Learning Outcomes and Practice Changes after a Postgraduate Course in Office Orthopedics

John P. Geyman, MD, and Michael J. Gordon, PhD
Seattle, Washington

Although many studies of continuing medical education have demonstrated information transfer and some have resulted in practice changes, little is known about the relationships between physician knowledge, intentions, and actual changes in clinical behavior. This paper presents the results of a study of immediate and delayed learning outcomes, including a parallel exploration of physician intentions and practice changes for three tracer procedures, among 83 physicians attending a postgraduate course in office orthopedics. A high level of knowledge gain and retention was found, and considerable change in practice behavior occurred in self-reported use of torsional profiles, ordering of shoulder arthrograms, and discontinued use of corrective shoes. Discrepancies between physicians' cognitive knowledge, their intentions to apply that knowledge, and reported clinical behavior were consistent with a conservative approach to orthopedic management among the primary care physicians studied.

Despite growing recognition of the need for more effective approaches to continuing medical education (CME), most current CME efforts are principally concerned with information transfer. The limits of such narrow goals for CME, however, have been clearly demonstrated by an increasing number of studies of the process and outcome of CME activities.

In an excellent recent analysis of the effectiveness of CME, Lloyd and Abrahamson assessed the evidence regarding three possible end products of CME: physician competence, physician performance, and patient health status.[1] They found that about one half of the 47 evaluation studies published in the last 20 years showed improvement in at least one of these three outcomes of CME, but they observed that various methodological problems in many of these studies do not allow the conclusion that the improvements were caused by CME.[1]

That a knowledge-based approach to CME is not sufficient can be readily demonstrated by many studies.[2] One example is a report by Ashbaugh and McKean that 94 percent of deficiencies in surgical practice identified by audits of over 5,000 patient records were in the area of performance, while only 6 percent of deficiencies were on the basis of lack of knowledge.[3] The complexity of relationships between CME, physician knowledge and behavior, and patient care outcomes is increasingly well recognized, and many are calling for extension of CME evaluation to include changes in physician behavior.[1,4,5]

Several links can be hypothesized in the chain of influence between a physician's clinical knowl-

From the Department of Family Medicine, School of Medicine, University of Washington, Seattle, Washington. Requests for reprints should be addressed to Dr. John P. Geyman, Department of Family Medicine, RF-30, School of Medicine, University of Washington, Seattle, WA 98195.

0094-3509/82/070131-06$01.50

edge and clinical practice behavior. Among these links are (1) acceptance of the knowledge as valid and useful in the physician's patient care context, and (2) a judgment that the anticipated benefits of a new clinical approach will outweigh the anticipated tangible and intangible costs required to make the change. Both of these conditions and perhaps others are necessary before knowledge can be said to lead to the *intention* to change behavior. Further, several studies have demonstrated that only a portion of individuals who intend to make purposeful changes in their behavior actually do make such changes within months of the decision.[6-8]

This paper presents the results of a study of immediate and delayed learning outcomes, with a parallel exploration of physician behavioral intentions, among physicians attending a five-day CME course in office orthopedics. Self-reported changes in practice behavior for three specific tracer procedures were also related to participants' knowledge and intentions as surveyed on three testing occasions.

Methods

The study sample included all physician enrollees (n = 83) in a CME course on Office Orthopedics for Primary Care Physicians, jointly sponsored by the Departments of Orthopaedics and Family Medicine at Sun Valley, Idaho, between February 23 and March 1, 1980. The course included a combination of didactic lectures and smaller group workshops, both supplemented with audiovisual aids.

Written clinical exercises were completed as part of a 20-minute precourse and postcourse assessment. These exercises required participants to read brief case descriptions and select from a list of possible diagnostic and therapeutic steps those that were judged as appropriate. The exercises were scored as a 46-item true-false test.

In an effort to distinguish between cognitive knowledge and intentions to apply this knowledge in one's practice, physicians with last names from A to L were given "academic" instructions and physicians with last names from M to Z were given "pragmatic" instructions in the precourse test administration. These two sets of instructions were as follows:

"Academic" Instructions. Unless a single best answer is called for, there may be more than one correct answer to these questions. Please indicate your responses by circling the letter opposite each appropriate answer.

"Pragmatic" Instructions. In most areas of medical practice, there is no one way of approaching or managing a given clinical problem. What the individual physician chooses to do varies with the circumstances of each situation, depending on many factors including the physician's training and experience, knowledge of the patient, cost, and related factors. Given these variations, please indicate *your own approach* to each of the following clinical situations. Unless a single best answer is called for, there may be more than one correct answer to these questions. Please indicate your responses by circling the letter opposite each appropriate answer.

At the conclusion of the course, both physician groups were administered the same test but with the alternate set of instructions. Six months after the course, all physician participants were again asked to complete the same exercises, applying the same instructions given to their group at the end-of-course test administration.

The "pragmatic" instructions asked physicians to reveal ways in which they departed from academically defined optimal case management. It was concluded that accurate data of this kind demanded a strong guarantee of anonymity. The alphabetic division of physicians into two groups, while not strictly random, made it possible to maintain accurate group membership on each testing occasion without the need to identify physicians by name. In the analysis the two groups of physicians were treated as randomly equivalent samples.

To further investigate the relationships between knowledge, intentions, and behaviors, three specific clinical actions were selected for study:

1. *Use of torsional profiles* in recording results of the physical examination involving the lower extremities of infants and children presenting with possible deformities (eg, femoral anteversion, tibial torsion, metatarsus varus)
2. *Use of corrective shoes* for management of such deformities
3. *Ordering of arthrograms of shoulder* in patients presenting with recurrent or chronic shoulder pain of undetermined cause

All three clinical activities were systematically

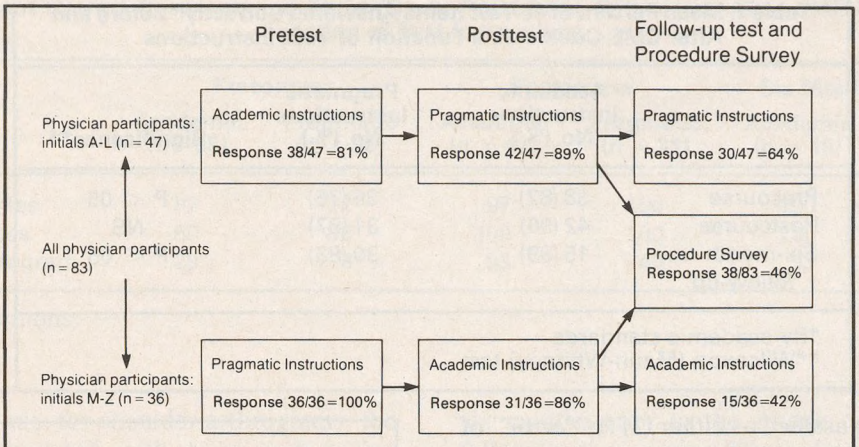

Figure 1. Study design and response rates

addressed in the course presentation and in the clinical exercises. All course participants received instruction in the indications for and methods of torsional profiles. Evidence was clearly presented refuting the clinical utility of corrective shoes for any of the conditions considered. Expanded indications of arthrography of the shoulder were emphasized in order to permit more specific diagnosis and therapy, particularly for recurrent or chronic shoulder pain of unknown cause and for confirmation of certain diagnoses when surgery is being considered (eg, rotator cuff tear).

Questions included in the clinical exercises at the pre- and postcourse test administration and at the six-month follow-up provided information about physician knowledge and intentions regarding these procedures. In addition, a procedure survey was included with the six-month follow-up to estimate the use of these procedures. The procedure survey asked physicians to compare the frequency with which they employed these procedures one year previously (ie, before the course) and at the present time. In each instance the physicians were given three options for the frequency of each procedure: never, occasionally, and usually.

This set of planned study procedures resulted in some puzzling results. Physician subjects were therefore contacted again for more detailed explanations of their seemingly inconsistent responses. The study design and response rates for the study are summarized in Figure 1. The clinical exercises presented under different instructions and on multiple occasions permit comparisons between academic knowledge and behavioral intentions before and after the course. Because of a falloff in response rate at the follow-up administration, the follow-up data cannot be interpreted with the same degree of confidence as the pre- and postcourse assessments.

Results

Table 1 summarizes the scores of respondents on the clinical exercises at the start, conclusion, and at six-month postcourse follow-up. It can be noted that rather high scores were attained under both instruction sets at the beginning of the course, that substantial score gains occurred by the end of the course, and that some erosion of those gains took place later, though not to pretest levels. Of particular interest are the differences observed on each testing occasion between the scores attained under the "academic" and "pragmatic" instructions.

At each test administration physicians following the academic instructions obtained higher scores than those following the pragmatic instructions. These differences were statistically significant for the precourse and follow-up administrations ($P < .05$). Differences cannot be attributed to sampling bias, since the group receiving academic instructions at pretest received pragmatic instructions at the follow-up and vice versa.

The six-month follow-up scores show almost no decrement from postcourse levels under academic instructions but show a statistically significant decrement under pragmatic instructions ($P < .05$).

Under the assumption that the academic responses are the correct ones, each discrepant an-

Table 1. Mean Percent of 46 Test Items Answered Correctly* Before and After CME Course as a Function of Test Instructions

	Academic Instructions No. (%)	Pragmatic Instructions No. (%)	Significance**
Precourse	38 (82)	36 (75)	$P < .05$
Postcourse	42 (90)	31 (87)	NS
Six-month follow-up	15 (89)	30 (83)	$P < .05$

*By academic standards
**Wilcoxon (Mann-Whitney) test

swer could be classified as either (1) an "error" of undermanagement or (2) an "error" of overmanagement. Comparison among the error rates displayed in Figure 2 shows a consistent pattern in which errors of undermanagement are more commonly made than errors of overmanagement at each administration. In fact, the lower test scores obtained under the pragmatic instructions are accounted for almost exclusively by errors of undermanagement. Especially intriguing is the finding that errors of undermanagement among physicians following either instruction set declined at the immediate postcourse administration only to reappear at six-month follow-up. In contrast, errors of overmanagement declined at the postcourse administration and declined even further over the long term.

Table 2 displays precourse, postcourse and six-month follow-up results on the clinical exercises related specifically to the three tracer procedures. This table shows gains in academic knowledge during the course to nearly the 100 percent level and recall of virtually all of the new cognitive knowledge at the follow-up testing six months later. Responses to the three tracer procedures were more frequently "in error," by academic standards on all three testing occasions for physicians responding to pragmatic instructions. Pragmatic responses did change from pretest to six-month follow-up in a direction consistent with academic advice, but for all three procedures these changes were smaller under pragmatic instructions than under academic instructions.

Precourse knowledge and pragmatic intentions regarding use of torsional profiles and use of corrective shoes were already close to academic standards. Therefore, only limited increases in scores were possible on these items.

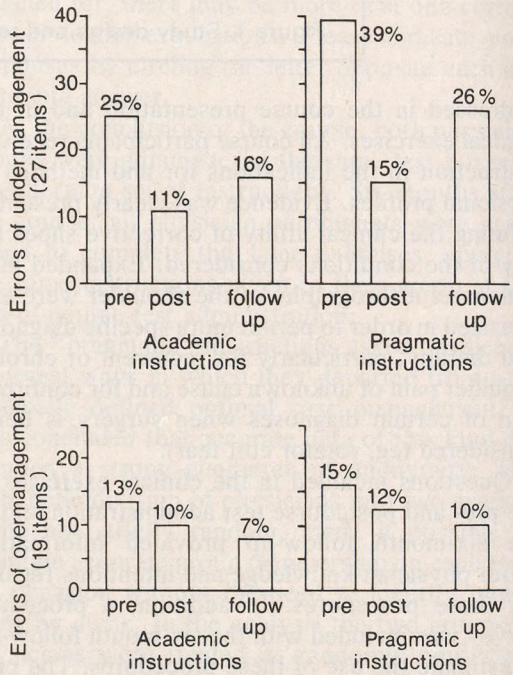

Figure 2. Percent errors of undermanagement and overmanagement under academic and pragmatic test instructions on three test administrations

In contrast, academic knowledge of a correct indication for a shoulder arthrogram was infrequent at the precourse testing (29 percent correct) and dramatically changed at the immediate postcourse and six-month follow-up test administrations (93 percent correct on both occasions). Under pragmatic instructions, only 6 percent of respondents reported an intention to order a shoulder arthrogram for the condition described. Immediately following the course, this percentage rose to 40 percent and remained at 37 percent six

Table 2. Percent Correct Responses on Three Tracer Procedures Before and After CME Course as a Function of Test Instructions

Instructions	Precourse		Postcourse		Six-Month Follow-Up	
	Academic (n =38)	Pragmatic (n = 36)	Academic (n = 42)	Pragmatic (n = 31)	Academic (n = 15)	Pragmatic (n = 30)
Torsional profiles	82	72	97	100	100	87
Corrective shoes	80	86	100	100	100	90
Shoulder arthrogram	29	6*	93	40**	93	37**

z test of proportions:
*P < .02
**P < .001

months later. Thus, for shoulder arthrograms, the course was associated with large and lasting changes of both knowledge and behavioral intentions but also with a puzzling discrepancy between retained knowledge and intentions to apply this knowledge.

To investigate this discrepancy, all enrollees were again contacted by mail. They were presented with the case described in previous tests, informed of the discrepancy found, and were asked several questions to elicit their management approach and rationale. Responses from 30 attendees were useful in understanding the apparent discrepancy. With respect to the care of a hypothetical patient with several months of gradually increasing shoulder pain without a previous history of shoulder trauma, arthritis, or related medical problems, three quarters of respondents considered the shoulder arthrogram to be a potentially useful diagnostic test but only in limited circumstances. None would use this study early in the workup of such a patient. All would first employ a full trial of conservative therapy (physical therapy and/or anti-inflammatory drug therapy), reserving arthrography for persistent and refractory shoulder pain when rotator cuff tear could be a serious diagnostic possibility. Some respondents would obtain consultation at that point and leave the decision for or against a shoulder arthrogram to the consulting orthopedic surgeon. Hence, even in this instance, there was considerable evidence that the CME experience increased the awareness of the potential role of shoulder arthrography, even though the respondent's pragmatic responses were less dramatic in this direction.

Figure 3 depicts the physician's self-reported behavioral changes in the use of the three tracer procedures at two points in time: approximately six months prior to the course, and six months following the course. These data provide some evidence that changes in cognitive knowledge and behavioral intentions regarding a procedure may lead to changes in physician behavior. The self-reported behavioral changes must be regarded with some caution, however, for several reasons: the data are retrospective, asking participants to describe behavior of a year ago; the categories of "never," "occasionally," and "usually" are crude and undefined; and responses to the procedural survey are strongly cued by the simultaneous administration of clinical exercises associated with the CME course taken six months previously.

Comment

It would be surprising to find that a five-day CME course did not result in some increased knowledge for participants. This course did demonstrate an overall gain in knowledge by attendees; in addition, the course was associated with a lesser change in pragmatic views about the applicability of what was formerly "known." More significantly, this study demonstrates that changes in physicians' cognitive knowledge may not lead to immediate behavioral change because that knowledge is either insufficiently compelling to the practicing physician or is outweighed by other factors.

Analysis of "errors" of under- and overmanagement (Figure 2) indicates that practicing physicians appeared to adhere quite closely to academic advice to avoid ineffective diagnostic or therapeutic interventions, but they seemed less inclined to include certain additional steps academics have determined should be useful. This finding again points to the practicing physician's independent evaluation of knowledge received in CME rather than his passive acceptance of it.

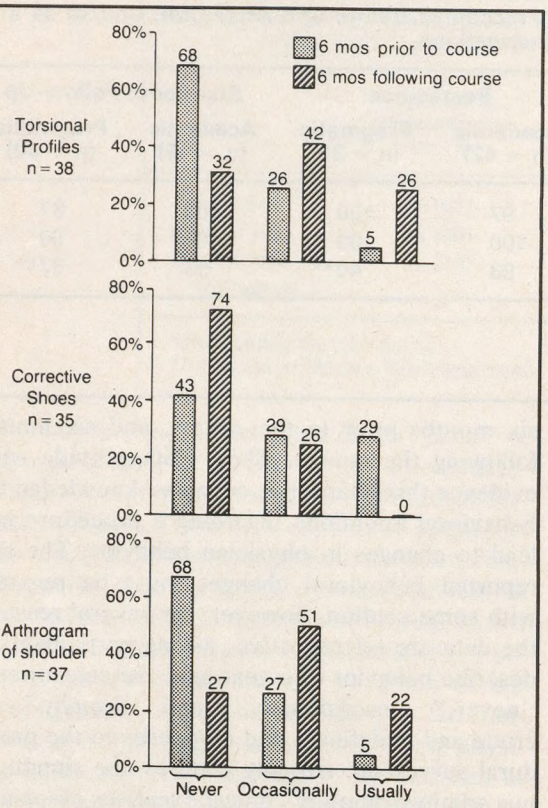

Figure 3. Percent of participants employing three tracer procedures six months prior to and six months following CME course

The analysis of the three tracer procedures leads to the conclusion that there was no significant forgetting of new knowledge between the postcourse test and the follow-up six months later, but that for each procedure, some physicians seem to have reconsidered the applicability of the information and here returned to a mode of practice closer to their precourse approach.

The apparent conservatism of physicians attending this CME course in actually applying new information, as exemplified by the shoulder arthrogram, bears further comment. It is likely that they remain unconvinced that rotator cuff tears are sufficiently common *in their practices* to warrant the cost and risks of this procedure without a full course of conservative management. This seems to illustrate the classic dilemma for many common problems encountered daily in primary care practices—the epidemiology and natural history of common diseases are still relatively unknown in *the general population* compared with the knowledge of diseases seen in referral practices and specialty clinics. That this conservatism is well grounded until further research is done in primary care settings is supported by the excellent analysis recently reported by Ellenberg and Nelson of the differences of the natural history of febrile seizures in population-based studies and hospital/referral clinic studies.[9] They found that population-based studies were in close agreement in showing low rates of unfavorable sequelae, whereas clinic-based studies showed wide variation and often high frequencies of recurrence of seizures.

There are undoubtedly many reasons why physicians' practice behaviors may not change fully in response to newly acquired medical knowledge. These may include force of habit, continued bias, medical fashion, economic disincentives, skepticism, time constraints, pressures by patients, and inadequate effort in revising one's practice methods. In general, some of these barriers (eg, economic disincentives) are not likely to have been responsible for the results noted here, and evidence has been found that the appropriate exercise of discretion also plays a role. Despite the evidence that continuing medical education apparently can have measurable effects on knowledge, intention, and behavior, greater understanding of the process by which new information is translated into practice will probably come from closer examination of the apparent inconsistencies between physicians' knowledge, intentions, and behavior.

References

1. Lloyd JS, Abrahamson S: Effectiveness of continuing medical education: A review of the evidence. Evaluation Health Professions 2(3):251, 1979

2. Stein LS: The effectiveness of continuing medical education: Eight research reports. J Med Educ 56:103, 1981

3. Ashbaugh DG, McKean RS: Continuing medical education: The philosophy and use of audit. JAMA 236:1485, 1976

4. Council on Health Manpower, Council on Medical Eduation, and Council Medical Services: Continuing Competence of Physicians: A Progress Report (Board of Trustees Report D). Chicago, American Medical Association, 1974

5. Report of the Second National Invitational Conference of the Alliance for Continuing Medical Education: Priorities for Action. Darien, Conn, Miller and Fink Corp, 1977

6. Beach LR, Davidson AJ: Error patterns in the prediction of fertility behavior. J Appl Soc Psychol (in press)

7. Davidson AJ, Jaccard J: Variables that moderate the attitude-behavior relationship. J Pers Soc Psychol 37:1364, 1979

8. Westoff C, Ryder N: The national fertility study: The predictive value of reproductive intentions. Demographics 14:431, 1977

9. Ellenberg JH, Nelson KB: Sample selection and the natural history of disease: Studies of febrile seizures. JAMA 243:1337, 1980