ORIGINAL RESEARCH

Telephone-Delivered Stepped Collaborative Care for Treating Anxiety in Primary Care: A Randomized Controlled Trial

Bruce L. Rollman, M.D., M.P.H.¹, Bea Herbeck Belnap, Dr. Biol. Hum.¹, Sati Mazumdar, Ph.D.², Kaleab Z. Abebe, Ph.D.¹, Jordan F. Karp, M.D.^{3,4}, Eric J. Lenze, M.D.⁵, and Herbert C. Schulberg, Ph.D.⁶

¹Division of General Internal Medicine, Center for Behavioral Health and Smart Technology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; ²Department of Biostatistics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA, USA; ³Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; ⁴Geriatric Research, Education & Clinical Center, VA Pittsburgh Healthcare System, Pittsburgh, PA, USA; ⁵Healthy Mind Lab, Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA; ⁶Department of Psychiatry, Weill Cornell Medical College, White Plains, NY, USA.

BACKGROUND: Collaborative care for depression is more effective in improving treatment outcomes than primary care physicians' (PCPs) usual care (UC). However, few trials of collaborative care have targeted anxiety.

OBJECTIVE: To examine the impact and 12-month durability of a centralized, telephone-delivered, stepped collaborative care intervention (CC) for treating anxiety disorders across a network of primary care practices.

DESIGN: Randomized controlled trial with blinded outcome assessments.

PARTICIPANTS: A total of 329 patients aged 18–64 referred by their PCPs in response to an electronic medical record (EMR) prompt. They include 250 highly anxious patients randomized to either CC or to UC, and 79 moderately anxious patients who were triaged to a watchful waiting (WW) cohort and later randomized if their conditions clinically deteriorated.

INTERVENTION: Twelve months of telephone-delivered CC involving non-mental health professionals who provided patients with basic psycho-education, assessed preferences for guideline-based pharmacotherapy, monitored treatment responses, and informed PCPs of their patients' care preferences and progress via the EMR.

MAIN MEASURES: Mental health-related quality of life (HRQoL; SF-36 MCS); secondary outcomes: anxiety (Hamilton Anxiety Rating Scale [SIGH-A], Panic Disorder Severity Scale) and mood (PHQ-9).

KEY RESULTS: At 12-month follow-up, highly anxious patients randomized to CC reported improved mental HRQoL (effect size [ES]: 0.38 [95 % CI: 0.13–0.63]; $P=0.003$), anxiety (SIGH-A ES: 0.30 [0.05–0.55]; $P=0.02$), and mood (ES: 0.45 [0.19–0.71] $P=0.001$) versus UC. These improvements were sustained for 12 months among African-Americans (ES: 0.70–1.14) and men (ES: 0.43–0.93). Of the 79 WW patients, 29 % met severity criteria for randomization, and regardless of treatment assignment, WW patients reported fewer anxiety and

mood symptoms and better mental HRQoL over the full 24-month follow-up period than highly anxious patients who were randomized at baseline.

CONCLUSIONS: Telephone-delivered, centralized, stepped CC improves mental HRQoL, anxiety and mood symptoms. These improvements were durable and particularly evident among those most anxious at baseline, and among African-Americans and men.

KEY WORDS: primary care; clinical trial; collaborative care; anxiety; depression; mental health.

J Gen Intern Med 32(3):245–55

DOI: 10.1007/s11606-016-3873-1

© Society of General Internal Medicine 2016

Collaborative care strategies are effective for treating a variety of mental health conditions within primary care settings,^{1,2} and interest in these programs has grown with the support for patient-centered medical homes under the Affordable Care Act (ACA).³ Collaborative care typically involves non-physician care managers who use evidence-based treatment protocols to proactively encourage provision of guideline-based care, promote patients' efforts to take a more active role in their care, and monitor their clinical response under the supervision of a primary care physician (PCP), with specialty back-up as necessary.^{1,4} Still, overcoming logistic challenges is critical to the efficient implementation of this model of integrated behavioral health care at scale in routine practice.^{5,6}

While dozens of trials have demonstrated the effectiveness of collaborative care for treating depression in primary care, comparatively few have addressed anxiety, despite their similar prevalence and adverse impact on health-related quality of life (HRQoL) and excess health services utilization.^{1,7–10} We previously completed a trial of collaborative care for patients with panic (PD) and/or generalized anxiety disorder (GAD) that utilized off-site non-behavioral health care managers who

Trial Registration: ClinicalTrials.gov Identifier: NCT00158327 (<http://www.clinicaltrials.gov/ct2/show/NCT00158327?term=rollman+anxiety&rank=1>)

Received May 12, 2016

Revised August 22, 2016

Accepted September 13, 2016

Published online October 6, 2016

communicated with patients via telephone and with PCPs through their practices' electronic medical record (EMR) system.^{7, 11} Yet while we demonstrated significant 12-month improvements in anxiety compared to PCPs' "usual care" (UC), which were similar to other collaborative care trials that utilized behaviorally trained care managers who met face-to-face with patients,^{9, 10, 12} these benefits largely accrued to the most symptomatically anxious, and we did not monitor patients to assess the durability of improvement. Moreover, we relied on research assistants to screen and obtain consent from patients in busy waiting rooms, and our trial was underpowered to determine the effectiveness of our approach for African-Americans.⁷

We designed the RELAX (REduce Limitations from AnX-ety) trial to build upon the strengths of our initial study^{7, 11} and other collaborative care trials,^{9, 10} while incorporating several pragmatic features that have not previously been evaluated in combination, in order to promote uptake of our intervention into routine practice. These include (1) using EMR-generated alerts to promote physician referrals to our study at the time of the patient encounter; (2) actively monitoring patients with moderate levels of anxiety for clinical deterioration before offering care management services; (3) addressing comorbid depressive symptoms; (4) monitoring the 12-month durability of our intervention; and (5) powering RELAX to assess the effectiveness of collaborative care among African-Americans.

METHODS

Study Setting

Implementing a protocol approved by the University of Pittsburgh institutional review board, we conducted our trial at six University of Pittsburgh Medical Center-affiliated primary care practices that shared a common EMR system (*Epic*). They included three urban faculty practices staffed by board-certified internists, and two suburban and one rural practice each staffed by family practitioners.

Participants

We developed an EMR "Best Practice Alert" (BPA) that exposed PCPs to a reminder about our study at the time of the clinical encounter.¹³ The BPA launched automatically for all patients aged 18–64 whenever anxiety, generalized anxiety, panic, or depression was on the electronic problem list or entered as an encounter diagnosis. If the patient agreed to referral, the PCP electronically "signed" the BPA that forwarded the patient's name and telephone number to a study recruiter, who then contacted the patient to confirm protocol eligibility.¹³

Physician Training

Prior to activating the BPA at each study site, we (BLR) informed PCPs about our trial at an in-service meeting, where

we provided brief education about the diagnosis and treatment of anxiety, and on how to refer patients to our trial. Later we distributed pens, mugs, and email newsletters with our study logo to PCPs to reinforce these messages and to promote referrals.

Subject Eligibility

Study recruiters telephoned referred patients to screen for anxiety disorders using the Primary Care Evaluation of Mental Disorders (PRIME-MD) anxiety module.¹⁴ Eligible patients needed to meet criteria for PD and/or GAD, could not be in active treatment with a mental health specialist (MHS), and had to have reliable telephone access and no psychotic illness, language or other communication barrier, or plans to leave the practice within the year. If this was confirmed, the recruiter administered the structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-A)¹⁵ and the Panic Disorder Severity Scale (PDSS)¹⁶ to determine symptom severity.

To remain protocol-eligible, patients were required to have at least a moderately severe level of anxiety (SIGH-A score ≥ 14 or a PDSS score ≥ 7) and no alcohol dependence, as determined by the Alcohol Use Disorders Identification Test (AUDIT-C),¹⁷ followed by the Rapid Alcohol Problems Screen¹⁸ for AUDIT-C scores ≥ 3 . Afterwards, the recruiter mailed a consent form to the patient and telephoned again to review it and obtain the patient's consent on a recorded line. If consent was provided, the research assistant asked the patient to sign and return the form, and then proceeded to administer our baseline assessment.

Randomization Procedure

Following the baseline assessment, we randomized patients who scored SIGH-A ≥ 20 or PDSS ≥ 14 to either our intervention or UC control group at a 1:1 ratio in blocks of eight, according to a computer-generated random assignment sequence prepared in advance by our study statistician (SM) and concealed until after the call. We assigned the remaining eligible patients to a watchful waiting (WW) cohort. Afterwards, we informed all patients of their treatment assignment via telephone and by letter, and notified their PCPs via an EMR message.

Usual Care

For ethical reasons,¹⁹ we informed UC patients, along with their PCPs, of their anxiety severity. However, we provided no treatment advice unless we detected suicidality on a follow-up assessment.

Intervention

Over the course of the trial, we employed four non-behaviorally trained college graduates with mental health research experience to serve as our telephone care managers. We first trained them in a basic understanding of anxiety, our treatment algorithm, and how to use our study registry to track

patients,¹¹ and reinforced this training in weekly case review meetings with our study clinicians.

Following randomization, our care managers telephoned their assigned patients to conduct a detailed mental health assessment, provide basic psychoeducation about PD and GAD, and assess their treatment preferences. As described previously,¹¹ all patients could choose any of the following: (1) a workbook for self-managing PD²⁰ or GAD²¹ (we encouraged those with both PD and GAD to choose the workbook for the condition that most distressed them); (2) a trial of anxiolytic pharmacotherapy, primarily a selective serotonin reuptake inhibitor, selected by patient preference and adjusted per patient response; or (3) referral to a community MHS. In addition, our care managers recommended such general lifestyle adjustments as social engagement, exercise, and adequate sleep.

The care managers presented their patients to the study PCP, psychiatrist, and psychologist-study coordinator in weekly 60-min case review meetings. To efficiently focus these sessions, we developed an electronic registry that could rapidly identify (1) all newly randomized patients, (2) those whose anxiety scores remained at a moderate level or higher, and (3) those the care manager had not been able to contact within a specified period.

We typically recommended a trial of pharmacotherapy or a dosage adjustment when the patient was symptomatically anxious and interested or already using pharmacotherapy.¹¹ In these cases, we advised the care manager to recommend a specific medication name and dosage to the PCP and patient. We also recommended referral to a MHS when the patient had complex psychosocial issues (e.g., impending divorce) or failed to recover despite three pharmacotherapy trials of adequate dose and duration.

Following case review, the care manager forwarded these treatment recommendations to the patient's PCP via the EMR for their consideration. The PCP was always free to reject this advice, as they were required to prescribe and approve all medications.

During the acute phase of treatment, the care manager telephoned patients every other week to review workbook assignments (e.g., read a chapter weekly and practice the techniques imparted); monitor pharmacotherapy; administer the PDSS, Generalized Anxiety Disorder Severity Scale (GADSS),²² or 9-item Patient Health Questionnaire (PHQ-9)²³ to assess treatment response; and inform the patient of new treatment recommendations generated at our case review sessions. Depending upon the patient's treatment choice(s), symptoms, and motivation, these contacts lasted 15–45 min and continued for 2 to 4 months. Afterwards, the patient transitioned to the "continuation phase" of care, during which the care manager contacted him/her every 1–2 months until the end of our 12-month intervention.

Watchful Waiting

We assigned patients who failed to meet severity criteria for randomization at baseline (SIGH-A 14–19 or PDSS 7–13) to

our WW cohort. Because all cases of clinical deterioration following randomization among UC patients in our previous anxiety trial had occurred by the 4-month follow-up,⁷ we later randomized all WW patients whose symptoms met our threshold for randomization on either their 2- or 4-month assessment.

Assessments

We blinded our telephone assessors to patients' randomization status, and used their ratings to determine the effectiveness of our intervention. However, given the nature of our intervention, neither patients nor their PCPs were blinded to the patient's treatment assignment. We used audiotapes, manuals, and practice interviews to train our assessors and establish inter-rater agreement for the SIGH-A, PDSS, and GADSS on an annual basis, or sooner whenever there was a change in personnel. Assessors also used a computer-assisted telephone interview (CATI) system that guided them in standardized administration of all study instruments. We further promoted quality assurance by digitally recording assessment contacts and conducting periodic spot-checks to confirm that responses were rated accurately and corresponded with those entered into our study database.

We collected information on patients' self-reported race and sociodemographic characteristics, and administered the SF-36²⁴ to determine HRQoL; the SIGH-A,¹⁵ PDSS,¹⁶ and GADSS²² to measure anxiety; and the PHQ-9²³ to assess mood symptoms at baseline and at 2-, 4-, 8-, and 12-month conclusion of our intervention on all subjects, and at 18 and 24 months on those enrolled in the first 2 years of our 3-year recruitment phase. Lastly, we abstracted each patient's electronic medical record to collect information on their comorbid medical conditions and health services utilization, and our care managers' electronic registry to document the number of intervention contacts and other process measures of care.

Data and Safety Monitoring

We programmed our data management system to identify those whose blinded SIGH-A or PDSS score increased $\geq 25\%$ above their baseline. If indicated following a review, we wrote to the treating PCP and offered to provide the names of local MHSs and additional treatment advice. Whenever our care managers or blinded assessors encountered suicidality, either spontaneously expressed or on routine administration of the PHQ-9, our CATI system automatically launched our suicide risk management protocol that provided triage advice (e.g., contact the study psychiatrist).²⁵ An independent data and safety monitoring board appointed by our funding agency also monitored the progress and safety of our trial.

Statistical Analyses

We powered RELAX to test our primary hypotheses within African-Americans on the SF-36 MCS, our primary outcome

measure, at 12-month follow-up. For a 2 (race) $\times$ 2 (gender) $\times$ time (baseline and 12-month) ANOVA with a continuous measure outcome and assuming a two-tailed $\alpha = 0.05$ and 15 % 12-month attrition rate, we needed to randomize 48 African-Americans to have 86 % power to detect a Cohen's d effect size (ES) of 0.50 or larger on the SF-36 MCS. We assumed minority patients would constitute 25 % of our randomized study cohort, and African-Americans would constitute 80 % of all minority patients (20 % overall). Therefore, we estimated we needed to randomize 240 highly anxious patients to randomize 48 African-Americans (SIGH-A ≥ 20 or PDSS ≥ 14). Based on our earlier anxiety trial,⁷ we assumed two-thirds of all enrolled patients would meet severity criteria for randomization. Thus we estimated we needed to enroll 360 subjects.

We compared baseline sociodemographic, clinical, and functional status measures by baseline anxiety severity and randomization status using t tests for continuous data and chi-square analyses for categorical data. All reported P values are two-tailed, with significance levels for all tests at $P \leq 0.05$.

Using established criteria,^{9,26} we defined a 40 % reduction from the baseline level of anxiety as a significant treatment response. We examined the impact of the intervention on our continuous measures for HRQoL, anxiety, and depression using repeated measures mixed-effect models to account for between-subject variations and to permit inclusion of patients with one or more missing follow-up assessments,²⁷ and calculated standard intent-to-treat ES differences between intervention and usual care groups of highly anxious patients at 12-month follow-up with 95 % confidence intervals (CI). In these models, we considered intercept and time as random effects, and group and group $\times$ time interaction as fixed effects. We then repeated these analyses on patients who completed their 24-month assessment in order to examine the durability of our intervention.

Given the sensitivity of the PDSS, GADSS, and PHQ-9 for panic, generalized anxiety, and mood symptoms, respectively, we conducted secondary analyses involving participants who met DSM-IV criteria for PD, GAD, or depression on their baseline PRIME-MD. Finally, we used Wilcoxon non-parametric rank-sum tests to examine differences in health services utilization among treatment conditions, given the non-normal underlying distribution of our data. All analyses were performed with the SAS statistical program using the PROC MIXED function for calculation of ES changes and scores (SAS Institute Inc., Cary, NC).

RESULTS

Patient Recruitment and Follow-Up

Over a 36-month period (01/2005–12/2007), PCPs referred 1338 patients to our trial in response to the EMR prompt, and 329 met all eligibility criteria and provided informed consent to enroll. Following their baseline assessment, 250 (76 %) met severity criteria for randomization (“high anxiety”), and we assigned the other 79 to our WW cohort (“moderate anxiety”) (Fig. 1).

Baseline sociodemographic and clinical characteristics of our highly anxious patients were similar by randomization status. However, compared to WW patients, those who were highly anxious were more likely to be female, unemployed, and unmarried, with lower levels of education, social support, and HRQoL, and higher levels of mood and anxiety symptoms (Table 1). Twenty-three WW patients later experienced clinical deterioration and met the severity criteria for randomization. They were more likely to be African-American, have lower levels of education, and score higher on their baseline SIGH-A than WW patients whose conditions did not deteriorate (Table 2).

12-Month Effectiveness of Collaborative Care

Intent-to-treat analyses involving the 250 highly anxious patients randomized at baseline demonstrated small to medium ES improvements for intervention patients versus UC on our primary outcome measure, the SF-36 MCS (ES: 0.38; 95 % CI: 0.13–0.63), and all secondary measures of anxiety and mood (Table 3). The number needed to treat (NNT) to achieve one additional ≥ 40 % decline in symptoms ranged from 3 to 5, depending on the measure (Table 4). Pre-planned subgroup analyses involving African-Americans and men showed similar ES improvements and NNTs (Table 3). Finally, we observed no differences in symptom improvements among the 23 WW patients with clinical deterioration by randomization status (Fig. 1), and adding these patients to those randomized at baseline did not change our findings (not shown).

Intervention Durability

Twelve months following the conclusion of our intervention, highly anxious African-Americans and men randomized to our intervention continued to report significant ES improvements on all primary and secondary measures of anxiety and mood (Table 3). Figure 2 portrays these trajectories over the full 24-month course of follow-up and illustrates that patients assigned to WW at baseline continually reported improved mental HRQoL and lower levels of anxiety than those randomized at baseline.

Care Processes Following Randomization

Care managers had a median of 14 telephone contacts with intervention patients over the 12-month intervention (range: 0–32), and compared to UC, intervention patients were more likely to be using anxiolytic pharmacotherapy (97 % vs. 85 %; $P = 0.002$) and less likely to go to the emergency room two or more times over the first 12 months following enrollment (8 % vs. 2 %; $P = 0.05$). However, the number of visits to PCPs, MHSs, or hospitalizations by randomization status was similar (Table 1). Finally, the number of visits to PCPs, MHSs, and ERs, and the use of pharmacotherapy within our high anxiety intervention cohorts were similar by race and gender (Table 2).

Figure 1 Flowchart of participants. Abbreviations: HRS-A, Hamilton Rating Scale for Anxiety¹⁵; PDSS, Panic Disorder Severity Scale¹⁶; PQ, Patient Questionnaire portion of Primary Care Evaluation of Mental Disorders (PRIME-MD)¹⁴; GAD, generalized anxiety disorder; PD, panic disorder. * Referred by PCP 01/2005 to 12/2007.

DISCUSSION

The RELAX trial confirms that a centralized, telephone-delivered collaborative care strategy provided by non-mental health professionals is more effective than PCPs' usual care at reducing anxiety and improving HRQoL among highly anxious patients referred by their PCP for treatment. It also extends our understanding of the effectiveness and durability of collaborative care among African-Americans and those with comorbid depression.

Furthermore, RELAX promotes the implementation of collaborative care into routine practice by demonstrating the effectiveness and efficiency of (1) EMR-generated alerts at identifying and then facilitating PCP referrals of distressed patients for further evaluation; and (2) focusing care managers' efforts on patients who are most likely to benefit from their attention.

RELAX corroborates the primary findings from our earlier trial that relied on hired research assistants to screen patients

Table 1 Baseline Sociodemographic and Clinical Characteristics by Baseline Anxiety Severity and Randomization Status

	High anxiety intervention* (N = 124)	High anxiety usual care* (N = 126)	P	High anxiety* (N = 250)	Moderate anxiety* (N = 79)	P
Age, mean (SD)	45.0 (10.5)	44.2 (11.9)	0.61	44.6 (11.2)	42.1 (11.1)	0.09
Male, % (N)	27 (33)	25 (32)	0.83	26 (65)	38 (30)	0.04
White, % (N)	77 (96)	78 (98)	0.95	78 (194)	85 (67)	0.17
African-American, % (N)	21 (26)	21 (27)	0.93	21 (53)	14 (11)	0.15
Other, % (N)	2 (2)	1 (1)	0.55	1 (3)	1 (1)	0.96
> High school education, % (N)	68 (84)	72 (91)	0.44	70 (175)	85 (67)	0.01
Married, % (N)	57 (71)	54 (68)	0.60	56 (139)	70 (55)	0.03
Living alone, % (N)	19 (23)	19 (24)	0.92	19 (47)	18 (14)	0.83
Working, part-time/full-time, % (N)	56 (69)	54 (68)	0.79	55 (137)	75 (59)	0.002
Social Support Scale, mean (SD)‡	64.1 (24.6)	65.7 (25.2)	0.61	64.9 (24.9)	74.3 (20.5)	<0.001
Anxiety disorder, % (N)						
Generalized anxiety	46 (57)	47 (59)	0.34	46 (116)	48 (38)	<0.0001
Panic	3 (4)	7 (9)		5 (13)	25 (20)	
Generalized anxiety and panic	51 (63)	46 (58)		48 (121)	27 (21)	
Major depression, % (N)†	90 (111)	90 (114)	0.80	90 (225)	68 (54)	<0.0001
SF-36 MCS, mean (SD)§ ‡‡	27.4 (10.5)	28.7 (9.9)	0.33	28.0 (10.2)	37.8 (8.9)	<0.0001
SF-36 PCS, mean (SD)§ ‡‡	45.6 (12.1)	45.3 (11.7)	0.85	45.5 (11.9)	52.0 (9.0)	<0.0001
SIGH-A, mean (SD)¶ ††	28.4 (7.3)	28.1 (6.5)	0.71	28.3 (6.9)	17.6 (4.0)	<0.0001
SIGH-A ≥ 20, % (N)	94 (116)	95 (120)	0.56	94 (236)	16 (13)	<0.0001
PDSS, mean (SD)¶ ††	12.8 (6.8)	12.4 (6.4)	0.64	12.6 (6.6)	7.6 (4.5)	<0.0001
PDSS ≥ 14, % (N)	56 (69)	50 (63)	0.37	53 (132)	5 (4)	<0.0001
GADSS, mean (SD)# ††	15.9 (3.1)	15.7 (3.2)	0.60	15.8 (3.2)	11.0 (2.7)	<0.0001
PHQ-9, mean (SD)** ††	15.2 (5.1)	15.0 (5.1)	0.73	15.1 (5.1)	9.6 (4.1)	<0.0001

Abbreviations: GAD, generalized anxiety disorder; GADSS, Generalized Anxiety Disorder Severity Scale; PRIME-MD, Primary Care Evaluation of Mental Disorders; PD, panic disorder; PDSS, Panic Disorder Severity Scale; PHQ-9, 9-item Patient Health Questionnaire; SIGH-A, Structured interview guide for the Hamilton Anxiety Rating Scale; SF-36 MCS, Medical Outcomes Study Short Form Mental Component Scale; SF-36 PCS, Medical Outcomes Study Short Form Physical Component Scale

* “High anxiety” patients scored either a SIGH-A ≥ 20 or a PDSS ≥ 10, and “moderate anxiety” patients who were assigned to our “watchful waiting” cohort scored either 14–19 on the SIGH-A or 7–9 on the PDSS, and not in the “high anxiety” range on either measure

† Determined using Primary Care Evaluation of Mental Disorders (PRIME-MD)

‡ Range 11–55

§ Range 0–100

¶ Range 0–56

¶ Range 0–28

Range 0–24

** Range 0–27

†† Higher scores indicate more severe symptoms

‡‡ Higher scores indicate better health-related quality of life

for study eligibility in practice waiting rooms (12-month SF-12 MCS ES: 0.38).⁷ Moreover, perhaps because of our recruitment of a larger and more symptomatic study cohort and our use of the more sensitive GADSS for assessing generalized anxiety symptoms than SIGH-A,²² we are now able to demonstrate the effectiveness of our intervention for GAD. Yet, perhaps most important to improving the efficiency, cost-effectiveness, and subsequent adoption of our treatment strategy into routine practice, our findings confirm post hoc analyses from our previous trial, which suggested that patients with just moderately elevated levels of anxiety tend to improve regardless of active care management.⁶

The ES improvements we obtained were similar to those described in a meta-analysis of collaborative care that involved 813 anxious primary care patients (ES: 0.33; 0.19–0.47)¹, as well as the Coordinated Anxiety Learning and Management (CALM) study by Roy-Byrne et al., the largest trial of collaborative care for anxiety.⁸ CALM randomized 1004 patients with PD, GAD, social anxiety, or post-traumatic stress disorder to either their PCPs’ usual care or an intervention that included anxiolytic pharmacotherapy, a computer-assisted program to optimize

delivery of cognitive behavioral therapy,²⁸ and both face-to-face and telephone contact from a study care manager. At 18-month follow-up, intervention patients reported ES improvements in HRQoL (SF-12 MCS: 0.39) and mood (PHQ-8: 0.24)⁸ similar to those in RELAX. However, neither CALM nor previous collaborative care trials for anxiety^{9,10} have utilized an EMR prompt to promote referrals, incorporated a WW cohort, separately reported the impact of their treatment strategy on African-Americans, specifically addressed the challenge of comorbid depression, or assessed study subjects for a full year following their intervention.

Our use of automated EMR-generated alerts to facilitate identification of anxious patients enabled us to enroll study patients from multiple practice locations simultaneously, and without burdening busy clinical staff or deploying research assistants in practice waiting rooms.^{7,29} Furthermore, it enabled us to exceed our randomization target within our trial’s scheduled 3-year recruitment phase (114 %; 273/240). However, since PCPs may fail to recognize and enter the anxiety disorder on the electronic problem list, we also programmed our BPA to fire when depression was entered into the EMR.

Table 2 Baseline Predictors of Clinical Deterioration at 2- and 4-Month Follow-Up Among Moderately Anxious Patients Assigned to the Watchful Waiting (WW) Cohort at Baseline

	Watchful waiting randomized* (N =23)	Watchful waiting not randomized* (N = 56)	P
Age, mean (SD)	44.7 (10.2)	41.1 (11.3)	0.18
Male, % (N)	26 (6)	43 (24)	0.16
White, % (N)	74 (17)	89 (50)	0.08
African-American, % (N)	26 (6)	9 (5)	0.05
Other, % (N)	0 (0)	2 (1)	0.52
> High school education, % (N)	70 (16)	91 (51)	0.02
Married, % (N)	65 (15)	71 (40)	0.59
Living alone, % (N)	22 (5)	16 (9)	0.55
Working, part-time/full-time, % (N)	74 (17)	75 (42)	0.92
Social Support Scale, mean (SD)	67.3 (20.7)	77.2 (19.8)	0.06
PRIME-MD diagnosis, % (N)			0.28
GAD	35 (8)	54 (30)	
PD	35 (8)	21 (12)	
GAD/PD	30 (7)	25 (14)	
Major depression, % (N)†	78 (18)	64 (36)	0.23
SF-36v2 MCS, mean (SD)§ ‡‡	37.3 (7.9)	38.0 (9.3)	0.75
SF-36v2 PCS, mean (SD)§ ‡‡	49.8 (9.4)	52.9 (8.7)	0.18
SIGH-A, mean (SD) ††	19.8 (5.5)	16.8 (2.8)	0.02
PDSS, mean (SD)¶ ††	8.4 (3.9)	7.3 (4.8)	0.28
GADSS, mean (SD)# ††	11.4 (2.8)	10.8 (2.7)	0.37
PHQ-9, mean (SD)** ††	9.8 (4.1)	9.5 (4.2)	0.75

Abbreviations: GAD, generalized anxiety disorder; GADSS, Generalized Anxiety Disorder Severity Scale; PRIME-MD, Primary Care Evaluation of Mental Disorders; PD, panic disorder; PDSS, Panic Disorder Severity Scale; PHQ-9, 9-item Patient Health Questionnaire; SIGH-A, structured interview guide for the Hamilton Anxiety Rating Scale; SF-36 MCS, Medical Outcomes Study Short Form Mental Component Scale; SF-36 PCS, Medical Outcomes Study Short Form Physical Component Scale

* At baseline, “watchful waiting” patients scored either 14–19 on the SIGH-A or 7–10 on the PDSS and were not in the “high anxiety” range on either measure. Later, we randomized those who scored ≥ 20 on the SIGH-A or ≥ 10 on the PDSS at either their 2- or 4-month telephone follow-up assessment

†Determined using Primary Care Evaluation of Mental Disorders (PRIME-MD)

‡Range 11–55

§Range 0–100

|| Range 0–56

¶Range 0–28

#Range 0–24

**Range 0–27

††Higher scores indicate more severe symptoms

‡‡Higher scores indicate better health-related quality of life

Perhaps as a result, we enrolled a larger proportion of subjects with major depression (85 %; 279/329) than expected, given the approximately 50 % comorbidity between mood and anxiety disorders.³⁰ Still, given our focus on anxiety, we required all study patients to meet criteria for PD and/or GAD on the PRIME-MD and to express at least a moderate level of anxiety symptoms. Given the widespread adoption of EMR systems into large practice networks,³¹ these alerts have important implications for (1) implementing evidence-based recommendations for screening primary care³² and specialty populations³³ for common mental health conditions, and (2) increasing participation of minorities and other groups traditionally underrepresented in clinical research.^{34–36}

Although our care managers frequently discussed the use of our anxiety workbooks with patients during calls, fewer than one in five patients self-reported completing them, despite multiple exhortations from our care managers, particularly when the patient remained symptomatic (Table 1). Moreover, we have no information on whether those who claimed to have read it actually did so. Although the overall ES improvements in anxiety and mood symptoms in RELAX are typical for collaborative care,¹ they are smaller than those generated by

experienced mental health professionals meeting face-to-face with clients (ES: 0.57; 0.29–0.84).³⁷ More recently, Internet-delivered computerized cognitive behavioral therapy (CCBT) provided under the direction of a “clinical helper” has been shown to be as effective as in-person CBT (ES: 0.88 1.08).^{38,39} Thus, replacing patient workbooks with proven effective CCBT programs could potentially enable a more effective and standardized version of collaborative care to be provided at scale.

Limitations

First, we did not systematically screen patients for a mental health disorder. Thus our reliance on EMR-generated alerts to identify patients for treatment is limited to settings with systems capable of generating these alerts and to clinician recognition and entry of the proper diagnostic (or medication) codes into the EMR. Second, a large proportion of patients failed to complete the workbook, which may have limited the effectiveness of our intervention. Third, we lacked claims to confirm patients’ use of services or to calculate any potential cost offset. However, we estimate that they are similar to CALM’s modest \$245 additional cost over an 18-month period (\$4.30/

Table 3 Effect Size Differences in Mean Rating Scale Scores Between Highly Anxious Patients Randomized to Collaborative Care Intervention or Usual Care Control Groups at 12- and 24-Month Follow-Up*

	12 Months (N = 208)†		24 Months (N = 138)†	
	Effect size (95 % CI)	P	Effect size (95 % CI)	P
SF-36 MCS				
All	0.38 (0.13–0.63)	0.003	0.17 (−0.07 to 0.42)	0.17
Male	0.64 (0.15–1.13)	0.01	0.73 (0.24–1.22)	0.003
Female	0.29 (0.00–0.58)	0.05	0.01 (−0.28 to 0.30)	0.95
African-American	0.47 (−0.08 to 1.01)	0.09	0.70 (0.16–1.24)	0.01
SIGH-A				
All	0.30 (0.05–0.55)	0.02	0.30 (0.05–0.54)	0.02
Male	0.72 (0.24–1.21)	0.004	0.78 (0.29–1.27)	0.002
Female	0.15 (−0.14 to 0.44)	0.32	0.13 (−0.15 to 0.42)	0.36
African-American	0.57 (0.03–1.11)	0.04	1.14 (0.60–1.68)	0.0001
PDSS (PD and PD/GAD cohorts only)‡				
All	0.42 (0.08–0.76)	0.02	0.19 (−0.15 to 0.53)	0.27
Male	0.87 (0.20–1.53)	0.01	0.75 (0.08–1.42)	0.03
Female	0.24 (−0.16 to 0.64)	0.23	0.03 (−0.37 to 0.43)	0.89
African-American	0.52 (−0.19 to 1.22)	0.15	0.71 (0.01–1.42)	0.05
GADSS (GAD and PD/GAD cohorts only)‡				
All	0.41 (0.15–0.66)	0.002	0.06 (−0.19 to 0.32)	0.63
Male	0.66 (0.16–1.15)	0.01	0.43 (−0.07 to 0.92)	0.09
Female	0.32 (0.02–0.62)	0.04	0.06 (−0.23 to 0.36)	0.67
African-American	0.90 (0.34–1.47)	0.002	0.82 (0.26–1.38)	0.005
SF-36 PCS				
All	0.06 (−0.19 to 0.31)	0.62	0.01 (−0.24 to 0.25)	0.96
Male	0.17 (−0.32 to 0.66)	0.49	0.19 (−0.29 to 0.68)	0.44
Female	0.14 (−0.14 to 0.43)	0.34	0.06 (−0.23 to 0.35)	0.69
African-American	0.02 (−0.52 to 0.56)	0.94	0.20 (−0.34 to 0.75)	0.46
PHQ-9 (Major depression cohort only)‡				
All	0.45 (0.19–0.71)	0.001	0.22 (−0.04 to 0.48)	0.10
Male	0.92 (0.39–1.46)	0.001	0.93 (0.40–1.46)	0.001
Female	0.32 (0.02–0.62)	0.04	0.03 (−0.28 to 0.33)	0.87
African-American	0.48 (−0.08 to 1.04)	0.09	0.78 (0.21–1.34)	0.01

Abbreviations: see Tables 1 and 2

* Positive scores indicate patients randomized to our intervention reported improvement in the measure vs. usual care at the time-point indicated

†Of the 250 highly anxious patients randomized at baseline, 208 completed our 12-month assessment and 138 completed our 24-month assessment

‡Determined using Primary Care Evaluation of Mental Disorders (PRIME-MD)

Table 4 Proportion of Patients Achieving ≥40 % Decline from Baseline Levels of Anxiety or Mood Symptoms at 12-Month Follow-Up

	High anxiety intervention		High anxiety usual care		P	NNT (95 % CI)*
	N	% (N)	N	% (N)		
SIGH-A						
All	99	53 (52)	108	32 (35)	0.003	5 (14, 3)
Male	28	71 (20)	27	33 (9)	0.005	3 (7, 2)
Female	71	45 (32)	81	32 (26)	0.10	8 (−41, 4)
African-American	21	38 (8)	19	16 (3)	0.16	4 (−24, 2)
PDSS (PD and PD/GAD cohorts only)†						
All	50	70 (35)	55	38 (21)	0.001	3 (7, 2)
Male	16	81 (13)	12	25 (3)	0.006	2 (4, 1)
Female	34	65 (22)	43	42 (18)	0.05	4 (96, 2)
African-American	10	60 (6)	12	17 (2)	0.07	2 (16, 1)
GADSS (GAD and PD/GAD cohorts only)†						
All	97	49 (48)	100	28 (28)	0.002	5 (12, 3)
Male	27	67 (18)	26	31 (8)	0.009	3 (9, 2)
Female	70	43 (30)	74	27 (20)	0.05	6 (226, 3)
African-American	21	48 (10)	17	6 (1)	0.01	2 (6, 2)
PHQ-9 (Major depression cohort only)†						
All	87	64 (56)	98	41 (40)	0.001	4 (10, 3)
Male	21	76 (16)	26	38 (10)	0.01	3 (9, 2)
Female	66	61 (40)	72	42 (30)	0.03	5 (39, 3)
African-American	19	58 (11)	19	26 (5)	0.05	3 (55, 2)

Abbreviations: see Tables 1 and 2

*Number needed to treat

†Determined using Primary Care Evaluation of Mental Disorders (PRIME-MD)

Figure 2 Observed main outcomes by baseline treatment assignment. The vertical line at 12 months signifies the end of the care manager-led intervention. The following 12 months were naturalistic follow-up to observe the durability of the intervention. * Statistically significant difference in outcome score for usual care vs. intervention, $P \leq 0.05$.

AFD)⁴⁰ and to the \$460 cost to provide a nurse-led, telephone-delivered, centralized collaborative care intervention for treating depression following coronary artery bypass graft surgery.⁴¹ Finally, since RELAX was not cluster-randomized, PCPs cared for patients in both treatment arms. While this could have produced a spillover effect that diminished outcome differences between treatment arms, we observed ES improvements similar to those in the cluster-randomized CALM³¹ and other trials.¹

In summary, a telephone-delivered, centralized collaborative care intervention provided by non-mental health professionals to highly anxious patients referred by their PCP in response to an EMR-generated prompt led to significantly improved mental HRQoL, anxiety, and mood symptoms compared to PCPs' usual care. These improvements were durable for up to 12 months, and were particularly evident among African-Americans. Future collaborative care trials should evaluate the effectiveness of replacing self-management workbooks with Internet-delivered CCBT programs that have the

potential to deliver more effective behavioral health treatment at scale.

Acknowledgments: Dr. Rollman had full access to all the data in the study and takes responsibility for the conceptualization of the described trials, their conduct, and data integrity as principal investigator, and for the preparation of this manuscript as its first and corresponding author.

Dr. Herbeck Belnap assisted with the implementation and conduct of both trials, their data collection, and with data analysis.

Drs. Abebe and Mazumdar, as trial statisticians, take responsibility for the accuracy of the data analysis.

All co-authors received a copy of this manuscript and provided Dr. Rollman with their comments on earlier drafts prior to its submission for editorial review.

We thank Fanyin He, Ph.D., for performing the primary analyses over the course of the trial under the direction of Dr. Mazumdar, and Diane M. Comer, B.A., at the Center for Research on Health Care's Data Center, for her data analyses and generating the graphics. We also thank our study staff and the many patients who volunteered to participate in the RELAX trial.

Corresponding Author: Bruce L. Rollman, M.D., M.P.H.; Division of General Internal Medicine, Center for Behavioral Health and Smart Technology, University of Pittsburgh School of Medicine, Suite 600, 230

McKee Place, Pittsburgh, PA 15213-2582, USA
(e-mail: rollmanbl@upmc.edu).

Compliance with Ethical Standards:

Role of the Funding Source: All work described herein was supported by grants from the National Institute of Mental Health (R01 MH09421 and R01 MH093501). The funding source had no role in the design, conduct, or reporting of our study, or in the preparation, review, or decision to submit this manuscript for publication.

Previous Presentations: Parts of this manuscript were presented at the annual national meeting of the Society for General Internal Medicine (Minneapolis, MN; April 29, 2010).

Conflict of Interest: Dr. Lenze has received research funding from the McKnight Brain Research Foundation, Barnes Jewish Foundation, Takeda, and Lundbeck. None of the other authors have any financial disclosures or potential conflicts of interest to report.

REFERENCES

1. Archer J, Bower P, Gilbody S, et al. Collaborative care for depression and anxiety problems. 2012; (4):CD004274; 10:CD006525.
2. Bao Y, Casalino LP, Pincus HA. Behavioral health and health care reform models: patient-centered medical home, health home, and accountable care organization. *J Behav Health Serv Res*. 2013;40:121-132.
3. Katon WJ, Unützer J. Health reform and the Affordable Care Act: The importance of mental health treatment to achieving the triple aim. *J Psychosom Res*. 2013;74:533-537.
4. Coleman K, Austin BT, Brach C, Wagner EH. Evidence on the chronic care model in the new millennium. *Health Aff*. 2009;28:75-85.
5. Katon W, Unützer J, Wells K, Jones L. Collaborative depression care: history, evolution and ways to enhance dissemination and sustainability. *Gen Hosp Psychiatry*. 2010;32:456-464.
6. Reed SJ, Shore KK, Tice JA. Effectiveness and value of integrating behavioral health into primary care. *JAMA Intern Med*. 2016;176:691-692.
7. Rollman BL, Belnap BH, Mazumdar S, et al. A randomized trial to improve the quality of treatment for panic and generalized anxiety disorders in primary care. *Arch Gen Psychiatry*. 2005;62:1332-1341.
8. Roy-Byrne P, Craske MG, Sullivan G, et al. Delivery of evidence-based treatment for multiple anxiety disorders in primary care: a randomized controlled trial. *JAMA*. 2010;303:1921-1928.
9. Roy-Byrne PP, Katon W, Cowley DS, Russo J. A randomized effectiveness trial of collaborative care for patients with panic disorder in primary care. *Arch Gen Psychiatry*. 2001;58:869-876.
10. Roy-Byrne PP, Craske MG, Stein MB, et al. A randomized effectiveness trial of cognitive-behavioral therapy and medication for primary care panic disorder. *Arch Gen Psychiatry*. 2005;62:290-298.
11. Rollman BL, Herbeck Belnap B, Reynolds C, Schulberg H, Shear M. A contemporary protocol for the treatment of panic and generalized anxiety in primary care. *Gen Hosp Psychiatry*. 2003;25:74-82.
12. Seekles W, Cuijpers P, Kok R, Beekman A, van Marwijk H, van Straten A. Psychological treatment of anxiety in primary care: a meta-analysis. *Psychol Med*. 2013;43:351-361.
13. Rollman BL, Fischer GS, Zhu F, Belnap BH. Comparison of electronic physician prompts versus waitroom case-finding on clinical trial enrollment. *J Gen Intern Med*. 2008;23:447-450.
14. Spitzer RL, Williams JBW, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. *JAMA*. 1994;272:1749-1756.
15. Shear MK, Vander Bilt J, Rucci P, et al. Reliability and validity of a structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-A). *Depress Anxiety*. 2001;13:166-178.
16. Shear MK, Brown TA, Barlow DH, et al. Multicenter collaborative panic disorder severity scale. *Am J Psychiatry*. 1997;154:1571-1575.
17. Gordon AJ, Maisto SA, McNeil M, et al. Three questions can detect hazardous drinkers. *J Fam Practice*. 2001;50:313-320.
18. Cherpitel CJ. A brief screening instrument for problem drinking in the emergency room: the RAPS4. *Rapid Alcohol Problems Screen J*. 2000;61:447-449.
19. Reynolds CF 3rd, Degenholtz H, Parker LS, et al. Treatment as usual (TAU) control practices in the PROSPECT Study: managing the interaction and tension between research design and ethics. *Int J Geriatric Psychiatry*. 2001;16:602-608.
20. Craske M, Barlow D. Mastery of Your Anxiety and Panic for Primary Care. Adapted from Mastery of Your Anxiety and Panic, 3rd ed. San Antonio, Tex: The Psychological Corporation; 1999.
21. Craske MG, Barlow DH. Mastery of your Anxiety and Worry. 2nd ed. New York: Oxford University Press; 2006.
22. Shear K, Belnap BH, Mazumdar S, Houck P, Rollman BL. Generalized anxiety disorder severity scale (GADSS): a preliminary validation study. *Depress Anxiety*. 2006;23:77-82.
23. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16:606-613.
24. Ware J, Kosinski M, Keller S. SF-36 physical and mental health summary scales: a user's manual. 2nd ed. Boston: New England Medical Center; 1994.
25. Herbeck Belnap B, Schulberg HC, He F, Mazumdar S, Reynolds CF 3rd, Rollman BL. Electronic protocol for suicide risk management in research participants. *J Psychosom Res*. 2015;78:340-345.
26. Barlow DH, Gorman JM, Shear MK, Woods SW. Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: a randomized controlled trial. *JAMA*. 2000;283:2529-2536.
27. Gibbons RD, Hedeker D, Elkin I, et al. Some conceptual and statistical issues in analysis of longitudinal psychiatric data. Application to the NIMH treatment of depression collaborative research program dataset. *Arch Gen Psychiatry*. 1993;50:739-750.
28. Craske MG, Rose RD, Lang A, et al. Computer-assisted delivery of cognitive behavioral therapy for anxiety disorders in primary-care settings. *Depress Anxiety*. 2009;26:235-242.
29. Rollman BL, Hanusa BH, Lowe HJ, Gilbert T, Kapoor WN, Schulberg HC. A randomized trial using computerized decision support to improve the quality of treatment for major depression in primary care. *J Gen Intern Med*. 2002;17:493-503.
30. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62:617-627.
31. Kern LM, Barron Y, Dhopeswarkar RV, Edwards A, Kaushal R, Investigators H. Electronic health records and ambulatory quality of care. *J Gen Intern Med*. 2013;28:496-503.
32. Siu AL, Force USPST, Bibbins-Domingo K, et al. Screening for depression in adults: US Preventive Services Task Force recommendation statement. *JAMA*. 2016;315:380-387.
33. Lichtman JH, Bigger JT Jr, Blumenthal JA, et al. Depression and coronary heart disease: recommendations for screening, referral, and treatment: a science advisory from the American Heart Association Prevention Committee of the Council on Cardiovascular Nursing, Council on Clinical Cardiology, Council on Epidemiology and Prevention, and Interdisciplinary Council on Quality of Care and Outcomes Research: endorsed by the American Psychiatric Association. *Circulation*. 2008;118:1768-1775.
34. Sears SR, Stanton AL, Kwan L, et al. Recruitment and retention challenges in breast cancer survivorship research: results from a multisite, randomized intervention trial in women with early stage breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2003;12:1087-1090.
35. Tanner A, Kim S-H, Friedman DB, Foster C, Bergeron CD. Promoting clinical research to medically underserved communities: current practices and perceptions about clinical trial recruiting strategies. *Contemp Clin Trials*. 2015;41:39-44.
36. Mendoza DB, Williams MT, Chapman LK, Powers M. Minority inclusion in randomized clinical trials of panic disorder. *J Anxiety Disord*. 2012;26:574-582.
37. Cuijpers P, Kok R, Beekman A, van Marwijk H, van Straten A. Psychological treatment of anxiety in primary care: a meta-analysis. *Psychol Med*. 2013;43:351-361.
38. Cuijpers P, Marks IM, van Straten A, Cavanagh K, Gega L, Andersson G. Computer-aided psychotherapy for anxiety disorders: a meta-analytic review. *Cogn Behav Ther*. 2009;38:66-82.
39. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS ONE*. 2010;5:e13196.

-
40. **Joesch JM, Sherbourne CD, Sullivan G, Stein MB, Craske MG, Roy-Byrne P.** Incremental benefits and cost of coordinated anxiety learning and management for anxiety treatment in primary care. *Psychol Med.* 2012;42:1937–1948.
 41. **Donohue JM, Beina BH, Men A, et al.** Twelve-month cost-effectiveness of telephone-delivered collaborative care for treating depression following CABG surgery: a randomized controlled trial. *Gen Hosp Psychiatry.* 2014;36:453–459.