Primary Care Physicians' Use of Lumbar Spine Imaging Tests

Effects of Guidelines and Practice Pattern Feedback

Donald K. Freeborn, PhD, Diana Shye, PhD, MPH, John P. Mullooly, PhD, Steve Eraker, MD, MPH, Jeffrey Romeo, MD, MPH

OBJECTIVE: **To reduce variability in primary care physicians' use of procedures for imaging the lumbar spine.**

DESIGN: **Controlled intervention using clinical practice guideline and practice pattern feedback.**

STUDY SAMPLE: **Sixty-seven internists and 28 family practitioners in a large, group-model HMO.**

MEASUREMENTS AND MAIN RESULTS: **Intervention group physicians received the clinical practice guideline for low back pain, followed after 4 months by three bimonthly feedback reports on their current use rates for lumber spine x-rays and computed tomography and magnetic resonance imaging scans of the lumbar spine. Control group physicians received neither the guideline nor the feedback reports. Automated radiology utilization data were used to compare intervention and control group physicians' changes in use rates and variability in use rates over the course of the study period. Neither the guideline alone nor the guideline plus feedback was associated with a significant decrease in use rates or in the variability in use rates for the lumbar spine imaging procedures under study.**

CONCLUSIONS: **Clinical practice guidelines and practice pattern feedback fail to achieve their goals when features of the practice setting and patient expectations and behavior are not identified and addressed.**

KEY WORDS: **low back pain; treatment; clinical practice guidelines; practice pattern feedback.**
J GEN INTERN MED 1997;12:619–625.

Low back pain affects about half the adults in the United States in a given year,[1] and is second only to colds as a reason for visits to primary care physicians.[2] In 1990, estimated direct medical care costs for low back pain were over $24 billion. Estimates for disability compensation and lost productivity brought the total costs associated with low back pain to approximately $100 billion.[3]

The cause of low back pain is often unclear, the correspondence between symptoms and anatomic findings is low, and up to 85% of patients with low back pain cannot be given a definitive diagnosis.[4] Recent trends toward increasing use of expensive imaging procedures such as computed tomography (CT) scans and magnetic resonance imaging (MRI) may be driven in part by physician uncertainty about the diagnosis of low back pain.[5] But lumbar spine imaging tests frequently reveal clinically irrelevant pathologic findings even in asymptomatic patients and thus may lead to unnecessary, expensive, and potentially harmful medical interventions.[6–8]

The majority of low back pain episodes can be treated conservatively in the primary care setting. Experts recommend early mobilization and nonprescription pain killers as the only necessary treatment for most patients, with lumbar spine imaging tests reserved for patients who are still limited by symptoms after a number of weeks of conservative treatment.[9,10] Yet despite this increasing consensus regarding the appropriate treatment of low back pain, wide geographic variations in diagnostic and treatment patterns have been found in the United States and other countries. These variations cannot be explained by differences in the patient populations studied.[7,11–15]

Variations in physicians' patterns of care for patients with low back pain are an example of a well-recognized broader phenomenon of unexplained practice pattern variations,[16–19] for which diagnostic uncertainty is one among a number of possible explanations.[20] This phenomenon has implications for efforts to improve medical care outcomes while containing costs,[20–22] and has prompted efforts to identify effective strategies for persuading physicians to adopt clinical behaviors that are consistent with state-of-the-art medical practice.

We conducted an exploratory analysis of low back pain incidence and treatment patterns in an HMO setting. We found a 6% to 7% annual incidence of low back pain among adult members in 1987. Three fourths of these low back pain episodes were initially treated by internal medicine and family practice physicians, and approximately two thirds of these patients received a nonspecific diagnosis (strains or sprains, or simply notation of the symptom of low back pain). Further analyses of data for 1991 showed that average use rates of tests for diagnosing low back pain were about 16% for lumbosacral spine x-rays, 5% for lumbar spine CT scans, and 1% for MRI scans (for internists and family practitioners combined). More striking than the absolute values of these use rates was the variation in rates among these primary care physicians: from 2% to 48% for x-rays, from 0% to 30% for CT scans, and from 0% to 9% for MRIs.

The principal motivation for the present study was concern about what this wide interphysician variation in

Received from Kaiser Permanente, Center for Health Research, Northwest Division, Portland, Ore.

Supported in part by a grant from the Kaiser Permanente Garfield Memorial Fund.

Address correspondence and reprint requests to Dr. Freeborn: Center for Health Research, 3800 N. Kaiser Center Dr., Portland, OR 97227.

use rates implied about quality of care. Our primary study objective was to determine whether dissemination of a clinical practice guideline, alone and together with practice pattern feedback, would reduce variability in use rates for lumbosacral spine x-rays and CT and MRI scans among primary care physicians. A secondary concern was the cost implications of increases in use rates in recent years for these imaging procedures—especially CT and MRI scans. Therefore, we were also interested in studying whether the intervention would reduce the overall use of these procedures.

METHODS

Study Setting

The study site was Kaiser Permanente Northwest Region, an established, not-for-profit, prepaid group-model HMO serving more than 400,000 members in Portland and Salem, Oregon, and southwest Washington. The HMO provides comprehensive outpatient and inpatient care to members, who are generally representative of the service area population. At the time this study was conducted, the HMO was organized into two administratively distinct medical areas within the service area, each having a number of ambulatory care medical facilities and an area hospital to which clinicians practicing in that area admitted their patients. In recent years, the HMO has increasingly relied on clinical practice guidelines to promote organizational goals of high-quality, cost-effective care. However, at the time this study began, the HMO did not have guidelines for the care of low back pain or for lumbar spine imaging test ordering, and the Agency for Health Care Policy and Research (AHCPR) low back pain guideline[10] had not yet been published.

Study Subjects

Subjects included all 67 general internal medicine physicians and 28 family practice physicians who were engaged exclusively in primary care practice during the year before and following the start of the intervention period, which began with the distribution of the low back pain clinical practice guideline in early May 1994.

Study Design

The design made use of the HMO's two administratively distinct medical areas. The 33 internists and 9 family practitioners in one administrative area (area A) were designated as the intervention group, while the 34 internists and 19 family practitioners in the other administrative area (area B) constituted the control group. The intervention group physicians received the low back pain guideline, followed by three feedback reports on their use rates for the lumbar spine x-rays and CT and MRI scans. Although it involved confounding between intervention effect and area effect, this design was considered necessary to avoid problems of contamination that could occur if physicians in the same medical office were individually randomized to intervention and control status.

The study design called for the research team to enlist several of the HMO's respected internal medicine and family practice physicians and relevant specialists to develop a guideline that focused on imaging test ordering by primary care physicians in the care of patients with low back pain. This plan had to be modified when we discovered that a group of specialist physicians in area A had been working for some time to develop a comprehensive process-of-care guideline for the care of patients with low back pain. This group wished to disseminate the guideline in both areas A and B. Negotiations ensued, and the research team agreed to use the comprehensive process-of-care guideline being developed by the specialist group for the research project. The specialist group agreed to delay disseminating the guideline in area B (the control group area) until after the completion of the research project. Recommendations for imaging test use were embedded at various junctures in the overall process-of-care guideline, which was summarized in a 1-page algorithm printed on a laminated removable page. The recommendations followed the approach advocated by Deyo et al.,[9] and subsequently outlined in the AHCPR low back pain guideline.[10]

Implementation of the intervention began with the mailing of the clinical practice guideline to the intervention group physicians in May 1994, under a cover letter from the co-principal investigators of the research team, one of whom was an internal medicine physician working in area A (the intervention area). Immediately following the mailing, members of the specialist group that developed the guideline attended internal medicine and family practice departmental meetings in the intervention group medical area to introduce and discuss the guideline. These meetings were well attended, and the primary care physicians expressed positive responses to the guideline. Other than these brief introductory presentations, no other educational activities about the guideline were carried out.

The "guideline-only" phase of the intervention continued for 4 months. The "guideline-plus-feedback" phase began with the mailing of the first feedback report to the intervention group physicians. Two more feedback reports were mailed at 2-month intervals. The physicians' use rates were tracked for a final 2 months following the mailing of the third feedback report.

The feedback reports were prepared by the research team with the help of the HMO's Medical Economics Department, using the HMO's radiology and appointment databases. Each report displayed individual use rates for the three procedures, ranked from highest to lowest for all the physicians (separately for internists and family practitioners). No physician identifiers were shown. However, to enable comparison with colleagues, each physician received a report on which his or her own rates had been highlighted with a marking pen.

The first feedback report covered the 3-month period during which the guideline alone had been implemented. The second report covered a 3-month period that included the last month of the guideline-only phase and the first 2 months of the guideline-plus-feedback phase. The third report covered the last 2 months of the guideline-plus-feedback phase. In addition to their use rates for the current period, the second and third feedback reports also included the rates of the immediately previous period (to allow the physicians to observe changes in their test-ordering rates). Although the reports were distributed only to the intervention group physicians in area A, they displayed the use rates for the study group physicians in area B as well.

Data Sources

The HMO's automated radiology information management database was the source of information on the numbers of imaging procedures ordered by the physicians (i.e., the numerator in the imaging test use rates). This database records all imaging procedures performed in the HMO's radiology department, including date of procedure and the identity of the patient and the ordering physician.

The HMO's automated appointment database was the source of information on the number of patient visits to intervention and control group physicians during the study period (i.e., the denominator in the imaging test use rates). This database records the data and appointment type of all visits to the HMO's providers (e.g., regularly scheduled office visit, same day appointment, walk-in visit, urgency care visit), and the identity of the provider and the patient (including patient age and gender).

Study Measures

Two study measures were used. The first was the use rate by internal medicine and family practice physicians (per 1,000 visits of patients aged 18 years and over) for lumbosacral spine x-rays and CT and MRI scans of the lumbar spine, indirectly standardized for visit type and patient age and gender, as indicators of case mix. Because the radiology management database did not provide information about the diagnosis for which imaging procedures were ordered, we identified these stratification variables as proxies for case mix, in the following manner. Using an outpatient utilization research database,[23] we carried out a linear regression analysis on the proportion of primary care physicians' total patient visits that were for low back pain, for two time periods (1980–1983 and 1984–1987). In both time periods, the proportion of total visits by patients in each of three age groups (18–39, 40–64, 65+), the proportion of patient visits by women, and the proportion of total visits of different appointment types (e.g., regularly-scheduled, walk-in, urgency care) accounted for approximately 40% of the variation in the proportion of physicians' total visits that were for low back pain. Therefore, both in the feedback reports that were distributed to the physicians and in the data analysis, use rates for the three imaging test procedures were indirectly standardized for these case-mix indicators.[24] We standardized separately for the preintervention and intervention periods, using all control and intervention group physicians as the reference population.

The second study measure was variability in internists' and family practitioners' use rates for lumbosacral spine x-rays and CT and MRI scans of the lumbar spine, as measured by the within-group variance of the individual indirectly standardized physician use rates.

Data Analysis

We compared the intervention and control group physicians with regard to the preintervention to postintervention changes in use rates and variability in use rates for the three lumbar spine imaging procedures. Individual physicians' rates for a given phase of the study period were compared with their own rates for the same period during the previous (preintervention) year, with separate comparisons for internal medicine and family practice physicians. We used two-sample, unweighted Student's *t* tests with the physician as the unit of analysis to compare use rate changes in the intervention and control groups. To compare intervention and control group differences in changes in the variance in use rates, we used the modified likelihood ratio test described by Morrison.[25]

RESULTS

Tables 1 and 2 compare, for internal medicine and family practice physicians, respectively, the mean individual use rates (indirectly standardized in the manner described above) for the three lumbar spine imaging procedures in the intervention and control group medical areas. The variability in use rates is indicated by the standard deviations of the use rates. For each phase of the intervention, rates for the intervention period are compared with rates for the corresponding time period in the previous year.

In general, the findings do not indicate any consistent pattern of reduction in either use rates or variability of use rates as a result of exposure to the intervention, for physicians from either specialty in either the intervention or control group. Among internal medicine physicians (Table 1) in the guideline-only phase of the intervention, rates for each imaging test procedure increased in comparison with the same period of the previous year among both the intervention and control group area physicians (with the exception of a slight drop in MRI scan use among intervention group physicians). In the guideline-plus-feedback phase, the findings were less consistent. X-ray use increased in the intervention group but decreased in the control group. Use of CT scans decreased in both groups, and MRI scan use was constant in the intervention group and decreased in the control group.

Table 1. Internal Medicine Physicians' Lumbar Spine Imaging Test Rates: Comparison of Preintervention and Postintervention Periods*

		Guideline Only Test Rate		Guideline-Plus-Feedback Test Rate	
Imaging Test	*n*	Preintervention (May–Aug 1993)	Intervention (May–Aug 1994)	Preintervention (Sept 1993–Feb 1994)	Intervention (Sept 1994–Feb 1995)
X-rays					
Intervention group	33	8.52 (5.63)	8.93 (5.53)	8.07 (6.15)	8.66 (4.76)
Control group	34	9.06 (4.77)	9.82 (5.49)	7.94 (4.13)	7.66 (4.70)
CT scans					
Intervention group	33	2.76 (2.75)	3.66 (3.24)	4.15 (3.15)	2.86 (2.08)
Control group	34	2.32 (2.24)	3.55 (2.81)	3.27 (2.52)	2.58 (2.25)
MRI scans					
Intervention group	33	0.21 (0.76)	0.19 (0.50)	0.35 (0.60)	0.35 (0.73)
Control group	34	0.21 (0.67)	0.33 (0.70)	0.31 (0.68)	0.16 (0.45)

**The rate is the number of tests per 1,000 visits of patients aged 18 years and over, indirectly standardized for type of contact and patient age and gender. Values are means (SD).*

Among family practitioners (Table 2), a fairly consistent pattern of reduction in rates for all three imaging procedures was evident among both intervention and control group physicians. Exceptions to this pattern were that x-ray use increased among intervention group physicians in the guideline-only phase and among both intervention and control group physicians in the guideline-plus-feedback phase. Also, MRI use increased in the control group during the guideline-plus-feedback phase. The standard deviations of the use rates indicate an inconsistent pattern of changes in variability in use rates.

Changes in the mean use rates (intervention period rate minus preintervention period rate) are contrasted for the intervention and control groups in Table 3 (for internists) and Table 4 (for family practitioners). These tables contrast the preintervention and postintervention variability in use rates, using the ratio of postintervention to preintervention within-group variances. In neither medical specialty were there any significant differences in mean use rates between intervention and control group physicians in the preintervention to postintervention change. Nor did intervention group physicians have greater reductions in use rate variability when compared with control group physicians.

Finally, we used paired Student's t tests to examine preintervention-postintervention changes in imaging test use rates for the guideline-only and guideline-plus-feedback periods *within* the intervention and control groups, separately for internal medicine and family practice physicians. A few significant differences were found, but they included increases as well as decreases. No consistent pattern of change was found among physicians of either specialty in either the control or intervention groups.

Table 2. Family Practice Physicians' Lumbar Spine Imaging Test Rates: Comparison of Preintervention and Postintervention Periods*

		Guideline Only Test Rates		Guideline-Plus-Feedback Test Rates	
Imaging Test	*n*	Preintervention (May–Aug 1993)	Intervention (May–Aug 1994)	Preintervention (Sept 1993–Feb 1994)	Intervention (Sept 1994–Feb 1995)
X-rays					
Intervention group	9	11.16 (5.32)	12.32 (6.53)	10.51 (5.47)	11.15 (6.94)
Control group	19	8.57 (7.84)	8.03 (5.89)	7.08 (4.87)	8.05 (5.15)
CT scans					
Intervention group	9	4.72 (3.51)	4.26 (4.75)	4.30 (4.38)	3.36 (5.10)
Control group	19	3.38 (2.30)	2.79 (2.82)	3.57 (2.06)	1.72 (1.54)
MRI scans					
Intervention group	9	0.25 (0.39)	0.16 (0.32)	0.36 (0.61)	0.15 (0.23)
Control group	19	0.51 (0.99)	0.42 (1.04)	0.17 (0.45)	0.40 (0.67)

**The rate is the number of tests per 1,000 visits of patients aged 18 years and over, indirectly standardized for type of contact and patient age and gender. Values are means (SD).*

Table 3. Internal Medicine Physicians: Intervention and Control Group Preintervention-to-Postintervention Changes in Mean Adjusted Imaging Test Use Rates and in Variability in Use Rates

	Preintervention-to-Postintervention Change in Use Rate*		Ratio of Postintervention to Preintervention Variances†	
Imaging Test	Guideline Only	Guideline Plus Feedback	Guideline Only	Guideline Plus Feedback
X-rays				
Intervention group	0.41 (5.17)	0.59 (5.25)	0.96	0.60
Control group	0.75 (4.97)	−0.28 (3.89)	1.32	1.30
Difference in change, *p* value	.75	.44	.82	.08
CT scans				
Intervention group	0.90 (2.55)	−1.29 (3.03)	1.39	0.44
Control group	1.23 (3.38)	−0.69 (3.26)	1.57	0.80
Difference in change, *p* value	.65	.44	.08	.34
MRI scans				
Intervention group	−0.02 (0.61)	0.01 (0.67)	0.43	1.48
Control group	0.13 (0.97)	0.15 (0.82)	1.09	0.44
Difference in change, *p* value	.46	.43	.004	.10

**Intervention period value minus preintervention period value; number of tests per 1,000 visits of patients aged 18 years and over, indirectly standardized for type of contact and patient age and gender.*
†Postintervention period variance/preintervention period variance.

DISCUSSION

Under certain circumstances, the implementation of clinical practice guidelines can be effective in changing physician practice patterns.[26] Research findings on practice pattern feedback are more equivocal than those on guidelines, but some feedback approaches appear to have at least moderate short-term effectiveness.[27–34] Yet carefully conducted implementations of both approaches have also failed to achieve their goals.[35–40] Our study findings fall into this latter category. In their responses to brief feedback questionnaires that we sent to the intervention group physicians following our study, most physicians reported that they had received the low back pain guideline, read it, and thought it useful. They also reported receiving and reading the feedback reports, though consensus about their usefulness was lower than for the guideline. Despite good dissemination of the intervention tools, we found that neither intervention was associated with a consistent pattern of reduction in use rates or in variability of use rates for any of the lumbar spine imaging tests.

A number of factors may have contributed to these negative findings. First, some technical aspects of our intervention may have diffused its impact on the targeted physicians. We originally intended to develop a brief guideline that delivered a succinct, focused message on lumbar spine test ordering. However, the guideline that we ultimately distributed was a 23-page document that dealt with the whole process of care for low back pain. Our guideline did include a 1-page laminated, removable algorithm that summarized the process of care, including recommendations for ordering imaging tests. However, the lack of a specific focus on this topic may have diffused the impact of the guideline on the single clinical behavior that we studied. Technical features of the feedback reports may also have limited their effectiveness. Because of the relatively short time period covered by each feedback report, the physicians' use rates were based on fairly small numbers of tests ordered. As a result, a small absolute change in the number of tests ordered by an individual physician could produce a dramatic change in his or her ranking. This may have reduced the relevance of the information.

In addition, the intervention strategy we tested was deliberately limited in scope to the dissemination of the guideline and the development and mailing of the feedback reports—"administrative" activities that can be relatively easily and inexpensively carried out in a large, group-model HMO. Aside from brief presentations by members of the group of specialists who developed the guideline, no other social influence or educational strategies were used to enlist the physician's adherence to any of the guideline's recommendations. A more intensive or comprehensive educational approach may be necessary to change physicians' usual practice patterns.

Finally, our findings may be partly due to features of this practice setting that were beyond the scope of our intervention to affect. Before beginning the intervention, we conducted focus groups with primary care physicians. In these groups, the physicians agreed with a conservative approach to ordering lumbar spine imaging tests that was consistent with other recommendations.[9,10] Yet they indicated that nonclinical factors sometimes influenced their imaging-test-ordering decisions. In particular, the physicians cited the effect of tensions arising from the diverse obligations of the primary care physician's role.

For example, to enlist patients' adherence to the

Table 4. Family Practice Physicians: Intervention and Control Group Preintervention-to-Postintervention Changes in Mean Adjusted Imaging Test Use Rates and in Variability in Use Rates

	Preintervention-to-Postintervention Change in Use Rate*		Ratio of Postintervention to Preintervention Variances†	
Imaging Test	Guideline Only	Guideline Plus Feedback	Guideline Only	Guideline Plus Feedback
X-rays				
Intervention group	1.16 (5.80)	0.63 (3.98)	1.51	1.61
Control group	−0.54 (6.31)	0.98 (4.01)	0.56	1.12
Difference in change, *p* value	.50	.83	.63	.78
CT scans				
Intervention group	−0.45 (3.14)	−0.94 (2.52)	1.83	1.36
Control group	−0.59 (3.15)	−1.86 (1.71)	1.50	0.56
Difference in change, *p* value	.92	.26	.20	.001
MRI scans				
Intervention group	−0.10 (0.24)	−0.21 (0.50)	0.69	1.48
Control group	−0.08 (1.58)	0.23 (0.62)	1.10	0.79
Difference in change, *p* value	.97	.08	<.0001	.01

**Intervention period value minus preintervention period value; number of tests per 1,000 visits of patients aged 18 years and over, indirectly standardized for type of contact and patient age and gender.*
†Postintervention period variance/preintervention period variance.

treatment recommendations, physicians felt the need to gain their confidence and trust by responding to their expectations and concerns. But as gatekeepers, the physicians were also expected to restrict unnecessary access and referrals—including those for lumbar spine imaging tests. At the same time, the physicians had to meet expectations of high patient care productivity, in an environment where the *time* required for conversations with patients is a constrained resource.[41] In this complex situation, ordering an imaging test—even one that the physician knows is not strictly medically indicated—may help resolve such tensions. The physicians' responses suggest that specific tools for changing physician behaviors, such as clinical practice guidelines and practice pattern feedback, may not be effective unless they are used with system-level, organizationally sponsored, quality-improvement efforts.

We acknowledge the significant contributions of Deborah Harris, who provided invaluable technical assistance at various stages of the project and was responsible for much of the data analysis. We also acknowledge the cooperation and important contributions of Paul Steele, who assisted with the data analysis and other aspects of the study. We also thank Dawn Hayami for her assistance, helpful advice, and ongoing support throughout the project.

REFERENCES

1. Sternbach RA. Survey of pain in the United States: the Nuprin pain report. Clin J Pain. 1986;2:49–53.
2. Cypress BK. Characteristics of physician visits for back symptoms: a national perspective. Am J Public Health. 1983;73:389–95.
3. Frymoyer JW, Cats-Baril WL. An overview of the incidences and costs of low back pain. Orthop Clin North Am. 1991;22:263–71.
4. White AA III, Gordon SL. Synopsis: workshop on idiopathic low-back pain. Spine. 1982;7:141–9.
5. Deyo RA. Magnetic resonance imaging of the lumbar spine. Terrific test or tar baby? N Engl J Med. 1994;331(2):115–6.
6. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic resonance scans of the lumbar spine in asymptomatic subjects. J Bone Joint Surg. 1990;72:403–8.
7. Cherkin DC, Deyo RA, Wheeler K, Ciol MA. Physician variation in diagnostic testing for low back pain. Who you see is what you get. Arthritis Rheum. 1994;37(1):15–22.
8. Wiesel SW, Tsourmas N, Feffer HL, Citrin CM, Patronas N. A study of computer-assisted tomography, I: the incidence of positive CAT scans in an asymptomatic group of patients. Spine. 1984;9:549–51.
9. Deyo RA. Newer thinking on diagnosis and therapy. Consultant. 1993;33(2):88–100.
10. AHCPR Low Back Pain Guideline Panel. Acute Low Back Pain Problems in Adults, Clinical Practice Guideline No. 14. Washington, DC; Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; 1995. AHCPR publication no. 95-0644.
11. Deyo RA, Cherkin DC, Conrad D, Volinn E. Cost, controversy, crisis: low back pain in the health of the public. Annu Rev Public Health. 1991;12:141–56.
12. Taylor VM, Deyo RA, Cherkin DC, Kreuter W. Low back pain hospitalization: recent United States trends and regional variations. Spine. 1994;19(11):1207–13.
13. Volinn E, Mayer J, Diehr P, Van Koevering D, Connell FA, Loeser JD. Small area analysis of surgery for low-back pain. Spine. 1992; 17:575–80.
14. Volinn E, Turczyn KM, Loeser JD. Patterns in low back pain hospitalizations: implications for the treatment of low back pain in an era of health care reform. Clin J Pain. 1995;10(1):643–70.
15. Walsh K, Cruddas M, Coggon D. Low back pain in eight areas of Britain. J Epidemiol Community Health. 1992;46:227–30.
16. Wennberg JE, Blowers L, Parker R, Gittlesohn AM. Changes in tonsillectomy rates associated with feedback and review. Pediatrics. 1977;59:821–6.
17. Wennberg J, Freeman JL, Clup SJ. Are hospital services rationed in New Haven or over-utilized in Boston? Lancet. 1987;1:1185–9.

18. Wennberg J, Freeman JL, Shelton RM, Bubolz TA. Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. N Engl J Med. 1989;321:1168–73.
19. Wennberg J, Gittelsohn A. Small area variations in health care delivery. Science. 1973;182:1102–8.
20. Eisenberg JM. Doctors Decisions and the Cost of Medical Care. Ann Arbor, Mich: Health Administration Press; 1986.
21. 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(9):784–94.
22. Epstein AM. Changing physician behavior: increasing challenges for the 1990s. Arch Intern Med. 1991;151:2147–9.
23. Greenlick MR, Freeborn DK, Pope CR. Health Care Research in an HMO. Baltimore, Md: The Johns Hopkins University Press; 1988.
24. Fleiss JL. Statistical Methods for Rates and Proportions. 2nd ed. New York, NY: Wiley; 1981.
25. Morrison DF. Multivariate Statistical Methods. New York, NY: McGraw-Hill; 1967.
26. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluation. Lancet. 1993; 342:1317–22.
27. Marton KI, Tul V, Sox HC. Modifying test-ordering behavior in the outpatient medical clinic. Arch Intern Med. 1985;145:816–21.
28. Berwick DM, Coltin KL. Feedback reduces test use in a health maintenance organization. JAMA. 1986;255:1450–4.
29. Applegate WB, Bennett MD, Chilton L, Skipper BJ, White RE. Impact of a cost-containment educational program on house-staff ambulatory care clinic charges. Med Care. 1983;21:486–96.
30. Braham RL, Ruchlin HS. Physician practice profiles: a case study of the use of audit and feedback in an ambulatory care group practice. Health Care Manage Rev. 1987;12:11–6.
31. Spiegel JS, Shapiro MF, Berman B, Greenfield S. Changing physician test ordering in a university hospital. Arch Intern Med. 1989;149:549–53.
32. Winickoff RN, Coltin KL, Morgan MM, Buxbaum RC, Barnett GO. Improving physician performance through peer comparison feedback. Med Care. 1984;22:527–34.
33. Fowkes FGR, Evans KT, Hartley G, et al. Multicenter trial of four strategies to reduce use of a radiological test. Lancet. 1986;2:367–9.
34. Eraker S, Romeo J, Freeborn D, Mullooly J, Nelsen J, Hayami D. Effect of practice pattern feedback on imaging test utilization in a health maintenance organization. Presented at a conference entitled “New Vision, New Ventures,” sponsored by Kaiser Permanente, San Francisco, Calif., Oct. 24–28, 1993.
35. Schroeder SA, Myers LP, McPhee SJ, et al. The failure of physician education as a cost containment strategy. JAMA. 1984;252:225–30.
36. Williams SV, Eisenberg JM. A controlled trial to decrease the unnecessary use of diagnostic tests. J Gen Intern Med. 1986;1:8–13.
37. Wones RG. Failure of low-cost audits with feedback to reduce laboratory test utilization. Med Care. 1987;25:78–82.
38. Schectman JM, Elinsky EG, Pawlson G. Effect of education and feedback on thyroid function testing strategies of primary care clinicians. Arch Intern Med. 1991;151:2163–6.
39. Schectman J, Kanwal N, Schroth W, Elinsky E. The effect of an education and feedback intervention on group-model and network-model health maintenance organization physician prescribing behavior. Med Care. 1995;33(2):139–44.
40. Lomas J, Anderson GM, Domnick-Pierre K, et al. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. N Engl J Med. 1989;321:1306–11.
41. Budrys G. Coping with change: physicians in prepaid practice. Sociol Health Illness. 1993;15(3):353–74.