ecp
EFFECTIVE CLINICAL PRACTICE

May/June 2000
Volume 3 Number 3

# Failure of a Continuous Quality Improvement Intervention To Increase the Delivery of Preventive Services

## *A Randomized Trial*

ORIGINAL ARTICLE

**LEIF I. SOLBERG, MD**
*HealthPartners Research Foundation*
*Minneapolis, Minn*

**THOMAS E. KOTTKE, MD**
*Mayo Clinic and Foundation*
*Rochester, Minn*

**MILO L. BREKKE, PhD**
*Brekke Associates, Inc.*
*Minneapolis, Minn*

**SANNE MAGNAN, MD, PhD**
*Blue Cross and Blue Shield of Minnesota*
*St. Paul, Minn*

**GESTUR DAVIDSON, PhD**
*Minneapolis, Minn*

**CAROLYN A. CALOMENI, RN, MPA**
*HealthPartners Research Foundation*
*Minneapolis, Minn*

*United HealthCare*
*Minneapolis, Minn*

**SHIRLEY A. CONN, RN, MSN**
*Blue Cross and Blue Shield of Minnesota*
*St. Paul, Minn*

*Minnesota Department of Health*
*Minneapolis, Minn*

**GAIL M. AMUNDSON, MD**
*HealthPartners*
*Minneapolis, Minn*

**ANDREW F. NELSON, MPH**
*HealthPartners Research Foundation*
*Minneapolis, Minn*

***Eff Clin Pract. 2000;3:105-115.***

**See related commentary on pages 153–155.**

*Edited by Lisa M. Schwartz, MD, MS*

**CONTEXT.** Although there has been enormous interest in continuous quality improvement (CQI) as a measure to improve health care, this enthusiasm is based largely on its apparent success in business rather than formal evaluations in health care.

**OBJECTIVE.** To determine whether a managed care organization can increase delivery of eight clinical preventive services by using CQI.

**DESIGN.** Primary care clinics were randomly assigned to improve delivery of preventive services with CQI (intervention group) or to provide usual care (control group).

**INTERVENTION.** Through leadership support, training, consulting, and networking, each intervention clinic was assisted to use CQI multidisciplinary teams to develop and implement systems for delivery of preventive services.

**SETTING.** 44 primary care clinics in greater Minneapolis–St. Paul.

**PATIENTS.** Patients 19 years of age and older completed surveys at baseline ($n$ = 6830) and at follow-up ($n$ = 6431). Medical chart audits were completed on 4777 patients at baseline and 4546 patients at follow-up.

**MAIN OUTCOME MEASURES.** The proportion of patients who were up-to-date (according to chart audit) and the proportion of patients who were offered a service if not up-to-date (according to patient report) for 8 preventive services.

**RESULTS.** Compared with the control group, based on the proportion of patients who were up-to-date, use of only one preventive service—pneumococcal vaccine—increased significantly in the intervention group (17.2% absolute increase from baseline to follow-up compared with a 0.3% absolute increase in the control group, $P$ = 0.003). Similarly, based on patient report of being offered a service if not up-to-date, delivery of only one preventive service—cholesterol testing—significantly increased in the intervention group compared with the control group (4.6% increase vs. 0.4% absolute decrease in the control group; $P$ = 0.006).

**CONCLUSION.** In this trial, CQI methods did not result in clinically important increases in preventive service delivery rates.

*The abstract of this paper is available at ecp.acponline.org.*

© 2000 American College of Physicians–American Society of Internal Medicine

Since publication of Berwick's seminal article on "continuous improvement as an ideal in health care,"[1] there has been enormous interest in using continuous quality improvement (CQI) to address issues of quality and cost in health care delivery. Although enthusiasm was undoubtedly spurred by earlier testimonials of success in other industries, CQI has been especially attractive in medicine because it uses the tools of science to improve health care.

However, initial enthusiasm for CQI has waned in the past few years because of the long lead time, large effort, and marginal benefits associated with many CQI projects. The response to these criticisms has consisted primarily of opinions, anecdotes, and case studies. In addition, the Institute for Healthcare Improvement launched the Breakthrough Series and has promoted new concepts in the hope of spurring larger and faster gains.[2, 3] To date, few formal studies have evaluated whether CQI is actually effective in improving the quality of health care.

We thought that HMOs would be a realistic choice to sponsor practice improvement initiatives in their contracted primary care practices. Therefore, we performed a randomized, controlled trial that was cosponsored by two of the three dominant competing HMOs in the Minneapolis–St. Paul area.[4, 5] The trial was called IMPROVE (IMproving PRevention through Organization, Vision, and Empowerment). It tested the hypothesis that a managed care organization, working through an iterative process of change called *continuous quality improvement*, can stimulate its private primary care clinics to develop better delivery systems for eight clinical preventive services: blood pressure monitoring, Pap smear, cholesterol monitoring, tobacco use cessation, breast examination, mammography, influenza vaccine, and pneumococcal vaccine. We chose these services on the basis of the goals of Healthy People 2000.[6]

We previously reported various aspects of the baseline data and showed that this CQI intervention increased the number of office systems to provide preventive services.[4, 7–13] This study examines whether the systems created by CQI intervention led to significantly higher rates of offering and performing preventive services.

## Methods

### Overview

**Figure 1** shows overall trial design and patient participation. Most of the methods used in this trial have been reported previously.[4, 7–11] An institutional review board approved the study protocol, and all patients gave written informed consent for record audits.

### Clinic Recruitment and Randomization

Eligible primary care clinics had to be part of a medical group that contracted with one or both of the sponsoring HMOs and were located within 50 miles of the Twin Cities.[4] Seventy-one medical groups with 164 separate clinic sites met these criteria and were jointly recruited by leaders of the two HMOs. Forty-four clinics agreed to participate in the study.

After completion of baseline clinic surveys, clinics were matched on the basis of three measures potentially predictive of ability to improve the delivery of preventive services[7]: size (the number of primary care clinicians treating adults), readiness to systematically improve preventive services, and existing quality-based organizational culture.

Each clinic's readiness to improve was measured by combining the standardized scores from three provider attitude survey measures, with an equally weighted standardized measure of the prevention system components in place.[8, 10] Organizational culture was measured by using a weighted sum of scores on seven Baldrige criteria, which were obtained from a survey of all personnel at each clinic.[11] Each of the three variables was categorized as high or low, depending on whether each clinic's score was above or below the median for that variable. This allowed us to group all 44 clinics into one of eight unique cells. A panel of three judges then used similarity of the three measures to create pairs of clinics. Randomization was performed within the matched pairs.

### Patient Selection and Recruitment

Use of preventive services was assessed in different samples of patients at baseline and at follow-up (**Figure 1**). We recruited a cross-section of patients who visited participating clinics during 5-week intervals before the intervention (baseline) and near its end (follow-up). At each clinic, approximately 175 patients who were selected from appointment lists for randomly identified days during the study interval were taken consecutively to fill each of the five age–sex cells. The size of each age–sex cell was defined by the power analysis described below.

All patients were asked to complete a survey soon after clinic visits. At baseline, the patient survey was completed by 3451 patients from control clinics (85% response rate) and 3379 patients from intervention clinics (86% response rate). At follow-up, 3136 patients from control clinics (82%) and 3295 patients from intervention clinics (83%) completed the patient survey.

All patients who completed the survey were asked to give written informed consent for medical record audits. Overall, 5517 patients at baseline and 5047 patients at follow-up consented to a record audit. Of these 10,564

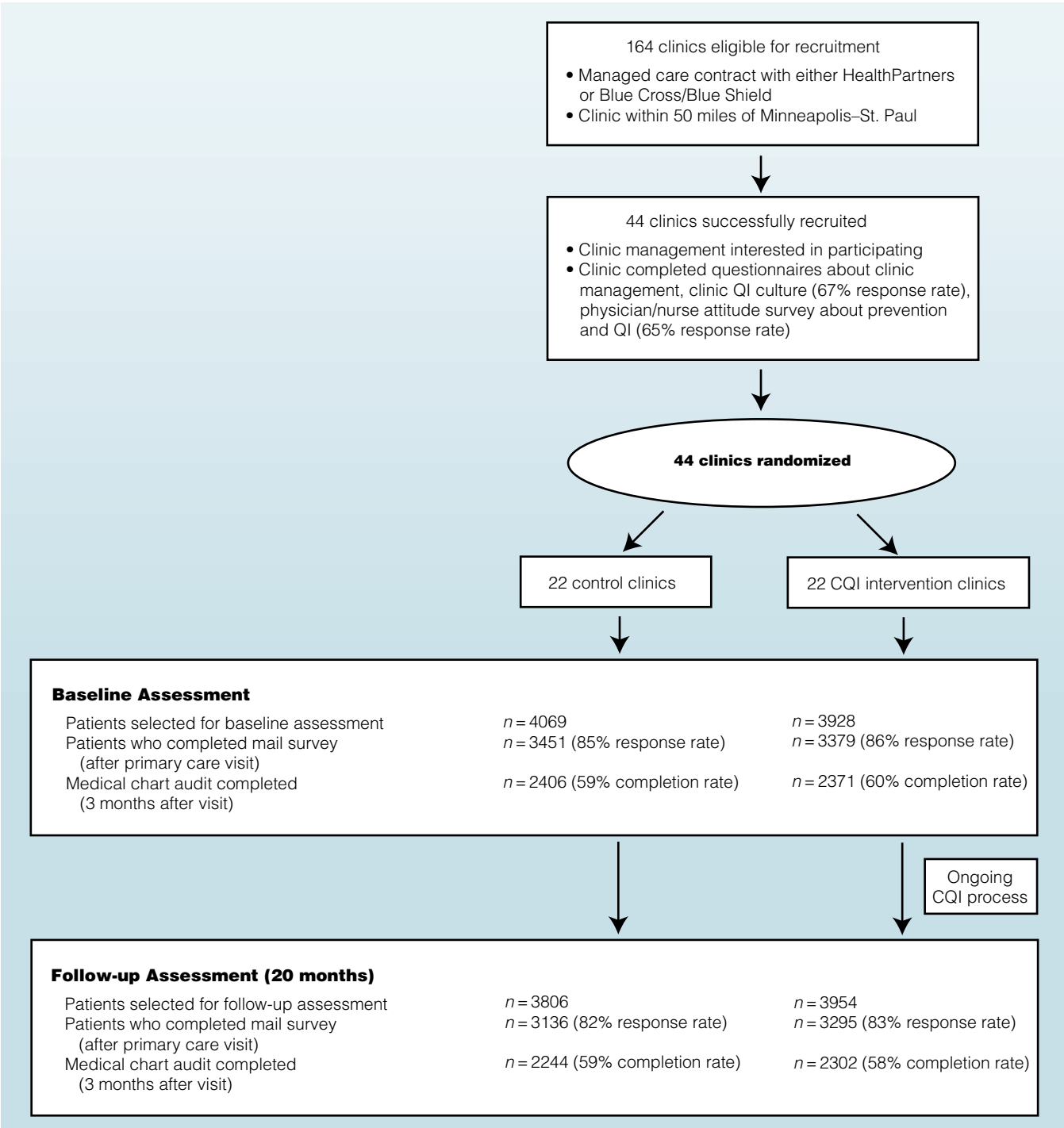


**FIGURE 1. Flow diagram of the randomized, controlled trial. Use of preventive services was assessed on different samples of patients at baseline and at follow-up. At each clinic, 175 patients who had primary care visits during a 5-week period were selected for evaluation. Patients reported whether they had received eight preventive services, and medical charts were reviewed to see whether patients received these services in the 3 months after their primary care visits.** CQI = continuous quality improvement; QI = quality improvement.

records, 9323 were obtainable and were reviewed (4777 at baseline and 4546 at follow-up). Thus, medical chart audit was completed on approximately 60% of selected patients at baseline and follow-up. Chart audit rates were similar for patients from control and intervention clinics.

## Intervention

The intervention was conducted for 22 months (September 1994 through June 1996) and had three major components: leadership involvement, training, and networking and consultation.

Before participating, management at each clinic agreed 1) to establish clinic guidelines for the study's preventive services and 2) to identify and support a management sponsor and an internal multidisciplinary improvement team for those services. IMPROVE physicians contacted the sponsor every 6 months, and clinic teams were trained to keep their leadership informed.

IMPROVE staff provided the leader and facilitator of each clinic's team with an initial 6-hour conference overview of CQI methods and prevention systems. This was followed by six 4-hour workshops over 6 months for experiential training in an improvement process to build prevention systems. Workshops were reinforced with modular manuals that contained many examples of the necessary process improvement steps and prevention systems.

During and after training, IMPROVE staff provided consultation by telephone (every 6 weeks) and by visit (every 3 months). They also organized and conducted bimonthly breakfast meetings and semiannual workshops and sent quarterly newsletters to clinic staffs.

The CQI change process that was taught and encouraged used a seven-step cyclical process that was commonly in use when this trial began. It recommended that each clinic's team 1) identify the problem, 2) collect enough data to understand their current care process, 3) analyze that data to understand the problems and their root causes in the current process, 4) develop alternate solutions, 5) generate and develop specific recommendations, 6) implement the recommendations, and 7) evaluate and improve the process in an iterative cycle through the above steps as needed.

Rather than asking the teams to come up with their own ideas for solutions, we provided them with evidence for a systems approach to improving preventive services (with an emphasis on integrating prevention into each patient visit). They were also provided with considerable information and examples of about 10 processes that are commonly missing in most practices but that we consider important components in an overall prevention system (**Appendix Table**).[4, 10, 14.]

The principal task for each clinic's team was to decide how these improvement and system concepts and examples might best fit into a system and be applied in their setting. Thus, the IMPROVE intervention provided training and consultation to help clinics undertake a CQI-based change process that would lead to new care systems for providing higher rates of preventive services.

Each contact with clinics in the intervention group was recorded. Evaluation surveys were conducted at each training session, at the end of the intervention, and again in June 1997 after 1 year of no contact with the intervention clinics. The leader, facilitator, and one team member of each clinic team recorded the time they spent working on the project.

## Outcome Measures

To assess the efficacy of the CQI intervention, we used two primary outcome measures: up-to-date (chart audit) and offered if not up-to-date (patient report). *Up-to-date* provides an estimate of the proportion of patients that received each preventive service because chart audits allowed a 3-month interval after each primary care visit, thereby providing a long enough interval after the visit for the service to be done and recorded. Our second outcome measure, *offered if not up-to-date*, helps to separate clinician behavior from patient factors. Some patients may be appropriately offered preventive services but may not actually receive the service because of various patient-related factors. In addition, we believe that use of two outcome measures from different sources protects against the inaccuracies inherent in both these types of measurement.

### *Up-to-Date*

*Up-to-date* is defined as the patients who were up-to-date for each service within 3 months of an index primary care visit on the basis of chart audit. The primary purpose of the chart audit was to assess both documented need for preventive services and the rate of performance of needed services within 3 months of the visit. This measure may underestimate actual rates. To assess any carryover effect on preventive services that were not targeted by the intervention, the chart audit also assessed documentation of tetanus immunization in the past 10 years and measurement of high-density lipoprotein cholesterol in the past 5 years for all persons 19 years of age and older. The chart audit was designed, pretested, and piloted in a demonstration clinic. After training and testing, auditors collected chart data from baseline and follow-up samples during a common time period in 1996. Multiple estimates of interauditor reliability (Cohen's $\kappa$ statistic) at each clinic ranged from 0.96 to 0.998.[14]

### *Offered if Not Up-to-Date*

*Offered if not up-to-date* is defined as the patients who reported being offered each of the eight preventive services at their primary care visit if receipt of the service was not up-to-date prior to that visit. This includes patients who were offered the service and chose not to have it as well as those who chose to receive the service. Data for this outcome measure were obtained from the Patient Recent Visit Survey, which has been described elsewhere.[9, 12, 13] Its primary purpose was to determine each patient's need for the selected preventive services and to

determine whether needed services had been offered during the visit. By asking patients directly about services received, issues that are unlikely to receive consistent chart documentation (e.g., clinician recommendations on tobacco cessation) can be addressed.

### Statistical Analysis

We conducted an intention-to-treat analysis, using all 22 clinics in each group. Three null hypotheses were tested: no improvement in the control group, no improvement in the intervention group, and no difference in the magnitude of improvement between both groups. Calculations for each clinic were based on the number of respondents in the appropriate age–sex cell for each preventive service. The sample sizes—22 clinics in each group and 175 patients per clinic—were determined by the number required to detect a minimum between-group difference of 15% for rates of delivering preventive services with a power of 0.80 and a type I error of 0.05 (two-tailed).

Outcome measures were calculated as a grand mean of the rates at which each service was performed in the intervention and control groups. To account for clustering of patients within a clinic, multilevel logistic regression analysis was used on our patient-level data; regression parameters were estimated by the penalized quasi-likelihood criterion, with second-order terms.[16] Parameter estimates were obtained, and tests of hypotheses were performed using MLn/MLwiN software packages.[17, 18] To assess the potential for confounding by differences in patient populations among clinics and between study groups, all rates were adjusted for the effects of clinics as well as for patient age, sex, health status, socioeconomic status, ethnicity, and purpose for visit. Because these adjusted rates were nearly identical to the unadjusted rates, only the latter are presented in this paper. All *P* values are two-tailed.

Additional analyses were performed for three subgroups of intervention clinics: subgroup A (13 clinics that implemented new prevention systems at least 3 months before follow-up), subgroup B (8 clinics with the most complete prevention systems at follow-up), and subgroup C (3 clinics with the best overall systems). Subgroups A and B were compared with their matched pairs among the control clinics, and subgroup C was compared with the overall results of the control clinics.

## Results

### Baseline Characteristics

**Table 1** shows the percentages of patients who were up-to-date for each preventive service at baseline. None of these percentages (with the exception of blood pressure) differed significantly between the intervention and control clinics, and all of them are higher than national figures. **Table 2** shows that the 22 intervention clinics and the 22 control clinics were identical in terms of randomization variables and other clinic characteristics. No measures differed significantly between groups, specifically measures used to assess attitudes, presence of prevention systems, or previous experience with CQI. **Table 3** shows that the patient populations of the intervention and control clinics were also identical.

**TABLE 1**
**Preventive Services Assessed and Proportion of Patients Up-to-Date at Baseline**

| PREVENTIVE SERVICE | TARGET GROUP (DENOMINATOR) | PROPORTION OF PATIENTS WHO WERE UP-TO-DATE AT BASELINE* | |
|---|---|---|---|
| | | INTERVENTION CLINICS | CONTROL CLINICS |
| Blood pressure measurement (past year) | >19 yr of age | 95.7% | 98.6% |
| Cholesterol check (past 5 yr) | >19 yr of age | 55.3% | 55.2% |
| Tobacco use status (prior visit) | Smokers >19 yr of age | 75.7% | 76.5% |
| Papanicolaou smear (past 2 yr) | Women >19 yr of age | 71.0% | 69.4% |
| Breast examination (past 2 yr) | Women >19 yr of age | 56.6% | 51.5% |
| Mammography (past 2 yr) | Women >49 yr of age | 56.8% | 54.4% |
| Influenza shot (past yr)† | >64 yr of age | 62.1% | 62.5% |
| Pneumococcal vaccine (ever) | >64 yr of age | 30.3% | 28.6% |

*$P > 0.2$ for all comparisons except blood pressure measurement ($P < 0.002$) and breast examination ($P = 0.18$).
†*Because index visits occurred in the spring, influenza shots were not required and were not available for chart audit. The proportion of patients who were up-to-date for influenza shots was obtained by self-report.*

**TABLE 2**

**Baseline Characteristics of the Intervention and Control Clinics***

| CHARACTERISTIC | INTERVENTION CLINICS (*n* = 22) | CONTROL CLINICS (*n* = 22) |
|---|---|---|
| **Randomization strata variables** | | |
| Adult primary care physicians, *n* | 6.7 ± 4.2 | 8.7 ± 6.1 |
| Prevention readiness score† | 49.7 ± 4.6 | 50.2 ± 5.5 |
| Quality culture score‡ | 614.6 ± 80.6 | 618.4 ± 73.0 |
| **Other** | | |
| Total physicians, *n* | 12.0 ± 9.9 | 13.0 ± 9.9 |
| Midlevel practitioners, *n* | 1.5 ± 2.3 | 1.4 ± 1.3 |
| Nurses, *n* | 13.2 ± 12.7 | 14.5 ± 12.0 |
| Other staff, *n* | 34.1 ± 40.1 | 26.8 ± 19.7 |
| Had a CQI team | 50% | 50% |
| Age of clinic, *yr* | 10.1 ± 7.5 | 12.4 ± 10.7 |
| Patients seen per week, *n* | 762 ± 693 | 679 ± 431 |
| Prepaid patients overall | 42.2% | 47.4% |
| Blue Plus | 9.5% | 9.1% |
| HealthPartners | 8.2% | 11.6% |

**Data are presented as the mean ± SD unless otherwise specified as a percentage. P >0.2 for all comparisons. CQI = continuous quality improvement.*
*†Higher scores are associated with greater readiness.*
*‡Weighted summary scores from the Quality Systems Inventory for Primary Care Clinics, a survey for assessing the seven Baldrige quality criteria. Theoretical range is from 167 to 1000. Higher scores are associated with a better clinic climate for quality.*

## Intervention Process

All 22 intervention clinics established improvement teams, and all but 1 participated in all training sessions. Teams from an average of 8 to 12 clinics attended each of the subsequent sessions. All sessions received excellent evaluations. Clinic teams met an average of 24 times for 60 to 90 minutes over the 22 months of the study. More than 570 telephone contacts were made with the IMPROVE consultants (28% of which were initiated by the clinics), and IMPROVE consultants made 86 clinic visits. Team leaders spent an average of 8.6 hours per month on the project, facilitators spent 11.7 hours per month, and team members spent 5.0 hours per month. However, few teams finished even one complete CQI cycle.[7]

Two clinics reported implementing new systems 7 months after the start of the intervention (9%), 5 had done so at 1 year (23%) and 10 at 15 months (45%). At the time of the follow-up survey and audit (20 months), 13 of the intervention clinics (59%) reported implementation. Six clinics had not implemented systems at the end of the intervention.

**TABLE 3**

**Baseline Characteristics of Patients in the Intervention and Control Clinics***

| CHARACTERISTIC | INTERVENTION PATIENTS (*n* = 3379) | CONTROL PATIENTS (*n* = 3451) |
|---|---|---|
| Mean age ± SD, *yr* | 48.4 ± 1.3 | 48.6 ± 2.3 |
| Female | 68.4% | 68.7% |
| Low socioeconomic status† | 16.3% | 17.3% |
| White | 94.7% | 92.4% |
| Patients who rated their general health now as "somewhat worse" or "much worse" than 1 year ago | 15.7% | 16.8% |
| Patients who reported that their last visit was mainly for a checkup or physical examination | 19.6% | 19.1% |

**All data were obtained from the patient survey. P > 0.2 for all comparisons.*
*†No insurance or medical assistance, or household income below the poverty line.*

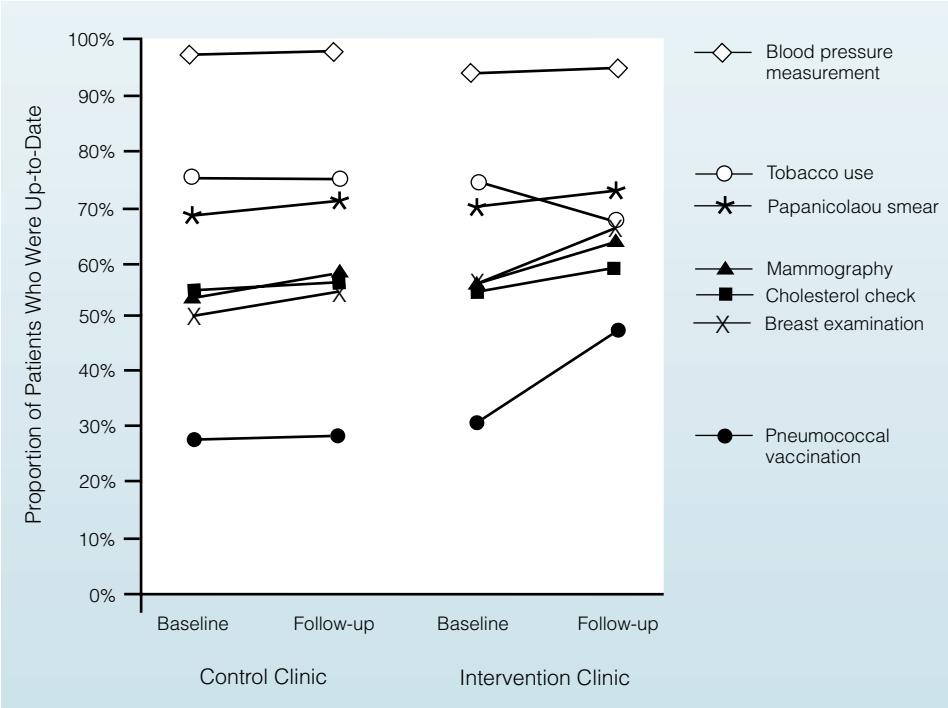


**FIGURE 2. Proportion of patients who were up-to-date for various preventive services.**

Clinicians and nurses were asked to report on the presence and function of each of the 10 defined processes (**Appendix**). This showed that the number of processes involving the identified services doubled at follow-up in the intervention clinics but had not changed in the control clinics.[7] However, only 25 processes out of the potential 80 (10 processes for 8 services) were in place, and the critical process of cueing or reminding the clinician about missing prevention services during a visit was only in place for a few services in a few of the intervention clinics.

A complete, functioning system to integrate preventive services into each office visit would have the following three components: 1) a way for nursing or reception staff to identify which services the visiting patient needed, usually by some combination of chart review before the visit and patient answers to direct questions or completion of a questionnaire at the beginning of the visit; 2) a way to summarize the preceding information on a form in the chart; and 3) a way to remind the clinician about each service needed (e.g., by using Post-it notes). Eight to 17 intervention clinics had components 1 and 2 in place for the eight preventive services. However, only 9 had component 3 in place for Papanicolaou smears and only 1 to 5 had component 3 in place for each of the other seven services.

At intervention's end, 94% of the 114 clinic team members who responded to a survey reported being very satisfied or satisfied with their experience; 91% reported that their time was well spent. One year after the intervention's end, with no intervening contact, we interviewed leadership of the 20 clinic teams who were able to meet with us. Seventeen teams (77% of 22) were still active. Since the intervention ended, 8 teams (36%) had added other preventive services to their systems, 20 (91%) had added other processes to their systems, 2 (9%) took on other problems, and 15 (68%) had collected more data and overcome additional barriers.

## Primary Outcomes

**Figure 2** shows the proportion of patients who were up-to-date by chart audit for seven preventive services at baseline and follow-up. Flu shots could not be included in this analysis because index visits occurred in the spring. Of the remaining selected preventive services, the intervention group showed statistically significant increases in the proportion of patients who were up-to-date for four services: cholesterol level (4.6% absolute increase), mammography (8.1% increase), breast examination (10.9% increase), and pneumococcal vaccine (17.2%). None of these services significantly increased in the control group.

However, only one of the seven selected preventive services—pneumococcal vaccine—had a significantly greater increase in the intervention group (17.2% absolute increase from baseline to follow-up compared with a 0.3% absolute increase in the control group; $P$ =

0.003). Of note, tetanus immunization (which was not one of the eight selected services) also increased more in the intervention group (12% in the intervention group vs. 3.3% in the control group; $P = 0.03$). This suggests that the intervention had a carryover effect.

**Figure 3** shows the proportion of patients at baseline and follow-up who reported being offered various preventive services if not up-to-date. Similar to the chart audit findings, the intervention and control groups did not differ significantly at baseline. In the intervention group, small, statistically significant increases were observed in four preventive services (blood pressure [5.4% increase], cholesterol [4.6% increase], pneumococcal vaccine [5.2% increase], and asking about tobacco use [15% increase]). In the control group, such differences were seen in two preventive services (blood pressure [5.5% increase] and asking about tobacco use [8.6% increase]). However, cholesterol monitoring was the only service for which the change in the two groups differed significantly (4.6% increase in the intervention group compared with 0.4% absolute decrease in the control group; $P = 0.006$).

Subgroup analyses showed that neither the 13 clinics reporting implementation before follow-up data collection nor the 8 clinics making the most complete systems changes brought about greater-than-average improvements in actual delivery of service. When the 3 clinics that had the best prevention systems were compared with control clinics, patients were more likely to be up-to-date at follow-up than at baseline for only one service (influenza shots), and only mammograms were more likely to be offered if not up-to-date.

## Discussion

The IMPROVE project was a randomized, controlled trial conducted to learn whether an external intervention sponsored by their contracting HMOs could lead private primary care clinics to use CQI techniques to improve their rate of delivery of important adult preventive services. Except for two small differences between the intervention and control clinics, CQI failed to produce any significantly greater improvement in the intervention clinics during the trial. In addition, the subgroup analyses suggest that this minimal intervention effect was not caused by an inadequate proportion of clinics implementing changes.

There are four possible explanations for the failure of this first large-scale controlled trial of a CQI-based intervention. The clinics recruited could have been atypical and more resistant to the intervention; the measurement and analysis of the preventive services rates may have been inadequate; delivery of the intervention may have been inadequate; and CQI may be an inappropriate mechanism for making preventive services improvements.

### Atypical Clinics

Clinics in the Minneapolis–St. Paul area are not necessarily typical. They are larger and more organized and are subject to an unusual degree of turmoil.[18] Nevertheless, they are also probably more interested in improving quality, as suggested by the relative ease of recruiting enough clinics for this complex, time-con-

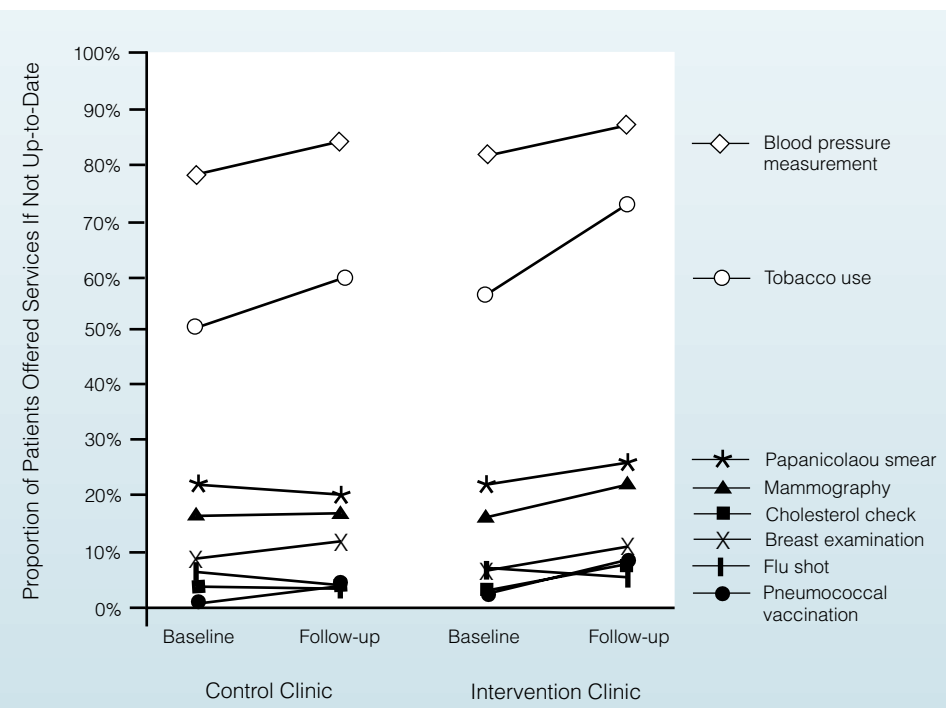


**FIGURE 3. Proportion of patients who were not up-to-date and were offered preventive services.**

suming trial and by clinics' statements that they volunteered because they wanted to improve preventive services and learn how to use CQI. The relatively high rates at which patients were receiving preventive services at baseline confirms clinic interest, but it also limits the extent to which further improvements may be possible. The clinics in this study may have been unusually good prospects for demonstrating the possibility of improvements, but their already better-than-average rates made further improvement more difficult.

### Inadequate Evaluation

Any negative study raises concerns about the possibility of type II error. Type II error can occur if the evaluation measures were unsatisfactory, if the sample sizes were too small to demonstrate differences, or if confounding variables obscured any differences. Although the validity of both patient report and chart audit have been questioned, the fact that rates were 12% to 38% lower on chart audit than on patient report (except for mammography) is consistent with the literature, which suggests that patients tend to overreport clinicians' actions and charts tend to underreport them.[20–24] Nevertheless, the direction and size of the changes were similar for both methods used in this trial, and neither method showed much difference between study groups in rate improvement.

Our sample size calculations gave us a power of 0.80 for detecting a 15% difference between intervention and control clinics in rates of Papanicolaou smears (the service that required the largest sample size). Consequently, the power should be greater for detecting such differences for all other services. In addition, we also inflated our calculated sample size by 20% in anticipation of survey return rates of less than 100%, resulting in slightly greater power. Perhaps most important, however, the actual differences observed were mostly so small that they would not be clinically important enough to warrant the amount of effort expended, even if they had been statistically significant.

It is unlikely that our results are due to confounding by patient or clinic characteristics. First, we found no clinically or statistically important differences between intervention and control clinics on any of the variables measured. Second, it is well known that standard errors tend to be underestimated in the presence of hierarchical data. The risk for erroneously rejecting the null hypothesis (type I error) is increased when this clustering is not taken into account through the use of an appropriate hierarchical estimator. When we used standard logistic regression techniques instead of our hierarchical estimator, we found no real differences in our conclusions, which suggests that the assessed variables did not cause important confounding.

### Inadequate Implementation of the Intervention

The generally enthusiastic participation and intensive, prolonged efforts of most of the clinic improvement teams suggest that they considered the approach and assistance to be useful and well carried out. Although our process evaluation showed that the teams did not complete or repeat the improvement cycle, were slow to implement changes, and usually implemented incomplete changes, this may have been the result of inexperience and of using a possibly inefficient, ineffective version of the CQI improvement process.

### Inadequacy of Continuous Quality Improvement

We believe that this trial demonstrated the relative ineffectiveness of the particular CQI approach that was tested. The literature offers only one other report of a completed randomized, controlled trial involving CQI. Goldberg and coworkers[25] showed that using CQI teams to improve compliance with Agency for Healthcare Research and Quality guidelines for depression and hypertension was similarly ineffective. Although that study put less emphasis on training, consulting, and networking, the outcome was similar: no real change in performance. One other trial is in progress.[26] Because of a growing perception that the traditional view of CQI does not work well in health care, leaders of quality improvement have developed a revamped approach that emphasizes rapid cycles and testing implementation rather than the more technical, studious approach tested in these two trials.[2, 3, 27, 28]

## Conclusions

Research has shown that office systems are an important vehicle for improvement in the delivery of clinical preventive services.[13, 29–31] The clinical trial by Dietrich and colleagues[32] was particularly convincing, although subsequent efforts to replicate this accomplishment were less successful.[33–36] The critical unanswered question is this: How might these systems be established and maintained in typical practice settings without research-funded support? Our study raises questions about whether CQI is the right model for making these changes.

After studying CQI in hospitals, Shortell and coworkers[37] concluded that there is "a relative dearth of empirical studies that provide comparative information on CQI implementation or impact." Our study and that of Goldberg and colleagues[25] suggest that early versions of CQI may have little effect in health care settings. However, additional studies of other versions and other settings are badly needed.

## Take-Home Points

- **Although there has been great interest in using CQI to improve the quality of health care, formal evaluations of its efficacy are lacking.**
- **We conducted a randomized, controlled trial of 44 primary clinics in the Minneapolis–St. Paul, Minnesota, area to test whether a CQI intervention could improve the delivery of eight preventive services.**
- **The CQI intervention clinics established improvement teams, participated in training sessions about CQI methods, had telephone contact with CQI consultants, and met to implement the iterative process of CQI to create better systems for the delivery of preventive services.**
- **As applied in our trial, a CQI-based intervention in primary care clinics produced minimal increases in the rates of preventive service delivery (measured by chart review or patient report).**

**References**


1. Berwick DM. Continuous improvement as an ideal in health care. N Engl J Med. 1989;320:53-6.
2. Nolan T. Accelerating the pace of improvement: an interview with Thomas Nolan. Jt Comm J Qual Improv. 1997;23:217-22.
3. Kilo CM. A framework for collaborative improvement: lessons from the Institute for Healthcare Improvement's Breakthrough Series. Qual Manag Health Care. 1998;6:1-13.
4. Solberg LI, Kottke TE, Brekke ML, Calomeni CA, Conn SA, Davidson G. Using continuous quality improvement to increase preventive services in clinical practice—going beyond guidelines. Prev Med. 1996;25:259-67.
5. Solberg LI, Isham G, Kottke TE, et al. Competing HMOs collaborate to improve preventive services. Jt Comm J Qual Improv. 1995;21:600-10.
6. U.S. Dept. of Health and Human Services. Healthy People 2000: national health promotion and disease prevention objectives. Washington, DC: U.S. DHHS, publication number PHS91-50212;1990.
7. Solberg LI, Kottke TE, Brekke ML. Will primary care clinics organize themselves to improve the delivery of preventive services? A randomized controlled trial. Prev Med. 1998;27:623-31.
8. Solberg LI, Brekke ML, Kottke TE. How important are clinician and nurse attitudes to the delivery of clinical preventive services? J Fam Pract. 1997;44:451-61.
9. Kottke TE, Solberg LI, Brekke ML, Cabrera A, Marquez MA. Delivery rates for preventive services in 44 midwestern clinics. Mayo Clin Proc. 1997;72:515-23.
10. Solberg LI, Kottke TE, Brekke ML, Conn SA, Magnan S, Amundson G. The case of the missing clinical preventive services systems. Eff Clin Pract. 1998;1:33-8.
11. Solberg LI, Brekke ML, Kottke TE, Steel RP. Continuous quality improvement in primary care: what's happening? Med Care. 1998;36:625-35.
12. Kottke TE, Solberg LI, Brekke ML, Cabrera A, Marquez M. Will patient satisfaction set the preventive services implementation agenda? Am J Prev Med. 1997;13:309-16.
13. Solberg LI, Brekke ML, Kottke TE. Are physicians less likely to recommend preventive services to low-SES patients? Prev Med. 1997;26:350-7.
14. Solberg LI, Kottke TE, Conn SA, Brekke ML, Calomeni CA, Conboy KS. Delivering clinical preventive services is a systems problem. Ann Behav Med. 1997;19:271-8.
15. Fleiss JL. Statistical Methods for Rates and Proportions. 2d ed. New York: J Wiley; 1981:212-36.
16. Goldstein H. Multilevel Statistical Models. 2d ed. New York: Halsted Pr; 1995.
17. Woodhouse G, ed. Multilevel Modelling Applications: A Guide for Users of MLn. London: Institute of Education, University of London; 1996.
18. Goldstein H, Rasbash J, Plewis I, et al. A User's Guide to MLwiN. London: Institute of Education, University of London; 1998.
19. Magnan S, Solberg LI, Giles K, Kottke TE, Wheeler JW. Primary care, process improvement, and turmoil. J Ambulatory Care Manage. 1997;20:32-8.
20. Boyer GS, Templin DW, Goring WP, et al. Discrepancies between patient recall and the medical record. Potential impact on diagnosis and clinical assessment of chronic disease. Arch Intern Med. 1995;155:1868-72.
21. Hiatt RA, Perez-Stable EJ, Quesenberry C Jr, Sabogal F, Otero-Sabogal R, McPhee SJ. Agreement between self-reported early cancer detection practices and medical audits among Hispanic and non-Hispanic white health plan members in northern California. Prev Med. 1995;24:278-85.
22. Ward J, Sanson-Fisher R. Accuracy of patient recall of opportunistic smoking cessation advice in general practice. Tob Control. 1996;5:110-3.
23. Montano DE, Phillips WR. Cancer screening by primary care physicians: a comparison of rates obtained from physician self-report, patient survey, and chart audit. Am J Public Health. 1995;85:795-800.
24. Solberg LI. Practical implications of recall bias. Tob Control. 1996;5:95-6.
25. Goldberg HI, Wagner EH, Fihn SD, et al. A randomized controlled trial of CQI teams and academic detailing: can they alter compliance with guidelines? Jt Comm J Qual Improv. 1998; 24:130-42.
26. Philbin EF, Lynch LJ, Rocco TA Jr, et al. Does QI work? The Management to Improve Survival in Congestive Heart Failure (MISCHF) study. Jt Comm J Qual Improve. 1996;22:721-33.
27. Berwick DM. Developing and testing changes in delivery of care. Ann Intern Med. 1998;128:651-6.
28. Alemi F, Moore S, Headrick L, Neuhauser D, Hekelman F, Kizys N. Rapid improvement teams. Jt Comm J Qual Improv. 1998;24:119-29.
29. McPhee SJ, Detmer WM. Office-based interventions to improve delivery of cancer prevention services by primary care physicians. Cancer. 1993;72:1100-12.
30. Leininger LS, Finn L, Dickey L, et al. An office system for organizing preventive services: a report by the American Cancer Society Advisory Group on Preventive Health Care Reminder Systems. Arch Fam Med. 1996;5:108-15.
31. Cohen SJ, Halvorson HW, Gosselink CA. Changing physician behavior to improve disease prevention. Prev Med. 1994;23:284-91.
32. Dietrich AJ, O'Connor GT, Keller A, Carney PA, Levy D, Whaley FS. Cancer: improving early detection and prevention. A community practice randomised trial. BMJ. 1992;304:687-91.
33. Dietrich AJ, Tobin JN, Sox CH, et al. Cancer early-detection services in community health centers for the underserved. A randomized controlled trial. Arch Fam Med. 1998;7:320-7.
34. Manfredi C, Czaja R, Freels S, Trubitt M, Warnecke R, Lacey L. Prescribe for health. Improving cancer screening in physi-

cian practices serving low-income and minority populations. Arch Fam Med. 1998;7:329-37.
35. Williams RB, Boles M, Johnson RE. A patient-initiated system for preventive health care. A randomized trial in community-based primary care practices. Arch Fam Med. 1998;7:338-45.
36. Ruffin MT 4th. Can we change physicians' practices in the delivery of cancer-preventive services? Arch Fam Med. 1998;7:317-9.
37. Shortell SM, Levin DZ, O'Brien JL, Hughes EF. Assessing the evidence on CQI: is the glass half empty or half full? Hosp Health Serv Adm. 1995;40:4-24.

**Acknowledgments**

The authors gratefully acknowledge the consistent, active, and cooperative support of both of the sponsoring HMOs (Blue Plus and HealthPartners). IMPROVE staff have been strongly committed, especially Katherine Giles, Kathy Schaivone, Carolyn Calomeni, Shirley Conn, Kathleen Conboy, Marty Campbell, Ann Sisel, Jerry Amundson, and Carol Westrum. The authors are also grateful to Eugene Nelson and Arnold Kaluzny for their insightful suggestions. However, the greatest appreciation goes to the staffs from the 2 demonstration and 44 trial clinics, especially the members of the individual process improvement teams:

Apple Valley Medical Center; Aspen Medical Group, Bloomington; Aspen Medical Group, West St. Paul; Aspen Medical Group, West Suburban; Chanhassen Medical Center, Chisago Medical Center; Creekside Family Practice; Douglas Drive Family Physicians; Eagle Medical; East Main Physicians; East Side Medical Center; Family Medical Center, Hennepin County; Forest Lake Doctors Clinic; Fridley Medical Center; Hastings Family Practice; HealthPartners St. Paul Clinic; Hopkins Family Practice; Hudson Physicians; Interstate Medical Center; Kasson Mayo Family Practice Clinic; Metropolitan Internists; Mork Clinic, Anoka; North Clinic; North St. Paul Medical Center; Northport Medical Center; Palen Heights Medical Center; Ramsey Clinic, Amery; Ramsey Clinic, Baldwin; Ramsey Clinic, Osceola; Ramsey Family Physicians; Richfield Medical Group; River Valley, Cottage Grove; River Valley, Hastings; River Valley, Woodbury; River Valley Clinic, Farmington; River Valley Clinic, Northfield; River Valley Medical Center, St. Croix Falls; Silver Lake Clinic; Southdale Family Practice; Southdale Internal Medicine; Southwest Clinic; St. Croix Valley Clinic; Stillwater Clinic; Sundance Medical Center; United Family Medical Center; and Valley Family Practice.

**Grant Support**

By the Agency for Healthcare Research and Quality (RO1 HS08091).

**Correspondence**

Leif I. Solberg, MD, HealthPartners Research Foundation, 8100 34th Avenue S, P.O. Box 1524, Minneapolis, MN 55440-1524; telephone: 612-883-5017; fax: 612-883-5022; e-mail: leif.i.solberg@healthpartners.com.

**APPENDIX TABLE**

**10 Key Processes for Implementing a System to Deliver Preventive Services**

| PROCESS | DESCRIPTION |
|---|---|
| Provide guidelines | Developing, obtaining broad buy-in, and updating a specific set of preventive services for defined age, sex, or risk groups |
| Screen | Obtaining information in a standard way about all patients to identify the specific prevention needs of each one |
| Summarize | Organizing and updating the information obtained in the screening process so that it can easily be reviewed by those who need to know the current prevention status of a particular patient |
| Cue | Reminding clinic staff and clinicians to undertake a prevention system task during office visits |
| Follow-up | Communicating to patients the results of preventive services, along with appropriate information and recommendations |
| Facilitate resource access | Selecting, gathering, organizing, and maintaining information on patient education and referrals that is needed by patients and clinic personnel |
| Counsel | Helping patients make needed changes in behavior |
| Track and recall | Reminding patients outside of the office encounter about their ongoing needs for specific preventive services |
| Activate patients | Encouraging patients to take greater responsibility for their own preventive services and behavior changes |
| Organize prevention visits | Providing all of a patient's necessary preventive services during a single visit that is designed and organized for that purpose |