ on-site mental health professionals.[17] The quality improvement intervention tested in this study attempted to fill this gap by training primary care professionals to more effectively identify and treat depression. In particular, the intervention trained office nurses to supplement the primary care physician's efforts to provide antidepressant medication treatment or referral to mental health counseling. The development of the intervention is presented in more detail elsewhere.[18]

In a prospective study, we evaluated the incremental cost-effectiveness of this brief quality improvement intervention for primary care patients beginning a new treatment episode for major depression relative to usual care. We compared the quality-adjusted life years (QALYs) of patients in practices who received the intervention (enhanced care) to patients in practices who did not (usual care). We estimated costs from a societal perspective by calculating the costs of the intervention, health care utilization, patient time, and transportation. We hypothesized that over a 12-month period of time, the brief intervention would be cost-effective compared to the cost-effectiveness (CE) ratios of common primary care interventions and commonly cited CE ratio thresholds for intervention implementation.

METHODS

Design

We used a randomized block design described previously to compare outcome differences between enhanced and usual care.[18] A 2-stage stratification plan was used to randomly assign practices to enhanced or usual care. The first stage divided 12 practices identified by numerical code into metropolitan versus nonmetropolitan location in order

to identify primary care sites with more (metropolitan) or less (nonmetropolitan) access to off-site mental health specialty care. The second stage paired 8 metropolitan practices and 4 nonmetropolitan practices based on the participating physicians' baseline (i.e., pre-intervention) proclivity to treat depressed patients in accordance with Agency for Healthcare Research and Quality (AHRQ) depression treatment guideline recommendations.[19,20] Baseline physician concordance with AHRQ treatment guidelines (concordance defined as patient on guideline-recommended dose of antidepressant medication and/or in specialty care counseling) was determined by reviewing treatment logs completed by each participating physician for 20 consecutive depressed patients.[18] One practice from each of the 6 pairs was then randomly selected to participate in the enhanced care intervention and the other practice provided usual care.

Intervention

Primary care teams assigned to the enhanced care condition completed the intervention training before subject recruitment began. Physicians and nurses completed a series of 4 academic detailing telephone calls over a 2-month period, the goal of which was to systematically engage providers with the content of Agency for Health Care Policy and Research Guidelines.[20,21] The intervention presented pharmacotherapy and psychotherapy as equally efficacious treatments.[18] The nurses completed an additional 8-hour face-to-face training session, conducted by the research team, designed to teach them to assess, educate, and monitor depressed patients during the acute and continuation stages of their depression episode.[18] Administrative staff in both enhanced and usual care practices completed an 8-hour training session on recruiting eligible patients using a 2-stage screening process.

At the initial visit, the nurse assessed depression symptoms, provided information to the patient about his/her preferred treatment, asked patients to complete an individualized assignment to increase or maintain their readiness to engage in active treatment, and arranged a time to talk with the patient during the next week. This material was summarized and documented on a short checklist attached to the front of the patient's chart prior to the physician visit. The nurses used a similar protocol to conduct 15-minute telephone or in-person follow-up discussions with patients during the next 5 to 7 weeks, averaging a total of 5.2 (SD = 1.9) contacts with each patient participating in the acute treatment stage. The continuation stage intervention was implemented on average 9 months after subjects' index visits to facilitate the re-initiation or adjustment of treatment in patients who were symptomatic (i.e., reporting 3 or more depression symptoms) at that time-point. Patient components of the continuation intervention included periodic symptom/treatment monitoring by nurse care managers, who also encouraged

patients to participate in active treatment. Physician components of the continuation intervention included reviewing monthly patient symptom/treatment summaries and recommendations for treatment re-initiation/adjustment (see Appendix at www.blackwellpublishing.com/jgi).[22] Nurses completed an average of 4.0 (SD = 2.9) symptom/treatment monitoring contacts among patients participating in the continuation intervention prior to 12-month follow-up. Physicians in usual care practices were not informed about which patients were participating in the study, nor did their nurses receive the enhanced care training or contact depressed subjects on a regular basis. All primary care professionals were salaried employees of the practice, not the study; however, the time they spent providing the intervention was factored into the intervention costs as described below.

Sites

The 12 primary care practices were located in 10 states across the country. Each practice was a member of 1 of 3 practice research networks, had 2 primary care physicians (family physicians or internists) willing to participate in the study, an office nurse willing to deliver the nursing intervention as detailed in the protocol if the practice was randomized to the enhanced care condition, and practice coordinators (administrative staff) willing to screen patients for major depression as part of usual care; none of the practices engaged an on-site mental health professional to provide depression treatment.

Subjects

Subject recruitment, described at length elsewhere,[18] included a 2-stage screening process in 1996–1997 to identify a representative group of primary care patients beginning a new treatment episode for major depression, as defined in the *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV), before they saw the participating physician for the index visit (Fig. 1).

Patients were eligible for this analysis if they: i) reported 5 or more of the 9 criteria for major depression in the past 2 weeks using the Inventory to Diagnose Depression[23]; ii) screened negative for lifetime mania; iii) screened negative for lifetime alcohol dependence with current drinking; iv) did not meet DSM-IV criteria for bereavement-related depression; v) reported no antidepressant medication in the last month and no specialty mental health care in the last 6 months; and (vi) had sufficient SF-36 data at baseline, 6, and 12 months to calculate SF-36 quality-adjusted index scores. The sixth eligibility criterion eliminated 5 otherwise eligible subjects from this study. In virtually all cases, both screening and enrollment were completed before the patient saw the physician for the index visit. Study enrollment procedures, including management of suicidal intent, were approved by the Human Research Advisory
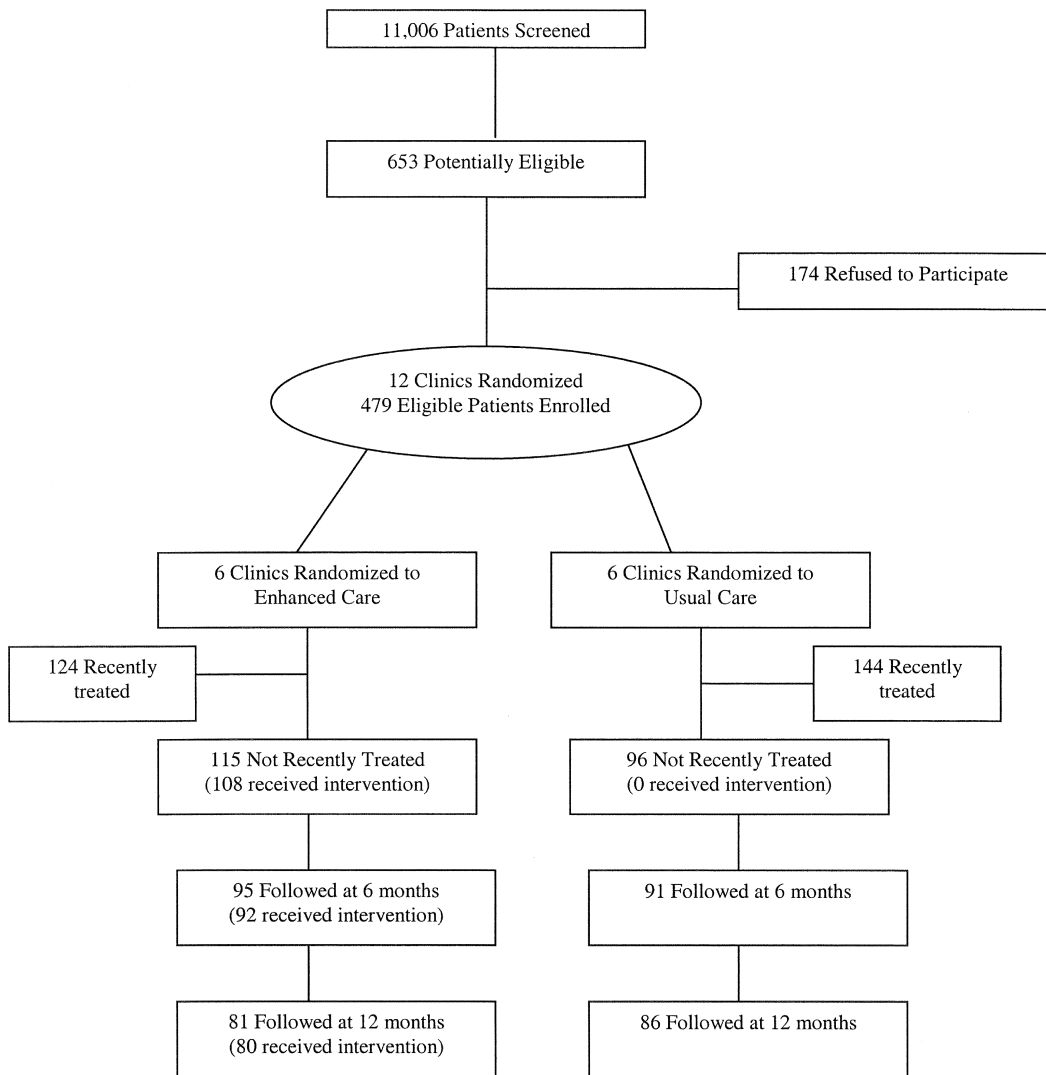
FIGURE 1. Subject recruitment and participation flow chart.

Committee of the University of Arkansas for Medical Sciences and the Colorado Multi-institutional Review Board. Although this study was one of the first investigations to attempt to enroll a representative (rather than referred) population of depressed primary care patients, we increased the representativeness of the sample by weighting subject responses to reflect the probability of being enrolled in the sample and the probability of completing the follow-up interviews. The enrollment and completed interview weights were based on demographic and clinical characteristics of enrolled and eligible subjects, and interview completed and not completed subjects, respectively.

Data Collection

Patients completed telephone interviews with the research team during the week following the index visit before any newly initiated intervention would be expected

to have a substantial effect. Structured telephone interviews were conducted again at 6 (88.2% of subjects completed) and 12 months (79.2% of subjects completed) following the index visit by the same interviewer. The interviewer was blinded to the subject's treatment condition except in 3 cases wherein the research team needed to contact the practice for assistance locating a subject.

Measurement of Quality-adjusted Life Years

We used a conversion formula reported by Brazier et al.[24] to calculate QALYs from baseline, 6-, and 12-month SF-36 data using visual analog scale (VAS) quality-adjustment weights. We chose the Brazier et al. method because it was based on a well-validated and commonly used generic health-related quality-of-life instrument, the SF-36. We chose the VAS adjustment weights because the VAS model satisfied the conventional diagnostic statistical tests and the standard gamble model did not.[24] Based on

Tukey's 1 degree of freedom test, we found that adding quadratic terms for physical functioning and the mental health index to the original conversion formula improved the fit of the regression model of SF-36 subscale scores predicting the visual analog scale scores. In the main analysis, we used the SF-36 to QALY conversion formula with the additional quadratic terms. In 1 of the sensitivity analyses, we used the original SF-36 to QALY conversion formula without the quadratic terms. In order to calculate the QALYs associated with each subject, we used an area under the curve calculation of the baseline, 6-, and 12-month quality-adjusted SF-36 data.[25]

Measurement of Costs

The cost of the intervention was determined from a societal perspective using a cost accountant approach.[26] As previously described, intervention costs included training and implementation costs for physicians, nurses, and office staff coordinators for the acute phase of the intervention.[18] Training costs included trainee time, airfare, meals and lodging, training manual costs, and trainer time. Implementation costs included primary care clinic staff time costs for patient screening, preparation for and delivery of the patient intervention, postsession record-keeping, communication among providers delivering the intervention, and posttraining supervision. The implementation costs of the acute and continuation stage components of the intervention across all study subjects over the 12-month period were $130 per capita. The implementation plus training costs were $247 per capita.

We also estimated the differences in total outpatient health care expenditures, time, and transportation costs between enhanced and usual care patients over a 12-month period (Table 1). We did not include hospital costs in the analysis because they were infrequent and expensive events in our sample of primary care depressed patients and our sample size was too small to provide reliable inferences with hospital costs in the analysis. Outpatient health care expenditures for emergency room visits, primary

care and specialty mental health care visits, and psychotropic medication during the time of the intervention were estimated from patient report at the 6- and 12-month interviews. Preintervention health care expenditures were estimated for the 6 months prior to the baseline interview. All costs were adjusted to reflect year 2000 dollars. Costs for an emergency room visit were estimated at $500.[27] Outpatient visit costs were estimated using 1999 average Medicare payment rates adjusted to year 2000 dollars. For primary care physical health visits we used $35.51, for primary care mental health visits we used $55.78, for specialty mental health psychiatrist visits we used $92.83, and for specialty mental health nonpsychiatrist visits we used $87.34. Psychotropic medication costs were priced at the lowest average generic wholesale price per medication dosage as reported in the 2000 *Red Book of Prescription Drugs*.

Time costs were estimated from patient reports of travel times to and from the clinic and waiting time. For employed patients, time costs were calculated using self-reported wage rates. For unemployed patients, we used the Federal minimum hourly wage for the year 2000 ($5.15) as a proxy of their time costs. Transportation costs were calculated from patient-reported round-trip miles to and from the location of services at a mileage rate of $0.325/mile.

Incremental Cost-effectiveness Ratio

The numerator in the incremental CE ratio is the difference in cost between enhanced care (EC) and usual care (UC). The denominator is the difference in QALYs between enhanced and usual care.

$$\frac{\text{cost (EC)} - \text{cost (UC)}}{\text{QALY (EC)} - \text{QALY (UC)}}$$

Costs and health effects were not discounted because the time horizon of the study was only 12 months.[16] For the main analysis, we defined total costs as intervention implementation costs, outpatient healthcare costs (emergency room, outpatient primary care and specialty care, and psychotropic medication), time costs (travel time to

Table 1. Descriptive Overview of Average Costs per Patient by Cost Domain for Enhanced and Usual Care Groups in Year 2000 U.S. Dollars

	6 Months Pre-intervention		0–6 Months Post-intervention		6–12 Months Post-intervention	
	EC	UC	EC	UC	EC	UC
PC (ph)	135	156	111	142	107	160
PC (mh)	82	102	150	86	94	93
SC (mh)	0	0	183	39	155	32
MH meds	24	20	245	103	215	112
Emergency room	382	399	300	246	249	259
Patient transport	30	60	57	42	38	74
Patient time	104	139	145	125	143	137

EC, enhanced care; UC, usual care; PC (ph), primary care visits for physical health problems; PC (mh), primary care visits for mental health problems; SC (mh), specialty mental health care visits; MH meds, mental health (psychotropic) medications.

and from medical services, waiting time, and time with the clinician), and transportation costs.[16] We did not include hospital costs because they typically constitute the largest single item dollar cost and affect only a small proportion of subjects and, therefore, are difficult to estimate with precision. We did not include productivity costs (lost wages for partial and full workdays missed for health-related reasons) because the enhanced productivity or functioning associated with an intervention is often thought of as part of the nonmonetized denominator of a cost-effectiveness ratio. In addition, it is difficult to separate the health-related quality-of-life (HRQL) effects of being ill from effects on role function and other experiences associated with the use of time.[16]

Covariates

Sociodemographic covariates included: age and annual household income adjusted by family size as continuous variables; and gender, minority status, marital status (currently married or not), education (high school graduate or not), employment status (working full/part time or not), and health insurance status (insured or not) as dichotomous variables. Clinical covariates included continuous measures of physical health comorbidity, depression severity, and a dichotomous measure of dysthymia during the past year. Physical health comorbidity was measured by summing the number of 14 chronic physical conditions the subject reported at baseline and coding as 0, 1, or 2 or more physical health comorbidities. Depression severity was measured using a modified Center for Epidemiologic Studies – Depression (CES-D) scale. The CES-D was modified (mCES-D) in order to include more items that reflected the DSM-IV criteria for major depression. The modifications are described in more detail elsewhere.[15] Briefly, 7 items from the original 20-item CES-D that did not directly parallel DSM-IV criteria for major depression were removed, and 10 items were added, using the original CES-D question format. Scoring was standardized on a 0- to 100-point scale with higher scores reflecting more severe depressive symptoms. The α coefficient for the mCES-D scale was 0.91. Dysthymia was measured using World Health Organization–Composite International Diagnostic Interview criteria for dysthymia within the past year.[28] Antidepressant medication and mental health counseling acceptance were collected at baseline using a 4-point Likert-type scale (definitely acceptable, probably acceptable, probably not acceptable, and definitely not acceptable). These results were dichotomized into acceptable and not acceptable in order to simplify the interpretation of these parameter estimates.

Statistical Analysis

We conducted intent-to-treat analyses, comparing enhanced versus usual care outcomes (costs and QALYs).

We estimated random effects models, in which patients were nested within practices, to test for intraclass correlation at the practice level for costs and QALYs. Results indicated minimal intraclass correlation (ICC) and little variation across practices with respect to cost and QALYs (ICC = 0.025 and 0.0007, respectively). Therefore, ordinary least squares regression analysis was used to estimate the impact of the intervention on costs and QALYs.

We evaluated the incremental effect (enhanced minus usual care) of the intervention on costs and QALYs using the perspective of a typical patient in our sample. The model for predicting costs was the same as described above for predicting QALYs, except that we added preintervention costs as a covariate in the cost model. Because of the skewed distribution of health care costs, we used a 2-part log transformation of health care costs. The distribution of the log-transformed cost residuals was not normal but homoscedastic, and we used the appropriate smearing retransformation to calculate predicted costs.[29] In addition, because the distribution of the retransformed incremental cost estimate is not well defined, we used a nonparametric bootstrap with replacement method to generate 95% confidence intervals for each mean incremental cost estimate.[30]

Typical standard error estimation methods do not apply to CE ratios because the possibility of having a zero or near zero denominator is nonnegligible, and the independence of cost and effectiveness estimates is not certain.[31] Therefore, we used a nonparametric bootstrap with replacement method to generate a joint distribution of costs and QALYs.[31] We solved simultaneous multiple regression equations for costs and QALYs with each bootstrap sample. Using this bootstrap with replacement method and 1,000 replications, we generated a bivariate incremental cost and QALY plot,[32] and acceptability curve representing the probability that the mean incremental CE ratio will be less than CE thresholds ranging from $0 to $100,000 per QALY saved.[33]

We conducted 4 sensitivity analyses. In the first sensitivity analysis, we controlled for preintervention health care costs by using stratification block variables describing the 6 paired enhanced/usual care clinics instead of preintervention patient-reported costs. The reason for conducting this sensitivity analysis was to facilitate comparison with a cost-effectiveness analysis (CEA) from a companion study (Partners in Care), which used the block variable approach.[27] In the second sensitivity analysis, we added the training costs to the main analysis intervention implementation costs. While training costs are often excluded from primary care depression intervention cost-effectiveness analyses,[27,34,35] we included this sensitivity analysis in order to reflect all of the start-up costs associated with implementing the intervention. In the third sensitivity analysis, we calculated QALYs using the original SF-36 conversion formula as described by Brazier et al.[24] without the quadratic terms (described in the Measurement of Quality-adjusted Life Years section above).

This sensitivity analysis was done in order to facilitate comparisons with CEAs that do not include the quadratic terms. In the fourth sensitivity analysis, we varied the outpatient health care cost estimates listed in the Methods section (emergency room visits, primary care and specialty mental health care visits, and psychotropic medication) by plus and minus 50%. This sensitivity analysis was done to determine the robustness of our conclusions to the cost estimates used in the analysis.

The authors are not aware of any conflict of interest associated with the conduct of this research study or the preparation of this manuscript. Specifically, we are not aware of any personal or financial relationships that might bias this work. The study sponsors were not involved in the study design, data collection, analysis, or interpretation, manuscript preparation, or decision to submit this manuscript for publication. The authors had full access to the data and accept full responsibility for the integrity of the data and the data analysis.

RESULTS

Patient Description

As shown in Figure 1 and described in more detail elsewhere,[18] 5.9% (653/11,006) of patients screened were potentially eligible for the study, and 26.6% of these patients (174/653) did not complete the baseline interview where final eligibility of this sample was determined.

As shown in Table 2, the 96 subjects in usual care practices and 115 subjects in the enhanced care practices were sociodemographically and clinically comparable with 3 exceptions. Subjects in usual care were older (46.6 vs 40.2 years, $P = .002$), less depressed on the mCES-D scale (50.9 vs 57.6, $P = .01$), and had more physical health comorbidities (63.5% vs 46.1% had 2 or more physical health comorbidities, $P = .03$) than the enhanced care subjects. Compared to the usual care group, the enhanced care patients reported greater use of antidepressant medications (65% vs 32%, $P < .001$) during the 12 months following baseline. There were no differences in baseline SF-36 VAS scores or in antidepressant acceptability between the usual and enhanced care patients. All significant baseline differences between enhanced and usual care patients were controlled for in multivariate analyses.

Intervention Effects on QALY

The effect of the intervention on QALYs, controlling for the sociodemographic and clinical covariates listed above was significant ($\beta = 0.04$, $P < .05$) (Table 3). In the bootstrap with replacement sample, the mean incremental outcome effect of the intervention relative to usual care was 0.041 QALYs (95% confidence interval [95% CI], 0.040 to 0.042). Similarly, the decrease in depression severity from baseline to 12 months was 7.7 units greater in the enhanced versus usual care groups ($P < .05$). In addition, a growth curve

Table 2. Sociodemographic and Clinical Comparison by Treatment Group

	Usual Care N = 96	Enhanced Care N = 115
Mean age, y (SD)	46.6* (14.2)	40.2 (14.7)
Mean household income, adjusted (SD)[†]	13.6 (23.9)	9.7 (10.7)
Mean mCES-D (SD)	50.8* (19.2)	57.6 (18.5)
Mean VAS SF-36 (SD)[‡]	0.446 (0.160)	0.453 (0.127)
Female, %	85.4	83.5
Caucasian, %	81.3	87.0
Married, %	45.8	48.7
HS graduate, %	79.2	79.1
Employed, %	61.5	62.6
Insured, %	79.2	85.2
Dysthymia, %	8.3	11.3
Physical health comorbidities, %		
None	15.6*	27.0
One	20.8	27.0
Two or more	63.5	46.1
Report antidepressants acceptable, %	50.0	54.8
Report counseling acceptable, %	63.5	73.0

* P < .05 for separate comparisons of usual versus enhanced care.
[†] *Annual household income per member of the household divided by 1,000.*
[‡] *VAS SF-36 represents the visual analogue scale conversion of the SF-36 to a quality-adjusted index score.*
mCESD, modified Center for Epidemiologic Studies–Depression scale.

analysis of depression-specific outcomes demonstrated that the intervention improved the probability of depression remission and role functioning over time.[22]

Intervention Effect on Costs

The effect of the intervention on costs, controlling for the sociodemographic and clinical covariates listed above, was also statistically significant ($\beta = 0.40$, $P = .006$). In the bootstrap with replacement sample, the mean incremental cost effect of the intervention relative to usual care was $634 (95% CI, $618 to $650).

Cost-effectiveness Ratios

Using the above data from the bootstrap sample, the mean incremental CE ratio for the intervention versus usual care was $15,463 per QALY (634/0.041). We also created an incremental cost-per-QALY space diagram (Fig. 2) and confirmed that the ratios tended to cluster in the upper right-hand quadrant, also known as the trade-off quadrant.[33,36] The trade-off aspect of the upper right-hand quadrant of Figure 2 is dependent on the CE ratio threshold that is considered acceptable for implementing the intervention (in general, the higher the CE ratio, the less likely

Table 3. Regression Results Predicting QALYs and Cost*

Variable	β for QALY	β for Cost[†]
Age	−0.002[‡]	0.002
Adjusted income	0.0005	0.004
Baseline depression (mCES-D)	−0.003[§]	0.008
Gender (0 = F, 1 = M)	−0.04	−0.11
Ethnicity, caucasian (0 = N, 1 = Y)	0.004	0.29
Marital status, married (0 = N, 1 = Y)	0.005	0.16
High school education (0 = N, 1 = Y)	−0.02	0.11
Employment status (0 = N, 1 = Y)	0.10[§]	−0.20
Health insurance status (0 = N, 1 = Y)	−0.05	0.85[§]
Dysthymia (0 = N, 1 = Y)	−0.02	0.45
Physical health comorbidity (0 = none, 1 = 1, 2 = 2+)	−0.06[§]	−0.11
Report antidepressants acceptable (0 = N, 1 = Y)	−0.006	0.11
Report counseling acceptable (0 = N, 1 = Y)	−0.01	0.22
Intervention (0 = N, 1 = Y)	0.04[‡]	0.40[‖]
Baseline cost (ln)	NA	0.28[§]

* *QALY and cost regression equations included weighted subject responses to account for the probability of enrollment and attrition over time.*
[†] *ln (cost); β is the beta value representing the relationship between the independent variable and QALY or cost controlling for the effect of the other independent variables.*
[‡] P < .05.
[§] P < .001.
[‖] P < .01.
QALY, quality-adjusted life year; mCES-D, modified Center for Epidemiologic Studies–Depression scale.

the intervention will be implemented). Commonly cited CE ratio thresholds are depicted by the diagonal lines in Figure 2 and represent $20,000 per QALY and $50,000 per QALY. The $20,000-per-QALY threshold is similar to the CE ratio estimate for treating mild hypertension in middle-aged males, and the $50,000-per-QALY threshold is similar to the CE ratio estimate for treating end-stage renal failure with peritoneal dialysis.[37]

To enhance the interpretability of the cost-per-QALY space diagram, we used the bootstrap data to generate an incremental CE ratio acceptability curve, which plots the probability that the mean incremental CE ratio will fall below CE thresholds ranging from $0 to $100,000 per QALY (Fig. 3). For example, the probability that the mean CE ratio was less than the $20,000 per QALY threshold was 0.65, and the probability that the mean CE ratio was less than the $50,000 per QALY threshold was 0.91.

Sensitivity Analyses

In the first sensitivity analysis, we controlled for pre-intervention service utilization by using stratification block dummy variables describing the 6 paired enhanced/usual care clinics instead of preintervention costs. The mean
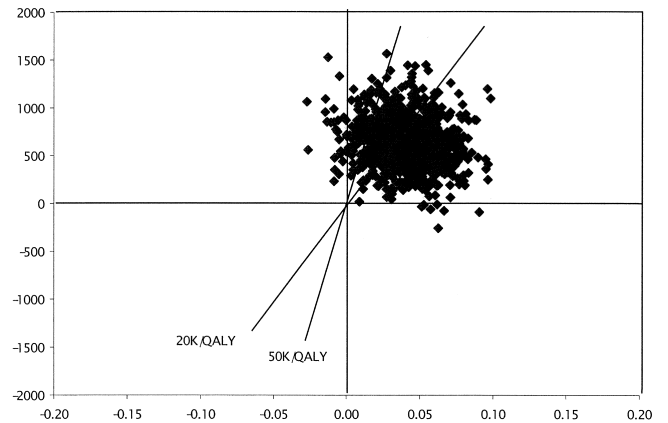
FIGURE 2. Bivariate incremental cost per quality-adjusted life year (QALY) plot, depicting a bootstrap sample (N = 1,000) with replacement of the mean incremental (enhanced minus usual care) cost-effectiveness ratios from this study. The diagonal lines through the center of Figure 2 represent the cost-effectiveness ratio thresholds of $20,000 and $50,000 per QALY.

incremental CE ratio using the geographic block variables was $11,341 per QALY (465/0.041). In the second sensitivity analysis, we included the training cost for the intervention. The resulting mean incremental CE ratio was $17,951 per QALY (736/0.041). In the third sensitivity analysis, we calculated QALYs using the exact SF-36 conversion formula as described by Brazier et al.[24] without the quadratic terms. Without the quadratic terms, the mean incremental QALY score was 0.042. The resulting mean incremental CE ratio was $15,095 per QALY (634/0.042). In the fourth sensitivity analysis we varied each of the health care cost estimates listed in the Methods section by ±50% individually and collectively. The resulting mean incremental cost-effectiveness ratios were all less than
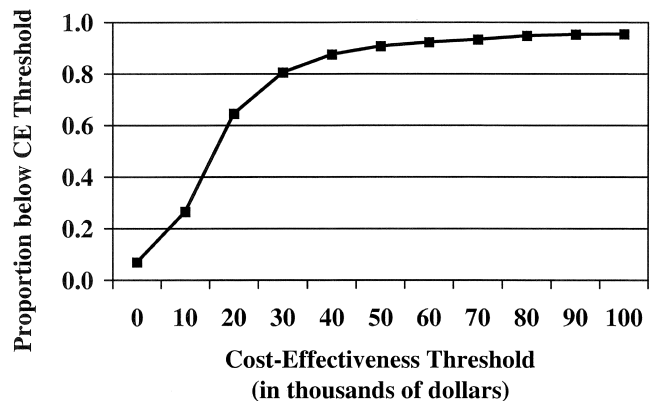
FIGURE 3. Acceptability curve: probability that the mean cost-effectiveness (CE) ratio will fall below CE thresholds ranging from $0 to $100,000 per quality-adjusted life year (QALY). The proportion below the CE threshold is calculated by using the bivariate plot shown in Figure 2 and determining the proportion of subjects below the CE threshold as it is varied from $0 to $100,000 per QALY.

$20,000 per QALY. The largest mean incremental CE ratio was $19,976 per QALY (819/0.041) using 50% greater cost estimates for the emergency room visits, primary care and specialty mental health care visits, and psychotropic medication. Worst and best case scenarios were estimated using the 95% confidence intervals from the bootstrap analyses reported above. The worst-case scenario was $21,075 per QALY (843/0.040), based on the upper confidence limit using 50% greater cost estimates and the lower confidence limit for the SF-36 conversion formula with quadratic terms. The best-case scenario was $10,372 per QALY (446/0.43) using the lower confidence limit for the geographic block variables and the upper confidence limit for the original SF-36 conversion formula.

DISCUSSION

The purpose of this study was to estimate the incremental cost-effectiveness of a depression quality improvement intervention in a representative sample of patients from community primary care practices. This article provides a unique contribution to the cost-effectiveness of depression treatment literature for 2 reasons. First, the QALYs used in this analysis are derived from a generic HRQL instrument, the SF-36, as recommended by the Panel on Cost-Effectiveness in Health and Medicine.[16] In contrast, other depression CE analyses were based on utility weights assigned to depression-specific outcomes,[34,35,38–40] or based solely on depression-specific outcomes such as successfully treated cases[41] or clinician ratings of improvement.[42] Second, the QALYs used in this analysis are based on a total of 9,000 possible SF-36 health states.[24] In contrast, the depression treatment CE analysis that used the SF-12 as a generic HRQL measure limited itself to only 6 SF-12 depression-specific health states.[27]

Among patients beginning a new treatment episode, the mean incremental QALY saved in the main analysis was 0.041, which is in the range considered clinically significant for patients with chronic physical health problems.[43] The main analysis mean CE ratio was $15,463, which is less than (more favorable than) the commonly used CE thresholds of $20,000 and $50,000 per QALY.[44,45] In sensitivity analyses, the range of CE ratios was from $11,341 to $19,976 per QALY. All of the sensitivity analyses except the worst-case scenario resulted in CE ratios less than the more conservative CE ratio threshold of $20,000 per QALY.[46] The consistency of these results supports the robustness of the model. In addition, we would expect the intervention costs to decrease as more patients receive the intervention, because the marginal cost of the intervention (cost of treating an additional subject) would be less than the average intervention cost used to calculate the above CE ratios.

Our estimates of the incremental cost-effectiveness of the depression intervention versus usual care for depression are consistent with previous estimates of the cost-effectiveness of depression interventions. The following CE ratios are adjusted to 2000 dollars to facilitate comparison with the results of our study. Lave et al. reported CE ratios for a nortriptyline protocol versus usual care in primary care settings of $10,632 to $17,857 per QALY and higher ratios for interpersonal psychotherapy versus usual care.[38] Simon et al. reported a CE ratio of $22,748 per QALY for a depression management program for high utilizers of medical care.[35] Schoenbaum et al. reported a CE ratio range of $10,143 to $22,986 per QALY for a Partners in Care quality improvement cognitive therapy intervention.[27]

In addition, our main analysis mean CE ratio ($15,463 per QALY) indicates that this intervention is a good or better healthcare "value" compared to other commonly implemented primary care interventions. For example, CE ratios (adjusted to 2000 U.S. dollars) for some common primary care interventions include: $2,271 for pneumococcal vaccine for the elderly; $8,313 for smoking cessation counseling; $14,015 for treatment of severe hypertension in men; $28,552 for treatment of mild hypertension in men; and $36,428 for chronic obstructive pulmonary disease rehabilitation.[47] Compared to these primary care CE ratios, the value of this primary care depression intervention is similar to that of the treatment of severe hypertension.

Although patients came from community primary care practices across the country, the health outcomes achieved by this brief intervention need to be replicated with a broader range of physicians and ethnically diverse patients to determine if the observed incremental QALYs are generalizeable to other health care settings. The advantages of converting the SF-36 to QALYs include the combination of physical and mental health symptoms and functioning from a well-validated and commonly used health status measure into a single quality-adjusted score as recommended for use in health care economic analyses.[16,48] A limitation of this SF-36 to QALY conversion model is the unrepresentative and relatively small British sample from which the SF-36 quality-adjustment weights were derived; however, available evidence suggests that these factors should not introduce substantial bias into the analysis.[49,50] An additional limitation of this SF-36 to QALY conversion model could be that the health-related quality-of-life data do not include all the relevant domains of depression symptoms and treatment. An alternative approach is to collect utility weights for current health from each subject over time. However, this approach may limit generalizability.

We recognize that patients do not provide perfect estimates of health care utilization; however, the use of administrative data to capture service use was not feasible, because such data did not exist for uninsured participants, and insured participants were enrolled in 65 different health plans. Previous methodological research for 6-month self-report of healthcare utilization found no associations between patient sociodemographic or health indicators and self-report health care utilization discrepancies.[51] In

addition, we found no preintervention differences in health care costs (see Table 1). Therefore, we had no reason to expect differential underreporting in the enhanced and usual care patients. We also did not account for all possible health care costs because we did not include diagnostic testing, nonpsychotropic medications, and inpatient service use. While the skewed cost distribution potentially reduced our ability to draw definitive conclusions about how the intervention affected costs, it was encouraging that when we repeated the analyses without the 2 highest- and 2 lowest-cost subjects, the overall results were unchanged, i.e., the mean CE ratio was less than $20,000 per QALY. We also recognize that the accuracy of the results may have been affected by missing data. To address this problem, we used nonresponse weights to account for the probability of enrollment and attrition over time; however, we cannot know the full extent to which this adjustment was successful.[52]

In summary, this study presents a CE analysis of a primary care depression intervention using a quality-adjusted generic effectiveness measure. The mean incremental CE ratio for this primary care depression intervention is very cost-effective relative to commonly delivered primary care interventions and commonly used CE ratio thresholds. On the basis of these results, this intervention should be implemented for depressed primary care patients beginning a new treatment episode.
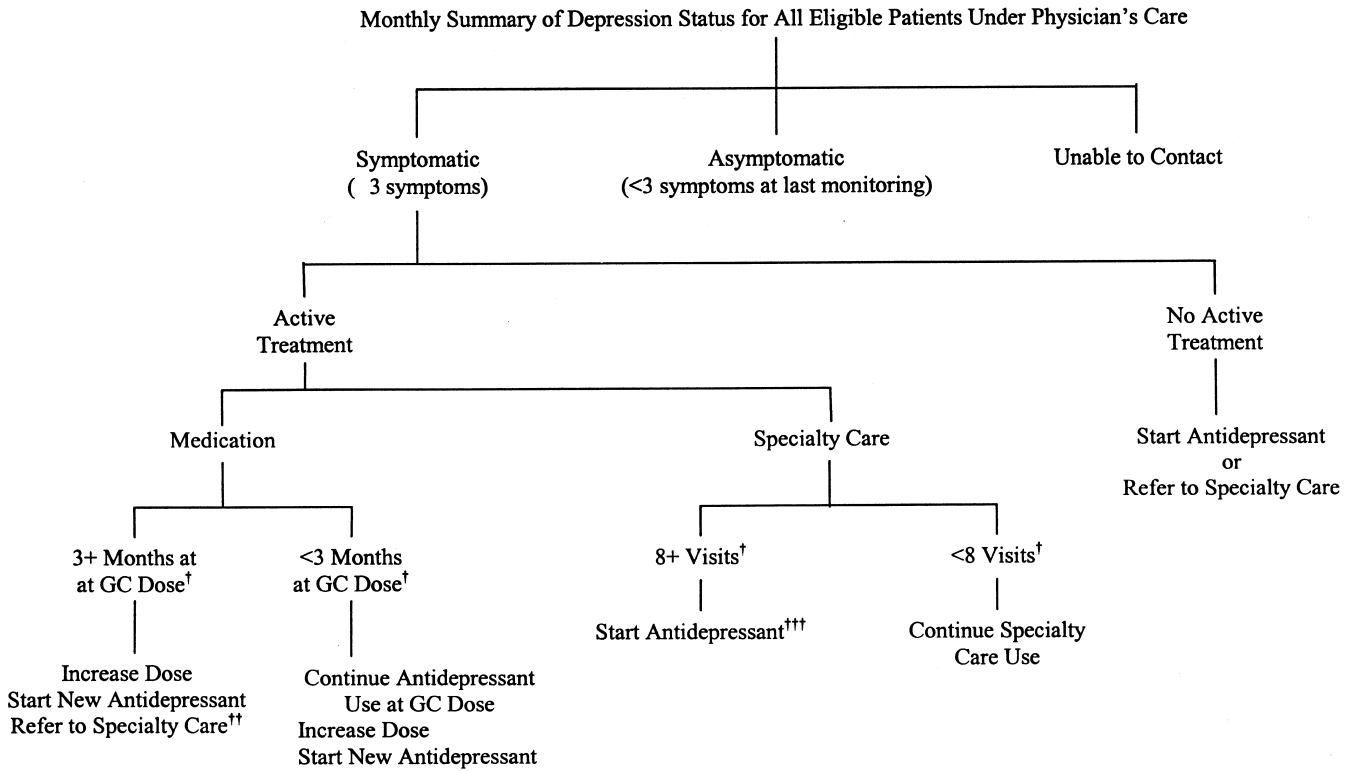

REFERENCES

1. Murray CJL, Lopez AD. The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 Projected to 2020. Cambridge, Mass: Harvard University Press; 1996.
2. Henry JA. Debits and credits in the management of depression. Br J Psychiatry. 1993;(suppl):33–9.
3. Schurman R, Kramer P, Mitchell J. The hidden mental health network: treatment of mental illness by nonpsychiatrist physicians. Arch Gen Psychiatry. 1985;42:89–94.
4. Miranda J, Munoz R. Intervention for minor depression in primary care patients. Psychosom Med. 1994;56:136–41.
5. Katon W, Von Korff M, Lin E, et al. Collaborative management to achieve treatment guidelines. Impact on depression in primary care. JAMA. 1995;273:1026–31.
6. Katon W, Robinson P, Von Korff M, et al. A multifaceted intervention to improve treatment of depression in primary care. Arch Gen Psychiatry. 1996;53:924–32.
7. Schulberg HC, Block MR, Madonia MJ, et al. Treating major depression in primary care practice. Eight-month clinical outcomes. Arch Gen Psychiatry. 1996;53:913–9.
8. Goldberg HI, Wagner EH, Fihn SD, et al. A randomized controlled trial of CQI teams and academic detailing: can they alter compliance with guidelines? Jt Comm J Qual Improv. 1998;24:130–42.
9. Katon W, Von Korff M, Lin E, et al. Stepped collaborative care for primary care patients with persistent symptoms of depression: a randomized trial. Arch Gen Psychiatry. 1999;56:1109–15.
10. Lin EH, Simon GE, Katon WJ, et al. Can enhanced acute-phase treatment of depression improve long-term outcomes? A report of randomized trials in primary care. Am J Psychiatry. 1999;156:643–5.
11. Simon GE, VonKorff M, Rutter C, Wagner E. Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. BMJ. 2000;320:550–4.
12. Schulberg HC, Block MR, Madonia MJ, et al. Randomized trial of a depression management program in high utilizers of medical care. Arch Fam Med. 2000;9:345–51.
13. Wells K, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care. JAMA. 2000;283:212–20.
14. Hunkeler EM, Meresman JF, Hargreaves WA, et al. Efficacy of nurse telehealth care and peer support in augmenting treatment of depression in primary care. (see comments). Arch Fam Med. 2000;9:700–8.
15. Rost KM, Nutting P, Smith J, Werner J, Duan N. Improving depression outcomes in community primary care practice: a randomized trial of the QuEST intervention. J Gen Intern Med. 2001;16:143–9.
16. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness in Health and Medicine. New York, NY: Oxford University Press, Inc.; 1996:425.
17. Williams JW, Rost K, Dietrich AJ, Ciotti MC, Zyzanski SJ, Cornell J. Primary care physicians' approach to depressive disorders. Effects of physician specialty and practice structure. Arch Fam Med. 1999;8:58–67.
18. Rost K, Nutting P, Smith J, Werner J. Designing and implementing a primary care intervention trial to improve the quality and outcome of care for major depression. Gen Hosp Psychiatry. 2000;22:66–77.
19. Depression Guideline Panel. Depression in Primary Care: Detection and Diagnosis. Rockville, Md: Agency for Health Care Policy and Research; 1993:1.
20. Depression Guideline Panel. Depression in Primary Care: Treatment of Major Depression. Rockville, Md: Agency for Health Care Policy Research (AHCPR); 1993:2.
21. Depression Guideline Panel. Treatment of Major Depression: Detection and Diagnosis. Rockville, Md: Agency for Health Care Policy Research (AHCPR); 1993:1.
22. Rost K, Nutting P, Smith JL, Elliott CE, Dickinson M. Managing depression as a chronic disease: a randomized trial of ongoing treatment in primary care. BMJ. 2002;325:934–40.
23. Zimmerman M, Coryell W. The validity of a self-report questionnaire for diagnosing major depressive disorder. Arch Gen Psychiatry. 1988;45:738–40.
24. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. J Clin Epidemiol. 1998;51:1115–28.
25. Ganiats T, Browner D, Kaplan R. Comparison of two methods of calculating Quality-adjusted Life Years. Qual Life Res. 1996;5:162–4.
26. Wolff N, Helminiak TW, Tebes JK. Getting the cost right in cost-effectiveness analyses. Am J Psychiatry. 1997;154:736–43.
27. Schoenbaum M, Unutzer J, Sherbourne C, et al. Cost-effectiveness of practice-initiated quality improvement for depression: results of a randomized controlled trial. JAMA. 2001;286:1325–30.
28. World Health Organization. Composite International Diagnostic Interview, Version 2.0. Geneva, Switzerland: World Health Organization; 1996.
29. Duan N. Smearing estimate: a non-parametric retransformation method. JASA 1983;78:605–10.
30. Efron B. Better bootstrap confidence intervals. JASA 1987;82:171–85.

31. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ. 1997;6:327–40.

32. Anderson JP, Bush JW, Chen M, Dolenc D. Policy space areas and properties of benefit-cost/utility analysis. JAMA. 1986;255: 794–5.

33. Hunink MGM, Bult JR, de Vries J, Weinstein MC. Uncertainty in decision models analyzing cost-effectiveness: the joint distribution of incremental costs and effectiveness evaluated with a nonparametric bootstrap method. Med Decis Making. 1998;18:337–46.

34. Simon GE, Katon WJ, VonKorff M, et al. Cost-effectiveness of a collaborative care program for primary care patients with persistent depression. Am J Psychiatry. 2001;158:1638–44.

35. Simon GE, Manning WG, Katzelnick DJ, Pearson SD, Henk HJ, Helstad CS. Cost-effectiveness of systematic depression treatment for high utilizers of general medical care. Arch Gen Psychiatry. 2001;58:181–7.

36. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making. 1998;18:68S–80S.

37. Kaplan RM, Anderson JP. The general health policy model: an integrated approach. In: Spilker B. ed. Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia, Pa: Lippincott-Raven; 1996:309–21.

38. Lave J, Frank R, Schulberg H, Kamlet M. Cost-effectiveness of treatments for major depression in primary care practice. Arch Gen Psychiatry. 1998;55:645–51.

39. Revicki DA, Brown RE, Keller MB, Gonzales J, Culpepper L, Hales RE. Cost-effectiveness of newer antidepressants compared with tricyclic antidepressants in managed care settings. J Clin Psychiatry. 1997;58:47–58.

40. Kamlet MS, Paul N, Greenhouse J, Kupfer D, Frank E, Wade M. Cost utility analysis of maintenance treatment for recurrent depression. Control Clin Trials. 1995;16:17–40.

41. Von Korff M, Katon W, Bush T, et al. Treatment costs, cost offset, and cost-effectiveness of collaborative management of depression. Psychosom Med. 1998;60:143–9.

42. Forder J, Kavanagh S, Fenyo A. A comparison of the cost-effectiveness of sertraline versus tricyclic antidepressants in primary care. J Affect Disord. 1996;38:97–111.

43. Bombardier C, Ware J, Russell IJ, Larson M, Chalmers A, Read JL. Auranofin therapy and quality of life in patients with rheumatoid arthritis: results of a multicenter trial. Am J Med. 1986;81:565–78.

44. Azimi N, Welch H. The effectiveness of cost-effectiveness analysis in containing costs. J Gen Intern Med. 1998;13:664–9.

45. Kaplan RM, Bush JW. Health-related quality of life measurement for evaluation research and policy analysis. Health Psychol. 1982; 1:61–80.

46. Laupacis A, Feeny D, Detsky A, Tugwell P. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. Can Med Assoc J. 1992;146:473–81.

47. Kaplan RM, Anderson JP. A General Health Policy Model: update and applications. Health Serv Res. 1988;23:203–35.

48. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996;313:275–83.

49. Kaplan RM. Value judgment in the Oregon Medicaid Experiment. Med Care. 1994;32:975–88.

50. Patrick DL, Sittampalam Y, Somerville SM, Carter WB, Bergner M. A cross-cultural comparison of health state values. Am J Public Health. 1985;75:1402–7.

51. Ritter PL, Stewart AL, Kaymaz H, Sobel DS, Block DA, Lorig KR. Self-reports of health care utilization compared to provider records. J Clin Epidemiol. 2001;54:136–41.

52. Young AS, Klap R, Sherbourne CD, Wells KB. The quality of care for depressive and anxiety disorders in the United States. Arch Gen Psychiatry. 2001;58:55–61.

APPENDIX: PHYSICIAN COMPONENTS OF CONTINUATION INTERVENTION

Monthly Summary of Depression Status for All Eligible Patients Under Physician's Care

Symptomatic (3 symptoms)   Asymptomatic (<3 symptoms at last monitoring)   Unable to Contact

Active Treatment   No Active Treatment

Medication   Specialty Care   Start Antidepressant or Refer to Specialty Care

3+ Months at at GC Dose[†]   <3 Months at GC Dose[†]   8+ Visits[†]   <8 Visits[†]

Start Antidepressant[†††]   Continue Specialty Care Use

Increase Dose
Start New Antidepressant
Refer to Specialty Care[††]

Continue Antidepressant Use at GC Dose
Increase Dose
Start New Antidepressant

GC = guideline-concordant

[†]   in current treatment episode
[††]   if not already in Specialty Care
[†††]   if not already on Antidepressant