

Understanding Team-based Quality Improvement for Depression in Primary Care

Lisa V. Rubenstein, Louise E. Parker, Lisa S. Meredith, Andrea Altschuler, Emmeline dePillis, John Hernandez, and Nancy P. Gordon

Objective. To assess the impacts of the characteristics of quality improvement (QI) teams and their environments on team success in designing and implementing high quality, enduring depression care improvement programs in primary care (PC) practices.

Study Setting/Data Sources. Two nonprofit managed care organizations sponsored five QI teams tasked with improving care for depression in large PC practices. Data on characteristics of the teams and their environments is from observer process notes, national expert ratings, administrative data, and interviews.

Study Design. Comparative formative evaluation of the quality and duration of implementation of the depression improvement programs developed by Central Teams (CTs) emphasizing expert design and Local Teams (LTs) emphasizing participatory local clinician design, and of the effects of additional team and environmental factors on each type of team. Both types of teams depended upon local clinicians for implementation.

Principal Findings. The CT intervention program designs were more evidence-based than those of LTs. Expert team leadership, support from local practice management, and support from local mental health specialists strongly influenced the development of successful team programs. The CTs and LTs were equally successful when these conditions could be met, but CTs were more successful than LTs in less supportive environments.

Conclusions. The LT approach to QI for depression requires high local support and expertise from primary care and mental health clinicians. The CT approach is more likely to succeed than the LT approach when local practice conditions are not optimal.

Key Words. Quality improvement, depression, primary care, health care provider teams

Primary care (PC) practices have been widely criticized for poor performance in caring for depression and are searching for effective methods for improving depression care quality. Team-based quality improvement (QI) methods such

as Continuous Quality Improvement (CQI) are an attractive option (Berwick 1989; Wagner et al. 2001). These methods place a high value on participation in QI design and implementation by a health care organization's or practice's own health care professionals and staff. Social psychology and organizational theory predicts that individuals will perform best when they have control over decisions that affect their work. On the other hand, clinical QI for complex problems like depression requires substantial technical knowledge and resources. In practice, health care organizations often turn to experts to design QI interventions. There is little available information on the effects of expert involvement. This paper assesses the degree to which local clinician participation in QI intervention design versus delegation of design to regional experts affects the quality and longevity of QI intervention programs for depression in primary care. The paper also evaluates what additional characteristics of QI teams and their organizational environments predict implementation of a high-quality, enduring depression intervention program. Understanding how best to structure QI teams to promote their success has the potential to improve the currently variable results of team-based QI (Shortell, Bennett, and Byck 1998).

Learning how QI teams function requires qualitative data collection (Shortell et al. 1995). We collected and analyzed qualitative data on QI teams in action. Our data are drawn from a comparative formative evaluation of the QI process in the Mental Health Awareness Project (MHAP), an ongoing National Institute of Mental Health (NIMH) funded randomized quasi-experiment comparing usual care for depression to care following team-based QI. In addition to the data reported here, the study collected quantitative data to evaluate the QI effects on depression care

This research was supported by a grant to Dr. Rubenstein from the National Institute of Mental Health (U01-MH50732), the John T. MacArthur Foundation, and the VA Health Services Research and Development (HSR&D) Center for the Study of Healthcare Provider Behavior, Sepulveda, CA.

Address correspondence to Lisa V. Rubenstein, M.D., M.S.P.H., Senior Natural Scientist, RAND Health Program, 1700 Main Street, Santa Monica, CA 90401-3297. Dr. Rubenstein is Professor of Medicine, VA Greater Los Angeles Healthcare System, and University of California, Los Angeles. Louise E. Parker, Ph.D., and Lisa S. Meredith, Ph.D., are Behavioral Scientists with RAND Health Program in Santa Monica. Andrea Altschuler, Ph.D., is a Program Manager with Kaiser Permanente, Division of Research, Oakland, CA. Emmeline dePillis, Ph.D., is an Assistant Professor of Management, University of Hawaii at Hilo, HI. John Hernandez, Ph.D., is a Senior Director, E-Business Products and Services, Quintiles Transnational, Inc., San Francisco, CA. Nancy P. Gordon, Sc.D., is a Research Scientist, Kaiser Permanente, Division of Research, Oakland, CA.

process and outcomes in randomly selected QI versus usual care control practices.

The study's QI process used two alternative approaches to structuring the QI teams, termed the Central Team (CT) approach and the Local Team (LT) approach. The LT approach to designing a depression QI intervention program emphasized meetings in the local primary care practice involving a multidisciplinary team and a QI facilitator, with some expert input. The CT approach emphasized delegation of planning to regional experts, with some input from local primary care practice clinical leaders. Both types of teams depended upon local clinicians to implement the programs. The two approaches thus varied the degree of centralization of design decisions. Centralization indicates the locus of authority and decision making in an organization as more or less consolidated (Zaltman, Duncan, and Holbek 1984). We assumed that a purely centralized top-down intervention would be unsuccessful (Zaltman, Duncan, and Holbek 1984), and instead designed the CT to enable local practitioners to have input into, and potentially veto power over, the expert-proposed designs and materials. Likewise, we thought it would be unnecessarily restrictive to depend solely on LT members' existing knowledge about depression QI, and provided LTs and CTs with organizational priorities for depression improvement, a four-hour depression improvement seminar, and sets of relevant literature and materials. Our analyses are based on the experiences of five QI teams (two CTs and three LTs) charged with implementing depression guidelines in large primary care practices (Depression Guideline Panel 1993).

We structured our qualitative analyses around hypotheses. We hypothesized that the number of mental health specialists and primary care nonphysician staff available to primary care practices would affect organizational priority setting as well as CT and LT intervention program quality and duration. Access to these resources is critical for implementing the collaborative care model shown in previous studies to substantially improve depression outcomes (Katon et al. 1995). We hypothesized that practices with sicker patients might experience more competing demands and be less likely to devote resources to improving depression care (Rost et al. 2000; Schulberg et al. 1996). We expected that the CTs would produce more evidence-based intervention programs, both because of the CTs' greater expertise and the lesser influence from diverse local clinicians. Finally, we expected the two types of teams to charge the project for their time and for intervention program costs at about the same levels.

We also considered the potential effects of variations in teams and their environments on the likelihood of achieving a successful QI team process. We hypothesized that team and environmental factors we identified in the QI, organizational, and psychology literatures would influence the quality and duration of implementation of QI intervention programs. The LTs, because they involve greater local clinician participation, might be more successful than CTs in gaining buy-in from local practices and hence in designing enduring programs (Gatchel 1980; Miller and Monge 1986; Tannenbaum et al. 1974).

METHODS

We conducted the MHAP study in, and partnered with, two not-for-profit staff model managed care organizations—one government-funded (Veteran's Administration [VA]) and the other community-based (Kaiser Permanente [KP] Medical Care Program of Northern California)—to implement and evaluate the QI programs. We asked primary care practices within each organization to participate in the study if they were not currently participating in any depression QI projects, and if we could match the practice on size, academic affiliation, patient population, and urban versus suburban location with at least two other practices. No practices refused to participate. Kaiser's practices were based in geographically separate medical facilities. Each of three VA study practices was within a single academically affiliated VA facility. This VA randomly assigned patients at enrollment to one of three primary care practices, or firms, without overlapping staff. We enrolled a total of six experimental group practices and three usual care group practices, and randomized them within blocks to the LT, CT, or Usual Care condition. At KP, we designed a single CT led by KP regional experts who were not affiliated with either of two practices randomly assigned to the CT condition. This resulted in three LTs (two at KP and one at VA) and two CTs (one at KP and one at VA). At VA, we deviated from random assignment in order to assign the CT condition to the team to which the regional primary care expert in guideline implementation and depression belonged, and then randomly assigned the remaining two practices.

We designed the CTs and LTs to operate in partnership with their organizations and with the research project, consistent with an action research framework (Elden and Chisholm 1993). Before we approached individual primary care practices to request their participation in the study, we asked each organization to set strategic priorities for depression QI

using an expert panel method (Rubenstein et al. 1995). Following the expert panel, study organizations chose a primary care clinician (PCC) and mental health specialist (MHS) LT leader from each participating LT practice. Organizations also chose a PCC and an MHS CT leader with regionally recognized expertise in depression and quality improvement for each CT. In addition, they chose a PCC and a MHS liaison to the CT from each of the two KP-CT practices. The CT practice liaisons consulted on design but were primarily responsible for implementation. Finally, study organizations collaborated with researchers to identify experienced QI facilitators for LTs and experienced project managers for CTs. A recognized CQI expert provided additional training to QI facilitators. Team leaders reviewed the LT or CT manual defining team membership and activities and chose the remaining team members. These pre-QI team activities took two months.

The planning process for each CT and LT began with a 30-minute introduction by the MHAP research team liaison, an organizational psychologist, followed by a 4-hour interactive training session led by the study PI during the following month. The training session focused on the evidence base for depression care improvement and provider behavior change. The session also introduced teams to the Depression Improvement Resource Center (DIRC), which contained a selection of 158 categorized and labeled materials, and a kit containing key selections. Thereafter, contacts between QI teams and researchers occurred through the MHAP study liaison, who attended about one-third of team meetings and made follow-up telephone calls to teams about every three months for a year. She provided teams with problem-solving support but not design guidance.

MHAP provided funds for LTs and CTs to meet for a total time of 16 hours over two to three months to plan the QI intervention programs. At the end of this period, the QI manuals instructed each team to submit a written proposal to primary care practice leadership and to the appropriate local approval and funding mechanisms. At KP this occurred through a pre-existing Clinical Innovations Program. At the VA, teams submitted proposals to the facility's Performance Improvement Council. Proposals were to include any additional planning team time needed for implementation.

After approval, enrolled practices implemented the QI interventions. During intervention implementation researchers provided quarterly summaries of each PC practice's depression-related care to each team, based on administrative data from their organization.

Evaluation

We collected computer administrative data on patient age and sex distributions for each CT and LT practice and telephone interview and administrative data on staffing. Site-based observers attended all QI planning meetings, generated written meeting transcripts, and conducted semistructured interviews with team leaders at 12 and 24 months after the beginning of planning. The CT and LT leaders reviewed and corrected QI team intervention summaries prepared by the observers after the first year of implementation.

We tested relevance of literature review findings on team, organizational environment, and program factors thought to be linked to QI success by comparing the list we generated from the literature to an independent list generated by team leaders and members during a participant panel held 18 months after planning began. Following the conference, we asked the 18 QI team leaders or members who had attended to rank each factor on a scale of 1 (most important) to 15 (least important). Twelve of the 18 (66 percent) returned their surveys.

Six national mental health experts not involved in the MHAP study rated the quality of QI team intervention proposals. Raters included two psychiatrists, two psychologists, and two internists with expertise both in depression and in improving care for depression in primary care. These individuals rated each QI program strategy along three dimensions on 9-point Likert scales (1 = low, 9 = high) as follows: (1) existence of evidence to support the strategy (evidence exists versus no evidence exists), (2) the nature of the evidence, and (3) likelihood the expert would include the strategy in a state-of-art depression program. The experts also rated the QI intervention programs, as a whole, along five dimensions: (1) evidence basis, (2) difficulty of implementation, (3) potential for long-term sustainability, (4) expected magnitude of improvement, and (5) overall quality. Half of the raters assessed programs as proposed and half of them rated programs as implemented.

Our analyses are qualitative (Geller 1989; Jick 1990). We created summed indices for selected variables (Miles and Huberman 1994) as numeric representations of qualitative data. Strategies are the specific intervention components, such as provider feedback, proposed by teams. Strategy Ratings (SR) range from a low of 1 to a high of 81, and represent the product of the expert ratings of a strategy's basis in evidence and of the likelihood that the expert would have included that strategy in designing an intervention, each rated on a 9-point Likert scale. The Evidence Based Index (EBI) is the sum of the SRs. The Overall Program Quality Index (OPQI) is the sum of the expert

ratings of each team's intervention program considered as a whole across all five dimensions, both as planned and as implemented, and ranges from 10 to 90.

We used a second set of indices to summarize each team's structure and environmental support. We rated the CT and LT approaches *as designed*, and then each QI team's structure and environment *as implemented*. The three qualitative researchers reviewed all transcripts and independently rated the QI teams (see Appendices 1 and 2, available from the authors) for those factors based on transcript data, and developed consensus ratings. Five of the authors of this paper rated the remaining factors, also by consensus. We rated all factors as Low, Moderate, Moderately High, or High. We carried out cross-case analysis using predictor–outcome matrices (Miles and Huberman 1994) by treating the quality of the team's intervention program (measured by the EBI and the OPQI) and its longevity as outcomes and team structure and environment as predictors.

RESULTS

Table 1 summarizes the characteristics of the patients and staff at the practices participating in the intervention. All six practices were large, with 30 to 40 primary care clinicians in each. Veteran's Administration practices had more mental health specialists and fewer support staff per primary care clinician than did KP practices, and included older and sicker patients.

In the priority-setting process, high-level management at each organization indicated the importance of increasing provider and patient knowledge about depression. Both organizations also endorsed increased access to depression evaluation and care. The VA, but not KP leadership, endorsed screening for depression in primary care and referring all detected patients to mental health specialists. Only KP endorsed improved management of depression in primary care. QI teams reacted positively to receiving, and indicated they would aim for, the priorities endorsed by management, even when they disagreed with them. For example, VA QI teams disagreed with management's goal of referring all depressed patients to mental health, but preferred knowing about this issue up front.

The QI team process followed the protocols outlined in the manual with a few exceptions. One team (VA-CT) developed its proposal in less than 10 hours of meeting time, as opposed to the recommended 16. Only the VA teams conducted pilot test cycles and used the resulting information to improve their intervention programs. All three LTs and both CTs requested additional

Table 1: Characteristics of Intervention Practices

Characteristics	Central Teams		Local Teams			
	KP-CT		VA-CT	KP-LT #1	KP-LT #2	VA-LT
	Practice #1	Practice #2				
PRIMARY CARE						
Active patients (%)*						
Male	45	45	96	44	44	95
Age 65+	20	12	36	7	23	37
3 or more chronic illnesses‡	8	7	32	7	12	31
Primary care providers (PCPs, FTEEs)§						
MDs	34	31	3	19	41	3
Residents	0	0	26	0	0	27
Nurse practitioners/physician assistants	5	5	4	3	4	4
MENTAL HEALTH						
Mental health clinicians (FTEEs)§						
Psychiatrists based in primary care clinic	0	0	1	0	0	.5
Psychologists based in primary care clinic	0	0	0	1	0	0
Psychiatrists, psychiatry residents, and psychologists	21	13	16**	7	25	16**
Licensed clinical social workers	0	0	.33	0	0	.33
TOTAL MENTAL HEALTH AND SUPPORT STAFF PER PCP[∇]						
PC support staff (a) per PCP (b)	1.8	1.8	.9	1.0	.9	.7
Total MH providers (c) per PCPs (b)	.8	.5	1.7**	.4	.4	1.7**

*Of those visiting at least once during calendar year 1997 to primary care clinicians.

**Mental health specialists are those of the clinic as a whole, and are not divided by individual firm.

§Full-time employee equivalents (FTEEs) supported by the clinic for direct outpatient clinical care or supervision of outpatient clinical care

∇Ratio of support staff and of mental health providers to primary care providers, defined as follows: (a) Licensed Vocational Nurses, nurses, assistants, and clerks; (b) MDs, residents stationed in clinic (about 4/session at VA only), and NPs/PAs; (c) psychiatrists, psychiatry residents, psychologists, and Licensed Clinical Social Workers for the clinic as a whole, not divided by teams.

‡Data based on self-reports from patient screening surveys.

resources or used materials from the DIRC. All teams both orally presented, and submitted in writing, their proposed interventions to their organizations' quality improvement bodies within the specified time period.

Table 2 focuses on QI team depression improvement interventions. The table shows the individual strategies included in each team's depression improvement intervention program, the expert rating for each strategy (SR), EBI summarizing the SRs, and the OPQI reflecting expert ratings of each program considered as a whole. The table also indicates which strategies were planned, planned and implemented, or subsequently implemented though not planned initially. Overall, team intervention strategies addressed most key elements of the collaborative care model (Von Korff et al. 1997), including patient and provider education, detection, assessment, and case management. Two teams planned, but did not implement, strategies for collaboration with mental health specialists, the remaining key element of collaborative care. CTs within each organization had higher ratios of implemented to planned strategies (CT mean 89 percent versus LT mean 68 percent) and the higher EBI ratings. The LTs had both the highest and the lowest OPQI scores. The VA-CTs and -LTs had lower EBI scores than their KP counterparts.

In terms of costs, KP-LT #2 designed the least ambitious intervention program and was least costly. The KP-CT team members charged \$7,018; KP-LT #1 charged \$6,147; and KP-LT #2 charged \$1,859, all for team member time. Charges to the KP Clinical Innovations Program for program implementation show a similar pattern. The KP-CT applied for and received \$101,762 (to cover two primary care practices) and KP-LT #1 applied for and received \$64,741. The KP-LT #2 did not apply for implementation resources. At the VA, charges to the grant were for CT project management (\$7,730), CT computer support (\$1,760), LT intervention support (\$1,760), and LT computer support (\$200). Support from the VA Performance Improvement Council was in-kind, and not measured. Overall, the 10 VA and KP leaders willing to estimate their time indicated spending between 60 and 882 hours on the project over the two years of planning and implementation. For KP leaders, these estimates indicate that more than three-fourths of the time spent was not charged.

Results of our QI team participant panel agreed substantially with the results of our literature review in terms of the factors that might most affect the success of the QI process. Panelists generated 64 percent (16 of 25) of factors we had identified from the literature (Appendix 3, available from the authors) and ranked multidisciplinary team membership, support from mental health specialty, and team leader interest in depression or flexible problem solving during implementation (a tie) as the three most important factors. We termed the factors identified by the panel or literature review as positive factors for QI.

Table 2: Central Team (CT) and Local Team (LT) Practice Quality Improvement Intervention Components, Implementation of Components, and Expert Ratings of Evidence Base

	Planned and Implemented Components						Strategy Rating (Product of Evidence and Likelihood of Use, 1-81, higher score is better)
	Central Teams			Local Teams			
	KP-CT	VA-CT	KP-LT #1	KP-LT #2	VA-LT		
<i>QI Team Intervention Plan Components Planned or Implemented by Any CT or LT</i>							
Provider education and decision support							
• <i>Classes</i>							
a. Orientation led by a local clinician	X			X	X		6.0
b. Multiple sessions led by a local clinician		X					11.9
c. Multiple sessions led by an outside expert			X	X			13.4
• <i>Written materials</i>			X	X			18.4
• <i>Face-to-face detailing</i>		X		X			26.1
• <i>Retrospective provider feedback</i>		X		X			10.5
Patient education							
• <i>Patient education classes</i>							
a. Single session overview				X			3.3
b. Eight session class	X			X			17.7
• <i>Written materials</i>	X	X	X	X	X		17.1
Detection							
• <i>Nurses flag charts for suspected depression</i>			P	P	P		11.2
• <i>Annual screening policy with computer reminders</i>							
a. Disseminated but not actively enforced					X		4.6
b. Monitoring and enforcement activities carried out		X					5.3
Assessment							
• <i>Provider depression assessment worksheet</i>	X	X					7.6
• <i>Provider reminders</i>							

Table 3 shows how QI teams varied in the extent to which they manifested positive support factors for QI. Positive factors could vary across teams by design, because of poor adherence to the design, or because of natural differences. Factors are listed as high, moderately high, moderate, or low, based on study records or process notes. The following factors are not listed in Table 3 because they were rated as uniformly "High" both as designed and as implemented: clinician majority on teams, organizational mandate to participate, and multiple stakeholders have a voice in planning. Two factors (flexible problem-solvers during implementation and leadership by respected local peers) were omitted because we did not collect sufficient data about them.

As shown in Table 3, the LT design included more positive factors, but LTs varied more than CTs in the extent to which they manifested characteristics we had tried to engender through our designs for team structure, protocols, and materials. For CT's as implemented, the only substantial deviations from expectations were lower support from clinical practice leadership in one KP-CT practice, lower involvement of pharmacists and higher use of CQI methods.

Intervention planning took an average of four and a half months. Full intervention implementation occurred an average of six and a half months after the end of planning. All intervention programs except KP-LT #2's were active more than six months after full intervention implementation, but only the KP-CT and KP-LT #1 programs were active more than one year after full implementation. The two VA team interventions depended heavily on the computer medical record, which displayed screening test results and was the basis for summary data for feedback. One year after full implementation the software for the computer record was changed, making the system inaccessible to the teams. In the context of simultaneous facility integration, this was enough to end the active phase of the interventions at the VA.

Team leaders and members often participated in depression improvement activities after the end of the study at intervention sites. The KP-LT #2 leader and pharmacist participated in a subsequent depression medication case-management intervention at their facility. Frustrated with lack of coordination of mental health consultations, the VA-LT worked with psychiatry to initiate a new, prompt psychiatric consultation system that persisted after the full intervention ended. At KP-LT #1, the intervention case manager became the behavioral health specialist required by a newly adopted KP primary care practice redesign model. In one of the KP-CTs the practice continued the intervention case manager's position after the innovations funding stopped, but also hired another behavioral health specialist. Ultimately, the two positions came into conflict and the case manager left.

Table 3 also shows the relationship between each team's positive factors for *QI* and that team's outcomes in terms of developing a high-quality, long-lasting program. The KP-LT #1 had the highest score for positive *QI* factors, followed by KP-CT, VA-CT, VA-LT, and KP-LT #2, in that order. Outcome scores for program quality and longevity followed the same order. Aggregating the scores from Table 3 across teams (not shown on the table), CTs scored about the same as LTs on positive factors (1.54 CT versus 1.62 LT) and better on outcomes (2.33 CT versus 1.00 LT). The two CTs implemented their interventions in practices with equivalent positive environmental factors (1.33 KP-CT versus 1.33 VA-CT). Two of the three LT's had more environmental support than CTs and one (VA-LT) had less. In the two PC practices with the lowest environmental support factors for *QI* (KP-CT Practice A and VA-LT) the CT but not the LT produced an enduring program. KP teams scored better than VA teams on positive factors (1.67 KP versus 1.48 VA) and outcomes (2.11 KP versus 1.17 VA).

Two teams experienced outstanding success in developing a high-quality program that remained active for more than a year after full implementation (KP-CT and KP-LT #1), one team had moderate success (VA-CT), and two teams had low success, with KP-LT #2 having the lowest ratings for program quality and duration of implementation. No team that did not have high ratings on two of three of the *QI* team leadership measures (interest in depression, content expertise, and participation) succeeded. No LT depression improvement program succeeded in a practice without high ratings on either support from mental health specialty or clinical practice leadership.

DISCUSSION

Over the last decade, health care organizations in the United States have commonly used team-based *QI* methods to address specific quality of care problems. This paper explored the advantages and disadvantages of a more centralized approach to team-based *QI* that invests in technical expertise (CT) compared to a more local approach that invests in local clinical team interaction and input (LT). The paper also identified theoretically based positive factors for *QI* and evaluated their impact on *QI* team success in developing and implementing high-quality depression improvement programs. Success in program design and implementation is a sine qua non for success in impacting clinical care, yet evaluation of the process of *QI* is rare in the abundant literature on team-based *QI*.

MEAN SCORE ACROSS ENVIRONMENTAL FACTORS	2.00	1.33	1.33	3.00	2.00	1.67	0.00
OVERALL MEAN SCORE ACROSS TEAM AND ENVIRONMENTAL FACTORS	1.75	1.64	1.54	2.45	2.00	1.27	1.00
Team outcomes							
Plan evidence based index (EBI)		High	High		High	Mod	Low
Overall plan quality index (OPQI)		High	Low		High	Low	Low
Longevity of implementation		High	Mod high		High	Low	Mod high
MEAN OUTCOME SCORE		3.00	1.67		3.00	0.33	0.67

§Factors that differed by design between LTs and CTs are in bold in the “As Designed” column; *SR = study records; PN = process notes; P = Plans; ND = No Data

†Table as displayed omits the following factors that were rated as “High” both as designed and as implemented: clinician majority on teams, organizational mandate to participate, and multiple stakeholders have a voice in planning; **Mental Health Specialty support was moderately high in Facility A and high in Facility B.

+Overall resource support in KP-CT combines support from the organization and support from the individual facilities.

We had hypothesized that LTs, with their greater local autonomy and involvement, would do better than CTs. We found instead that overall, CTs were more likely to be successful in program design and implementation than LTs, except when the local environment was highly supportive and team leaders had high levels of expertise and commitment (e.g., KP-LT #1). For example, one of the KP-CT practices had low support from clinical practice leadership, and the VA-CT had only moderate support from mental health specialty; these teams did much better than the two LTs with similar profiles (e.g., KP-LT #2 and VA-LT).

As a group, CTs produced higher quality intervention program designs than LTs. Despite LT access to the same depression improvement evidence base as CTs, study results support our prior hypothesis that CTs would produce more evidence-based program designs than LTs. Even the VA-CT, which was prohibited by organizational priorities from carrying out the program strategy rated highest by the experts (nurse case management), developed a highly evidence-based program by including a variety of other moderately ranked strategies. The CTs by design had expert, interested leaders from primary care and mental health who came from a wider, regional pool than did LT leaders. CT leaders also derived their authority from high-level central management. In relatively unsupportive local environments greater local input may have worked against high-quality design by favoring strategies requiring less effort and less organizational change. These less supportive practices may also attract and employ fewer individuals with interest and expertise in depression, further disadvantaging a local approach.

Research shows that to improve depression outcomes in primary care the intervention program must incorporate key elements of the collaborative care model, including provider and patient education, collaboration with mental health specialists, and case management (Simon et al. 2000; Von Korff et al. 1997). Simpler knowledge-based solutions such as providing written guidelines, reminders (Goldberg and Gater 1996) and education (Gerrity et al. 1999) are not sufficient. Yet the collaborative care model is an innovation that has not diffused, or spread, across primary care practices through natural channels. We have used team-based QI as a method for assisting the diffusion of this model (Rogers 1995; Zaltman, Duncan, and Holbek 1984). Other studies have tested other, related diffusion methods. The Partners in Care (PIC) study (Rubenstein et al. 1999; Wells et al. 2000) used a centralized diffusion method that relied on researcher experts to design the new care model, and then trained CT-like QI teams to implement the model across multiple practices within their organizations. This method succeeded in improving care and outcomes.

A decentralized Continuous Quality Improvement (CQI) approach resembling our LT approach, but without team exposure to a depression care evidence base and tool kit or to organizational priorities, had no measurable effect on patients (Horwitz et al. 1996).

The finding that, at least under some circumstances, more centralized QI approaches perform better than less centralized approaches is surprising. Organizational and psychology theories suggest that, in general, successful uptake of innovations is more likely with participatory diffusion methods, such as CQI, that depend primarily on expected users of the innovation (Sandstrom, De Meuse, and Futrell 1990). The CQI method as usually used is a decentralized approach. On the other hand, CQI experts and innovation diffusion theory predict poor results from CQI for depression due to the level of technical expertise and organizational change required for success (Rogers 1995; Shortell, Bennett, and Byck 1998). Our findings, and those of previous studies of depression QI, suggest that it may be preferable to delegate some authority to experts when an innovation is difficult to design and implement.

We had hypothesized that CTs and LTs would charge about the same amounts for their programs. We thought CTs might design more elaborate programs while LTs might expend more on team process. In fact, LTs were less expensive as a group, but varied based on the programs they designed. The least expensive LTs developed less expensive and less evidence-based programs, while KP-LT #1, with the highest-rated program, spent about as much as the KP-CT.

The VA practices had high competing demands, relatively low access to primary care support staff, and relatively high access to mental health specialists compared to KP. Based on our prior hypotheses, we would therefore expect less organizational commitment in VA than KP to manage depression within primary care, and more emphasis on referral. The results of the priority-setting process support our hypotheses. The VA instructed its teams to refer all depressed patients to a mental health specialty, and provided no support for case management. The VA teams therefore did not include case management strategies in their proposals, resulting in overall lower expert ratings of VA QI programs.

We found that health care professionals who were approached by their organizations to participate in team-based QI were generally enthusiastic and competent participants. Study sites were not volunteers, but rather were invited to participate in the study because they had not spontaneously carried out depression improvement activities. Yet all team leaders volunteered significant time over-and-above their paid hours, and all produced and implemented

reasonable programs within the constraints of their environments. Four of the five QI intervention programs ended because the larger organization made a major change in a key depression care element (i.e., the computer system or case management) without involving the QI teams, rather than because the teams themselves lost momentum.

The small number of QI teams studied limits the generalizability of our conclusions, and the naturalistic model used leaves room for many unexpected events to affect success. We could only test some combinations of team type, leadership, and environmental factors. We did not evaluate organizational culture or QI using assessment instruments (Shortell et al. 2000). We evaluated charges, rather than costs. We rated factors supporting team success by review of data and consensus, but did not carry out formal content analysis. The study's qualitative researchers are carrying out formal qualitative analysis; the purpose of this paper, however, was to link what we observed to existing concepts in quality improvement and organizational psychology. Finally, we do not know which, if any, QI team programs will result in better clinical outcomes for patients.

In summary, we conclude that QI teams led by primary care and mental health clinicians and supported by example materials, literature, and previously identified organizational priorities can design evidence-based models for depression improvement in primary care practices. The CTs and LTs work equally well in supportive local practice environments, but CTs work better than LTs when local practice environment support is low. CTs are likely to produce more evidence-based designs than LTs, and to implement more of their planned intervention components. Primary care practices without access to regional expertise but with high interest or concern about depression may wish to make use of the LT approach. To avoid wasting QI resources, however, the LT approach should be used only when a team leader with interest or expertise in depression and time to participate is available, and when support from local practice leadership and/or support from mental health specialists is high.

ACKNOWLEDGEMENTS

The authors acknowledge the invaluable contributions of Carole Oken, M.A., Karen Swanson, M.P.H., Christy Klein, B.A., Mary Abdun-Nur, M.A., Maria Orlando, Ph.D., Jim Chiesa, Ph.D., and Mary Vaiana, Ph.D. The VA Center for the Study of Healthcare Provider Behavior, a VA Health Services Research and

Development (HSR & D) Center of Excellence, substantially supported the analyses reported here. The MHAP study is part of the Quality Improvement for Depression (QID) project, a four-study collaboration funded by the National Institute for Mental Health and the MacArthur Foundation (Kenneth Wells, M.D., M.P.H., Lisa Rubenstein, M.D., M.S.P.H., Kathryn Rost, Ph.D., and Daniel Ford, M.D., M.P.H., Principal Investigators). The study could not have taken place without the support of Kaiser Permanente in Northern California and of the VA Greater Los Angeles Healthcare System, Sepulveda Ambulatory Care Center, Sepulveda CA.

REFERENCES

- Berwick, D. M. 1989. "Continuous Improvement as an Ideal in Health Care." *New England Journal of Medicine* 320 (1): 53.
- Depression Guideline Panel. 1993. *Clinical Practice Guideline, Depression in Primary Care, Detection and Diagnosis*, vol. 1., AHCPR Publication no. 93-0550. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service Agency for Health Care Policy and Research.
- Elden, M., and R. F. Chisholm. 1993. "Emerging Varieties of Action Research: Introduction to the Special Issues." *Human Relations* 46 (2): 121-42.
- Gatchel, R. J. 1980. "Perceived Control: A Review and Evaluation of Therapeutic Implications." In *Advances in Environmental Psychology*, edited by A. Bauman and J. E. Singer, pp. 1-22. Hillsdale, NJ: Lawrence Erlbaum.
- Geller, E. S. 1989. "Applied Behavioral Analysis and Social Marketing: An Integration for Environmental Preservation." *Journal of Social Issues* 45 (1): 17-3.
- Gerrity, M. S., S. A. Cole, A. J. Dietrich, and J. E. Barrett. 1999. "Improving the Recognition and Management of Depression: Is There a Role for Physician Education?" *Journal of Family Practice* 48 (12): 949-57.
- Goldberg, D., and R. Gater. 1996. "Implications of the World Health Organization Study of Mental Illness in General Health Care for Training Primary Care Staff [See Comments]." *British Journal of General Practice* 46 (409): 483-5.
- Horwitz, C. R., H. I. Goldberg, D. P. Martin, E. H. Wagner, S. D. Fihn, D. B. Christensen, and A. D. Cheadle. 1996. "Conducting a Randomized Controlled Trial of CQI and Academic Detailing to Implement Clinical Guidelines." *Joint Commission Journal on Quality Improvement* 22 (11): 734-50.
- Janis, I. L. 1989. *Crucial Decisions: Leadership in Policymaking and Crisis Management*. New York: Free Press.
- Jick, T. D. 1990. *Mixing Qualitative and Quantitative Methods: Triangulation in Action*. Beverly Hills: Sage.
- Katon, W., M. Von Korff, E. Lin, E. Walker, G. E. Simon, T. Bush, P. Robinson, and J. Russo. 1995. "Collaborative Management to Achieve Treatment Guidelines. Impact on Depression in Primary Care." *Journal of the American Medical Association* 273 (13): 1026-31.

- Miles, M. B., and A. M. Huberman. 1994. *Qualitative Data Analysis*. Thousand Oaks, CA: Sage.
- Miller, K. I., and P. R. Monge. 1986. "Participation, Satisfaction, and Productivity: A Meta-Analytic Review." *Academy of Management Journal* 29 (4): 727-53.
- Rogers, E. M. 1995. *Diffusion of Innovations*, 4th ed. New York: Free Press.
- Rost, K., P. Nutting, J. Smith, J. C. Coyne, L. Cooper-Patrick, and L. V. Rubenstein. 2000. "The Role of Competing Demands in the Treatment Provided Primary Care Patients with Major Depression." *Archives of Family Medicine* 9(2): 150-4.
- Rubenstein, L. V., A. R. Fink, E. M. Yano, B. Simon, B. Chernof, and A. S. Robbins. 1995. "Increasing the Impact of Quality Improvement on Health: An Expert Panel Method for Setting Institutional Priorities." *Joint Commission Journal on Quality Improvement* 21 (8): 420-32.
- Rubenstein, L. V., M. Jackson-Triche, J. Unutzer, J. Miranda, K. Minnium, M. L. Pearson, and K. B. Wells. 1999. "Evidence-Based Care for Depression in Managed Primary Care Practices." *Health Affairs (Millwood)* 18 (5): 89-105.
- Sandstrom, E., K. P. De Meuse, and D. Futrell. 1990. "Work Terms: Applications and Effectiveness." *American Psychologist* 45 (2): 120-33.
- Schulberg, H. C., M. R. Block, M. J. Madonia, C. P. Scott, E. Rodriguez, S. D. Imber, J. Perel, J. Lave, P. R. Houck, and J. L. Coulehan. 1996. "Treating Major Depression in Primary Care Practice: Eight-Month Clinical Outcomes." *Archives of General Psychiatry* 53 (10): 913-9.
- Shortell, S. M., C. L. Bennett, and G. R. Byck. 1998. "Assessing the Impact of Continuous Quality Improvement on Clinical Practice: What It Will Take to Accelerate Progress [See Comments]." *Milbank Quarterly* 6 (4): 593-624.
- Shortell, S. M., R. H. Jones, A. W. Rademaker, R. R. Gillies, D. S. Dranove, E. F. Hughes, P. P. Budetti, K. S. Reynolds, and C. F. Huang. 2000. "Assessing the Impact of Total Quality Management and Organizational Culture on Multiple Outcomes of Care for Coronary Artery Bypass Graft Surgery Patients." *Medical Care* 38 (2): 207-17.
- Shortell, S. M., J. L. O'Brien, J. M. Carman, R. W. Foster, E. F. Hughes, H. Boerstler, and E. J. O'Connor. 1995. "Assessing the Impact of Continuous Quality Improvement/Total Quality Management: Concept Versus Implementation." *Health Services Research* 30 (2): 377-401.
- Simon, G. E., M. Von Korff, C. Rutter, and E. Wagner. 2000. "Randomized Trial of Monitoring, Feedback, and Management of Care by Telephone to Improve Treatment of Depression in Primary Care." *British Medical Journal* 320 (7234): 550-4.
- Tannenbaum, A. S., B. Kavicic, M. Rosner, M. Vianello, and G. Weiser. 1974. *Hierarchy in Organizations*. San Francisco: Jossey-Bass.
- Von Korff, M., J. Gruman, J. Schaefer, S. J. Curry, and E. H. Wagner. 1997. "Collaborative Management of Chronic Illness." *Annals of Internal Medicine* 127 (12): 1097-102.

- Wagner, E. H., R. E. Glasgow, C. Davis, A. E. Bonomi, L. Provost, D. McCulloch, P. Carver, and C. Sixta. 2001. "Quality Improvement in Chronic Illness Care: A Collaborative Approach." *Joint Commission Journal on Quality Improvement* 27 (2): 63–80.
- Wells, K. B., C. Sherbourne, M. Schoenbaum, N. Duan, L. Meredith, J. Unutzer, J. Miranda, M. F. Carney, and L. V. Rubenstein. 2000. "Impact of Disseminating Quality Improvement Programs for Depression in Managed Primary Care: A Randomized Controlled Trial." *Journal of the American Medical Association* 283 (2): 212–20.
- Zaltman, G., R. Duncan, and J. Holbek. 1984. *Innovations and Organizations*. Malabar, FL: Robert E. Krieger.