Psychological Medicine, 2005, **35**, 13–24. © 2004 Cambridge University Press
DOI: 10.1017/S0033291704002624 Printed in the United Kingdom

Randomized trial of a population-based care program for people with bipolar disorder

GREGORY E. SIMON*, EVETTE J. LUDMAN, JÜRGEN UNÜTZER, MARK S. BAUER, BELINDA OPERSKALSKI AND CAROLYN RUTTER

Center for Health Studies, Group Health Cooperative, Seattle, WA, USA; UCLA Neuropsychiatric Institute, Los Angeles, CA, USA; Providence Veterans Affairs Medical Center and Brown University, Providence, RI, USA

ABSTRACT

Background. Despite the availability of efficacious medications and psychotherapies, care of bipolar disorder in everyday practice is often deficient. This trial evaluated the effectiveness of a multi-component care management program in a population-based sample of people with bipolar disorder.

Method. Four hundred and forty-one patients treated for bipolar disorder during the prior year were randomly assigned to continued usual care or usual care plus a systematic care management program including: initial assessment and care planning, monthly telephone monitoring including brief symptom assessment and medication monitoring, feedback to and coordination with the mental health treatment team, and a structured group psychoeducational program – all provided by a nurse care manager. Blinded quarterly assessments generated week-by-week ratings of severity of depression and mania symptoms using the Longitudinal Interval Follow-Up Evaluation.

Results. Participants assigned to the intervention group had significantly lower mean mania ratings averaged across the 12-month follow-up period ($Z = 2{\cdot}44$, $p = 0{\cdot}015$) and approximately one-third less time in hypomanic or manic episode (2·59 weeks *v.* 1·69 weeks). Mean depression ratings across the entire follow-up period did not differ significantly between the two groups, but the intervention group showed a greater decline in depression ratings over time (Z statistic for group-by-time interaction = 1·98, $p = 0{\cdot}048$).

Conclusions. A systematic care program for bipolar disorder significantly reduces risk of mania over 12 months. Preliminary results suggest a growing effect on depression over time, but longer follow-up will be needed.

INTRODUCTION

Since the discovery of lithium's effectiveness over 30 years ago, a number of medications have proven safe and effective for acute and maintenance treatment of bipolar disorder (McElroy & Keck, 2000; APA, 2002; Suppes *et al.* 2002). Despite the availability of several effective medical treatments, typical care for bipolar disorder falls far short of optimal. At any time, over half of people with bipolar disorder are receiving no active treatment (Johnson & McFarland, 1992; Unützer *et al.* 1998). Non-adherence to mood stabilizer medications is frequent (Svarstad *et al.* 2001) with 60–80% of patients interrupting or discontinuing medication over 12 months (Johnson & McFarland, 1996; Unützer *et al.* 2000; Svarstad *et al.* 2001). Frequency of follow-up visits and laboratory monitoring fall far short of recommended rates (Unützer *et al.* 2000). Given these frequent shortcomings, it is not surprising that outcomes in everyday practice do not approach those seen in controlled

* Address for correspondence: Dr Gregory Simon, Center for Health Studies, 1730 Minor Ave, #1600 Seattle, WA 98101, USA.
(Email: simon.g@ghc.org)

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

efficacy trials or specialized bipolar disorder clinics (O'Connell *et al.* 1991; Gitlin *et al.* 1995; Goldberg *et al.* 1995; Coryell *et al.* 1997; Keck *et al.* 1998; Silverstone *et al.* 1998).

Several specific structured psychoeducational programs or psychotherapies have shown promising results in pilot studies or randomized efficacy trials (Perry *et al.* 1999; Huxley *et al.* 2000; Scott *et al.* 2001; Bauer, 2001*a*, *b*; Colom *et al.* 2003; Lam *et al.* 2003; Otto *et al.* 2003). These programs typically incorporate cognitive and behavioral interventions to promote early recognition of mood episodes, reduce exposure to triggers or precipitating stresses, increase medication adherence, and develop effective self-management skills. To date, these programs have not moved beyond efficacy trials into everyday practice.

The chasm between research evidence and everyday practice in the management of bipolar disorder is similar to that seen in other chronic medical illnesses (Institute of Medicine, 2001). Low rates of adherence to pharmacotherapy and premature treatment discontinuation are also seen in treatment of chronic medical conditions such as diabetes, hypertension, or heart failure (Institute of Medicine, 2001). Effective behavioral interventions to promote self-management are rarely provided. Wagner and VonKorff have described a general model to improve management of chronic illnesses (Wagner *et al.* 1996; VonKorff *et al.* 1997). Key elements of this systematic approach include: outreach to promote systematic follow-up, information systems to support follow-up efforts, systematic application of evidence-based guidelines, and organized psychoeducational programs to promote effective self-care. This population-based approach considers the needs of all those suffering from a particular chronic condition rather than those choosing to present for care. Recent trials have demonstrated the effectiveness of this model for management of depression in primary care (Katon *et al.* 1995; Hunkeler *et al.* 2000; Simon *et al.* 2000; Wells *et al.* 2000), but it has not yet been applied to more severe mental disorders.

We describe here a randomized trial evaluating the effectiveness of a multi-component care program in a sample of patients with bipolar disorder systematically sampled from all those treated in a prepaid health plan. The intervention included key elements of the chronic care model described above (Wagner *et al.* 1996; VonKorff *et al.* 1997). We hypothesized that, compared with usual care, the intervention program would lead to greater intensity and continuity of treatment and fewer symptoms of mania and depression averaged across the follow-up period. Here we report analyses (specified *a priori*) based on the first 12 months of the 24-month study.

METHOD

Study methods are described in detail in an earlier publication (Simon *et al.* 2002) and are summarized here.

Sample and recruitment

Participants were enrolled between August 1999 and October 2000 at four group-model behavioral health clinics of Group Health Cooperative (GHC), a managed care organization in Washington State. GHC provides general medical and specialty mental health care on a prepaid (capitated) basis, with specialty mental health staffing levels similar to those at other group-model managed care organizations (Dial *et al.* 1998). All mood stabilizer drugs (including newer anticonvulsants) and all atypical antipsychotic drugs were available without formulary restriction. Treated prevalence of bipolar disorder in the GHC population is approximately 0·4% (Unützer *et al.* 1998).

Using computerized visit and hospital discharge records, we identified health plan members aged 18 and over with any diagnosis of bipolar spectrum disorder (bipolar disorder type I or type II, schizoaffective disorder, or cyclothymia) during the prior 12 months. A subsequent clinical assessment (see below) limited the sample to those with confirmed bipolar disorder type I or II. Potentially eligible members were invited (by letter followed by phone call) to an in-person baseline assessment that included relevant modules of the Structured Clinical Interview for DSM-IV (SCID) (First *et al.* 1997) (current and lifetime depression, current and lifetime mania/hypomania, alcohol abuse/dependence, drug abuse/dependence, and psychosis screen). If SCID did not confirm a diagnosis of bipolar disorder type I or II, the principal investigator consulted with the

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

treating psychiatrist and reviewed medical records to identify past mood episodes overlooked by the structured interview. All those with diagnoses of bipolar disorder type I or II confirmed by either SCID or record review were invited to participate in the randomized trial. The only exclusion criterion was cognitive impairment severe enough to preclude informed consent (in the judgment of the treating physician). Medical or psychiatric co-morbidity was not reason for exclusion. Each potential participant was advised that he or she might be offered additional treatment services, but willingness to accept or participate in any specific treatment was not a condition for enrollment.

Treatment assignment

One to seven days following the baseline interview, consenting participants were randomly assigned to continued usual care or usual care plus a multi-component intervention program described below. Assignment was performed by the study database administrator using a concealed table of computer-generated random numbers (with no blocking or stratification). Participants were notified of treatment assignment in a subsequent phone call. An individual physician might have patients in both treatment conditions. Participants assigned to usual care received no additional services, but no services normally available (inside or outside of GHC) were withheld.

Intervention program

The program included several elements adapted from systematic care programs for unipolar depression (Katon *et al.* 1995; Hunkeler *et al.* 2000; Katzelnick *et al.* 2000; Simon *et al.* 2000; Wells *et al.* 2000) or bipolar disorder (Bauer *et al.* 1997; Ludman *et al.* 2000; Simon *et al.* 2000; Bauer, 2001*a*; Bauer & McBride, 2003): structured initial assessment and care planning, structured monthly telephone monitoring of mood symptoms and medication use, feedback to and coordination with the mental health treatment team, and a structured group psychoeducational program – all provided by a nurse care manager. All intervention services were intended as additions to ongoing mental health care and usual mental health providers (psychiatrists, psychotherapists) maintained primary responsibility for all treatment decisions. Intervention services were offered for 24 months following randomization.

During an initial in-person assessment the care manager reviewed each patient's treatment history and developed a collaborative treatment plan including current medications, expected frequency of follow-up visits, typical early warning signs of mood episodes, coping strategies for responding to warning signs, and identification of a care partner (family member or significant other).

Following the initial assessment, care managers called each patient at least once per month to complete self-reported measures of mood symptoms (the Internal State Scale of Bauer and colleagues (Bauer *et al.* 1991, 2000), a structured clinician rating of current symptoms (depression, mania, and psychosis), and a structured evaluation of current medication use and side-effects. Care managers provided general support, encouragement to attend group sessions (see below), and brief education regarding medications, side-effects, and warning signs of relapse. As appropriate, care managers also facilitated in-person follow-up care and provided supplemental phone contacts, crisis intervention, outreach and coordination with family members or significant others, and assistance with barriers to treatment. Following each phone contact, care managers sent structured feedback to treating mental health providers regarding mood symptoms, current medication use, side-effects, and algorithm-based suggestions regarding medication adjustments, laboratory testing, and follow-up visits. Care managers communicated with treating psychiatrists and psychotherapists via printed reports, voice messages, electronic messages, live telephone calls, or in-person conversations depending on clinical urgency. Telephone monitoring and care management were supported by a web-based computer application integrating contact tracking, structured assessment, detailed algorithms to generate pharmacotherapy recommendations, and standardized feedback reports to providers. For example, a scenario of manic or mixed symptoms in a patient taking lithium with current blood level of 0·3 and no significant side-effects would lead to a recommendation to 'Consider increasing dose of lithium'. The same clinical picture with a lithium level greater than

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

0·8 or intolerable side-effects would lead to a recommendation to 'Consider medication switch or augmentation for lithium'. The complete algorithm considered approximately 7000 clinical scenarios defined by symptom pattern, current medication dose, blood levels (where applicable), and side-effects. Clinic psychiatrists and psychotherapists received no special training other than general information regarding the intervention program and the role of the care manager.

The group psychoeducational program was adapted from Bauer and McBride's Life Goals Program (Bauer & McBride, 2003). This manualized program includes five weekly group sessions (Phase 1) followed by twice-monthly sessions for the duration of the intervention (Phase 2). Phase 1 included structured education regarding the nature of bipolar illness, triggers and early symptoms of mood episodes, and self-management strategies for triggers and early symptoms. Phase 2 used a structured problem-solving format to focus on accomplishment of specific life goals. Participants created and updated personalized self-management plans describing triggers, warning signs, and coping strategies.

Care managers were registered nurses with at least 5 years of clinical psychiatric experience. Training for the telephone monitoring program included 4 hours of didactic instruction and 4 hours of demonstration or role-play. Training for the group program included 10–12 hours of didactic instruction, 6 hours observing groups sessions, and 6 hours leading group sessions under supervision. Care managers received 60 min of supervision per week from a psychiatrist (G.S.) and/or psychologist (E.L.). Caseloads averaged approximately 95 patients per full-time equivalent. Care managers were trained (by E.L.) in techniques of motivational interviewing (Miller & Rollnick, 1991; Rollnick and Miller, 1995), a collaborative method of addressing barriers to positive behavior change.

Clinical assessments

Baseline SCID interviews were used to generate ratings of mania and depression severity using the 6-point Psychiatric Status Rating or PSR scale (Keller *et al.* 1987). Ratings of 1 or 2 represent remission or minimal symptoms, ratings of 3 or 4 represent clinically significant subthreshold symptoms, a rating of 5 represents a current episode of hypomania or moderate major depression, and a rating of 6 represents a current episode of mania or severe depression.

All participants were contacted for blinded in-person follow-up assessments every 3 months. Assessments included relevant modules of the SCID (current depression, current mania/hypomania, alcohol abuse/dependence, drug abuse/dependence) and a weekly timeline follow-back rating of depressive and manic/hypomanic symptoms using the Longitudinal Interval Follow-Up Evaluation (LIFE; Keller *et al.* 1987). This assessment yielded separate ratings of depression and mania severity for each week of follow-up using the 6-point PSR scale.

All interviewers were mental health clinicians with at least 1 year of clinical experience in the assessment of mood disorders. Interviewers received 8 hours of additional training from the principal investigator including role-play of at least three assessments. On completion of training, each interviewer observed at least five assessments by a previously certified interviewer and completed at least five interviews under direct observation prior to certification. Each interviewer also completed at least five additional observed interviews during the course of the project. In order to preserve blinding, participants were repeatedly advised not to discuss specific treatment services with blinded follow-up interviewers. A sub-study evaluated the accuracy of 3-month LIFE recall using blinded telephone re-appraisals at random points between LIFE assessments (Rutter & Simon, in press). Accuracy of recall was generally good, with some tendency of depressed participants to over-recall episodes of depression.

Protection of human subjects

All study procedures were approved by Institutional Review Boards at GHC and UCLA. Participants provided written informed consent at three stages: prior to the baseline assessment, prior to enrollment in the randomized trial, and (for those in the intervention group) prior to the initial intervention visit.

Pharmacy and utilization data

Computerized pharmacy and visit registration data were used to examine mental health visits and psychotropic drug prescription fills/refills

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

Table 1. *Baseline characteristics of participants assigned to intervention or usual care groups*

	Intervention ($n=212$)	Usual care ($n=229$)	Test statistic
Mean age (S.D.)	44·1 (13·4)	44·3 (12·9)	$t=0·10$, $p=0·92$
Female (%)	144 (68)	157 (69)	$\chi^2=0·02$, $p=0·89$
Caucasian (%)	184 (87)	206 (90)	$\chi^2=1·07$, $p=0·30$
Currently employed (%)	138 (65)	142 (62)	$\chi^2=0·45$, $p=0·50$
Living with spouse or partner (%)	102 (48)	125 (55)	$\chi^2=1·85$, $p=0·17$
Type 1 bipolar disorder (%)	165 (78)	171 (75)	$\chi^2=0·60$, $p=0·44$
Mean baseline depression PSR (S.D.)	3·44 (1·48)	3·54 (1·45)	$t=0·73$, $p=0·46$
No. with depression PSR $\geqslant 3$ (%)	158 (74)	167 (73)	$\chi^2=0·15$, $p=0·70$
No. with depression PSR $\geqslant 5$ (%)	71 (34)	83 (36)	$\chi^2=0·37$, $p=0·54$
Mean baseline mania PSR (S.D.)	2·77 (1·72)	2·78 (1·60)	$t=0·02$, $p=0·98$
No. with mania PSR $\geqslant 3$ (%)	101 (48)	116 (51)	$\chi^2=0·40$, $p=0·53$
No. with mania PSR $\geqslant 5$ (%)	33 (16)	35 (15)	$\chi^2=0·01$, $p=0·94$
No. with current psychotic symptoms* (%)	33 (16)	30 (13)	$\chi^2<0·55$, $p=0·46$
No. with current alcohol abuse/dependence (%)	8 (4)	7 (3)	$\chi^2=0·17$, $p=0·68$
No. with current drug abuse/dependence (%)	9 (4)	7 (3)	$\chi^2=0·45$, $p=0·50$
No. with psychiatric hospital stay in previous year (%)	17 (8)	28 (12)	$\chi^2=2·13$, $p=0·14$
No. disabled for >30 of previous 90 days‡ (%)	63 (30)	75 (33)	$\chi^2=0·47$, $p=0·49$

* Rating of 'definite' for any item of the SCID psychotic screen module.

‡ Self-reported inability to fulfill usual work or household responsibilities because of illness for more than 30 of previous 90 days (Chyba & Washington, 1993).

PSR, psychiatric status rating.

throughout the follow-up period. At each blinded follow-up assessment, participants were asked about out-patient visits or psychotropic drug prescriptions from non-GHC providers or pharmacies. Refill and interview data were used to calculate the total days' supply of individual drugs and drug classes used during follow-up.

Data analyses

Primary analyses compared intervention and usual care participants' PSR scores for the first 52 weeks of follow-up using a repeated-measures linear model with relevant baseline characteristics (age, sex, depression severity, mania severity, psychotic symptoms, substance abuse, and recent psychiatric hospitalization) included as covariates. All analyses were based on original treatment assignment (intent to treat) regardless of treatment actually received. Test statistics were calculated using generalized estimating equations (GEEs) as implemented by the SAS software package version 8 (SAS Inc., Cary, NC, USA). GEEs are a semi-parametric estimation approach, based on specification of the mean and covariance, but not the full distribution, of modeled outcomes (Diggle *et al.* 1995). All models were estimated using an independence working correlation structure, with variance based on the empirical covariance matrix. Our primary hypothesis was that intervention participants would show lower mean mania and depression scores across the full follow-up period (i.e. main effect of treatment assignment) with secondary tests of differences in course of symptoms over time (i.e. group-by-time interaction). *A priori* power calculations indicated that a sample size of 250 per group would allow 80% power to detect a 15% difference in risk of moderate or severe mood episodes.

RESULTS

Study sample

Of 785 patients initially contacted, 509 (65%) attended a baseline assessment. At this assessment, 417 met criteria for bipolar disorder type 1 or 2 by SCID and an additional 33 had diagnosis confirmed by record review and consultation with the treating physician, leaving 59 who were found ineligible because bipolar diagnosis was not confirmed. Of the 450 found eligible, 441 (98%) enrolled in the randomized trial. Compared with those attending the baseline assessment, non-participants did not differ significantly in age, sex, or likelihood of psychiatric hospitalization or psychiatric emergency room use in the last year. Participants assigned to intervention and usual care groups did not differ significantly in any demographic or clinical characteristics assessed at baseline (Table 1).

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

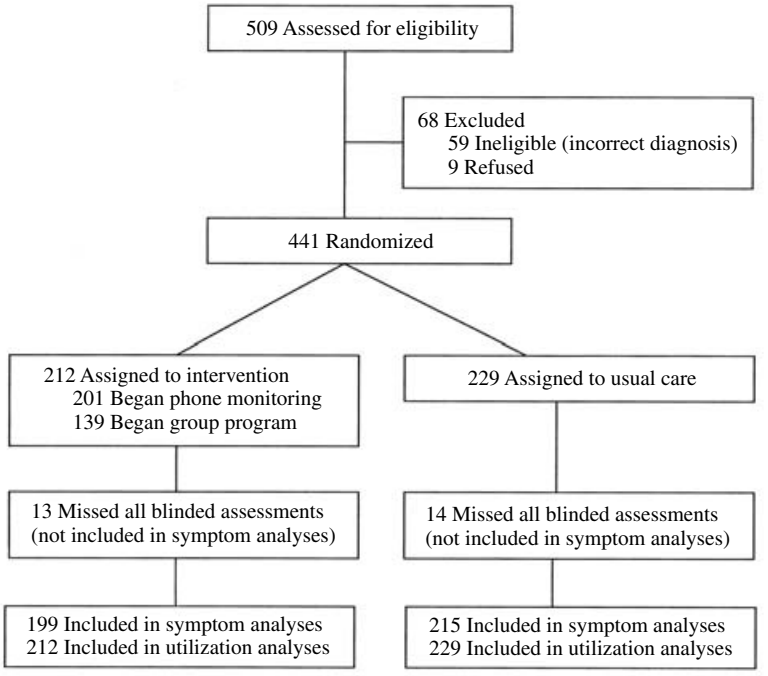


FIG. 1. Randomized trial flow diagram.

Of those randomized, 94% completed at least one blinded outcome assessment and 82% completed all four. Those missing at least one assessment did have higher mania PSR scores at baseline (2·97 *v.* 2·55, $t=2·20$, $p=0·03$) and a higher prevalence of current substance use disorder at baseline (10·1% *v.* 2·9%, $\chi^2=10·2$, $p=0·006$). Fig. 1 shows progress of participants through the randomized trial.

Participation in the intervention program

Of participants assigned to the intervention group, 95% completed the initial treatment planning visit and entered the telephone monitoring program. The mean number of telephone monitoring contacts over 12 months was 9·4 (S.D. = 3·1). Approximately 65% attended at least one group session, 59% completed the five-session Phase 1 program, and 52% continued Phase 2 group participation for at least 12 months. Costs for direct intervention services and clinical supervision were calculated using actual personnel costs, fringe benefit costs, and overhead costs. Average cost for the first 12 months of the program was $521 per patient.

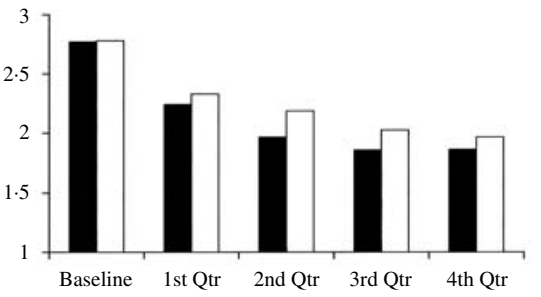


FIG. 2. Mean psychiatric status ratings for manic symptoms during first 12 months of follow-up for patients assigned to intervention (■, $n=212$) or usual care (□, $n=229$) groups. Baseline value reflects single baseline assessment, while each follow-up bar reflects mean of all assessments during that 3-month follow-up period. Scale ranges from 1 (no symptoms of mania) to 6 (severe manic episode).

Effects on clinical outcomes

As shown in Fig. 2, intervention group patients reported fewer symptoms of mania through 12 months of follow-up. Primary analyses compared mania PSR scores averaged over 12 months using a repeated-measures model with adjustment for age, sex, baseline severity of

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

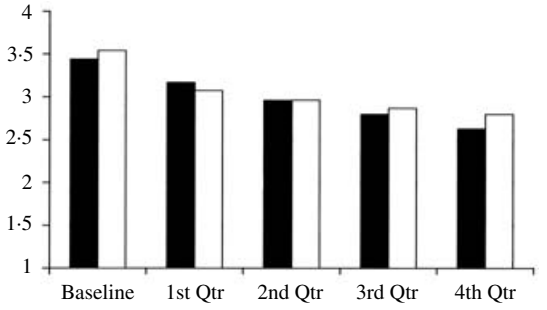


FIG. 3. Mean psychiatric status ratings for depressive symptoms during first 12 months of follow-up for patients assigned to intervention (■, $n=212$) or usual care (□, $n=229$) groups. Baseline value reflects single baseline assessment, while each follow-up bar reflects mean of all assessments during that 3-month follow-up period. Scale ranges from 1 (no symptoms of depression) to 6 (severe depressive episode).

depressive and manic symptoms, baseline psychotic symptoms, baseline substance use disorder, and psychiatric hospitalization in the year prior to randomization. This model found significantly lower mania scores in the intervention group over 12 months ($Z=2{\cdot}24$, $p=0{\cdot}025$) with no significant group by time interaction ($Z=0{\cdot}32$, $p=0{\cdot}70$). Secondary analyses examined categorical definitions of mania severity. The total time with any significant symptoms of mania (i.e. mania PSR of 3 or more) was 12·3 weeks (S.D. = 13·3) in the intervention group *versus* 14·3 weeks (S.D. = 14·8) in the usual care control group, but this difference was not statistically significant at the 5% level (adjusted difference = −2·1 weeks, 95% CI −4·3 to +0·01). Time spent in full hypomanic or manic episode (i.e. mania PSR of 5 or more) was significantly less in the intervention group (mean 1·7 weeks, S.D. = 3·8 in the intervention group *v.* 2·6 weeks, S.D. = 5·3 in the usual care control group; adjusted difference = −0·9 weeks, 95% CI −1·7 to −0·1). Additional secondary analyses examined the proportion of patients with onset of a new hypomanic or manic episode (after excluding the 68 participants meeting criteria for mania or hypomania at baseline). The proportion with a new manic or hypomanic episode was 23% (39 of 169) in the intervention group compared with 32% (58 of 182) in the usual care group ($\chi^2=3{\cdot}39$, df = 1, $p=0{\cdot}07$).

As shown in Fig. 3, intervention group patients reported somewhat more depressive symptoms during the first quarter of follow-up and somewhat less during the third and fourth quarters. Primary analyses compared mean PSR depression scores throughout the follow-up period using a repeated-measures model with adjustment for the baseline characteristics listed above. In this model, there was no significant difference between the intervention and usual care groups averaged across the full follow-up period ($Z=0{\cdot}23$, $p=0{\cdot}82$), but there was a significant group-by-time interaction ($Z=2{\cdot}02$, $p=0{\cdot}043$), indicating that depressive symptoms in the intervention group showed a larger decline over time than those in the usual care group. Secondary analyses examined the proportion of patients with onset of a new depressive episode during follow-up (after excluding the 154 participants meeting criteria for depressive episode at baseline). Over the full follow-up period, the proportion experiencing a new depressive episode was 54% (73 of 134) in the intervention group compared with 54% (74 of 136) in the usual care group ($\chi^2=0{\cdot}0$, df = 1, $p=0{\cdot}99$). During the second half of the follow-up period, the proportions experiencing new depressive episodes were 33% (42 of 128) in the intervention group and 39% (52 of 133) in usual care ($\chi^2=0{\cdot}29$, df = 1, $p=0{\cdot}30$).

Results for both mania and depression ratings were similar when the sample was stratified according to baseline severity of mania (PSR $\leqslant 2$ *v.* PSR $\geqslant 3$), baseline severity of depression, or baseline diagnosis (bipolar disorder type I *versus* type II).

Effects on treatment received during follow-up

The intervention and usual care groups did not differ significantly in use of individual psychotherapy, but intervention patients made significantly more psychiatric visits for medication management (Table 2). The two groups did not differ significantly in use of in-patient or emergency room care, but statistical power was not adequate to detect moderate differences in these relatively rare events. The two groups did not differ significantly in use of mood stabilizer or antidepressant drugs when measured either as proportion treated for more than 270 days over the 12-month period (shown in Table 2) or as mean number of days of use over 12 months (data not shown). Those assigned to the intervention group were significantly more likely to

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

Table 2. *Treatment received during 12 months following randomization* (*not including visits to study nurse care managers*)

	Intervention ($n=212$)	Usual care ($n=229$)	Test statistic*
Mean (s.d.) of medication management visits	7·19 (12·80)	5·11 (4·59)	$Z=2·44$, $p=0·01$
Mean (s.d.) of individual psychotherapy visits	6·11 (11·36)	5·54 (11·44)	$Z=0·36$, $p=0·72$
Mean (s.d.) of psychiatric emergency room visits	0·05 (0·29)	0·09 (0·39)	$Z=1·19$, $p=0·23$
Mean (s.d.) of psychiatric inpatient days	0·50 (2·35)	0·60 (3·21)	$Z=0·02$, $p=0·98$
No. with any psychiatric hospitalization (%)	12 (5·7)	17 (7·4)	$\chi^2=0·56$, $p=0·29$
No. with >270 days of mood stabilizer (%)†	108 (51)	118 (52)	$\chi^2=0·02$, $p=0·90$
No. with >270 days of antidepressant (%)	67 (32%)	78 (34)	$\chi^2=0·30$, $p=0·58$
No. with >90 days of typical antipsychotic (%)	9 (4)	7 (3)	$\chi^2=0·44$, $p=0·50$
No. with >90 days of atypical antipsychotic (%)‡	48 (23)	35 (15)	$\chi^2=3·90$, $p=0·05$

* Based on Mann–Whitney test for visit counts.
† Includes lithium, divalproex, carbamazepine, topiramate, lamotrigine, and gabapentin.
‡ Includes olanzapine, risperidone, quetiapine, ziprasidone, and clozapine.

use one or more atypical antipsychotics for at least 90 days.

DISCUSSION

A systematic multi-component care program was generally well-accepted in this population-based sample of people with bipolar disorder. Compared with usual care, the intervention program led to significant reductions in severity of mania symptoms and time spent in full or partial manic episodes over the first 12 months of the program. While average effects on mania symptom scores were small, manic or hypomanic episodes were reduced by one-third. Effects on depression were less clear, with the intervention group reporting more severe depressive symptoms during the first quarter of follow-up and less severe depressive symptoms during the third and fourth quarters.

Design of this study incorporated several elements intended to promote generalizability to everyday practice. We approached all patients treated for bipolar disorder in a large managed care organization (including approximately 10% covered by Medicare, Medicaid, or other low-income programs). All services were provided through existing health plan mental health clinics, and participants continued ongoing treatment relationships. While we offered recommendations and support to improve the quality of pharmacotherapy, treating psychiatrists remained responsible for all prescribing decisions.

By design, this study sample differed significantly from those usually enrolled in clinical trials of acute- or maintenance-phase treatments for bipolar disorder. Our intent was to reflect the typical cross-section of outpatients treated for bipolar disorder. By design, participants showed a full range of mood disorder symptoms at baseline in terms of type (pure depression, pure mania, mixed) and severity. Participants were not selected based on any treatment event (e.g. hospitalization). While all had received a diagnosis of bipolar disorder in the prior 12 months, some were receiving no active treatment at time of recruitment. Given this approach, our analyses focus on the overall severity of mood symptoms over time rather than dichotomous measures (remission or relapse) appropriate for more homogeneous samples. Compared with other heterogeneous out-patient cohorts such as that of the Stanley Foundation Bipolar Network (Suppes *et al.* 2001; Post *et al.* 2003), this sample showed a similar distribution of type 1 and type 2 illness, generally similar severity of illness at baseline and during follow-up, and similar levels of impairment or disability. Rates of hospitalization were low, but this may reflect close management of inpatient care in this insurance system.

We must, however, acknowledge some important limitations. First, this intervention would probably require adaptation for use in other settings such as community private practice or public-sector mental health clinics. An ongoing study by Bauer and colleagues will evaluate a similar systematic care program among more severely ill patients treated in the Veterans Affairs health system (Bauer, 2001*a*; Bauer *et al.* 2001). Second, we cannot determine how the

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

added attention of the research program affected the behavior of patients or providers in the usual care control group. However, any positive effect on patients' self-management or on frequency of follow-up care would tend to reduce differences between the intervention and control groups. Third, our primary outcome measure, week-by-week mood recall over 3-month intervals, might have been more accurate had we used more frequent assessments or required participants to maintain mood diaries. These methods, however, would exclude those unable or unwilling to comply and might introduce an intervention effect through more frequent self-monitoring and supportive contact. Fourth, we cannot determine which elements of the intervention (patient education, peer support, telephone monitoring, recommendations to provider, etc.) might account for any intervention effects. By design, our intervention integrated structured group psychotherapy with monitoring to improve quality of pharmacotherapy. Among intervention patients, those who completed five or more group sessions did have better clinical outcomes than those who did not, but group participants also had less severe symptoms at baseline. Consequently, we cannot say whether better outcomes in group participants reflect benefits of the group program or the fact that those with the worst prognosis may have been less likely to participate. Furthermore, we cannot separate specific components of the program (medication recommendations to providers, knowledge gained in group sessions) from non-specific elements such as greater attention and support. Fifth, spillover or contamination effects (i.e. exposure to the intervention program causing clinicians to alter treatment of usual care patients) might blunt any differences between the intervention and usual care groups. Similar studies of organized care programs for depression, however, have found no significant spillover or training effects (Lin *et al.* 1997, 2001). Finally, this program might prove difficult to implement in solo or small-group practices, and important elements of the program would not be reimbursable under current fee-for-service arrangements.

Participants in both intervention and control groups showed significant improvement in mood symptoms over time. This was somewhat unexpected, because participants were approached at a relatively random time during treatment rather than at time of treatment-seeking or acute episode. Consequently, we would not expect the usual care group to show regression to the mean. The overall tendency to improvement over time may reflect a generally favorable prognosis in this sample or the beneficial effect of increased attention from the research program (Rucci *et al.* 2002). Alternatively, those who were experiencing increased symptoms at time of recruitment may have been more likely to participate.

Our findings regarding reduction in severity of mania symptoms are consistent with those of other psychoeducational or psychotherapeutic interventions for bipolar disorder. For example, Perry and colleagues (Perry *et al.* 1999) found that a brief psychoeducational intervention for patients in remission from bipolar disorder significantly reduced risk of manic relapse over the next 18 months. Psychoeducational interventions emphasizing self-monitoring and early intervention may have greater effects for preventing manic relapse than for treating or preventing depression. In contrast, Lam and colleagues (Lam *et al.* 2003) found that a 12- to 20-session cognitive therapy program significantly reduced risk of relapse in bipolar disorder, with significant effects on both manic and depressive relapses. Colom and colleagues (Colom *et al.* 2003) reported that a 21-session psychoeducational group program reduced the risk of both manic and depressive relapse.

We observed no overall effect on depression across the first 12 months of follow-up, but we do find that depression scores in the intervention group showed a larger decline over time. Intervention patients reported higher depression scores than usual care patients in the first quarter and lower scores in the final two quarters. This pattern could represent a true effect of the intervention program if more aggressive pharmacotherapy of mania or closer self-monitoring led to increased reports of depressive symptoms during the first months of the intervention program. Alternatively, higher early depression scores in the intervention group could reflect a failure of randomization or simple random fluctuation. We should emphasize that these possible explanations are all *post hoc* speculations.

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

Our 12-month data are not consistent with our original hypothesis of a reduction in mean severity of depressive symptoms averaged across the follow-up period. While these data suggest an increasing intervention effect on depressive symptoms over time, data from the second year will be necessary to draw any final conclusions.

The effects of the intervention on use of mental health services include expected and unexpected findings. Consistent with the intent of the care management program, medication monitoring visits were more frequent in the intervention group. While we saw no effect on use of traditional mood stabilizer drugs, intervention patients more often used atypical antipsychotics, several of which have demonstrated efficacy for treatment of mania (Delbello *et al.* 2002; Sachs *et al.* 2002; Tohen *et al.* 2002). Because we measured medication consumed rather than medication prescribed, we cannot determine whether greater use of atypical antipsychotics without greater use of mood stabilizers reflects physicians' or patients' treatment choices.

Future research should address the targeting of this or other care improvement programs to those with greatest clinical need. While a single baseline rating did not identify those with greater likelihood of benefiting from the intervention, more detailed history of mood episodes might be more predictive. If final results indicate that benefits are limited to reducing manic episodes, then services might be focused on those with histories of frequent or severe manic relapse. Furthermore, this type of intervention might only prove cost-beneficial or cost-saving in a population at high risk for hospitalization.

We should emphasize that improving long-term outcomes of bipolar disorder is not a simple task. A recent multi-center efficacy trial of long-term pharmacotherapy was not able to demonstrate that lithium or divalproex was superior to placebo in preventing mood episodes (Bowden *et al.* 2000). Bipolar disorder is a complex and heterogeneous condition requiring individualized treatment adjusted over time. We regard our findings as important and encouraging news regarding the benefits of organized care programs for people with severe mood disorders.

ACKNOWLEDGEMENTS

Funded by grant R01 MH59125 from the National Institute of Mental Health. The funding agency provided consultation regarding study design, but was not involved in data collection, data analysis and interpretation, or preparation of this manuscript.

DECLARATION OF INTEREST

None.

REFERENCES

APA (2002). Practice guideline for the treatment of patients with bipolar disorder (revision). *American Journal of Psychiatry* **159**, 1–50.

Bauer, M. (2001*a*). The collaborative practice model in bipolar disorder: design and implementation in a multi-site randomized controlled trial. *Bipolar Disorders* **3**, 233–244.

Bauer, M. (2001*b*). An evidence-based review of psychosocial interventions for bipolar disorder. *Psychopharmacology Bulletin* **35**, 109–134.

Bauer, M., Crits-Christoph, P., Ball, W., Dewees, E., McAllister, T., Alahi, P., Cacciiola, J. & Whybrow, P. (1991). Independent assessment of manic and depressive symptoms by self-rating. Scale characteristics and implications for the study of mania. *Archives of General Psychiatry* **48**, 807–812.

Bauer, M. & McBride, L. (2003). *Structured Group Psychotherapy for BipolarDisorder: the Life Goals Program* (2nd edn). Springer: New York.

Bauer, M., McBride, L., Shea, N., Gavin, C., Holden, F. & Kendall, S. (1997). Impact of an easy-access VA clinic-based program for patients with bipolar disorder. *Psychiatric Services* **48**, 491–496.

Bauer, M., Vojta, C., Kinosian, B., Altshuler, L. & Glick, H. (2000). The Internal State Scale: replication of its discriminating abilities in a multi-site, public sector sample. *Bipolar Disorders* **2**, 340–346.

Bauer, M., Williford, W., Dawson, E., Akiskal, H., Altshuler, L., Fye, C., Gelenberg, A., Glick, H., Kinosian, B. & Sajatovic, M. (2001). Principles of effectiveness trials and their implementation in VA Cooperative Study #430, 'Reducing the efficacy-effectiveness gap in bipolar disorder'. *Journal of Affective Disorders* **67**, 61–78.

Bowden, C., Calabrese, J., McElroy, S., Gyulai, L., Wassef, A., Petty, F., Pope, H. J., Chou, J., Keck, P. J., Rhodes, L., Swann, A., Hirschfeld, R. & Wozniak, P. (2000). A randomized, placebo-controlled 12-month trial of divalproex and lithium in treatment of outpatients with bipolar I disorder. Divalproex Maintenance Study Group. *Archives of General Psychiatry* **57**, 481–489.

Chyba, M. & Washington, L. (1993). Questionnaires from the National Health Interview Survey, 1985–89. National Center for Health Statistics: Washington, DC.

Colom, F., Vieta, E., Martinez-Aran, A., Reinares, M., Goikolea, J., Benabarre, A., Torrent, C., Comes, M., Corbella, B., Parramon, G. & Coromina, J. (2003). A randomized trial on the efficacy of group psychoeducation in the prophylaxis of recurrences in bipolar patients whose disease is in remission. *Archives of General Psychiatry* **60**, 402–407.

Coryell, W., Winokur, G., Solomon, D., Shea, T., Leon, A. & Keller, M. (1997). Lithium and recurrence in a long-term follow-up of bipolar affective disorder. *Psychological Medicine* **27**, 281–289.

Delbello, M., Schweiers, M., Rosenberg, H. & Strakowski, S. (2002). A double-blind, randomized, placebo-controlled study of quetiapine as adjunctive treatment for adolescent mania. *Journal*

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

of American Academy of Child Adolescent Psychiatry **41**, 1216–1223.

Dial, T., Bergsten, C., Haviland, M. & Pincus, H. (1998). Psychiatrist and nonphysician mental health provider staffing levels in health maintenance organizations. *American Journal of Psychiatry* **155**, 405–408.

Diggle, P., Liang, K. & Zeger, S. (1995). *Analysis of Longitudinal Data*. Oxford University Press: New York.

First, M., Spitzer, R., Gibbon, M. & Williams, J. (1997). *Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I), Clinician Version*. American Psychiatric Press: Washington.

Gitlin, M., Swendsen, J., Heller, T. & Hammen, C. (1995). Relapse and impairment in bipolar disorder. *American Journal of Psychiatry* **152**, 1635–1640.

Goldberg, J., Harrow, M. & Grossman, L. (1995). Course and outcome in bipolar affective disorder: a longitudinal follow-up study. *American Journal of Psychiatry* **152**, 379–384.

Hunkeler, E., Meresman, J., Hargreaves, W., Fireman, B., Berman, W., Kirsch, A., Groebe, J., Braden, P., Getzell, M., Feigenbaum, P., Hurt, W., Peng, T. & Salzer, M. (2000). Efficacy of nurse telehealth care and peer support in augmenting treatment of depression in primary care. *Archives of Family Medicine* **9**, 700–708.

Huxley, N., Parikh, S. & Baldessarini, R. (2000). Effectiveness of psychosocial treatments in bipolar disorder: state of the evidence. *Harvard Review of Psychiatry* **8**, 126–140.

Institute of Medicine (2001). *Crossing the Quality Chasm*. National Academies Press: Washington, DC.

Johnson, R. & McFarland, B. (1992). Treated prevalence of severe mental illness among HMO members. *Hospital Community Psychiatry* **45**, 919–924.

Johnson, R. & McFarland, B. (1996). Lithium use and discontinuation in a health maintenance organization. *American Journal of Psychiatry* **153**, 993–1000.

Katon, W., VonKorff, M., Lin, E., Walker, E., Simon, G., Bush, T., Robinson, P. & Russo, J. (1995). Collaborative management to achieve treatment guidelines: impact on depression in primary care. *JAMA* **273**, 1026–1031.

Katzelnick, D., Simon, G., Pearson, S., Manning, W., Helstad, C., Henk, H., Cole, S., Lin, E., Taylor, L. & Kobak, K. (2000). Randomized trial of a depression management program in high utilizers of medical care. *Archives of Family Medicine* **9**, 345–351.

Keck Jr., P., McElroy, S., Strakowski, S., West, S., Sax, K., Hawkins, J., Bourne, M. & Haggard, P. (1998). 12-month outcome of patients with bipolar disorder following hospitalization for a manic or mixed episode. *American Journal of Psychiatry* **155**, 646–652.

Keller, M., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P. & Andreasen, N. C. (1987). The Longitudinal Interval Follow-up Evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* **44**, 540–548.

Lam, D., Watkins, E., Hayward, P., Bright, J., Kerr, N., Parr-Davis, G. & Sham, P. (2003). A randomized controlled trial of cognitive therapy for relapse prevention for bipolar affective disorder. *Archives of General Psychiatry* **60**, 145–152.

Lin, E., Katon, W., Simon, G., VonKorff, M., Bush, T., Rutter, C., Saunders, K. & Walker, E. (1997). Achieving guidelines for treatment of depression in primary care: is physician education enough? *Medical Care* **35**, 831–842.

Lin, E., Simon, G., Katzelnick, D. & Pearson, S. (2001). Does physician education on depression management improve treatment in primary care? *Journal of General Internal Medicine* **16**, 614–619.

Ludman, E., VonKorff, M., Katon, W., Lin, E., Simon, G., Walker, E., Unützer, J., Bush, T. & Wahab, S. (2000). The design, implementation, and acceptance of a primary care based intervention to prevent depression relapse. *International Journal of Psychiatric Medicine* **30**, 229–245.

McElroy, S. & Keck, P. J. (2000). Pharmacologic agents for the treatment of acute bipolar mania. *Biological Psychiatry* **48**, 539–557.

Miller, W. & Rollnick, S. (1991). *Motivational Interviewing: Preparing People to Change Addictive behavior*. Guilford Press: New York.

O'Connell, R., Mayo, J., Flatow, L., Cuthbertson, B. & O'Brien, B. (1991). Outcome of bipolar disorder on long-term treatment with lithium. *British Journal of Psychiatry* **159**, 123–129.

Otto, M., Reilly-Harrington, N. & Sachs, G. (2003). Psychoeducational and cognitive-behavioral strategies in the management of bipolar disorder. *Journal of Affective Disorders* **73**, 171–181.

Perry, A., Tarrier, N., Morriss, R., McCarthy, E. & Limb, K. (1999). Randomised controlled trial of efficacy of teaching patients with bipolar disorder to identify early symptoms of relapse & obtain treatment. *British Medical Journal* **318**, 149–153.

Post, R., Denicoff, K., Leverich, G., Altshuler, L., Frye, M., Suppes, T., Rush, A., Keck, P. J., McElroy, S., Luckenbaugh, D., Pollio, C., Kupka, R. & Nolen, W. (2003). Morbidity in 258 bipolar outpatients followed for 1 year with daily prospective ratings on the NIMH life chart method. *Journal of Clinical Psychiatry* **64**, 680–690.

Rollnick, S. & Miller, W. (1995). What is motivational interviewing? *Behavioural and Cognitive Psychotherapy* **23**, 325–334.

Rucci, P., Frank, E., Kostelnick, B., Fagliolini, A., Mallinger, A., Swartz, H., Thase, M., Siegel, L., Wilson, D. & Kupfer, D. (2002). Suicide attempts in patients with Bipolar I disorder during acute and maintenance phases of intensive treatment with pharmacotherapy and adjunctive psychotherapy. *American Journal of Psychiatry* **159**, 1160–1164.

Rutter, C. & Simon, G. (in press). A Bayesian method for estimating the accuracy of recalled depression. *Applied Statistics*.

Sachs, G., Grossman, F., Ghaemi, S., Okamoto, A. & Bowden, C. (2002). Combination of a mood stabilizer with risperidone or haloperidol for treatment of acute mania: a double-blind, placebo-controlled comparison of efficacy and safety. *American Journal of Psychiatry* **159**, 1146–1154.

Scott, J., Garland, A. & Moorhead, S. (2001). A pilot study of cognitive therapy in bipolar disorders. *Psychological Medicine* **31**, 459–467.

Silverstone, T., McPherson, H., Hunt, N. & Roman, S. (1998). How effective is lithium in the prevention of relapse in bipolar disorder? A prospective naturalistic follow-up study. *Australian and New Zealand Journal of Psychiatry* **32**, 61–66.

Simon, G., Ludman, E., Ununtzer, J. & Bauer, M. (2002). Design and implementation of a randomized trial evaluating systematic care for bipolar disorder. *Bipolar Disorders* **4**, 226–236.

Simon, G., VonKorff, M., Rutter, C. & Wagner, E. (2000). A randomized trial of monitoring, feedback, and management of care by telephone to improve depression treatment in primary care. *British Medical Journal* **320**, 550–554.

Suppes, T., Kennehy, E., Swann, A., Bowden, C., Calabrese, J., Hirschfeld, R., Keck, P. J., Sachs, G., Crismon, M., Toprac, M., Shon, S. & the Texas Consensus Conference Panel on Medication Treatment of Bipolar Disorder (2002). Report of the Texas Consensus Conference Panel on medication treatment of bipolar disorder 2000. *Journal of Clinical Psychiatry* **63**, 288–299.

Suppes, T., Leverich, G., Keck, P., Nolen, W., Denicoff, K., Altshuler, L., McElroy, S., Rush, A., Kupka, R., Frye, M., Bickel, M. & Post, R. (2001). The Stanley Foundation Bipolar Treatment Outcome Network II. Demographics and illness characteristics of the first 261 patients. *Journal of Affective Disorders* **67**, 45–59.

Svarstad, B., Shireman, T. & Sweeney, J. (2001). Using drug claims data to assess the relationship of medication adherence with hospitalization and costs. *Psychiatric Service* **52**, 805–811.

Tohen, M., Baker, R., Altshuler, L., Zarate, C., Suppes, T., Ketter, T., Milton, D., Risser, R., Gilmore, J., Breier, A. & Tollefson, G. (2002). Olanzapine versus divalproex in the treatment of acute mania. *American Journal of Psychiatry* **159**, 1011–1017.

Unützer, J., Simon, G., Pabiniak, C., Bond, K. & Katon, W. (1998). The treated prevalence of bipolar disorder in a large staff-model HMO. *Psychiatric Services* **49**, 1072–1078.

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press

Unützer, J., Simon, G., Pabiniak, C., Bond, K. & Katon, W. (2000). The use of administrative data to assess quality of care for bipolar disorder in a large staff model HMO. *General Hospital Psychiatry* **22**, 1–10.

VonKorff, M., Gruman, J., Schaefer, J., Curry, S. & Wagner, E. (1997). Collaborative management of chronic illness. *Annals of Internal Medicine* **127**, 1097–1102.

Wagner, E., Austin, B. & VonKorff, M. (1996). Organizing care for patients with chronic illness. *Milbank Quarterly* **74**, 511–544.

Wells, K., Sherbourne, C., Schoenbaum, M., Duan, N., Meredith, L., Unützer, J., Miranda, J., Carney, M. & Rubenstein, L. (2000). Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *Journal of the American Medical Association* **283**, 212–230.

https://doi.org/10.1017/S0033291704002624 Published online by Cambridge University Press