Changing Physician Performance: A Systematic Review of the Effect of Continuing Medical Education Strategies

Davis, David A. MD; Thomson, Mary Ann BHSc; Oxman, Andrew D. MD; Haynes R. Brian, PhD MD

From the Faculty of Medicine, University of Toronto (Ontario) (Dr Davis); the Faculty of Health Sciences, McMaster University, Hamilton, Ontario (Ms Thomson and Dr Haynes); and the National Institute of Public Health, Oslo, Norway (Dr Oxman).

Abstract

Objective: To review the literature relating to the effectiveness of education strategies designed to change physician performance and health care outcomes.

Data Sources: We searched MEDLINE, ERIC, NTIS, the Research and Development Resource Base in Continuing Medical Education, and other relevant data sources from 1975 to 1994, using continuing medical education (CME) and related terms as keywords. We manually searched journals and the bibliographies of other review articles and called on the opinions of recognized experts.

Study Selection. We reviewed studies that met the following criteria: randomized controlled trials of education strategies or interventions that objectively assessed physician performance and/or health care outcomes. These intervention strategies included (alone and in combination) educational materials, formal CME activities, outreach visits such as academic detailing, opinion leaders, patient-medicated strategies, audit with feedback, and reminders. Studies were selected only if more than 50 percent of the
subjects were either practicing physicians or medical residents.

Data Extraction: We extracted the specialty of the physicians targeted by the interventions and the clinical domain and setting of the trial. We also determined the details of the educational intervention, the extent to which needs or barriers to change had been ascertained prior to the intervention, and the main outcome measure(s).

Data Synthesis: We found 99 trials, containing 160 interventions, that met our criteria. Almost two thirds of the interventions (101 of 160) displayed an improvement in at least one major outcome measure: 70 percent demonstrated a change in physician performance, and 48 percent of interventions aimed at health care outcomes produced a positive change. Effective change strategies included reminders, patient-mediated interventions, outreach visits, opinion leaders, and multifaceted activities. Audit with feedback and educational materials were less effective, and formal CME conferences or activities, without enabling or practice-reinforcing strategies, had relatively little impact.

Conclusion: Widely used CME delivery methods such as conferences have little direct impact on improving professional practice. More effective methods such as systematic practice-based interventions and outreach visits are seldom used by CME providers.

(JAMA. 1995;274:700-705)

IN MOST developed nations, health care systems are facing important changes, driven by political and economic forces and by the general concern for competent and consistent quality care. [1] Responses to these forces include altering the training of physicians across the continuum of medical education. Undergraduate medical education reform, stressing interdisciplinary care, problem-based learning, increased critical appraisal of research evidence, and community-based training, is part of the response to health care reform. [2] Graduate education reforms have also begun to focus on increased community-based experiences, with an emphasis on generalist education. These changes in medical education are aimed at producing a new generation of clinicians whose practice more completely reflects the goals of health care reform. [3]

Even if effective, changes aimed at undergraduates and residents necessarily demand a lengthy incubation period because of the slow percolation of graduates into the health care system. The contributions of continuing medical education (CME), a major facilitator of change in practitioner behavior, are undermined by difficulties in CME delivery, which is often disjointed, driven by the largesse of drug companies and the requests of physicians, and plagued by doubts about and relatively little evidence for its own effectiveness and efficacy. [4,5] Similar deficiencies can be found in the adoption of new information and techniques. For example, despite convincing evidence about the effectiveness of thrombolytic therapy in the management of myocardial infarction, even current texts and review articles frequently fail to reflect this treatment. [6] Thus, the effectiveness of educational methods aimed at improving physician performance and the health status of their patients appear inadequate to respond to the urgent demands of health care reform. [7]

Keeping in mind the context of larger reform issues, this review of educational interventions
as tools in a CME system has the following objectives. First, it attempts to increase the knowledge base about the effectiveness of broadly defined educational interventions targeted at changing the performance of practicing physicians and the health status of their patients, to build on earlier work in this area [4,5] and a recent publication [8] that includes nonphysician health professional educational interventions and refines methodological practices in the search and review process, to extend others' similar syntheses, [9-11] and to describe which educational interventions have been tested in what populations of physicians in which settings and practice domains. Second, we analyze the extent to which preintervention design elements affect outcomes (for example, assessment of practice needs or analysis of barriers to change). Third, we attempt to further the theoretical knowledge base of changing physician behavior. Finally, we outline areas in which knowledge is as yet weak or inconclusive, thus shaping a research and development agenda.

METHODS

We used search and data extraction strategies similar to those of our earlier reviews [4,5,8] and briefly outlined here.

Data Sources

From 1975 to 1994, we searched the computer bibliographic databases of MEDLINE on a monthly basis, and those of ERIC, NTIS, CINAHL, HEALTHLINE, and EMBASE semiannually to select trials, studies, and descriptions in the broad domain of continuing health professional education, using keywords described in an earlier publication. [4] These keywords were combined with a sensitive search strategy [12] to identify randomized controlled trials (RCTs). In addition, manual searches of journals and the bibliographies of retrieved articles and the input of recognized experts in the field were used to retrieve articles. These articles were then reviewed by a health sciences librarian and/or a member of the study team, additional keywords or descriptors were added, and a summary was entered into a computerized and retrievable format using Reference Manager (Research Information Systems, Carlsbad, Calif). This final common source is called the Research and Development Resource Base in CME (RDRB-CME). [13]

Selection Criteria

To study the effects of the CME interventions that had been most rigorously tested and to gain as much comparability across studies as possible, we retrieved from the RDRB-CME those studies that met the following criteria: (1) RCTs or trials that used alternatives for randomization, such as alternating allocation to either control or experimental arms; (2) replicable educational interventions directed at changing clinical behavior or health outcomes; (3) study populations of health professionals in which the majority were either practicing physicians or medical residents; and (4) objective measurements of physician performance or health care outcomes. Inclusion criteria were pilot tested by all authors on five studies and then applied by at least two of three of us (M.A.T., A.D.O., D.A.D.) independently.

Educational Interventions
We defined an educational intervention as any attempt to persuade physicians to modify their practice performance by communicating clinical information. We excluded strategies that were purely administrative (such as the reorganization of charts in a computerized or problem-based format) or applied incentives or coercion (financial rewards or sanctions). The details of the educational interventions were modified from a recent study [8] and included the following: (1) educational materials (including noninteractive printed, audiovisual, and computer-produced information); (2) formal CME programs such as conferences, seminars, workshops, small group sessions, traineeships (in-depth, mini-residencies, or preceptorships), and teleconferences; (3) outreach visits, including academic detailing (visits by physician educators such as pharmacists) or counter-detailing directly to physicians by pharmacists and others, or practice facilitation by nurses and other professionals; (4) local opinion leaders or educational influentials; (5) patient-mediated interventions such as patient education materials; (6) audit with feedback, including chart review with peers or supervisors; and (7) reminders. We also included combinations of these activities.

Data Extraction

Several categories of data were extracted: (1) the discipline of the targeted physician population, (2) the intended behavioral change, (3) the setting of the trial, (4) the degree to which a learning or performance gap had been identified by the authors, (5) the intervention, as outlined previously, and (6) the clinical setting of the trial. Further, (7) we attempted to determine the degree to which each intervention was planned or based on an analysis of practice deficiency or need. Specifically, we assessed whether national policy or guidelines had been developed and used; if local consensus had been sought, using the target audience or representatives of the target audience in planning; whether barriers to change were identified and targeted by the intervention; and finally, if gaps in performance were determined by the study authors, followed by interventions targeted at practitioners whose performance was judged to be suboptimal. These last data extraction criteria were developed by one of us (D.A.D.) and pilot tested by another (M.A.T.) on five studies, then applied in all trials.

Finally, the outcome measures used by the study authors were extracted, and a determination was made regarding the degree to which each study was positive, negative, inconclusive, or mixed. Positive studies were those that demonstrated a statistically significant change in at least one major outcome measure at the level of P<.05; negative studies showed no important change, despite sufficient statistical power to do so; inconclusive studies failed to demonstrate a change but lacked the statistical power to demonstrate or exclude a clinically important difference; and mixed studies were those that improved physician performance but not health care outcomes. These data were extracted by one of the authors (D.A.D. or M.A.T.) and checked by at least one of the other authors.

RESULTS

At the time of preparation of this review, the RDRB-CME contained more than 6000 articles and monographs on the broadly defined topic of CME. Ninetynine studies met our criteria for inclusion and form the substrate for this review. [14-112] Within these 99 studies, there were 160
separate interventions apart from placebo or nonintervention controls.

Types of Physician Studies

Fifty-one of the 99 studies included internists, 35 focused on family physicians or general practitioners, and 15 trials did not specify physician type or background. Smaller numbers of trials studied obstetrician/gynecologists (n=5), pediatricians (n=4), emergency physicians (n=3), surgeons (n=2), and anesthesiologists (n=2). Many studies included more than one category of physician. Residents were a part of the study physician population in one third (n=34) of the studies, either in conjunction with practicing physician-supervisors or by themselves.

Settings

Three quarters (n=75) of the studies took place in outpatient settings: 27 in private office practices and 48 in ambulatory care centers or clinics. Of the latter, 35 authors described their settings as teaching centers, nine as managed care settings, [17,19,35,71,74,84,86,110,111] and three as veterans centers [54,70,92]; one study failed to characterize its location precisely. Twenty-six trials took place either in hospitals (20 on wards, [14,20,27,31,37,41,56,57,80,85,89,91,93-96,105,106,108,112] three in emergency departments, [51,78,79] two on labor floors, [28,53] and one in an intensive care unit [77]) or in the long-term care setting. [16] Two of the trials each took place in two settings. [85,91]

Interventions

Of the 160 noncontrol interventions in these studies, 99 (62 percent) showed an improvement in at least one major outcome measure in physician performance or health care; 53 (33 percent) failed to demonstrate a change in either of these areas. Eight interventions (5 percent) displayed mixed results. [16,25,35,36,49,74,109] Many interventions analyzed both physician performance and health care outcomes. Of the 145 interventions that focused on physician performance, 101 (70 percent) demonstrated change in at least one major measure; 44 (30 percent) did not. Of the 46 interventions targeted at changing health care outcomes, 22 (48 percent) succeeded in doing so. The size of the effects in positive studies was small to moderate. The impact of specific interventions is described by single, two-method, and multifaceted interventions.

Single Interventions

Eighty-one single-intervention strategies were used in these trials, 49 (60 percent) of which demonstrated change in at least one major outcome measure, 30 (37 percent) failed to do so, and two studies [74,109] effected change in physician performance without altering health care outcomes.

Outreach visits, including academic detailing, demonstrated positive changes in interventions in which they were used alone. [15,92] Academic detailers in single-intervention and other studies included pharmacists, [15,84] standardized patients, [83] study investigators, [81] or nurse-facilitators. [39] Opinion leader strategies displayed positive effects in a few of the outcomes measured in the three trials in which they were used as a single intervention. [93-95]
mediated methods, such as patient educational materials or patient reminders, showed a positive change in seven interventions [21,54,68,75,87,103] and negative effects in two. [67]

Physician reminders was an effective single-method intervention: 22 of 26 interventions effected change, [17,22,26,29,34,52,63-66,71,75,78,79,86,87,98-102,104] and four interventions in two studies failed to do so. [67,108]

More variable results were displayed by audit with feedback: here, we found 10 positive outcomes [19,31,42,46,58,59,70,72,85,99] and 14 negative outcomes. [19,38,41,80,85,92,97,111,112] Outcomes from audit with feedback were more consistently effective when feedback was delivered in the form of chart review (eg, in studies by Everett et al, [41] Martin et al, [57] Pinkerton et al, [76] and Restuccia [85]).

Relatively short (1 day or less) formal CME events such as conferences generally effected no change; six interventions demonstrated negative or inconclusive effects. [23,38,44,45,50,96] One study, [74] using an in-depth traineeship or mini-residency, demonstrated changes in physician performance but not health care outcomes. Educational materials (eg, printed monographs or audiovisual programs) demonstrated a positive effect in four interventions, [55,58,77] but failed to demonstrate an effect in seven instances. [15,30,32,41,82,83,90]

Two-Method Interventions

Thirty-nine interventions used two educational methods, and 25 (64 percent) of these were positive, [27,30,33,37,39,41,48,55,57,58,65,68,70,83,84,89,96,102,106,110] 12 (31 percent) were negative or inconclusive, [19,33,38,40,44,47,60,76,82,97,105,107] and three (<1 percent) yielded mixed results. [16,49] When effective single methods, as described previously, were used in pairs, positive changes were generally noted. Examples included patient-mediated strategies combined with reminders, which demonstrated positive changes in three studies [27,68,102] and negative results in one. [105] In contrast, the combination of two less effective interventions, formal CME activities and educational materials, resulted in seven positive [33,37,39,48,55,106] and five negative or inconclusive studies. [19,33,60,76,82]

Multifaceted Interventions

Thirty-nine interventions used three or more educational strategies. Of these, 31 (79 percent) were positive, [14,18,20,26,28,33,39,44,53,56,60-62,69,73,77,81,91,103,109] including positive changes in health care outcomes. [103,109] A further five (13 percent) were negative or inconclusive, [23,43,53,88] and three (8 percent) [25,35,36] displayed mixed results.

Effects of Needs Analysis

Several levels of needs analysis were reported in these studies, indicating in progressive fashion the degree to which study authors considered or defined a clinical need, performance gap, or barrier to performance change. First, 12 studies (<1 percent) did not mention the clinical need on which the intervention was based; five (42 percent) of these were positive studies. Second, about one fifth (n=34) of the interventions provided one to several references in clinical care and identified a general clinical area requiring change; 18 (53 percent) of these showed a positive
change, 14 (41 percent) did not, and two (6 percent) displayed mixed results. Third, one quarter (n=41) of the interventions were based on already-developed clinical guidelines generally approved by a national body, producing a positive change in 25 (61 percent) of these 41 interventions, negative results in 15 (37 percent), and mixed results in one instance. Fourth, consensus processes to achieve agreement on the part of local health professionals was used by roughly one fourth (n=45) of the interventions, effecting a positive change in 26 (58 percent), no change in 16 (36 percent), and positive physician performance but no corresponding health care outcome change in three instances (6 percent). Finally, less than one fifth (n=28) of the interventions were targeted at specific behaviors identified by a gap analysis technique (such as an audit to determine suboptimal performers) or addressed specific barriers to change; 25 (89 percent) of these interventions exhibited a positive change, three (11 percent) did not.

Analysis by Domain of Behavior Change Targeted

Interventions in these studies targeted two broad domains of behavioral change: roughly one third (n=54) in the area of health promotion and disease prevention and two thirds (n=107) in clinical disease management. One study [48] involved two areas of behavioral change.

The category of disease prevention and health promotion yielded two subcategories, prevention and screening, in which there were 36 interventions, and lifestyle education and counseling (including smoking cessation), which accounted for 18 interventions. The former subcategory yielded 30 positive results [21,26,29,34,39,48,52,59,66,68,69,73,75,87,98,99,102,104,110] and six negative or inconclusive results. [67,97] Examples of effective interventions here included patient-mediated strategies, [21,87] reminders, [26,29,34,52,66,67] and multifaceted interventions. [69] Of the 18 lifestyle or educational counseling interventions, 10 produced positive results, [30,33,49,55,83,109] five were negative, [30,33,50,83] and three were mixed. [35,36,109] Among the effective interventions, the use of peer discussion (eg, Davidoff et al [37]) and "practice rehearsal" (eg, Kottke et al [49]), allowing physicians to rehearse certain communication skills in the setting of a formal CME event, produced positive results.

Under the category of clinical management, four interventions were found in the area of diagnosis. Of these, two were positive [78,79] and two negative. [88,89] Reminders appeared as effective change agents. [78,79] We found 24 interventions in the area of resource utilization, principally radiology and laboratory use, 17 of which produced significant change in outcomes [19,20,28,31,37,41,56-58,77,85,100,101] and seven did not. [19,41,85,112] Reminders [100,101] and audit with feedback (either alone [31,58] or with another intervention such as a conference [57]) appeared to be useful measures here, as did multifaceted interventions. [20,28,56,77] In the category of prescribing, there were 14 interventions, three quarters (n=11) of which were positive. [15,16,22,42,46,62,70,84,91,92] An effective change agent in this category was the academic detail visit. [15,16,62] Finally, interventions to produce enhanced procedural skills demonstrated negative results in three of four interventions. [24,74] Audit with feedback was ineffective in one trial [24] using three interventions, while a clinical traineeship produced a change in the rate of sigmoidoscopy in one study. [74]

We found 58 interventions that attempted to alter outcomes in general medical management,
32 (55 percent) of which achieved their goal and 26 (45 percent) did not. Examples of effective strategies include reminders, [63,64,71] two-method interventions (eg, reminders and patient-mediated strategies [27]), or intensive conferencing methods and educational materials [48] and multifaceted interventions. [14,23,51,60,61] In the subcategory of hypertension, many interventions yielded negative results. [38,40,43,82,11]

**COMMENT**

The evidence provided by these trials supports some clear conclusions. First, there is a robust body of research assessing the outcomes of physicians' clinical education, as evidenced by several thousand extant articles, including increasing numbers of the most scientifically rigorous variety, the RCT. In addition, this research has grown substantially over the years: in 1984, seven articles were found matching our criteria [4]; in 1992, a further 43 trials [5]; and, 3 years later, another 49. Second, while definitional problems still exist, there is a growing convergence about what constitutes educational interventions, including a broadening operational definition of the term CME. It is apparent that ongoing physician education may now comprise a variety of measures well beyond the traditional short-course model; academic detailing, opinion leaders, audit with feedback, reminders, and other systems may be components of broadly defined CME. Third, this review confirms the major finding of our 1992 review, [5] namely, physician performance may be altered by many of these CME interventions and, to a lesser extent, so may health care outcomes. These alterations are most often small, less often moderate, and rarely large.

The reason changes in health care outcomes lag behind those in physician performance have been explored elsewhere [5,8] and include patients' not accepting physician recommendations, the socioeconomic and educational status of patients, and, frequently, the limited effectiveness of the clinical interventions themselves.

**Study Limitations**

In interpreting this review, readers should note that some publications may have escaped our search strategy and the literature may not contain important findings derived from negative trials. Further, when reporting is incomplete or vague, data extraction becomes problematic, as was frequently the case in these studies. For example, many studies did not indicate the method of randomization and most did not describe in replicable detail the exact nature of the educational intervention. Because the interventions themselves were of such variability and their participants, settings, targeted behaviors, and outcomes were so varied, estimates of the average effect from meta-analysis would have little practical meaning. Similarly, between-trial comparisons of different interventions are of limited value. [113] Importantly, among "positive" studies are many in which only one or two outcome measures of several are positive. This review made no attempt to summarize the size of effects. Our focus on RCTs systematically excluded the results of qualitative research methods that would afford additional valuable insights into physician behavioral change.

**Comparisons With the Health Care Environment**
Do these trials represent actual physician educational experiences and settings? To a great extent they do not. Neither the complexity and method of the educational intervention nor the optimal circumstances provided in teaching centers or wards in which many of these studies were conducted match the learning and practice circumstances of most physicians. In addition, many of these trials experience a ceiling effect in studying volunteer physicians who are often performing at or near optimum levels. In this group, there may be only room for small changes that are difficult to detect. Moreover, these studies often use end points or outcome markers that may be selected more for their ease of determination (eg, blood pressure) than their clinical imperative, such as heart failure or stroke. Lastly, the types of physicians selected as the focus of these studies do not match the spectrum of today's US physicians. Few if any studies focused on the skills of surgeons, the clinical work of psychiatrists, the function of nephrologists, or other subspecialists. In contrast, the trials did represent prevention and screening, the prime initiatives of health care reform.

The Science and Practice of CME

In a manner analogous to the evidence-based care movement, this review begins to apply scientific and theoretical principles to the area of health professional performance change, and it begins to indicate a movement toward "evidence-based CME."

Particularly promising among the effective interventions are reminders, academic detailing, and patient-mediated strategies. This effect may support the PRECEDE model of Green et al, which suggests that the interventions that best succeed in changing performance and health care outcomes are those using practice-enabling strategies (office facilitators or patient educational methods, for example) or reinforcing methods (feedback or reminders) in addition to predisposing or disseminating strategies. To some extent, the effectiveness of opinion leaders may also support the extensive work of Rogers showing that the presence of an opinion leader permits colleagues to observe the outcomes of a particular innovation in an atmosphere conducive to its adoption on a community basis. Formal CME activities that include peer discussion and role-playing would appear to support the theories of Nowlen, who stresses the role of the group in adopting new information, and Bandura, whose Social Learning Theory stresses the importance of personal, environmental/situational, and behavioral factors. The sequenced provision of multifaceted interventions may mimic the process outlined by Fox et al in which learning and change take place through a series of "impactors" or learning resources.

The effect of the interventions is not, however, consistent across practitioners, settings, or behaviors. At least part of the variation may be attributable to an individual readiness to change: physician-learners may progress at their own rates, depending on their motivation, their knowledge of a problem, or the perception of a gap between current knowledge and skills and those needed. This was borne out by our needs analysis: when barriers to change were addressed or gaps were demonstrated and resources deployed to help the learner, change appeared to occur relatively frequently.

Whatever the theoretical basis of CME, there is clear evidence that its current provision in North America consists mostly of the less effective change strategies such as conferences.
compared with vanishingly small but more effective methods such as academic detailing or reminder systems. To base CME interventions on identified clinical needs, however, new linkages for CME providers will need to be found—in health services research, in hospitals, in provincial or state-generated data sources, from insurance carriers, and within managed care systems.

A full research agenda awaits the expanding interdisciplinary field of CME. This agenda needs to consider further intervention innovations and attend more closely to rigorous evaluations combined with more qualitative methods to elucidate physicians' perceptions regarding learning and change. To advance the process of describing research findings related to CME, more consistent reporting and design strategies and up-to-date systematic reviews of trials that address specific questions are required, along with collaboration between CME researchers and providers such as the Cochrane Collaboration on Effective Professional Practice. Additionally, there is a need for more coherent research approaches that build on prior work and a focused effort to disentangle the interplay of change, setting, type of health professional, and intervention, perhaps following the model of multicenter trials. Such cooperative evidence-based approaches will advance the effort to improve the performance of physicians and the health outcomes of their patients.

The RDRB-CME is supported in part by the Alliance for Continuing Medical Education, Northbrook, Ill, the Society of Medical College Directors of Continuing Medical Education, Washington, DC, and the University of Toronto Faculty of Medicine.

We are indebted to Anne Taylor-Vaisey, MLS, continuing education librarian at the University of Toronto Faculty of Medicine, for her preparation of this review's database.

Corresponding author: Office of Continuing Education, Faculty of Medicine, University of Toronto, 150 College St, Toronto, Ontario, Canada M5S 1AB (Dr Davis).

REFERENCES AND NOTES


24. Buntinx F. Does feedback improve the quality of cervical smears? a randomized controlled trial. Br J


52. Litzelman DK, Dittus RS, Miller ME, Tierney WM. Requiring physicians to respond to computerized reminders improves their compliance with preventive care protocols. J Gen Intern Med. 1993;8:311-317. [Context Link]


90. Sibley JC, Sackett DL, Neufeld V, Gerrard B, Rudnick KV, Fraser W. A randomized trial of continuing...


100. Tierney WM, McDonald CJ, Hui SL, Martin DK. Computer predictions of abnormal test results: effects on outpatient testing. JAMA. 1988;259:1194-1198. Bibliographic Links | [Context Link]


112. Wones RG. Failure of low-cost audits with feedback to reduce laboratory test utilization. Med Care. 1987;25:78-82. Bibliographic Links


Accession Number: 00005407-199509060-00017