

A Stepped-Wedge Evaluation of an Initiative to Spread the Collaborative Care Model for Depression in Primary Care

Leif I. Solberg, MD¹

A. Lauren Crain, PhD¹

Michael V. Maciosek, PhD¹

Jürgen Unützer, MD, MPH²

Kris A. Ohnsorg, RN, MPH¹

Arne Beck, PhD³

Lisa Rubenstein, MD⁴

Robin R. Whitebird, PhD, MSW¹

Rebecca C. Rossom, MD, MSCR¹

Pamela B. Pietruszewski, MA⁵

Benjamin F. Crabtree, PhD⁶

Kenneth Joslyn, MD⁷

Andrew Van de Ven, PhD⁸

Russell E. Glasgow, PhD⁹

¹HealthPartners Research Foundation, Minneapolis, Minnesota

²University of Washington Medical Center, Seattle, Washington

³Kaiser Permanente Colorado, Denver, Colorado

⁴RAND Corporation, Santa Monica, California

⁵Institute for Clinical Systems Improvement, Minneapolis, Minnesota

⁶Robert Wood Johnson Medical School, New Brunswick, New Jersey

⁷Private practice, Minneapolis, Minnesota

⁸University of Minnesota, Minneapolis, Minnesota

⁹University of Colorado, Denver, Colorado



Conflicts of interest: authors report none.

CORRESPONDING AUTHOR

Leif I. Solberg, MD
HealthPartners Institute for Education
and Research
PO Box 1524, MS#23301A
Minneapolis, MN 55440-1524
leif.i.solberg@healthpartners.com

ABSTRACT

PURPOSE Scale-up and spread of evidence-based practices is one of the most important challenges facing health care. We tested whether a statewide initiative, Depression Improvement Across Minnesota—Offering a New Direction (DIAMOND), to implement the collaborative care model for depression in 75 primary care clinics resulted in patient outcome improvements corresponding to those reported in randomized controlled trials.

METHODS Health plans provided a new monthly payment to participating clinics after a 6-month intensive training program with ongoing data submission, networking, and consultation. Implementation was staggered, with 5 sequences of 10 to 40 clinics every 6 months. Payers provided weekly contact information for members from participating clinics who were filling antidepressant prescriptions, and we conducted baseline and 6-month surveys of 1,578 patients about their care and outcomes.

RESULTS There were 466 patients in DIAMOND clinics who received usual care before implementation (UCB), 559 who received usual care in DIAMOND clinics after implementation (UCA), 245 who received DIAMOND care after implementation (DCA), and 308 who received usual care in comparison clinics (UC). Patients who received DIAMOND care after implementation reported more collaborative care depression services than the 3 comparison groups (10.9 vs 6.4-6.7, on a scale of 0 to 14, where higher numbers indicate more services; $P < .001$) and more satisfaction with their care (4.0 vs 3.4 on a scale 1 to 5, in which higher scores indicate higher satisfaction; $P \leq .001$). Depression remission rates, however, were not significantly different among the 4 groups (36.4% DCA vs 35.8% UCB, 35.0% UCA, 33.9% UC; $P = .94$).

CONCLUSIONS Despite the incentive of a supporting payment change and intensive training and support for clinics volunteering to participate, no difference in depression outcomes was documented. Specific unmeasured actions present in trials but not present in these clinics may be critical for successful outcome improvement.

Ann Fam Med 2015;13:412-420. doi: 10.1370/afm.1842.

INTRODUCTION

Despite 25 years of extensive efforts to improve care quality and extensive research efforts to identify the best methods for widespread implementation of evidence-based practices, we still have limited knowledge about how to facilitate such spread. One of the most striking examples of a gap between evidence and practice in need of such spread is in primary care for depression. At least 79 randomized controlled trials of the collaborative care model for depression have found improved patient outcomes, and there is evidence that the model is cost-effective and even cost-saving.¹⁻⁸ As a result, one might expect widespread implementation of collaborative care, but even care systems and clinics that participated in randomized trials have not usually continued the approach.^{9,10}

A key barrier to implementation of collaborative care is lack of reimbursement for the model's components. The model also requires major changes in traditional primary care and mental health practice, because it is built on a multidisciplinary team approach with a primary care physi-

cian, care manager, and consulting psychiatrist. When a multistakeholder quality improvement collaborative in Minnesota, the Institute for Clinical Systems Improvement (ICSI), developed a statewide initiative to overcome these implementation barriers by providing a new payment approach and the training, tools, and facilitation of practice change, it provided an important opportunity to learn about the potential for large scale uptake and spread.¹¹ The initiative was called DIAMOND (Depression Improvement Across Minnesota—Offering a New Direction).

This study was developed to answer the questions of local and national stakeholders, especially regarding initiative effects on patient outcomes. A staggered implementation (stepped-wedge) design served the needs of both the initiative for manageable implementation and the evaluation for a comparison group.

METHODS

Intervention

The DIAMOND initiative steering committee adapted the version of the collaborative care model tested in the largest randomized trial—the IMPACT (Improving Mood: Promoting Access to Collaborative Treatment) study.^{6,12} This model included 7 components: (1) consistent use of the 9-item Patient Health Questionnaire (PHQ-9) depression scale for monitoring depression severity^{13,14}; (2) systematic patient follow-up tracking and monitoring; (3) treatment intensification for patients not improving; (4) relapse prevention planning for patients achieving remission; (5) on-site care manager for educating, monitoring, and coordinating care; (6) scheduled weekly caseload review with a consulting psychiatrist; and (7) monthly descriptive data submissions.

Primary care clinics were recruited by ICSI staff from the medical group members of ICSI based on their interest and readiness. Eighty clinics signed up for 1 of 5 sequences of a potential 280 clinics. Additional clinics joined later but others dropped out, so only 75 clinics completed training, implemented the DIAMOND care model, and provided DIAMOND care for at least several years. More than 10,000 patients were enrolled in the initiative from March 2008 through May 2013.

Training was provided to clinic leaders, clinicians, and staff during the 6 months just before implementation, and a version of the care management tracking system used in the IMPACT study was used by 54% of participating clinics.¹⁵ The remainder set up their own electronic medical record-based versions. After implementation, ICSI staff conducted networking conference calls and individual site consultation.^{16,17} Clinics

could tailor their approach as long as they provided the model components described above.

The other major component of the DIAMOND intervention was new financial coverage for this care by all large commercial payers in the state for enrolled adult patients (PHQ-9 scores greater than 9, and specific codes for major depression or dysthymia). Payment amounts and arrangements varied by individual payer contracts with participating clinics, usually in a monthly bundled fee. Unfortunately, one-third of enrolled patients were covered by Medicaid or Medicare, neither of which provided payments for DIAMOND care.

Research Design

The study used a stepped-wedge design, with new patients recruited weekly for 30 months at all sites.¹⁸⁻²² Key outcome measures of care received, depression severity, work productivity, health status, and satisfaction were derived from patient surveys at baseline and 6 months later.

Participants

To standardize patient recruitment across clinics while minimizing impact on care, all but 1 commercial payer sent weekly lists of members recently initiating antidepressant medications in a participating clinic. These lists were supplemented with lists of patients billed for DIAMOND care. Patients were then called to assess study eligibility and willingness to participate. The recruitment process and comparison of participants and nonparticipants showed only small differences between patients in the sampling frame and enrolled study participants.¹⁸ The study was reviewed, approved, and monitored by the relevant institutional review boards.

Patient Survey

This survey assessed the main patient outcomes at enrollment and again 6 months later, using the PHQ-9 measure of depression severity,^{23,24} the Work Productivity and Activity Impairment (WPAI),²⁵ the World Health Organization question about health status,²⁶ a modified version of the Patient Assessment of Chronic Illness Care (PACIC),²⁷ and a question about satisfaction.

PHQ-9 scores (range, 0 to 27, with higher scores indicating greater depression severity) measured baseline and 6-month depression severity. Patients were classified as improved if the 6-month PHQ-9 score was 50% lower than baseline, and as in remission if the 6-month PHQ-9 score was less than 5.

The WPAI is a self-report measure of absence from work because of health problems, as well as of productivity impairment while at work. It was administered to patients who reported working for pay at

least part-time. The proportion of expected work time that was affected by health problems was calculated as work time missed plus time at work multiplied by impairment. Functional health status and satisfaction with care were each assessed by single items (range of 1 to 5, with higher scores indicating higher status and greater satisfaction).

The PACIC was revised with its developer so questions reflected collaborative care of depression.²⁷⁻²⁹ The revision reduced the original 20 questions to 14 reporting on the number of care elements received in the previous 6 months (scores range from 0 to 14).

Predictors of Patient Outcomes

Patients were placed into 1 of 4 treatment groups characterized by their clinic's involvement, as well as their own enrollment in DIAMOND care. Payer utilization data showing DIAMOND care charges and patient self-report identified those who had received DIAMOND care. The baseline survey date was used to determine whether patients were participating in the study before or after clinic implementation of the DIAMOND model. The resulting 4 treatment groups consisted of patients who received (1) usual care in DIAMOND care clinics before implementation (UCB), (2) usual care in DIAMOND clinics after implementation (UCA), (3) DIAMOND care after implementation (DCA), and (4) usual care in clinics that planned to but never implemented DIAMOND care (UC).

The effectiveness of DIAMOND care was gauged by comparing patient outcomes among DCA patients relative to the 3 comparison groups.

The timing of each patient's outcomes was calculated as the number of months that had elapsed between the date the first sequence of clinics implemented DIAMOND (March 2008) and the patient's baseline survey date. This secular trend covariate was included in the patient outcome analyses to quantify community-wide changes in the outcomes of depression treatment that were independent of those attributable to the DIAMOND model.

Analyses

Patient-Reported Care Processes and Outcomes

The likelihoods of depression response and remission at 6 months were predicted in generalized linear mixed models (binomial distribution, logit link) from treatment group, secular trend, and, if significant, the treatment group by secular trend interaction. Treatment group and secular trend were fixed effects, patients were nested within primary care clinics, a random clinic intercept was estimated, fixed-effects standard errors were calculated using empirical sandwich estimation, and each observation was weighted by the

multiplicative product of inverse probability weights to correct for the likelihoods of 6-month survey response and of receiving DIAMOND care. Model-estimated likelihoods are presented. Simple effects tests that compared model-predicted, 6-month outcomes of each comparison group with DCA guided interpretation of significant treatment effects.

Similarly, structured analyses assessed whether baseline to 6-month changes in depression symptoms, care processes, productivity loss, functional health status, and satisfaction with care (normal distribution, identity link) differed by treatment. Repeated outcomes were predicted from fixed effects of treatment group, outcome timing (baseline, 6 months), treatment by timing, secular trend, and, if significant, secular trend by timing and by treatment. The model structure differed from the 6-month outcome models in that repeated outcomes were nested within patients, and a random patient intercept was estimated. Simple effects tests by treatment group and by timing aided interpretation.

Patient Survey Retention and Treatment Selection

We assessed response bias in key survey outcomes by comparing patient characteristics of 6-month survey respondents ($n = 1,578$) with those of nonrespondents ($n = 770$). Characteristics that were associated with response status at $P < .20$ were considered for inclusion in a non-parsimonious response propensity model. The logistic model nested patients within clinics, predicted 6-month response from baseline characteristics, estimated a random clinic intercept, and used empirical sandwich estimation to calculate fixed-effects standard errors. Characteristics that predicted response at $P < .10$, did not diminish goodness of fit, and were not collinear with other characteristics were retained. Final model response propensities corresponded to a response likelihood $M = 0.68$ ($M_{\text{respondent}} = 0.70$, $M_{\text{nonrespondent}} = 0.63$). Higher response propensities were associated with higher likelihoods of depression response ($P < .005$) and remission ($P < .002$), lower PHQ-9 scores ($P < .001$) and WPAI scores ($P = .02$), and higher functional health status ($P < .001$), all at 6 months. Because response propensity was related to survey response and to key study outcomes, we attempted to ameliorate potential response bias by calculating stabilized inverse probability weights ($M = 1.04$; $M_{\text{respondent}} = 1.00$, $M_{\text{nonrespondent}} = 1.14$) for the patient outcome analyses.

Although UCB or UC patients could not receive DIAMOND care, patients treated after DIAMOND was implemented in their clinics could be in the UCA or DCA group. We also used a propensity approach to assess DIAMOND care treatment selection. Characteristics associated with DCA group membership were included in a non-parsimonious propensity model

to calculate stabilized inverse probability weights for UCA and DCA patients. The resulting weights were $M_{DCA} = 2.31$ and $M_{UCA} = 0.94$, and set to 1.00 for UCB and UC patients.

All analyses were conducted using SAS 9.2 software (SAS Institute) with P values $< .05$ deemed to be statistically significant.

RESULTS

Participation

Of the 24,065 names received from payers during the 3 years, 5,080 had incorrect contact information and 4,509 could not be reached within the 21-day window required to ensure baseline assessment before major effects from treatment. The participation rate of those contacted and potentially eligible was 36.4%. (Figure 1 displays further details about recruitment.)

Sixty-seven percent of patients completed the 6-month survey with similar retention rates

across the study groups: UCB = 67%; UCA = 65%; DCA = 72%; UC = 68%.

The characteristics of enrolled patients by treatment group are displayed in Table 1. The UCA and DCA patients differed on baseline characteristics, suggesting that patients with more severe depression may have been referred to DIAMOND care. Additional information about the frequency of various comorbidities among the patients in each study group is contained in the Supplemental Table 1 (available at <http://www.annfammed.org/content/13/5/412/suppl/DC1>), although none of these differed significantly among the 4 groups.

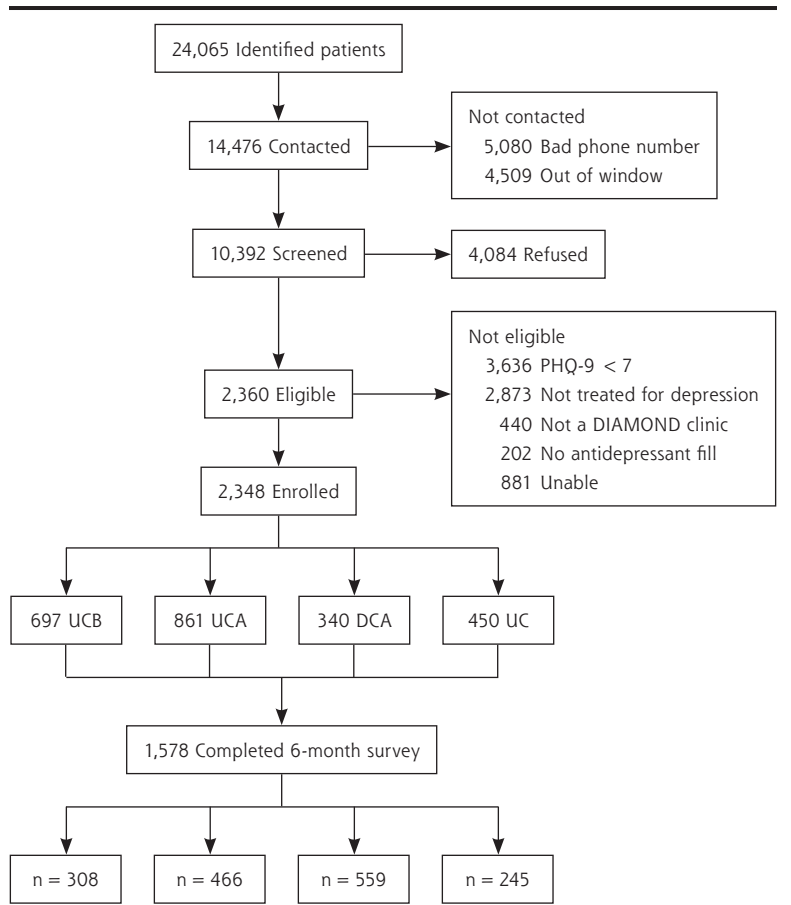
Clinic characteristics were obtained from all 75 clinics (Table 2). One-half of the clinics were from outside the metropolitan Twin Cities region, 39% had fewer than 6 primary care physicians, and 95% were part of medical groups containing at least 3 clinics.

Patient-Reported Depression Care Process and Outcomes

Table 3 displays patient reports of DIAMOND processes and response and remission rates for the 4 patient groups. Although patients in the DCA group reported receiving

67% more of the care processes that were part of the DIAMOND initiative than patients in any comparison group ($P < .001$), their 6-month response and remission rates were not statistically different. Baseline PHQ-9 scores were higher among DCA patients than among comparison groups ($P < .002$ vs UCB; $P < .01$ vs UC; $P < .0001$ vs UCA), but the slightly larger drop in PHQ-9 scores among DCA patients did not reach statistical significance ($P = .63$). Only 4 medical groups had enough patients across all of their clinics to permit statistical assessment of their outcomes by medical group, and none attained statistical significance. Informal nonquantitative review did not suggest that there were any unintended adverse outcomes. Additional information about the frequency by study group with which specific care processes were reported is available in Supplemental Table 2 (<http://www.annfammed.org/content/13/5/412/suppl/DC1>), and the relationship between care processes and outcomes can be

Figure 1. Study patient flow diagram.



DCA = DIAMOND care after implementation; DIAMOND = Depression Improvement Across Minnesota—Offering a New Direction; PHQ-9 = 9-item Patient Health Questionnaire depression scale; UC = usual care in comparison clinics; UCA = usual care in DIAMOND clinic after implementation; UCB = usual care before implementation.

Table 1. Baseline Survey Patient Characteristics by Treatment Group

Characteristic	All N = 2,348	DCA n = 340	UCA n = 861	UCB n = 697	UC n = 450	P Value
Sequence, %						
1	12.2	29.4	20.4	1.6		
2	24.0	32.1	37.4	18.9		
3	13.0	12.3	17.0	16.6		
4	14.5	12.1	15.1	24.3		
5	17.0	12.3	10.1	38.6		
Not sequenced	19.1				100.0	
Clinic unknown	0.2	1.8				
Female, %	72.7	75.9	70.9	73.6	72.2	.31
Age, mean (SD), y	44.4 (14.9)	44.4 (14.5)	44.5 (15.7)	44.2 (14.3)	44.3 (14.6)	.97
Insurance type, %						<.001
Commercial	65.3	67.7	61.3	68.3	66.4	
State program	25.1	25.9	27.1	22.5	24.7	
Medicare	6.9	4.4	10.2	5.2	5.3	
Other/unknown	2.7	1.8	1.4	4.1	3.6	
Education, %						.37
High school or less	31.0	31.2	31.4	31.5	28.9	
Some college	38.2	37.9	35.5	40.2	40.2	
College degree	31.0	30.9	33.1	28.3	30.9	
Marital status, %						<.005
Married	45.8	45.9	42.3	47.9	49.3	
Never married	23.4	23.8	28.0	20.5	18.7	
Other	31.2	33.3	29.8	31.5	32.0	
Hispanic, %	3.8	6.2	3.7	2.2	4.7	<.01
Race, %						<.001
White	88.7	83.8	90.4	91.2	85.3	
Black	5.1	8.2	4.3	3.0	7.6	
Other	6.2	7.1	5.5	5.8	7.1	
Antidepressant medication >4 wk, %	47.0	42.9	50.9	48.5	40.4	<.005
Depression care (medication or therapy) >4 wk, %	58.6	57.1	63.3	58.3	51.3	<.001
Baseline PHQ-9 score, mean (SD)	12.4 (4.5)	13.4 (4.7)	12.0 (4.4)	12.4 (4.5)	12.5 (4.5)	<.001
7-9	32.2	25.3	35.9	30.6	32.7	<.001
10-14	38.9	36.5	40.5	39.9	36.0	
15-19	19.6	25.9	15.8	20.4	21.1	
>20	9.3	12.4	7.8	9.2	10.2	
Previous depression care, %						<.005
None	37.4	44.7	36.7	35.7	36.0	
Once	23.5	21.8	25.1	20.2	26.7	
2 or more	36.0	31.5	33.8	41.5	34.9	
Depression treatment, %						.39
Medication	98.6	97.7	98.6	99.0	98.7	
Counseling	27.5	39.1	25.7	25.1	26.0	<.001
Group therapy	3.8	4.7	3.5	2.9	4.9	.25
Psychiatrist	5.5	6.2	4.8	5.9	5.8	.69
Other	4.0	5.6	4.0	4.2	2.4	.16
Employment, %						<.01
Employed for wages	56.9	60.3	52.4	60.3	58.0	
Self-employed	5.4	1.8	6.3	5.3	6.4	
Out of work	12.1	14.4	13.7	9.4	11.1	
Unable to work	9.3	7.1	9.9	9.3	10.2	
Other	16.1	16.3	17.3	15.6	14.3	
Poverty, %						.34
Below 2 times poverty level	34.9	34.3	37.2	32.8	34.0	
Above 2 times poverty level	65.2	65.7	62.8	67.2	66.1	

DCA = DIAMOND care after; DIAMOND = Depression Improvement Across Minnesota—Offering a New Direction; PHQ-9 = 9-item Patient Health Questionnaire depression scale; UC = usual care; UCA = usual care after; UCB = usual care before.

Table 2. Clinic Characteristics (n = 75)

Variable	No.	Percent or Mean No. (SD)	Range	Variable	No.	Percent or Mean No. (SD)	Range
Clinic				Sites in medical group		15.6 (11.8)	1-48
Location				1-2	4	5.3	
Metropolitan Twin Cities	38	50.7		3-5	16	21.3	
Nonmetropolitan	37	49.3		6-10	4	5.3	
Ownership				>10	50	66.7	
Health system	51	68.0		Patients' insurance			
Health plan	2	2.7		Commercial		52.4 (18.9)	4-80
Physicians	21	28.0		0%-10%	2	2.7	
No. of primary care physicians for adults		8.6 (7.8)	1-39	11%-25%	6	8.0	
1-2	8	10.7		>25%	64	85.3	
3-5	21	28.0		Medicare		23.8 (10.6)	10-38
6-10	33	44.0		0%-10%	12	16.0	
>10	13	17.3		11%-25%	35	46.7	
Any in medical group				>25%	25	33.3	
Psychiatrists	37	49.3		Medicaid		12.0 (11.3)	3-55
Mental health therapists	38	50.7		0%-10%	48	64.0	
No. of NPs/PAs for adults		2.1 (1.9)	0-8	11%-25%	20	26.7	
0	16	21.3		>25%	4	5.3	
1-2	32	42.7		Uninsured		3.9 (3.1)	0-10
>2	27	36.0		0%-10%	72	96.0	

NP = nurse practitioner; PA = physician assistant.

Table 3. Model-Predicted Depression Care Process and Outcome Measures, Adjusted for Secular Trend and Weighted by Survey Response and Treatment Selection Likelihoods

Treatment Model	No.	Care Process 6 mo	Response 6 mo (%)	Remission 6 mo (%)	PHQ-9	
					Baseline	6 mo
Usual care before	466					
Mean		6.5 ^a	46.1	35.8	12.2 ^b	7.8 ^c
SE		0.2	2.7	2.6	0.2	0.3
Cohen's <i>f</i> ^{2,d}		0.088				
Usual care after	559					
Mean		6.4 ^a	46.3	35.0	12.7	7.7 ^c
SE		0.2	2.7	2.6	0.2	0.3
Cohen's <i>f</i> ^{2,d}		0.127				
DIAMOND care after	245					
Mean		10.9	46.7	36.4	13.2	8.0 ^c
SE		0.3	4.4	4.2	0.3	0.4
Usual care	308					
Mean		6.7 ^a	46.4	33.9	12.3 ^e	7.8 ^c
SE		0.3	3.1	3.0	0.2	0.4
Cohen's <i>f</i> ^{2,d}		0.058				
<i>P</i> value ^f		<.001	.99	.94	<.06	.92

DIAMOND = Depression Improvement Across Minnesota—Offering a New Direction; PHQ-9 = 9-item Patient Health Questionnaire depression scale; SE = standard error.

^a *P* < .001 relative to DIAMOND care after treatment group.^b *P* < .01.^c *P* < .001 change from baseline to 6 months.^d Cohen's *f*² for comparison group relative to DIAMOND care after.^e *P* < .05.^f *P* values for treatment group effect at denoted measurement point.

found in Supplemental Table 3 (<http://www.annfammed.org/content/13/5/412/suppl/DC1>).

Patient-Reported Work Productivity, Health Status, and Satisfaction

Among the 59% of patients who reported working outside the home at the 6-month survey, the mean productivity loss of 27.6% (combined absenteeism and presenteeism) was less than at baseline (37.7%, *P* < .001), but there was no difference in rates of decline across treatment groups (Table 4). There was also no difference in health status at either time among the 4 treatment groups, although among each, health status improved from baseline (*M* = 2.99) to 6 months (*M* = 3.13) by about 5% (*P* = < .001). Finally, care satisfaction ratings by DCA patients were significantly higher than the comparison groups at both baseline and at 6 months (*P* < .001),

and satisfaction ratings increased more among DIAMOND care patients than comparison groups ($P < .05$).

DISCUSSION

This unique large-scale initiative to spread the collaborative care DIAMOND model for depression appears to have improved patient satisfaction, but it had little impact on other patient outcomes. Despite good evidence of implementation of practice systems important to collaborative care, and despite enrolled patients reporting receiving more desired care processes, patients receiving DIAMOND care had neither better depression outcomes nor better improvement in work productivity or health status.

These results were surprising and disappointing to the participants in the DIAMOND initiative, especially because measurements by the initiative from clinic data submissions had shown 6-month response and remission rates of 40% and 30% among all patients enrolled, and 66% and 48% among those 61% that could be measured again at 6 months, rates that are almost as good as those seen in intervention groups in the scientific trials. In the absence of any type of comparison group, however, initiative leaders had no way

of knowing that usual care among these clinics was as good as it was.³⁰

Despite our measures of system implementation and care provision, we are unable to definitively identify the reason the groups did not differ. One likely factor is the high quality of usual depression care in Minnesota, making it harder to further improve. Health Employer Data Information Set scores for all 3 of the main health plans in Minnesota for depression medication management show each in the top decile for both measures.³¹ Moreover, the DIAMOND clinics were all members of ICSI and had gained experience in quality improvement methods. In addition, the up-to-4-week delay in obtaining study baseline PHQ-9 scores may have reduced baseline scores and therefore the potential for improvement.

As previously reported, our measures of implementation suggest that most practice systems needed to provide collaborative care for depression were in place in these clinics at 1 year and were sustained at 2 years (although there was considerable variation).³² It is possible that some particularly crucial components of the intervention were either inadequately measured by our measures or were diluted by other less important components. We suspect that a particularly important

component is the extent to which treatment-to-target and treatment intensification by changing medications or adding psychotherapy occurs for patients who are not improving. Treatment optimization is dependent on how consistently such cases are followed, monitored, and highlighted during systematic case reviews with the consulting psychiatrist, how often those reviews result in recommendations for treatment changes, and how often these recommendations are then fully carried out. Our measurement systems were not specific enough to provide such information.

One large medical group had few patients in our study sample because its main insurer declined participation. This group separately analyzed their data and found both more medication changes and better depression outcomes among DIAMOND care patients than in nonrecipients.^{33,34} This group was one of the few with internal psychiatric

Table 4. Model-Predicted Secondary Outcomes, Adjusted for Secular Trend and Weighted by Survey Response and Treatment Selection Likelihoods

Treatment Model	No. ^a	Productivity Loss		Health Status		Satisfaction	
		Baseline	6 mo	Baseline	6 mo	Baseline	6 mo
Usual care before							
Mean	296	37.0	24.5 ^b	3.04	3.16 ^c	3.41 ^d	3.44 ^d
SE		1.4	1.6	0.04	0.05	0.05	0.05
Cohen's f^2						0.002	0.018
Usual care after							
Mean	311	36.7	26.9 ^b	3.00	3.10 ^e	3.45 ^d	3.37 ^d
SE		1.5	1.6	0.04	0.05	0.05	0.06
Cohen's f^2						0.001	0.001
DIAMOND care after							
Mean	135	39.8	31.0	2.97	3.13	3.74	3.95 ^e
SE		2.5	3.8	0.07	0.08	0.06	0.07
Usual care							
Mean	191	37.7	26.9 ^b	2.92	3.12 ^b	3.21 ^d	3.35 ^d
SE		1.8	2.1	0.05	0.06	0.06	0.07
Cohen's f^2						0.008	0.018
P value ^f		.70	.40	.27	.86	<.001	<.001

DIAMOND = Depression Improvement Across Minnesota—Offering a New Direction; SE = standard error.

^a For productivity loss analysis; numbers for other outcomes are the same as for primary outcomes.

^b $P < .001$ change from baseline to 6 months.

^c $P < .01$ change from baseline to 6 months.

^d $P < .001$ relative to DIAMOND care after.

^e $P < .05$ change from baseline to 6 months.

^f P values are for treatment group effect at denoted measurement point.

leaders very actively involved in the initiative committees, so they may have ensured more active treatment intensification. It is also possible that other unadjusted biases between patient groups confounded differences, but the consistency in results across the 3 different comparison groups makes such bias less likely.

At least 10 systematic reviews and/or meta-analyses of the 79 randomized controlled trials of collaborative care for depression have been published.^{1,3,4,8,35-40} They have almost all concluded that the model produces better outcomes, but the following reports have tried to identify the individual components of the model that were most important: staff assistance with case management and mental health specialist involvement³; revision of professional roles and provision of a care manager who delivers psychotherapy (not part of this initiative)³⁷; patient education and self-management, symptom monitoring, decision support for treatment adherence, patient registries, and mental health supervision of care managers⁸; professional background and method of supervision of care managers¹; systematic follow-up and service restructuring³⁵; and ensuring adequate doses of antidepressants.³³ Although most of these strategies were recommended to participating clinics, we lack detailed measures of the extent to which they were implemented.

This study shows the difficulties of widespread implementation of evidence-based practices that require major changes in roles and extensive financial and leadership support. In the clinical trials that provided the evidence for collaborative care, support for these major changes was usually provided by the research team, so those attempting to replicate trial results may have to replicate those supports. That this initiative was unable to produce results similar to clinical trials with volunteer clinics, despite the guidance and support of a skilled and experienced quality improvement collaborative and payment changes, highlights how difficult it really is to implement evidence-based care effectively.

To read or post commentaries in response to this article, see it online at <http://www.annfammed.org/content/13/5/412>.

Key words: quality; depression; primary care; health care delivery; collaborative care

Submitted March 19, 2015; submitted, revised, June 30, 2015; accepted July 7, 2015.

Supplementary materials: Available at <http://www.AnnFamMed.org/content/13/5/412/suppl/DC1>.

References

1. Archer J, Bower P, Gilbody S, et al. Collaborative care for depression and anxiety problems. *Cochrane Database Syst Rev*. 2012;10:CD006525.
2. Jacob V, Chattopadhyay SK, Sipe TA, Thota AB, Byard GJ, Chapman DP; Community Preventive Services Task Force. Economics of collaborative care for management of depressive disorders: a community guide systematic review. *Am J Prev Med*. 2012;42(5):539-549.
3. O'Connor EA, Whitlock EP, Beil TL, Gaynes BN. Screening for depression in adult patients in primary care settings: a systematic evidence review. *Ann Intern Med*. 2009;151(11):793-803.
4. Thota AB, Sipe TA, Byard GJ, et al; Community Preventive Services Task Force. Collaborative care to improve the management of depressive disorders: a community guide systematic review and meta-analysis. *Am J Prev Med*. 2012;42(5):525-538.
5. U.S. Preventive Services Task Force. Screening for depression in adults: U.S. preventive services task force recommendation statement. *Ann Intern Med*. 2009;151(11):784-792.
6. Unutzer J, Katon WJ, Fan MY, et al. Long-term cost effects of collaborative care for late-life depression. *Am J Manag Care*. 2008;14(2):95-100.
7. van Steenbergen-Weijenburg KM, van der Feltz-Cornelis CM, Horn EK, et al. Cost-effectiveness of collaborative care for the treatment of major depressive disorder in primary care. A systematic review. *BMC Health Serv Res*. 2010;10:19.
8. Williams JW Jr, Gerrity M, Holsinger T, Dobscha S, Gaynes B, Dietrich A. Systematic review of multifaceted interventions to improve depression care. *Gen Hosp Psychiatry*. 2007;29(2):91-116.
9. Katon W, Unutzer J. Collaborative care models for depression: time to move from evidence to practice. *Arch Intern Med*. 2006;166(21):2304-2306.
10. Smith JL, Williams JW Jr, Owen RR, Rubenstein LV, Chaney E. Developing a national dissemination plan for collaborative care for depression: QUERI Series. *Implement Sci*. 2008;3:59.
11. Solberg LI, Glasgow RE, Unutzer J, et al. Partnership Research: A Practical Trial Design for Evaluation of a Natural Experiment to Improve Depression Care. *Med Care*. 2010;48(7):576-582.
12. Unutzer J, Katon W, Callahan CM, et al; IMPACT Investigators. Improving Mood-Promoting Access to Collaborative Treatment. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA*. 2002;288(22):2836-2845.
13. Kroenke K, Spitzer RL. The PHQ-9: A new depression and diagnostic severity measure. *Psychiatr Ann*. 2002;32(9):509-515.
14. Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J Affect Disord*. 2004;81(1):61-66.
15. Unutzer J, Choi Y, Cook IA, Oishi S. A web-based data management system to improve care for depression in a multicenter clinical trial. *Psychiatr Serv*. 2002;53(6):671-673, 678.
16. Farley DO, Haims MC, Keyser DJ, Olmsted SS, Curry SV, Sorbero M. *Regional Health Quality Improvement Coalitions: Lessons Across the Life Cycle*. Santa Monica, CA: RAND Health;2003.
17. Mosser G. Clinical process improvement: engage first, measure later. *Qual Manag Health Care*. 1996;4(4):11-20.
18. Crain AL, Solberg LI, Unutzer J, et al. Designing and implementing research on a statewide quality improvement initiative: the DIAMOND study and initiative. *Med Care*. 2013;51(9):e58-e66.
19. Glasgow RE, Magid DJ, Beck A, Ritzwoller D, Estabrooks PA. Practical clinical trials for translating research to practice: design and measurement recommendations. *Med Care*. 2005;43(6):551-557.
20. Speroff T, O'Connor GT. Study designs for PDSA quality improvement research. *Qual Manag Health Care*. 2004;13(1):17-32.
21. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol*. 2006;6:54.
22. Handley MA, Schillinger D, Shiboski S. Quasi-experimental designs in practice-based research settings: design and implementation considerations. *J Am Board Fam Med*. 2011;24(5):589-596.

23. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613.
24. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med Care*. 2004;42(12):1194-1201.
25. Beck A, Crain AL, Solberg LI, et al. Severity of depression and magnitude of productivity loss. *Ann Fam Med*. 2011;9(4):305-311.
26. DeSalvo KB, Bloser N, Reynolds K, He J, Muntner P. Mortality prediction with a single general self-rated health question. A meta-analysis. *J Gen Intern Med*. 2006;21(3):267-275.
27. Glasgow RE, Wagner EH, Schaefer J, Mahoney LD, Reid RJ, Greene SM. Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). *Med Care*. 2005;43(5):436-444.
28. Glasgow RE, Whitesides H, Nelson CC, King DK. Use of the Patient Assessment of Chronic Illness Care (PACIC) with diabetic patients: relationship to patient characteristics, receipt of care, and self-management. *Diabetes Care*. 2005;28(11):2655-2661.
29. Schmittiel J, Mosen DM, Glasgow RE, Hibbard J, Remmers C, Bellows J. Patient Assessment of Chronic Illness Care (PACIC) and improved patient-centered outcomes for chronic conditions. *J Gen Intern Med*. 2008;23(1):77-80.
30. The DIAMOND Program: treatment for patients with depression in primary care. Bloomington, MN: ICSI; 2014.
31. NCQA's health insurance plan rankings 2012-2013 – private plan details. Washington, DC: National Center for Quality Assurance; 2012.
32. Solberg LI, Crain AL, Jaeckels N, et al. The DIAMOND initiative: implementing collaborative care for depression in 75 primary care clinics. *Implement Sci*. 2013;8(1):135.
33. Dejesus RS, Angstman KB, Cha SS, Williams MD. Antidepressant medication use among patients with depression: comparison between usual care and collaborative care using care managers. *Clin Pract Epidemiol Ment Health*. 2013;9:84-87.
34. Shippee ND, Shah ND, Angstman KB, et al. Impact of collaborative care for depression on clinical, functional, and work outcomes: a practice-based evaluation. *J Ambul Care Manage*. 2013;36(1):13-23.
35. Vergouwen AC, Bakker A, Katon WJ, Verheij TJ, Koerselman F. Improving adherence to antidepressants: a systematic review of interventions. *J Clin Psychiatry*. 2003;64(12):1415-1420.
36. Neumeyer-Gromen A, Lampert T, Stark K, Kallischnigg G. Disease management programs for depression: a systematic review and meta-analysis of randomized controlled trials. *Med Care*. 2004;42(12):1211-1221.
37. Craven MA, Bland R. Better practices in collaborative mental health care: an analysis of the evidence base. *Can J Psychiatry*. 2006;51(6) (Suppl 1):7S-72S.
38. Gunn J, Diggins J, Hegarty K, Blashki G. A systematic review of complex system interventions designed to increase recovery from depression in primary care. *BMC Health Serv Res*. 2006;6:88.
39. Christensen H, Griffiths KM, Gulliver A, Clack D, Kljakovic M, Wells L. Models in the delivery of depression care: a systematic review of randomised and controlled intervention trials. *BMC Fam Pract*. 2008;9:25.
40. Woltmann E, Grogan-Kaylor A, Perron B, Georges H, Kilbourne AM, Bauer MS. Comparative effectiveness of collaborative chronic care models for mental health conditions across primary, specialty, and behavioral health care settings: systematic review and meta-analysis. *Am J Psychiatry*. 2012;169(8):790-804.