By Roger A Forsyth, MD
Sudarmo Winarko, MD

Peer Discussion in a Family Practice

Abstract

Background

The combination of three techniques—peer discussion groups, peer lectures, and data feedback—was tested in a pilot program to maximize economy and quality of patient care as indicated by quantifiable measures.

Methods

In an office staffed by 15 adult primary care physicians in a group-model health maintenance organization (HM0), we implemented an ongoing peer-discussion process designed to increase use of recommended tests and to decrease use of tests not meeting criteria set by national authorities. Two similar offices served as controls. All groups were exposed to similar lectures and guidelines designed to increase use of recommended laboratory tests; only the study group was exposed to the intervention designed to reduce use of nonrecommended tests.

Results

Study and control groups showed similar increases in use of recommended tests. The study group had five consecutive monthly declines in overall number of diagnostic imaging procedures (-30%), laboratory tests (-35%), and incoming telephone calls requesting test results (-34%) ($p < 0.000001$). These reduced rates represented an annualized savings of >$400,000. In contrast, both control groups ordered more laboratory tests (13% and 19% increases) and received more calls (23% and 29% increases).

Conclusions

Peer discussion sessions and presentation of follow-up data can increase the use of recommended laboratory tests while reducing both the number of unnecessary tests and incoming telephone calls as well as health care costs.

Peer discussion group process

- Select a physician facilitator
- Suggest and rank discussion topics
- Present favored topic (volunteer)
- Involve each participant
- Reach consensus
- Gather baseline data
- Present blinded monthly follow-up

Introduction

Lectures and medical literature have been the primary methods used for physician education. Their efficacy has been questioned because they have not precluded substantial variation in practice patterns or suboptimal outcomes of testing and treatment. For this reason, new methods have been developed to enhance physician learning, change physician behavior, and improve treatment outcome.

Although guidelines were expected to change behavior by alerting physicians to recommendations of a panel of experts, results have varied.[1-3] The use of feedback of data to change behavior has also often been disappointing.[4-7]

Other approaches have included financial rewards, imposition of penalties, and systemic changes which force adoption of particular behavior. Examples include linking income to performance; requiring that all referrals to specialists be approved by a consultant; restricting the number of drugs included in the formulary; and use of order forms which exclude certain tests.[8-10] These approaches are frequently met with resistance.

Therefore, we set up a pilot program to improve patient care based on the success achieved by others in changing physician behavior through use of peer-based discussions, lectures, and data feedback. Evidence has shown that the magnitude and longevity of change is enhanced by simultaneous use of several methods of behavioral change instead of initiating changes one at a time.[11-15]

Methods

Study Scope and Participants

Fifteen physicians practicing in a freestanding adult primary care office of a group-model health maintenance organization (HMO) participated in a pilot program for enhancing the quality of patient care. Of these physicians, ten were Board-certified or Board-eligible in family practice, and five were Board-certified or Board-eligible in internal medicine. Two similar nearby 10-physician medical offices of the same group did not participate in the educational intervention and served as controls. Patient populations were similar in age and sex distribution, but ethnic distribution and mean education levels varied. Average panel sizes at the con-

ROGER A FORSYTH, MD, is a retired family practitioner from SCPMG, Pasadena. From 1981-2000, he served as Physician-in-Charge. He currently is clinical professor of family practice at the University of Southern California. E-mail: raforsuth@yahoo.com.
SUDARMO WINARKO, MD, is an internist who has been practicing primary care at Pasadena MOB since 1981. He previously served as the Quality Management coordinator for FP at LA Medical Center. Currently, he is the Physician-in-Charge and physician leader for the PEP at the Pasadena MOB as well as the Primary Care Service Line Director for the LA Medical Center. E-mail: sudarmo.x.winarko@IREmail.

trol offices (1400 patients) were smaller than at the study office (1750 patients).

Peer discussion process, topics, and analysis

The first step was to choose a physician to act as facilitator. Meetings were held for two hours twice monthly. The facilitator first asked each member of the group, in turn, to suggest a topic for discussion. Once the group ranked the topics, the top-ranked topic was chosen for peer discussion. One member of the group volunteered to research the subject and present it. Then each participant was given an opportunity to provide input on the topic. Where resistance to standard approaches was found, peer discussion allowed the group to find compromises which could increase compliance.

Correlation coefficients were calculated to help participants discard untenable justifications for avoiding adherence to accepted methods of care.

Acceptance was encouraged also by providing follow-up data for five months. The feedback was confidential but was accompanied by the blinded results for the group. By comparing individual data with group data, we hoped to encourage physicians with below-average performance to improve. Anonymity avoided embarrassing the participants. Before-and-after data showed that the efforts being made were improving the results.

The group voted to apply the process to improving the quality of physicians' professional lives. Each physician listed factors that were negatively affecting the quality of his or her professional life. Telephone calls requesting test results were voted the factor most amenable to change.

Through the discussion process, participants decided that telephone calls might be reduced by reducing the number of nonrecommended tests ordered and thereby reducing the need for patients to call for results. Pertinent data were gathered to verify whether volume of calls received was related to volume of tests ordered.

The lecture phase of the process included data considered impartial, eg, data published by the United States Preventive Services Task Force,[16] by the American Academy of Family Practice,[17] and by the Joint National Commission VI (JNC)[18] on treatment of hypertension; and a study from the Mayo Clinic on yield of positive results from commonly ordered tests.[19] This activity was designed to reinforce the fact that when previously known true positives, false-positives, and duplicative test results were excluded, the yield of positive results from administering nonrecommended tests would be very low. Norms were discussed, but application was at each physician's discretion without penalty for nonconformance.

Table 1. Correlation between laboratory test volume and various factors

Factor	Correlation Coefficient
Number of telephone calls	0.74
Panel size	0.03
Panel complexity	
% with diabetes	0.23
% with congestive heart failure	0.25
Quality of care for diabetic patients	
% with HgA_{1C} level measured	0.02
% with cholesterol level measured	0.42
% with HgA_{1C} <8.0 % of total hemoglobin	0.23
Patient satisfaction	0.20
Practice pattern	0.82

Results

Raw data for each physician in 1997 showed substantial variation in amount of time physicians spent delivering outpatient care. To make the data comparable, actual data on test volume for each physician was multiplied by a ratio (mean number of patients seen per physician in 1997 divided by number of patients seen by that physician in 1997), yielding volume of tests each physician would probably order had they spent the same number of hours delivering outpatient care in 1997. Similar corrections were made for all data, thereafter referred to as adjusted data.

High correlation was found between adjusted test volume and adjusted telephone volume (Table 1), convincing the group that reducing test volume would be a way to affect call volume.

The group named potential impediments to individual acceptance of group norms. These impediments included differences in size or complexity of patient panel, attentiveness to quality of care, interest in promoting patient satisfaction, and individual practice patterns which result from training, ability to tolerate ambiguity, and experience. A coefficient of correlation between each factor and test volume was calculated to find any evidence justifying a high rate of tests or calls.

The mean adjusted panel size was 2099 and ranged from 1331 to 2578.

Complexity of patient panel was measured by determining percentage of patients in the panel who had diabetes or congestive heart failure.

Each physician listed factors that were negatively affecting the quality of his or her professional life.

Two measures of quality of care are reported here as representative of those examined: 1) percentage of diabetic patients whose HgA_{1c} or cholesterol level was measured in 1997 (a measure of whether a physician was ordering recommended tests); and 2) percentage of diabetic patients with HgA_{1c} level <8.0% of total hemoglobin (a measure of whether the physician was achieving high-quality results).

The mean percentage of diabetic patients with HgA_{1c} level measured in 1997 was 58% (range, 51% to 74%); mean percentage of diabetic patients with cholesterol level measured in 1997 was 61% (range, 40% to 75%).

Patient satisfaction was determined by the 1997 score on 100 or more patient satisfaction surveys asking patients to rate their visit on a scale of 1 to 10 for eight indicators of patient satisfaction. Patient satisfaction scores ranged from 8.31 to 9.67.

Table 2. Cost of laboratory tests and diagnostic imaging procedures ordered in 1997 adjusted to a 206-workday year for 15 physicians

	Cost of laboratory tests	Cost of diagnostic tests	Total cost
High	$89,750	$100,904	$163,374
Mean	$40,578	$47,025	$87,603
Low	$20,875	$20,480	$43,775
Total	$608,645	$705,383	$1,314,028

Relation between practice patterns and laboratory test volume was evaluated by comparing laboratory test volume and diagnostic imaging volume. If pattern of practice determines volume of tests ordered, a physician who orders many laboratory tests would be likely to order many radiology procedures. The coefficient of correlation (0.82) showed that the pattern of practice was the only proposed factor that correlated with laboratory test volume.

Next, physicians were given feedback on their individual data and the blinded cost data for the group (Table 2). Adjusted cost comparisons for laboratory tests and radiology procedures ordered in 1997 (Figure 1) showed wide variation in costs of laboratory testing and diagnostic imaging. Laboratory costs ranged from $20,875 to $89,750 (mean cost, $40,576). Diagnostic imaging costs ranged from $20,480 to $100,904 (mean cost, $47,026).

To evaluate the possibility that changes in telephone or test volume at the study office might be part of a preexisting trend not due to the intervention, we compared volume in July, the baseline month, to volume in the month before the intervention. Neither number of laboratory tests nor number of calls were declining in the month before the study. Telephone calls per day averaged 1.58 calls in June vs 1.60 calls in July, and laboratory test volume per day averaged 46.8 tests per day in June vs 49.8 tests per day in July (range, 27.6 to 75.8 tests per day). Thus, no downward trend in tests or incoming calls was demonstrated before the study began.

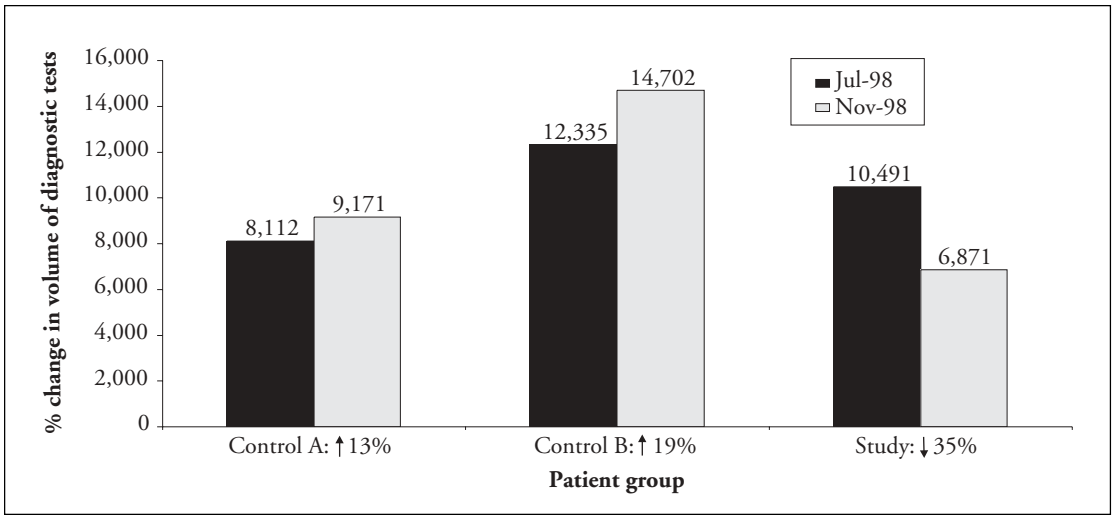

Figure 1. Percentage change in volume of diagnostic tests ordered (adjusted to a 206-workday year) by 15 physicians five months after implementing a physician peer discussion group.

If the peer discussion process were shown to change behavior, control and study offices, would be expected to show a similar trend in utilization before the study, but only the study office would be expected to show a reduced number of incoming telephone calls and tests after the study.

After baseline measurements were taken, rates of testing and incoming telephone calls at the study office declined every month. During the five-month study period, physicians at the study office had a mean 34% fewer incoming telephone calls requesting test results—a reduction of 0.5 telephone calls per day. The physician with the highest number of calls had 77% fewer incoming telephone calls, a mean reduction of 2.4 telephone calls per day.

The 35% five-month decline in number of laboratory tests ordered was equivalent to a mean reduction of 16.4 tests per day. Even the physician with the lowest utilization rate, 27.6 tests per day, ordered 31% fewer laboratory tests, a mean 8.6 fewer tests per day. The physician with the highest utilization rate, 75.8 tests per day, ordered 55% fewer tests, a mean of 41.7 fewer tests per day.

To evaluate the possibility of a general trend affecting volume, five-month data from the two control offices were examined. The data showed that the two control offices had 23% and 29% more incoming telephone calls (Figure 2) and ordered 13% and 19% more laboratory tests per day (Figure 1). These figures suggest that without the intervention, volume at the study office would have increased—as was occurring before the intervention and as continued to occur elsewhere during the study.

Comparing number of tests by clinic by month using the chi-square test of a contingency table shows that neither the observed change in test volume (chi-square = 1026.287, df = 2, $p < 0.000001$) nor the change in call volume (chi-square = 31.927, df = 2, $p < 0.000001$) was likely to have occurred by chance. However, comparing the observed number of calls to the expected number of calls using clinic-specific call-to-test ratios by using the chi-square goodness-of-fit test showed that these ratios did not significantly change (chi-square = 1.2027, df = 3, $p = 0.75$). Thus, reduced number of tests is the most likely explanation for the reduced number of calls.

Further validation of the 35% decrease in test volume being a result of the intervention is that the volume of radiologic procedures decreased by 30%. Because laboratory volume and radiology volume were highly correlated, these two volumes should increase or decrease together.

During the 12 months after discontinuing monthly

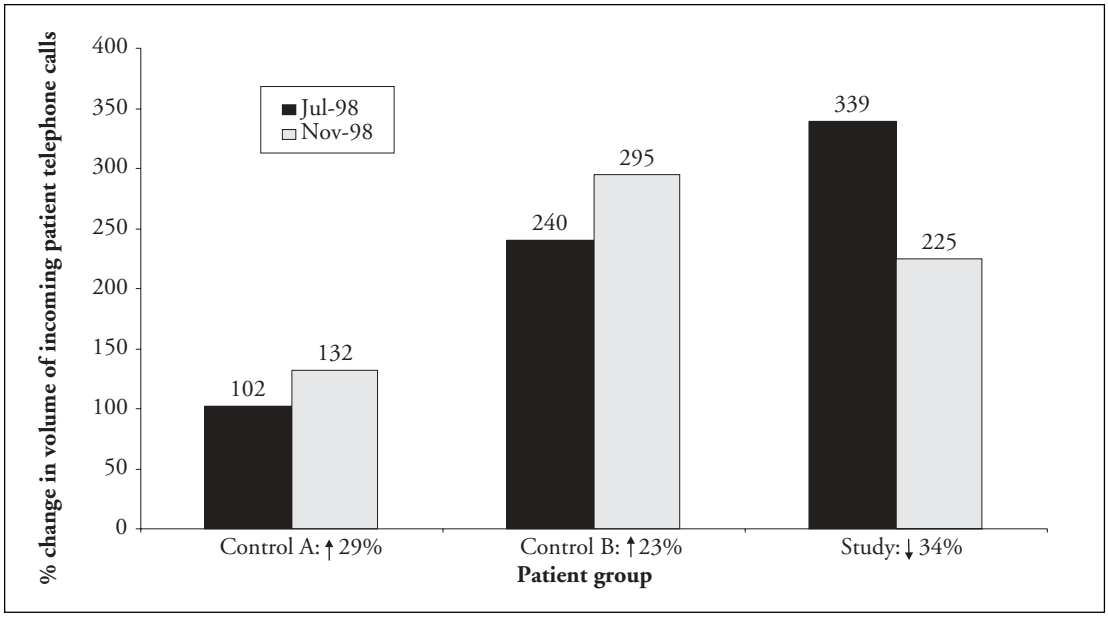

Figure 2. Percentage change in volume of incoming patient telephone calls, adjusted to a 206-workday year, five months after implementing a physician peer discussion group.

presentation of data to the study group, the rate of testing maintained a mean decline of 33.4% below baseline level. Mean number of incoming telephone calls maintained a decline of 33% in the ensuing year.

An eight-question survey distributed to participants in the study group rated their impression of effect of the intervention on their practices. Responses were scaled from 1 to 5, with 1 indicating complete disagreement and 5 indicating complete agreement with the statement. Results of the participant survey (Table 3) indicated strong support for the peer discussion process.

Table 3. Responses to eight-question survey asking 15 physicians to rate[a] effect of peer discussion on their medical practice

The peer discussion process	Mean Response Score
Helped me manage my practice	4.8
Improved the quality and efficiency of my practice	4.5
Did not compromise the quality of care delivered	4.5
Improved my professional satisfaction	4.3
Cut costs without regard to quality of care	1.8
Pressured me to reduce necessary care	1.8
Hindered my freedom to order tests	1.3

[a] on a scale of 1 to 5, where 1 = complete disagreement and 5 = complete agreement with statement.

Discussion

In this study to determine effectiveness of an educational process to improve the quality of care delivered in the medical offices, behavioral change was expected to increase by 1) using multiple educational techniques; 2) making the program peer-based (ie, not directed from the top downward); 3) identifying individual physicians' reservations about change and addressing these reservations by presenting the physicians with coefficients of correlation to confirm or refute the validity of specific reservations; and 4) using monthly (instead of annual or quarterly) follow-up with group participants to reinforce their behavioral change.

Lectures and data presented by peers seemed to foster in the physicians a willingness to participate, to voice reservations about guidelines, and to work toward a group norm acceptable to the individual.

The facilitator must be someone who can encourage group participation and who is comfortable with gathering data and making some simple calculations. The group must experience the process of voicing reservations, discussing relevant issues, and considering data which verify or refute the issues raised by individual physicians.

Feedback of data to physicians serves several functions. First, baseline data confirm whether change is needed. Such feedback permits physicians to identify their practice patterns in relation to those of other physicians, a comparison which can encourage setting goals. Presenting follow-up data reminds physicians of the need to change and rewards them by documenting change.

The study achieved a statistically significant ($p < 0.000001$) reduction in number of incoming telephone calls requesting test results and in laboratory tests ordered in the study group but not in the control groups, where both number of telephone calls and number of tests increased. Moreover, 12 months after behavioral reinforcement was discontinued, the reduction both in call volume and in test volume was largely maintained. This result indicated that the observed change was permanent and not a temporary study effect. That the reduction was seen throughout the study group—even for the physician with the lowest test utilization rate—we interpret as indicating pervasive use of laboratory tests beyond necessary levels. A high rate of laboratory utilization had been previously defended as indicating quality of care, yet the data we collected before the study showed very low correlation between number of tests ordered and measures of quality of care.

Conclusions

High volume of tests is associated with high rate of incoming telephone calls from patients requesting test results. The institution of a peer discussion process was associated with a statistically significant reduction in laboratory utilization and call volume.

Physicians who ordered a high volume of laboratory tests did not do more appropriate testing, achieve better disease control, or produce a higher level of patient satisfaction. Physicians made aware of these facts in a peer discussion setting—even physicians who have a relatively low level of test utilization—can achieve long-lasting reduction in their overall utilization of tests. This reduction is possible even with simultaneous implementation of programs designed both to increase use of recommended tests

and to improve quality of care. The authors attribute the success of this program to the emphasis on quality of care not quantity of tests. Physicians are likely to be more receptive to cost containment when they see that it does not affect quality of care. ❖

Related publication

Forsyth RA. HMO productivity: back to basics. Manag Care Interface 1998 Jan;11(1):80-4, 87.

References

1. Thompson RS, Kirz HL, Gold RA. Changes in physician behavior and cost savings associated with organizational recommendations on the use of "routine" chest X rays and multichannel blood tests. Prev Med 1983;12:385-96.
2. Wachtel TJ, O'Sullivan P. Practice guidelines to reduce testing in the hospital. J Gen Intern Med 1990;5:335-41.
3. Davis DA, Taylor-Vaisey A. Translating guidelines into practice. A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. CMAJ 1997;157:408-16.
4. Tierney WM, Miller ME, McDonald CJ. The effect on test ordering of informing physicians of the charges for outpatient diagnostic tests. N Engl J Med 1990;322:1499-504.
5. Wones RG. Failure of low-cost audits with feedback to reduce laboratory test utilization. Med Care 1987;25:78-82.
6. Berwick DM, Coltin KL. Feedback reduces test use in a health maintenance organization. JAMA 1986;255:1450-4.
7. Winkens RA, Pop P, Grol RP, Kester AD, Knottnerus JA. Effect of feedback on test ordering behavior of general practitioners. BMJ 1992;304:1093-6.
8. Lundberg GD. Laboratory request forms (menus) that guide and teach. JAMA 1983;249:3075.
9. Zaat JO, van Eijk JT, Bonte HA. Laboratory test form design influences test ordering by general practitioners in The Netherlands. Med Care 1992;30:189-98.
10. Axt-Adam P, van der Wouden JC, van der Does E. Influencing behavior of physicians ordering laboratory tests: a literature study. Med Care 1993;31:784-94.
11. van Walraven C, Goel V, Chan B. Effect of population-based interventions on laboratory utilization: a time-series analysis [published erratum appears in JAMA 2000 283:481]. JAMA 1998;280:2028-33.
12. Nardella A, Farrell M, Pechet L, Snyder LM. Continuous improvement, quality control, and cost containment in clinical laboratory testing. Enhancement of physicians' laboratory-ordering practices. Arch Pathol Lab Med 1994;118:965-8.
13. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. CMAJ 1995;153:1423-31.
14. Spiegel JS, Shapiro MF, Berman B, Greenfield S. Changing physician test ordering in a university hospital: an intervention of physician participation, explicit criteria, and feedback. Arch Intern Med 1989;149:549-53.
15. Solomon DH, Hashimoto H, Daltroy L, Liang MH. Techniques to improve physicians' use of diagnostic tests: a new conceptual framework. JAMA 1999;280:2020-7.
16. US Preventive Services Task Force. Guide to clinical preventive services: an assessment of the effectiveness of 169 interventions. Baltimore (MD): Williams & Wilkins; 1989.
17. American Academy of Family Practice Reference Manual 1997-1998. Summary of policy recommendations for periodic health examination. Kansas City (MO): American Academy of Family Physicians; 1998. p 55-68.
18. The sixth report of the Joint National Committee on prevention, detection, evaluation, and treatment of high blood pressure [published erratum appears in Arch Intern Med 1998;158:573]. Arch Intern Med 1997;157:2413-46.
19. Boland BJ, Wollan PC, Silverstein MD. Yield of laboratory tests for case-finding in the ambulatory general medical examination. Am J Med 1996;101:142-52.

It Should Be Unavoidable

Health education should not only be available, it should be unavoidable.

Sidney Garfield, MD, founder The Kaiser Permanente Health Plan
This "Moment in History" quote collected by Steve Gilford, KP Historian