A Randomized Trial of Medical Quality Assurance

Improving Physicians' Use of Pelvimetry

Mark R. Chassin, MD, MPP, MPH, Sally M. McCue, MBA

The capacity of educational programs to improve physician performance remains doubtful despite many evaluative efforts. This is especially true for programs sponsored by the federal government. We tested the efficacy of an educational program conducted by Professional Standards Review Organizations in reducing the inappropriate use of x-ray pelvimetry. This procedure may cause harm to the fetus, and there is little evidence that it is efficacious. We randomly assigned 120 hospitals in six Professional Standards Review Organizations to study and control groups. Physicians with delivery privileges at each study hospital participated in an educational program that discussed acceptable indications for x-ray pelvimetry. Pelvimetry use was similar in study and control hospitals before the program. However, after the program, pelvimetry was performed by physicians at study hospitals less than one third as often as by physicians at control hospitals. We conclude that educational programs can improve physician performance substantially and that such programs can be effectively conducted by federally sponsored physician organizations.

(*JAMA* 1986;256:1012-1016)

RESEARCHERS have been hard pressed rigorously to document substantial improvements in physician performance in response to educational programs. Previous randomized controlled trials have either been small, involving few physicians and hospitals,[1,2] have demonstrated at best modest improvements,[3,4] or have focused on interns and residents.[5,6] In addition to these limiting factors, studies that have used quasiexperimental designs[7-19] are frequently open to criticism because they fail adequately to isolate the educational program as the sole cause of any change that is measured.

It has been especially difficult to measure the effectiveness of governmental programs in influencing physician behavior. Most such programs are begun as demonstrations rather than randomized controlled trials, making their evaluation problematic at best.[20-23]

In early 1980, the Office of Professional Standards Review Organizations (PSROs) in the US Department of Health and Human Services was deciding how PSROs should perform review of ancillary services, then an area rarely examined by any of them. Because of the lack of data substantiating the effectiveness of physician education on the one hand and PSRO review on the other, the Office of PSROs initiated a randomized controlled trial of the effectiveness of an educational program conducted by PSROs to reduce the inappropriate use of x-ray pelvimetry. This procedure is generally considered to be of little usefulness because there is little evidence that it is efficacious[24-27] and a substantial amount of data suggesting that it is harmful.[28-35] The American College of Obstetricians and Gynecologists, the American College of Radiology, and the Bureau of Radiological Health of the Food and Drug Administration have all issued statements discouraging its use.[36-38]

To our knowledge, this study is the first randomized controlled trial of a federal health care review program. It involved 1483 physicians at 120 hospitals in six states.

METHODS

Six PSROs were selected to participate in the study. They were not chosen at random, but rather because of their interest in and their ability to implement programs such as the one conceived for this study. Of the six PSROs, two were large, statewide organiza-

From the Health Program, The Rand Corporation, Santa Monica, Calif (Dr Chassin), and the La Jolla Management Corporation, Rockville, Md (Ms McCue).

The opinions, conclusions, and proposals in the text are those of the authors and do not necessarily represent the views of the Department of Health and Human Services, The Rand Corporation, or the La Jolla Management Corporation.

Reprint requests to The Rand Corporation, 1700 Main St, Santa Monica, CA 90406-2138 (Dr Chassin).

tions: Arkansas and New Hampshire. The remaining four were smaller, representing parts of states: southwestern Pennsylvania, the city of Baltimore, the area around Wayne, NJ, and the area around Winston-Salem, NC. In Arkansas and New Hampshire, all hospitals performing routine deliveries were randomly assigned to study or to control groups after excluding hospitals with fewer than two pelvimetries per 100 deliveries in the previous three years. Each study hospital was to receive an intervention in the form of an educational program discussing appropriate indications for pelvimetry. This design was selected to maximize the comparability of study and control groups.

In striving to maximize comparability, however, we recognized that random assignment of hospitals within a PSRO might result in the assignment of some hospitals with almost identical obstetrical staffs to both study and control groups, thus risking contamination of the controls by interventions at some study hospitals. We attempted to minimize this risk by choosing a different experimental design for the remaining four PSROs. These PSROs were divided into two pairs on the basis of their demographic characteristics and preintervention pelvimetry rates. One PSRO of each pair was selected at random to be a study PSRO and the other to be a control. The two study PSROs, Baltimore and Winston-Salem, are separated geographically by large distances from the control PSROs, southwestern Pennsylvania and Wayne, thus minimizing the possibility of contamination. In the study PSROs, all hospitals performing routine deliveries received the intervention, while none of the hospitals in the control PSROs did.

In June 1980, three years of baseline data on monthly rates of pelvimetry were collected retrospectively at all hospitals in all six PSROs. The data on monthly frequency of pelvimetry were collected by abstracting the logs of the radiology departments of each hospital. Data on the monthly frequency of deliveries were collected by abstracting delivery room logs. The educational interventions took place in early 1981, and the study concluded in early 1982 when data were again retrospectively collected on the monthly frequency of pelvimetry and delivery in study and control PSROs from June 1980 through March 1982. These data collection activities were monitored for reliability by reabstracting each hospital's radiology and delivery room logs for 10% of the 58 months of data included in the study. Reabstractor counts were within 0.1% of original counts of deliveries and within 2% for pelvimetries in both study and control groups.

Table 1.—Selected Characteristics of Study and Control Hospitals

Characteristic	Study Group (N=64)	Control Group (N=56)	P
% Urban*	44	57	NS
% Teaching†	23	14	NS
% With obstetrical residencies	19	4	<.05
Average No. of beds	194	191	NS

*Urban indicates population located in counties of Standard Metropolitan Statistical Areas with at least 50 000 population in 1970.
†Teaching indicates presence of any approved residency program.

Each study hospital received an educational intervention sponsored by the PSRO. During the development of the study, obstetricians from three of the four study PSROs met and agreed on the content of the program. These physicians reached unanimous agreement on currently acceptable indications for pelvimetry, which were very few. A standardized set of materials was prepared for use during the educational program. These same obstetricians visited the study hospitals in their PSROs and delivered the standardized presentation at a medical or obstetrical staff meeting. In the fourth PSRO, a team of three obstetricians reviewed and approved the educational program and conducted the meetings. All physicians with delivery privileges were invited to the meetings. The program discussed the lack of evidence of the efficacy of pelvimetry, its possible harm, and the few acceptable indications for its use. Effective methods for monitoring the progress of labor, both invasive and noninvasive, were also discussed.

The program was entirely educational. Never was denial of reimbursement considered by the Office of PSROs as a method to influence physician behavior. Nor was this possibility mentioned in any of the standardized study materials sent to physicians and hospitals or in any of the presentations made at any of the study hospitals. In 16 of 64 study hospitals, it proved impossible to schedule a meeting with the obstetrical staff within the time period provided in the study. Each physician with delivery privileges at these hospitals received a packet of materials by mail with a letter from the obstetrician-educator that described the study and discussed current indications for pelvimetry. At the remaining 48 hospitals, 60% of the programs were conducted during regularly scheduled staff meetings; 40% of the time, a special meeting was called. At 81% of the meetings, physicians received continuing medical education credits for their participation. Every physician with delivery privileges who did not attend the meetings received a packet of study materials by mail.

For three months after the educational program, each PSRO monitored each study hospital's use of pelvimetry and reported back to the hospital, on a monthly basis, the hospital's previous rate of use of pelvimetry (per 100 deliveries) and the rate for the current month. This monitoring period was conceived as part of the overall intervention and was intended to reinforce the message of the educational program. The monitoring reports were sent to the hospital administrator, the chief of obstetrics, the chief of radiology, the chairman of the quality assurance committee, and the obstetrical nurse supervisor and were posted in the physicians' lounge. Control hospitals received neither the educational program nor the monitoring. These hospitals were thus not contacted at all by the study between the baseline data collection activity in June 1980 and the final data collection activity in April 1982.

RESULTS

Table 1 shows the number of hospitals involved in the study and some of their characteristics. The only statistically significant difference is that a greater proportion of hospitals in study areas had obstetrical residency programs (19% vs 4%; $P<.05$). This lack of balance occurred because one of the two paired PSROs that was assigned to the study group had a large proportion of these hospitals.

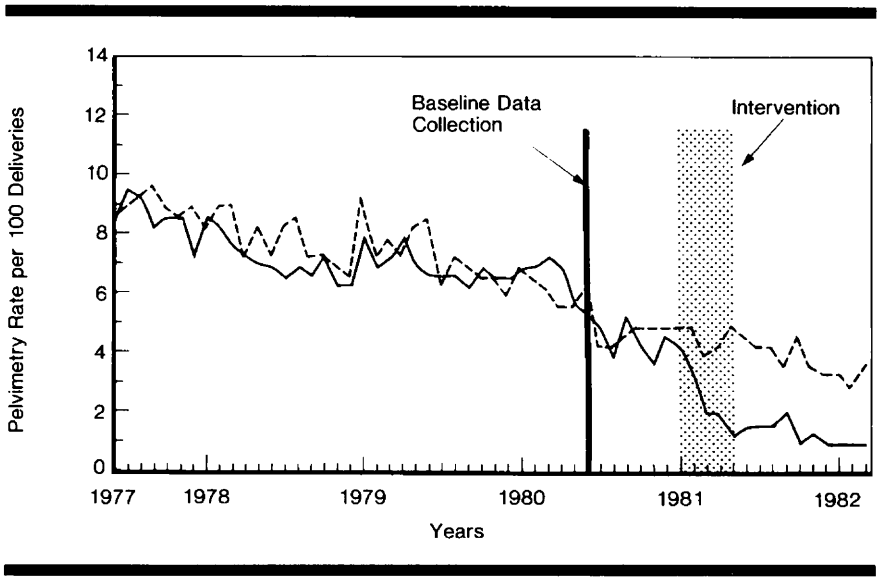
The Figure shows the entire 58 months of data for both study and control groups. Simple regression analysis of the first 36 months of data, representing the baseline period for both study and control groups, shows that pelvimetry rates were declining slowly during this period in both

groups. Statistical comparison of the rates of decline for the study and control groups in the baseline period shows no significant difference. While the Figure shows the average rate of use of pelvimetry in study and control areas, pelvimetry use varied substantially by hospital, ranging from 0 to 48 per 100 deliveries, with a mean of 7.3. Twenty-two percent of hospitals had pelvimetry rates greater than ten per 100 deliveries, and 10% were greater than 15.

The Figure also demonstrates that after the baseline data collection in June 1980, there was a small decrease in pelvimetry use in both study and control groups. Beginning with the start of the intervention period, however, and continuing through the end of the follow-up period, the study group exhibited a marked decline in pelvimetry use. No such decrease occurred in the control group.

Table 2 presents these data in numerical form. The average pelvimetry rate during the 36-month baseline period is compared with the average pelvimetry rate during the last six months of the follow-up period for study and control groups. The difference in pelvimetry rates between study and control groups in the baseline period, while very small, just achieves statistical significance at the 5% level, with 95% confidence limits for the difference ranging from 0.1 to 0.5 pelvimetries per 100 deliveries. The difference in pelvimetry rates between study and control groups in the postintervention period is significant, with 95% confidence limits for the difference between 2.3 and 2.9 pelvimetries per 100 deliveries. Both study and control groups experienced statistically significant declines in pelvimetry use after the educational intervention. That the study group experienced a greater decline in pelvimetry rate than the control group is also statistically significant.

We analyzed the data from the study hospitals in more detail to discover whether the intervention's effectiveness depended on precisely how it was carried out or on the characteristics of the hospitals that experienced it. To do this, we constructed simple regression models of pelvimetry rate over time for the 36-month baseline period for study hospitals in each of the groups listed in Table 3. We then used this model to estimate what the pelvimetry rate in that group of hospitals would have been during the last nine months of the follow-up period had the trend present in the baseline period not been disturbed by the educational intervention. We then compared this estimated rate with the rate actually observed at the end of the follow-up period.



Rates of pelvimetry over time in study and control groups. Solid line represents study group; dashed line, control group.

Table 2.—Results of Educational Program to Decrease Use of Pelvimetry

	Study Group	Control Group
Baseline period		
No. of pelvimetries	9704	9004
No. of deliveries	132 215	117 285
Pelvimetry rate/100 deliveries (mean ± SD)	7.34 ± 0.07*	7.68 ± 0.08
Postintervention period		
No. of pelvimetries	243	724
No. of deliveries	23 028	19 909
Pelvimetry rate/100 deliveries (mean ± SD)	1.06 ± 0.07†	3.64 ± 0.13‡
% Reduction in pelvimetry rate between periods§	86	53

*Difference between study and control groups in baseline period was significant at $P<.001$ (Z test).
†Difference between periods for study group was significant at $P<.00001$ (Z test).
‡Difference between periods for control group was significant at $P<.00001$ (Z test).
§Difference between study and control groups in amount pelvimetry rate reduced between periods was significant at $P<.00001$ (Z test).

Table 3 presents data on the effectiveness of the intervention, taking into account several of its operational characteristics. There was no difference in effect whether or not hospitals provided continuing medical education credits. The program was also equally effective whether presented at a regularly scheduled staff meeting or a specially called one. There was a difference in effect between those hospitals that experienced the educational program at a staff meeting and those in which the materials were mailed to physicians with delivery privileges. The onsite meetings were nearly twice as effective as the mailings. However, hospitals whose staffs received mailings experienced lower pelvimetry rates during the baseline period than hospitals at which meetings were held.

Table 3 also lists some of the characteristics of the study hospitals and how each group responded to the intervention. There was little difference in response based on hospital size, location, or teaching status. This occurred despite the fact that during the baseline period, hospitals that had more

Table 3.—Effectiveness of Intervention in Study Hospitals

	No. of Hospitals	Postintervention Pelvimetry Rate/100 Deliveries*		% Difference†
		Estimated	Actual	
Program characteristics				
Type				
Onsite	48	5.4	1.2	78
Mailed intervention	16	3.6	1.9	47
Meetings				
Regularly scheduled	29	4.6	1.0	78
Special	19	7.4	1.9	74
Continuing medical education credits				
Awarded	39	5.0	1.1	78
Not awarded	9	8.7	2.1	76
Hospital characteristics				
Location				
Urban	28	3.4	1.1	68
Rural	38	7.3	2.4	67
Type				
Teaching	15	2.4	0.9	63
Nonteaching	49	5.8	1.8	69
Obstetrical residency				
Yes	12	2.4	0.8	67
No	52	4.9	1.7	65
% of Obstetricians				
<50‡	31	4.9	1.7	65
≥50	31	3.5	1.2	66
Size				
>100 beds	36	3.8	1.2	68
≤100 beds	28	6.0	2.2	63

*Last nine months of follow-up period.
†All differences were significant at $P<.00001$ (Student's t test).
‡Data unavailable for two hospitals.

than 100 beds, had teaching programs, or were located in urban areas experienced more rapid rates of decline in their pelvimetry use than their smaller, rural, nonteaching counterparts. The effect of the intervention was also unrelated to the proportion of physicians with delivery privileges who were obstetricians.

Finally, although we considered it unlikely, we thought that some patients might be referred to radiologists' offices for pelvimetry if pressure was exerted in the study hospitals to reduce its use. Accordingly, we studied the use of pelvimetry in private radiologists' offices in one of the study PSROs. Ninety-one percent of offices performing pelvimetry agreed to participate. We found very little pelvimetry performed in offices: only 25 in the year before the intervention, compared with 963 in-hospital pelvimetries. During the postintervention period, only three pelvimetries per year were performed in private offices.

COMMENT

We have demonstrated that an educational program followed by feedback of data can markedly improve physician performance in a wide variety of hospital settings. Large reductions in pelvimetry use were recorded in small as well as large, urban as well as rural, and teaching as well as nonteaching hospitals. Similar results were found in sites as widely separated as Arkansas, New Hampshire, Maryland, and North Carolina. Furthermore, the effect of the intervention was long lasting, persisting up to ten months after the conclusion of the feedback period. To our knowledge, this study is the largest randomized controlled trial of physician education performed to date.

The study design did not permit us to achieve perfect comparability between study and control groups. In most important respects, however, the two groups were very similar. The baseline trends in pelvimetry rates were the same (Figure), although the mean baseline pelvimetry rates were very slightly different (Table 2). Most hospital characteristics were represented to the same degree in study and control groups, with the exception of proportion of hospitals with obstetrical residency programs. Nineteen percent of study hospitals had such programs, compared with only 4% of control hospitals. As shown in Table 3, however, the educational program was equally effective in both kinds of hospitals. Thus, the effect we observed was not due simply to a reduction in pelvimetry in hospitals with obstetrical residency programs.

The impact of the educational intervention on quality of care was significant. If one projects the baseline trend in pelvimetry rate for the study hospitals through the end of the study period (Figure) and assumes that the same number of deliveries that took place during the study period would have occurred in its absence, then one can estimate that 3080 pelvimetric examinations would have occurred had the intervention not been carried out. Only 1033 pelvimetries actually took place during this time. Thus, an estimated 2047 fetuses were spared the dose of ionizing radiation that accompanies x-ray pelvimetry as a result of this study.

One important reason that this study demonstrated a larger effect than many others may be the fact that the procedure we studied is generally agreed to be outmoded, with very few remaining indications for use. This consensus notwithstanding, we found pelvimetry performed at unexpectedly high rates, particularly in a subset of high-use hospitals. We believe that this finding may have important implications for other procedures where general agreement may exist as to their ineffectiveness or harmfulness. This study suggests that use of such procedures may be more widespread than is generally appreciated. Further, we have demonstrated that education coupled with feedback of data can be a powerful means through which to reduce the use of such procedures faster than would occur otherwise. For example, again projecting the baseline trend in pelvimetry use in the study hospitals, one can estimate that, in the absence of the study, it would have taken five additional years to attain the same rate of pelvimetry use that was reached by the end of the study period.

We do not believe that the methodology employed in this study need be restricted only to outmoded tests or procedures. As long as agreement

among physicians can be achieved on specific inappropriate uses of a particular health service, we believe that this study's methods may prove efficacious in decreasing the use of that service for those specific reasons.

Several factors may limit the generalizability of our findings. First, the PSROs that implemented this intervention were selected because they were interested in performing utilization review and quality assurance activities in the area of ancillary services. Organizations without similar interest or enthusiasm might not have been as effective.

Second, we studied a procedure with documented, though infrequent, adverse side effects. Physicians might not modify their use of other procedures so readily where the risks to patients and the threat of legal liability are less clear. Third, performing pelvimetry yields no economic reward to physicians ordering the test. The use of other procedures that do might be harder to change.

Our results are consistent with previous work that demonstrates that educational efforts to modify physician behavior are more likely to be successful if they are conducted face-to-face rather than by mail,[39,40] involve physician leaders,[40] and employ timely feedback of data.[40,41]

This study presents an example of a federal health care review program that was successfully tested on an experimental basis before implementation. In most instances, experiments, particularly those using randomized, controlled trial techniques, are superior to the more commonly employed demonstration methods since they allow for more conclusive evaluation of the effectiveness of the program than is ordinarily possible for demonstrations. We believe that experimental evaluation methods should be given greater consideration when new governmental health care programs are being planned or old ones significantly altered.

This study also demonstrates that physician organizations, in this case PSROs, can improve physician performance substantially by means of a carefully designed educational intervention. It seems warranted to pursue the use of this kind of intervention both in programs aimed at improving quality of care by changing physician behavior and in future research. Further studies will be required to define precisely which clinical areas are most amenable to this method and to devise the most cost-effective educational approaches to them.

This work was initiated while Dr Chassin was deputy director of the Office of Professional Standards Review Organizations, Health Care Financing Administration. The study was conducted by the La Jolla Management Corporation, under the direction of Ms McCue, under contract 500-80-0057 from the US Department of Health and Human Services.

We are deeply grateful to the executive directors and project coordinators at each of the six participating PSROs: the Arkansas Foundation for Medical Care, the Baltimore City PSRO, the New Hampshire Foundation for Medical Care, the Passaic Valley PSRO, the Piedmont Medical Foundation, and the Southwestern Pennsylvania PSRO. Without their enthusiasm and perseverance, this study could not have been completed. In addition, we want especially to recognize the contribution of the physicians who participated in the development of the educational materials and those in the study PSROs who conducted the educational programs; their efforts too were indispensable.

The efforts of William Nicholls, the project officer at the Health Standards and Quality Bureau of the Health Care Financing Administration, were equally important in shepherding the project through good times and bad.

Finally, we would also like to thank Robert Brook, MD, ScD, and R. E. Park, PhD, of The Rand Corporation for their perceptive comments on earlier drafts of this article.

References

1. Stross JK, Bole GG: Evaluation of a continuing education program in rheumatoid arthritis. *Arthritis Rheum* 1980;23:846-849.
2. Sibley JC, Sackett DL, Neufeld V, et al: A randomized trial of continuing medical education. *N Engl J Med* 1982;306:511-515.
3. Linn BS: Continuing medical education: Impact on emergency room burn care. *JAMA* 1980;244:565-570.
4. Sanazaro PJ, Worth RM: Concurrent quality assurance in hospital care. *N Engl J Med* 1978; 298:1171-1177.
5. Cohen DI, Jones P, Littenberg B, et al: Does cost information availability reduce physician test usage? *Med Care* 1982;20:286-292.
6. Everett GD, deBlois S, Chang P, et al: Effect of cost education, cost audits, and faculty chart review on the use of laboratory services. *Arch Intern Med* 1983;143:942-944.
7. Martin AR, Wolf MA, Thibodeau LA, et al: A trial of two strategies to modify the test-ordering behavior of medical residents. *N Engl J Med* 1980;303:1330-1336.
8. Schaffner W, Ray WA, Federspiel CF, et al: Improving antibiotic prescribing in office practice. *JAMA* 1983;250:1728-1732.
9. Eisenberg JM, Williams SV, Garner L, et al: Computer-based audit to detect and correct overutilization of laboratory tests. *Med Care* 1977; 15:915-921.
10. Griner PF: Use of laboratory tests in a teaching hospital: Long-term trends. *Ann Intern Med* 1979;90:243-248.
11. Craig WA, Uman SJ, Shaw WR, et al: Hospital use of antimicrobial drugs. *Ann Intern Med* 1978;89:793-795.
12. Geyman JP, Gordon MJ: Learning outcomes and practice changes after a postgraduate course in office orthopedics. *J Fam Pract* 1982;15:131-136.
13. McDonald CJ, Wilson GA, McCabe GP: Physician response to computer reminders. *JAMA* 1980; 244:1579-1581.
14. Lyle CB, Bianchi RF, Harris JH, et al: Teaching cost containment to house officers at Charlotte Memorial Hospital. *J Med Educ* 1979;54:856-862.
15. Wang VL, Terry P, Flynn BS, et al: Evaluation of continuing medical education for chronic obstructive pulmonary diseases. *J Med Educ* 1979; 54:803-813.
16. Eisenberg JM: An educational program to modify laboratory use by house staff. *J Med Educ* 1977;52:578-581.
17. Rhyne RL, Gehlbach SH: Effects of an educational feedback strategy on physician utilization of thyroid function panels. *J Fam Pract* 1979;8:1003-1007.
18. Schroeder SA, Kenders K, Cooper JK, et al: Use of laboratory tests and pharamceuticals. *JAMA* 1973;225:969-973.
19. Brook RH, Williams KN: Effect of medical care review on the use of injections. *Ann Intern Med* 1976;85:509-515.
20. *PSRO: An Initial Evaluation of the Professional Standards Review Organization*, publication 1978-260-943/1048. US Dept of Health, Education, and Welfare, 1978.
21. *Professional Standards Review Organization 1978 Program Evaluation*, publication HCFA-03000. US Dept of Health, Education, and Welfare, 1979.
22. *The Effect of PSROs on Health Care Costs: Current Findings and Future Evaluations.* Congressional Budget Office, 1979.
23. *1979 PSRO Program Evaluation.* US Dept of Health, Education, and Welfare, 1980.
24. Radomsky JW, Radomsky NA: Efficacy of pelvimetry. *J Can Assoc Radiol* 1979;31:43-44.
25. Laube DW, Varner MW, Cruikshank DP: A prospective evaluation of x-ray pelvimetry. *JAMA* 1981;246:2187-2188.
26. Kelly KM, Madden DA, Arcarese JS, et al: The utilization and efficacy of pelvimetry. *Am J Radiol* 1975;125:66-74.
27. Varner MW, Cruikshank DP, Laube DW: X-ray pelvimetry in clinical obstetrics. *Obstet Gynecol* 1980;56:296-300.
28. Gibson RW, Bross IDJ, Graham S, et al: Leukemia in children exposed to multiple risk factors. *N Engl J Med* 1968;279:906-909.
29. Bross IDJ, Natarajan N: Leukemia from low-level radiation: Identification of susceptible children. *N Engl J Med* 1972;287:107-110.
30. MacMahon B: Prenatal x-ray exposure and childhood cancer. *JNCI* 1962;28:1173-1191.
31. Bithell JF, Stewart AM: Pre-natal irradiation and childhood malignancy: A review of British data from the Oxford survey. *Br J Cancer* 1975; 31:271-287.
32. Stewart AM, Kneale GW: Age-distribution of cancers caused by obstetric x-rays and their relevance to cancer latent periods. *Lancet* 1970; 2:4-8.
33. Stewart AM, Kneale GW: Radiation dose effects in relation to obstetric x-rays and childhood cancers. *Lancet* 1970;1:1185-1187.
34. Kneale GW, Stewart AM: Mantel-Haenszel analysis of Oxford data: I. Independent effects of several birth factors including fetal irradiation. *JNCI* 1976;56:879-883.
35. Stewart A, Webb J, Hewitt D: A study of childhood malignancies. *Br Med J* 1958;1:1495-1511.
36. *American College of Obstetrics and Gynecology Newsletter*, 1979;23(October-November):10,11:2.
37. *American College of Radiology Bulletin*, October 1979, p 2.
38. *The Selection of Patients for X-ray Examinations: The Pelvimetry Examination*, publication (FDA) 80-8128. US Dept of Health and Human Services, 1980.
39. Evans CE, Haynes RB, Birkett NJ, et al: Does a mailed continuing education program improve physician performance? Results of a randomized trial in antihypertensive care. *JAMA* 1986;255:501-504.
40. Eisenberg JM: Physician utilization: The state of research about physicians' practice patterns. *Med Care* 1985;23:461-483.
41. Haynes RB, Davis DA, McKibbon A, et al: A critical appraisal of the efficacy of continuing medical education. *JAMA* 1984;251:61-64.