Original Research

A Meta-Analysis of Continuing Medical Education Effectiveness

MALIHEH MANSOURI MD;JOCELYN LOCKYER PHD.

Introduction: *We undertook a meta-analysis of the Continuing Medical Education (CME) outcome literature to examine the effect of moderator variables on physician knowledge, performance, and patient outcomes.*

Methods: *A literature search of MEDLINE and ERIC was conducted for randomized controlled trials and experimental design studies of CME outcomes in which physicians were a major group. CME moderator variables included the types of intervention, the types and number of participants, time, and the number of intervention sessions held over time.*

Results: *Thirty-one studies met the eligibility criteria, generating 61 interventions. The overall sample-size weighted effect size for all 61 interventions was r = 0.28 (0.18). The analysis of CME moderator variables showed that active and mixed methods had medium effect sizes (r = 0.33 [0.33], r = 0.33 [0.26], respectively), and passive methods had a small effect size (r = 0.20 [0.16], confidence interval 0.15, 0.26). There was a positive correlation between the effect size and the length of the interventions (r = 0.33) and between multiple interventions over time (r = 0.36). There was a negative correlation between the effect size and programs that involved multiple disciplines (r = −0.18) and the number of participants (r = −0.13). The correlation between the effect size and the length of time for outcome assessment was negative (r = −0.31).*

Discussion: *The meta-analysis suggests that the effect size of CME on physician knowledge is a medium one; however, the effect size is small for physician performance and patient outcome. The examination of moderator variables shows there is a larger effect size when the interventions are interactive, use multiple methods, and are designed for a small group of physicians from a single discipline.*

Key Words: *Continuing Medical Education (CME), physician knowledge, physician performance, patient health, meta-analysis*

Introduction

Increasing attention is being paid to continuing medical education (CME) as a mechanism to improve physician and patient outcomes,[1] with CME being described as "any and all the ways by which doctors learn after formal completion of their training."[2] According to Fox and Bennett,[3] CME is the systematic attempt to facilitate change in physicians' practices. In most developed countries, mandatory requirements for participation in CME started in the 1970s in the belief that if physicians were up-to-date, they could change and improve their practice, resulting in better physician performance and ultimately better patient care. On average, health professionals spend between 1 and 3 weeks per year at educational meetings.[4,5]

Despite tremendous efforts, the goals of CME to ensure that physicians are up-to-date and that the quality of health care delivery improves have not been met.[6,7] Studies continue to demonstrate considerable gaps between the real and ideal performance and patient outcomes. For example, studies in the United States[8] and the Netherlands[9] suggest that about 30% to 40% of patients do not receive care according to present scientific evidence. Furthermore, as much as 25% of care provided is not needed or is potentially harmful. This gap between the real and ideal performance of the health care system increases uncertainty about the role of CME.

Different scholars have acknowledged that CME outcomes can range along a continuum of participation, satisfaction, learning, performance, patient health, and population health.[10] Increasingly, there is pressure for CME outcomes to go beyond increasing knowledge and skills to improving

Dr. Mansouri: PhD candidate, Haskayne School of Business, University of Calgary; *Dr. Lockyer:* associate dean, Continuing Medical Education and Professional Development, and associate professor, Department of Community Health Sciences, University of Calgary, Calgary, Alberta.

Correspondence: Maliheh Mansouri MD, Haskayne School of Business, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada, T2N 1N4; e-mail: maliheh.mansouri@haskayne.ucalgary.ca.

 DOI: 10.1002/chp.088

physician competence and performance in practice and lead to better patient health. Although several systemic reviews have been conducted to measure the effectiveness of CME, most of the studies have failed to show a significant association between CME and health outcomes. For example, Beaudry's meta-analysis[11] of 41 studies found a strong effect size (d = standardized mean difference) between CME and knowledge ($d = 0.79$), a moderate effect size ($d = 0.55$) between CME and physician performance, and a small effect size ($d = 0.37$) between CME and patient health status. Davis and colleagues[12] in their systemic review identified 99 studies involving CME interventions and found that multifaceted interventions were more likely to be successful. They found that widely used CME delivery methods such as conferences had little direct effect on improving professional practice. Pippalla and colleagues' meta-analysis[13] of 43 studies showed that the combination of active and passive strategies resulted in a better effect on physicians' prescribing behavior than other interventions. Davis and associates' 7-study meta-analysis[7] failed to show a significant effect of didactic educational methods on the outcomes of CME ($d = 0.34$); however, they did find that interactive and mixed educational sessions were associated with a significant effect on physicians' performance ($d = 0.67$).

The results of CME systematic reviews and meta-analyses have led to speculation that moderator variables (eg, course length) may affect CME outcomes. Although these studies have focused on comparisons between different delivery methods (eg, didactic lectures vs. videoconferences), they have paid little attention to other aspects of CME despite the fact that CME interventions vary by content, the number and type of participants, the degree and type of interaction within the CME event, and the length and frequency of these interventions.

The purpose of this study is to determine the effect of moderator variables on CME outcomes. We specifically wanted to (1) determine the effect size of CME interventions on physician knowledge, physician performance, and patient outcomes and (2) examine the effect of moderator variables on the effectiveness of CME. From these findings, we suggest further avenues for research.

Methods

According to Hunter and Jackson (1982), "meta-analysis is a statistical technique that can be used to integrate research findings across studies."[14] Through meta-analysis, it is possible to examine the primary effects between the independent (CME interventions) and dependent variables (CME outcomes) as well as the effect of moderator variables. Moderator variables are those that influence the strength of the relationship between the independent and dependent variables.

Systematic database searches for our meta-analysis were conducted electronically using MEDLINE and the Educational Resources and Information Center (ERIC) because they are the most commonly used in medicine and educational research. We searched for CME outcome studies published between 1990 and 2004. The studies were required to meet 4 main criteria:

- Have a study design that was either a randomized controlled trial or before-and-after experimental designs of CME interventions in which participants were practicing physicians
- Focus on at least 1 of the 3 identified outcome effects of physician knowledge, physician performance, or patient outcome
- Provide an adequate description of the CME intervention, such as the type of intervention, the participants, and the duration of the intervention
- Be based on study design that reports quantitative analyses, including sample size and *t* scores, *d* scores, or F scores.

Our search terms included *continuing medical education*, *continuing education*, *education*, and *professional development* without restriction to language. The search strategy was supplemented by cross-reference checking for other articles.

Reviews of the CME literature suggest that there are 3 outcomes in CME: physician knowledge, physician performance, and patient outcomes. Variation in the results of the CME literature lead us to consider the role that moderator variables may be playing. Accordingly, we suggest that several moderator variables may affect the efficiency and effectiveness of CME (TABLE 1).

Each study was coded based on the independent variables (CME interventions), the 4 potential groups of moderator variables, and 3 groups of dependent variables or outcomes (physician knowledge, physician performance, and patient outcomes). To increase the reliability of the coding process, double coding by the original coder and random checks by another coder were used.

Our approach to the meta-analysis closely followed the method used by Hunter and Schmidt in which the mean effect size for a set of primary studies is estimated as well as the percentage of unreliability variance, usually sampling error and unreliability.[15] Sampling error is the only variance that is considered in this study because only 5 studies reported reliability estimates.[16–19] Mean effect sizes are expressed as correlations (*r*), consequently requiring the conversion of *t* scores, *d* scores, and F scores to *r*. The Pearson correlation effect size measures the correlation between the independent variable and the scores on the dependent variables.[20] Using the formula for transformation given by Rosenthal (p. 240, Table 16.1)[21] and Lipsey and Wilson (p. 201, Table B11),[22] we converted the scores to the Pearson correlation.

Confidence intervals (CIs) refer to the precision with which the expected mean effect is measured. Consistent with the random effects model, the heterogeneous form is employed here, which takes into account uncertainty in the distribution of observed effects.[23] Consequently, quality weightings based on study sample size were used as an adjustment mechanism in computing an average weighted effect size.[24] Because studies differ from one another in

DOI: 10.1002/chp

TABLE 1. Variables That May Affect CME Effectiveness

Type of interventions	Active interventions (eg, workshops and individual training) Passive interventions (eg, conferences and print-only interventions) Mixed interventions in which combinations of active or passive interventions, or both (eg, small-group activity and post-course feedback), are used
Participant types and numbers	Single discipline (eg, general practitioners only) Multiple disciplines (eg, general practitioners and pediatricians, physicians, and nurses) Number of participants
Time	Contact time (length of intervention) Assessment time (time interval between the educational intervention and the measurement of its effect)
Single versus multiple sessions	Single session (intervention takes place only once) Multiple sessions (interventions take place longitudinally)

many methodological and substantive ways, appropriate weights by study sample size were calculated to minimize the variance.[25] Such a weighting assumes that studies with larger samples have a smaller variance and in turn are better estimates of effect size. Shadish and Haddock[25] termed this a quality rating. We computed an estimate of the "fail-safe N" to assess the file drawer problem, which assumes that journals contain 5% of the studies that show type I errors, whereas the file drawers in research laboratories are filled with 95% of studies showing nonsignificant results.[26] According to Smith and Glass,[27] the standard of importance in meta-analysis is demonstrated by any positive effect size greater than 0.10. Cohen's[28] criteria for evaluating the magnitude of Pearson correlation effect sizes are broader: 0.10 = small, 0.24 = medium, and 0.37 = large. This study follows Cohen's criteria and classifies the outcomes as positive if 1 or more of the primary outcome measures related to physician knowledge, physician performance, or patient health care are demonstrated by a statistically significant change with a moderate or large effect size and negative if no such change occurred. Results in this study are reported as pooled effect sizes with 95% confidence intervals.

Results

Overall Effects of CME Interventions

The literature search, covering 1990 to 2004, identified a total of 3,233 articles. Of these, 83 articles initially appeared to meet the inclusion criteria. An examination of these articles for references missed by the database search provided 28 more articles. Of these 111 articles, 80 were eliminated because they did not meet all of the inclusion criteria. Thirty-one studies met all the inclusion criteria. Because multiple reports of single studies were treated separately, the 31 studies generated 61 interventions meeting study criteria (TABLE 2).

A simple count determined that 4 studies reported a negative effect size between the CME intervention and some outcomes.[29–32] The rest of the studies reported a moderate to large positive effect size. The overall sample-size weighted mean effect size for all 61 interventions was $r = 0.28$ (0.18) (95% CI, 0.20–0.39). A summary of the results, the descriptive characteristics of studies, and the effect size are presented in TABLE 3.

An analysis of the data to determine the mean effect size of the 61 CME interventions and the 3 different groups of outcomes (dependent variables) revealed the following:

- The mean effect size between CME and the physician knowledge over 15 studies was close to moderate: $r = 0.22$ (0.16); 95% CI, 0.15–0.28.[17–19,29,30,32–41]
- The mean effect size between CME and the physician performance over 19 studies was small: $r = 0.18$ (0.21); 95% CI, 0.08–0.28.[16,19,29–34,36,42–51]
- The mean effect size between CME and patient outcomes over 8 studies was small: $r = 0.14$ (0.21); 95% CI, 0.31–0.63.[17,18,36,46,52–55]

Type of Interventions

Three types of interventions were examined: active, passive, and mixed. The mean effect size for the active methods was $r = 0.33$ (0.33) (CI, 0.13–0.50); for passive interventions was $r = 0.20$ (0.16) (CI, 0.15–0.26); and for mixed interventions was $r = 0.33$ (0.26) (CI, 0.24–0.43). A comparison of mixed and single methods of intervention showed that the mixed methods are associated with a larger effect size ($r = 0.33$ [0.26]; CI, 0.24–0.43) than are single methods (like conference or online education), where the effect size was $r = 0.24$ (0.25) (CI, 0.12–0.35).

A further examination of the specific types of intervention is presented in TABLE 3. The largest effect sizes were found with multifaceted educational programs, longitudinal workshops, interactive small groups, and case discussion interventions. The lowest effect sizes were

DOI: 10.1002/chp

TABLE 2. Interventions Meeting Study Criteria

	Study	Sample Size	Population Type	Intervention	Contact Time, hours	Outcome	Assessment Time, Weeks	Effect Size
1	Bexell et al., 1996[16]	26	Single	Conference	21	Physician performance	12	0.10
2	Brown et al., 1999[17]	69	Mixed	Workshop	8	Physician knowledge	12	0.16
3	Brown et al., 1999[17]	69	Mixed	Workshop	8	Patient outcomes	12	−0.24
4	Butler et al., 1997[30]	15	Single	Conference	4	Physician knowledge	6	−0.01
5	Butler et al., 1997[30]	15	Single	Conference	4	Physician performance	6	0.06
6	Carney et al., 1995[33]	65	Mixed	Online education	2	Physician performance	9	0.26
7	Carney et al., 1995[33]	20	None	Lecture	None	Physician knowledge	9	0.31
8	Clark et al., 1998[53]	74	Single	Conference	5	Patient outcomes	22	0.25
9	Coonrod et al., 2000[34]	102	Single	Video	0.33	Physician performance	52	0.19
10	Coonrod et al., 2000[34]	102	Mixed	Video	0.33	Physician knowledge	52	0.32
11	Dolan et al., 1997[18]	96	Single	Lecture	2	Physician knowledge	Immediately	0.18
12	Dolan et al., 1997[18]	96	Single	Lecture	2	Patient outcomes	4	0.22
13	Doucet et al., 1998[35]	38	Single	Problem-based	6	Physician knowledge	Immediately	0.73
14	Doucet et al., 1998[35]	21	Single	Problem-based	6	Physician knowledge	12	0.52
15	Doucet et al., 1998[35]	49	Single	Lecture	2	Physician knowledge	Immediately	0.63
16	Doyne et al., 2004[42]	60	Mixed	Multifaceted educational program	None	Physician performance	52	0.20
17	Fender et al., 1999[51]	348	Single	Multifaceted educational program	None	Physician performance	52	0.22
18	Gask et al., 1998[36]	20	Single	Interactive small group	10	Physician knowledge	Immediately	0.37
19	Gask et al., 1998[36]	20	Single	Interactive small group	10	Physician performance	Immediately	0.28
20	Gask et al., 1998[36]	20	Single	Interactive small group	10	Patient outcomes	Immediately	0.36
21	Gerstein et al., 1999[29]	290	Single	Workshop	7	Physician knowledge	5	0.12
22	Gerstein et al., 1999[29]	290	Single	Workshop	7	Physician performance	5	0.25
23	Gerstein et al., 1999[29]	193	Single	Workshop	7	Physician performance	60	−0.36
24	Hendryx et al., 1998[43]	122	Mixed	Problem-oriented educational outreach visits	10	Physician performance	52	0.02
25	Hergenroeder et al., 2002[37]	33	Single	Video	0.33	Physician knowledge	20	0.40
26	Hergenroeder et al., 2002[37]	42	Single	Multifaceted educational program	2	Physician knowledge	16	0.66
27	Labelle et al., 2004[38]	24	Single	Workshop	17	Physician knowledge	52	0.17
28	Labelle. et al., 2004[38]	16	Single	Repeated workshop	24	Physician knowledge	52	0.46
29	Maiuro et al., 2002[49]	99	Mixed	Online education	2	Physician performance	3	0.27
30	Majumdar et al., 2003[54]	39	Single	Multidisciplinary educational outreach	24	Patient outcomes	27	0.14

(continued on next page)

TABLE 2. Continued

	Study	Sample Size	Population Type	Intervention	Contact Time, hours	Outcome	Assessment Time, Weeks	Effect Size
31	Markert et al., 2003[39]	558	Mixed	Conference	8	Physician knowledge	Immediately	0.44
32	Markert et al., 2003[39]	109	Mixed	Conference	40	Physician knowledge	Immediately	0.22
33	Mc Donald et al., 1984[44]	80	Single	Online education	None	Physician performance	108	0.04
34	Model et al., 1998[45]	80	Mixed	Lecture	3	Physician performance	52	0.85
35	Nilsson et al., 2001[31]	40	Single	Multifaceted educational program	4	Physician performance	52	–0.02
36	Pazirandeh et al., 2000[55]	134	Mixed	Lecture	None	Patient outcomes	27	0.02
37	Roter et al., 1998[46]	15	Single	Lecture	8	Physician performance	None	0.47
38	Roter et al., 1998[46]	15	Single	Lecture	8	Patient outcomes	None	0.49
39	Sanci et al., 2000[41]	108	Single	Multifaceted educational program	17	Physician knowledge	55	0.86
40	Sanci et al.. 2000[41]	108	Single	Multifaceted educational program	17	Physician knowledge	30	0.85
41	Sanci et al., 2000[41]	108	Single	Multifaceted educational program	17	Physician performance	30	0.85
42	Scardovi et al., 2003[40]	9	Single	Multifaceted educational program	36	Physician knowledge	8	0.73
43	Sharp et al., 2002[47]	58	Mixed	Conference	7	Physician performance	Immediately	0.15
44	Sharp et al., 2002[47]	58	Mixed	Conference	7	Physician performance	13	0.09
45	Socolar et al., 1998[19]	147	Mixed	Detailed comment and written feedback	10	Physician performance	27	0.17
46	Socolar et al., 1998[19]	147	Mixed	Detailed comment and written feedback	10	Physician knowledge	27	0.02
47	Van Eijk et al., 2001[50]	43	Mixed	Individual training	0.33	Physician performance	16	0.32
48	Van Eijk et al., 2001[50]	41	Mixed	Interactive small group	0.33	Physician performance	16	0.29
49	Veninga et al., 1999[32]	181	Single	Auditing and peer group discussions	4	Physician knowledge	27	–0.04
50	Veninga et al., 1999[32]	181	Single	Auditing and peer group discussions	4	Physician performance	27	0.19
51	Veninga et al., 1999[32]	204	Single	Auditing and peer group discussions	4	Physician knowledge	27	–0.05
52	Veninga et al., 1999[32]	204	Single	Auditing and peer group discussions	4	Physician performance	27	0.09
53	Veninga et al., 1999[32]	199	Single	Auditing and peer group discussions	4	Physician performance	52	0.46
54	Veninga et al., 1999[32]	199	Single	Auditing and peer group discussions	4	Physician knowledge	52	0.35
55	Veninga et al., 1999[32]	81	Single	Auditing and peer group discussions	4	Physician knowledge	27	0.34
56	Veninga et al., 1999[32]	81	Single	Auditing and peer group discussions	4	Physician performance	27	0.65
57	Weller et al., 2003[48]	46	Single	Outreach visit	None	Physician performance	27	0.41
58	Weller et al., 2003[48]	47	Single	Mail-out strategy	None	Physician performance	27	0.15
59	White et al., 1995[52]	103	Single	Lecture	2	Physician knowledge	Immediately	0.96
60	White et al., 1995[52]	103	Single	Lecture	2	Physician knowledge	27	0.70
61	White et al., 1995[52]	103	Single	Lecture	2	Patient outcomes	27	0.28

TABLE 3. Effects of Different Kinds of CME Interventions on Outcomes

Type of Treatment		Overall Effect sizes for Each Outcome		
Type of CME	Articles	Physician Knowledge	Physician Performance	Patient Outcomes
Auditing and peer group discussions	Veninga et al. [32]	$r = 0.01$ (0.06) (CI, –0.01–0.03)	$r = 0.24$ (0.20) (CI, 0.19–0.30)	
Case-based training	Doucet et al. [35]	***r* = 0.64 (0.08)** (CI, 0.54–0.73)		
Conference and lecture	Markert et al. [39]; Sharp et al. [37]; Dolan et al. [18]; Bexell et al. [16]; Clark et al. [53]; Butler et al. [30]; Doucet et al. [35]; Pazirandeh[55]; Modell et al. [45]; Roter et al. [46]; Carney et al. [33]; White et al. [52]	$r = 0.22$ (0.38) (CI, 0.13–0.31)	$r = 0.06$ (0.15) (CI, 0.02–0.09)	$r = 0.04$ (0.10) (CI, 0.01–0.07)
Detailed comment and written feedback	Socola et al. [19]	$r = 0.02$ (0.05) (CI, –0.09–0.13)	$r = 0.17$ (0.05) (CI, 0.10–0.24)	
Individual training	van Eijk et al. [50]		***r* = 0.32 (0.04)** (CI, 0.24–0.38)	
Interactive small group	Gask et al. [36]; van Eijk et al. [30]	***r* = 0.44 (0.13)** (CI, 0.36–0.51)	$r = 0.13$ (0.15) (CI, 0.06–0.19)	***r* = 0.35 (0.09)** (CI, 0.30–0.40)
Mail-out strategy	Weller et al. [48]		$r = 0.01$ (0.04) (CI, –0.01–0.04)	
Multidisciplinary educational outreach visits	Majumdar et al. [54]; Weller et al. [48]; Hendryx et al. [43]		$r = 0.02$ (0.06) (CI, –0.01–0.04)	$r = 0.02$ (0.04) (CI, –0.01–0.04)
Multifaceted educational program	Fender[52]; Majumdar et al. [34]; Hergenroeder et al. [37]; Nilsson et al. [31]; Scardovi et al. [40]; Doyne et al. [42]	***r* = 0.69 (0.34)** (CI, 0.59–0.79)	$r = 0.04$ (0.11) (CI, 0.00–0.08)	
Online education	Maiuro,et al. [49]; Carney et al. [33]		$r = 0.18$ (0.13) (CI, 0.11–0.24)	
Repeated workshop	Labelle et al. [38]	***r* = 0.61 (0.06)** (CI, 0.35–0.88)		
Video	Coonrod et al. [34]; Hergenroeder et al. [37]	$r = 0.05$ (0.13) (CI, 0–0.09)	$r = 0.02$ (0.05) (CI, –0.02–0.05)	
Workshop	Labelle et al. [38]; Gerstein et al. [29]; Brown et al. [17]	$r = 0.03$ (0.07) (CI, 0.02–0.05)	$r = 0.01$ (0.09) (CI, –0.02–0.05)	

Note: The *r* values in bold type indicate statistical significance.

found with conferences and lectures, mail-out strategies, and videotapes.

Participant Types and Number

The examination of participant type (eg, single discipline and multiple disciplines) showed a greater effect size for single disciplines. The mean effect size for interventions with participants from multiple disciplines was small: $r = 0.13$ (0.16); CI, 0.12–0.21. The mean effect size for an intervention with single-discipline participants was medium: $r = 0.30$ (0.27); CI, 0.16–0.32. There was also a negative correlation between the number of participants and effect size: $r = -0.13$.

Time

The contact time for each continuing education program ranged from 0.33 to 36 hours, with a mean of 7.95 contact hours. More than half of the programs were 4 hours or less. The correlation between contact hours and the general effect size was positive: $r = 0.33$.

The assessment time between the educational intervention and measurement of its impact was examined because it determines the sustainability of the learned behavior. The assessment times ranged from immediately after conducting a program to 52 weeks after the intervention. The correlation between the effect sizes and the length of time for outcome assessment was negative ($r = -0.31$). Separate correlations between the assessment interval and the dependent

DOI: 10.1002/chp

variables provide additional clarification about the role played by the assessment interval:

- The correlation between physician knowledge and the measurement interval was positive: $r = 0.04$.
- The correlation between the physician performance and the measurement interval was negative: $r = -0.34$.
- The correlation between patient outcomes and the measurement interval was negative: $r = -0.44$.

Single Versus Multiple Sessions Over Time

The correlation between the number of sessions held, and general effect size was positive: $r = 0.36$.

File Drawer Analysis

We computed an estimate of the fail-safe N to assess the file drawer problem (described earlier).[26] We found that 26 unreported studies averaging a null result would have to exist somewhere before the overall results of the current meta-analysis could be reasonably ascribed to sampling bias.

Discussion and Conclusion

Our meta-analysis examined 31 studies involving 61 CME interventions. We extended the systemic reviews and meta-analyses conducted by others. We believe that our findings are valid. The data are based on a large number of participants. The data have relatively narrow confidence intervals. Finally, there is a moderate file drawer requirement. Our work shows the following:

- The effect of CME on physician knowledge ($r = 0.22$), physician performance ($r = 0.18$), and patient outcome ($r = 0.14$) was small to moderate.
- The type of interventions, types and number of participants, the length of the intervention, and holding multiple sessions over time were all found to mediate the effects of CME on its outcomes.
- The time between the intervention and the assessment was negatively correlated with the effect size.

Comparing the overall effect size in this study ($r = 0.28$) shows that it is consistent with other work[7] but suggests a lower effect size than others found.[13] When we separated CME outcomes into the 3 different outcomes, we found a moderate effect size between CME and physician knowledge and a small effect size between CME and physician performance. We also obtained a small effect size between CME and patient outcomes over 8 studies that reported these outcomes.

Our examination of moderators provides new information about the role of these on physician and patient outcomes following a CME program. An examination of the first group of moderator variables, the types of interventions, shows that CME effectiveness increases as the intervention strategy becomes more active, which is consistent with results from other reviews.[2,12,13,56,57] Similar to previous studies, the results of our meta-analysis similarly confirmed that traditional passive approaches to CME are not associated with changes in physician performance or patient outcome, although they may increase knowledge and awareness of issues.[2,12] The results suggest that a combination of different interventions results in a better effect size than other methods. This finding is also consistent with those outcome studies that used mixed educational methods.[13,58–60] These results reinforce the need to adopt new, creative, efficient, and interactive approaches with more than 1 method when attempting to change physician behavior. Planning for physicians as adult learners needs to recognize that physicians learn independently of their teachers and in ways that are aligned with their previous experience.[61] As Slotnick[62] has noted, adults generally and physicians specifically learn in response to problems they perceive they have. The finding that a greater effect size is associated with interactive interventions such as case-based training, interactive small groups, and multifaceted educational program reinforces the need for adult learning approaches in CME.

Participant type and number were the second group of moderator variables. We found that CME interventions designed and run for single groups of participants (eg, 1 discipline) were associated with better outcomes. By way of explanation, it is possible that interventions with only a single group of participants (eg, pediatricians) are more focused and present materials that are more relevant to the practitioner group. The finding that there is an association between group size and outcomes is likely linked to the increased opportunity for active participation and for obtaining information directly linked to practice concerns.

The third moderator variables were time. This meta-analysis suggests that increasing the length of the intervention will have a positive effect on the results of CME. We did not get a significant correlation between CME outcomes and the measurement interval finding that as the time interval between the course and the measurement of outcomes increases, the effect size decreases. This suggests that new behaviors may need reinforcement for sustainability. It may also suggest that there is an optimal and realistic length of time to assess an intervention.

Our final moderator variable examined the impact of single versus multiple sessions of CME held over time. We found that educational interventions with a longer contact time and continuing contact were associated with a larger effect. Beer and colleagues[63] found the same results when they used training interventions to improve the quality of their technical centers. They found that there was a complex pattern in which change started and stopped. But each time the change restarted, they were further along the continuum of adoption, and participants returned toward a previous stage but with a higher level of awareness and readiness for change.[64] Using the result of this meta-analysis and Beer

DOI: 10.1002/chp

Lessons for Practice

- CME interventions are likely to have a small to moderate effect on physician knowledge, physician performance, and patient outcomes.
- Judicious use of moderator variables in course planning can improve the effect size of the CME intervention. The use of active and interactive teaching methods versus passive methods, education for a single group versus multiple groups, smaller versus larger groups, longer versus shorter sessions, and increasing the number of sessions all increase the effect size.

and colleagues' findings, we conclude that if CME interventions are continuous and periodic, with new interventions, there will be more chance for behavior change to occur.

Overall, although the results of the present study indicate that the relationship between CME and change in physician performance and patient outcomes are small and not always sustained, our examination of moderator variables suggests that the addition of specific known and proven moderator variables will improve the effects of CME.

There are limitations to the present meta-analysis. Meta-analysis combines data from studies that can be different from one another.[65] The technique considers only relatively direct evidence on a given topic and cannot be used to infer which characteristics of the studies on a given topic caused the differing results.[66] Our search was limited to articles in major journals, which generally publish studies with significant results; therefore, this may have generated a biased sample of studies.[24]

In many of the studies coded, specific demographic information and proportions were not provided on variables such as the type of participants and measurement interval. More original research is needed to determine the effect of potentially important moderator variables such as duration of intervention as well as more detailed and comprehensive reporting of the results. Based on the attempts to apply the techniques of meta-analysis to the CME literature, several recommendations for future evaluation in this field seem practical. It is suggested that individuals who plan, implement, and report evaluation studies in CME should provide the following elements:

- Detailed descriptive information about the characteristics of the participants and the continuing education program or activity
- Information about sample size, means, and standard deviations of comparison and experimental groups using accepted standards for reporting[67]
- Numeric data when graphs are used to display results
- Validity and reliability data about the measurement tools used for data collection and outcome assessment
- Data about the period of time between the continuing education intervention and the measurement of performance
- Information about other moderator variables such as personal and environmental factors that may also influence the effectiveness of CME.

References

1. Bennett N, Davis D, Easterling W, Friedmann P, Green J, Koeppen B, et al. Continuing medical education: A new vision of the professional development of physicians. Acad Med. 2000;75(12):1167–1172.
2. Davis D. Does CME work? An analysis of the effect of educational activities on physician performance or health care outcomes. Int J Psychiatry Med. 1998;28(1):21–40.
3. Fox R, Bennett N. Continuing medical education: Learning and change: Implications for continuing medical education. BMJ. 1998;316:466–468.
4. Goulet F, Gagnon R, Desrosiers G, Jacques A, Sindon A. Participation in CME activities. Can Fam Physician. 1998;44:541–548.
5. Nylenna M. Primary care physicians and their information-seeking behaviour. Scand J Primary Health Care. 2000;18(1):9–13.
6. Chan D, Leclair K, Kaczorowski J. Problem-based small-group learning via the Internet among community family physicians: A randomized controlled trial. MD Comput. 1999;16(3):54–58.
7. Davis D, O'Brien M, Freemantle N, Wolf F, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: Do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? JAMA. 1999;282(9):867–874.
8. Schuster M, McGlynn E, Brook R. How good is the quality of health care in the United States? Milbank Q. 1998;76(4):517–563.
9. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care. 2001;39(8 Suppl 2):II46-II54.
10. Davis D, Barnes M, Fox R. The Continuing Professional Development of Physicians: From Research to Practice. Chicago: American Medical Assocation Press; 2003.
11. Beaudry J. The effectiveness of continuing medical education: A quantitative synthesis. J Contin Educ Health Prof. 1989;9(4):285–307.
12. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance: A systematic review of the effect of continuing medical education strategies. JAMA. 1995;274(9):700–705.
13. Pippalla R, Riley D, Chinburapa V. Influencing the prescribing behavior of physicians: A meta-evaluation. J Clin Pharm Ther. 1995;20(4):189–198.
14. Hunter J, Schmidt F, Jackson G. *Meta-Analysis: Cumulating Research Findings Across Studies*. Beverly Hills, CA: Sage, 1982.
15. Hunter J, Schmidt F. *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage, 2004.
16. Bexell A, Lwando E, von Hofsten B, Tembo S, Eriksson B, Diwan V. Improving drug use through continuing education: A randomized controlled trial in Zambia. J Clin Epidemiol. 1996;49(3):355–357.
17. Brown J, Boles M, Mullooly J, Levinson W. Effect of clinician communication skills training on patient satisfaction: A randomized, controlled trial. Ann Intern Med. 1999;131(11):859–860.
18. Dolan N, Ng J, Martin G, Robinson J, Rademaker A. Effectiveness of a skin cancer control educational intervention for internal medicine housestaff and attending physicians. J Gen Intern Med. 1997;12(9):531–536.
19. Socolar R, Raines B, Chen-Mok M, Runyan D, Green C, Paterno S. Intervention to improve physician documentation and knowledge of child sexual abuse: A randomized, controlled trial. Pediatrics. 1998;101(5):817–824.

DOI: 10.1002/chp

20. Rosnow R, Rosenthal R. Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. Psychol Methods. 1996;1(4):331–340.
21. Rosenthal R. Parametric measures of effect size. In: Cooper H. and Hedges LV, Editors, *The handbook of research synthesis:* New York, Russell Sage Foundation, 1994:231–244.
22. Lipsey M, Wilson D. Practical Meta-Analysis. Thousand Oaks, CA: Sage; 2001.
23. Whitener E. Confusion of confidence intervals and credibility intervals in meta-analysis. J Appl Psychol. 1990;75(3):315–321.
24. Rosenthal R. Writing meta-analytic reviews. Psychol Bull. 1995; 118(2):183–192.
25. Shadish W, Haddock C. Combining estimates of effect size. In: Cooper H. and Hedges LV, Editors, *The handbook of research synthesis:* New York, Russell Sage Foundation, 1994:261–281.
26. Rosenthal R. *Meta-Analytic Procedures for Social Research:* Newbury Park, CA: Sage Publications, 1991.
27. Smith M, Glass G. Meta-analysis of psychotherapy outcome studies. Am Psychol. 1977;32(9):752–760.
28. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Erlbaum; 1988.
29. Gerstein HC, Reddy SS, Dawson KG, Yale JF, Shannon S, Norman G. A controlled evaluation of a national continuing medical education programme designed to improve family physicians' implementation of diabetes-specific clinical practice guidelines. Diabet Med. 1999;16(11): 964–969.
30. Butler R, Collins E, Katona C, Orrell M. Does a teaching programme improve general practitioners' management of depression in the elderly? J Affect Disord. 1997;46(3):303–308.
31. Nilsson G, Hjemdahl P, Hässler A, Vitols S, Wallén NH, Krakau I. Feedback on prescribing rate combined with problem-oriented pharmacotherapy education as a model to improve prescribing behaviour among general practitioners. Eur J Clin Pharmacol. 2001;56(11):843–848.
32. Veninga C, Wahlstrom R, Muskova M, Denig P, Berkhof J, Kochen M, et al. Evaluating an educational intervention to improve the treatment of asthma in four European countries. Am J Respir Crit Care Med. 1999;160(4):1254–1262.
33. Carney P, Dietrich A, Freeman D Jr, Mott L. A standardized-patient assessment of a continuing medical education program to improve physicians' cancer-control clinical skills. Acad Med. 1995;70(1):52–58.
34. Coonrod D, Bay R, Rowley B, Del Mar N, Gabriele L, Tessman T, et al. A randomized controlled study of brief interventions to teach residents about domestic violence. Acad Med. 2000;75(1):55–57.
35. Doucet M, Purdy R, Kaufman D, Langille D. Comparison of problem-based learning and lecture format in continuing medical education on headache diagnosis and management. Med Educ. 1998;32(6):590–596.
36. Gask L, Usherwood T, Thompson H, Williams B. Evaluation of a training package in the assessment and management of depression in primary care. Med Educ. 1998;32(2):190–198.
37. Hergenroeder AC, Chorley JN, Laufman L, Fetterhoff A. Two educational interventions to improve pediatricians' knowledge and skills in performing ankle and knee physical examinations. Arch Pediatr Adolesc Med. 2002;156:225–229.
38. Labelle M, Beaulieu M, Renzi P, Rahme E, Thivierge R. Integrating clinical practice guidelines into daily practice: Impact of an interactive workshop on drafting of a written action plan for asthma patients. J Contin Educ Health Prof. 2004;24(1):39–49.
39. Markert R, O'Neill S, Bhatia S. Using a quasi-experimental research design to assess knowledge in continuing medical education programs. J Contin Educ Health Prof. 2003;23(3):157–161.
40. Scardovi A, Rucci P, Gask L, Berardi D, Leggieri G, Ceroni G, et al. Improving psychiatric interview skills of established GPs: Evaluation of a group training course in Italy. Fam Pract. 2003;20(4):363–369.
41. Sanci LA, Coffey CM, Veit FC, Carr-Gregg M, Patton GC, Day N, et al. Evaluation of the effectiveness of an educational intervention for general practitioners in adolescent health care: Randomised controlled trial. BMJ. 2000;320:224–230.
42. Doyne EO, Alfaro MP, Siegel RM, Atherton HD, Schoettker PJ, Bernier J, et al. A randomized controlled trial to change antibiotic prescribing patterns in a community. Arch Pediatr Adolesc Med. 2004;158(6):577–583.
43. Hendryx M, Fieselmann J, Bock M, Wakefield D, Helms C, Bentler S. Outreach education to improve quality of rural ICU care. Am J Respir Crit Care Med. 1998;158(2):418–423.
44. McDonald C, Hui S, Smith D, Tierney W, Cohen S, Weinberger M, et al. Reminders to physicians from an introspective computer medical record: A two-year randomized trial. Ann Intern Med. 1984;100(1):130–138.
45. Modell M, Wonke B, Anionwu E, Khan M, Tai SS, Lloyd M, et al. A multidisciplinary approach for improving services in primary care: Randomised controlled trial of screening for haemoglobin disorders. BMJ. 1998;317(7161):788–791.
46. Roter D, Rosenbaum J, de Negri B, Renaud D, DiPrete-Brown L, Hernandez O. The effects of a continuing medical education programme in interpersonal communication skills on doctor practice and patient satisfaction in Trinidad and Tobago. Med Educ. 1998;32(2):181–189.
47. Sharp LK, Lipsky MS. Continuing medical education and attitudes of health care providers toward treating diabetes. J Contin Educ Health Prof. 2002;22(2):103–112.
48. Weller D, May F, Rowett D, Esterman A, Pinnock C, Nicholson S, et al. Promoting better use of the PSA test in general practice: Randomized controlled trial of educational strategies based on outreach visits and mailout. Fam Pract. 2003;20(6):655–661.
49. Maiuro R, Delate T. Can Internet-based education improve physician confidence in dealing with domestic violence? Fam Med. 2002;34(4): 287–292.
50. van Eijk MEC, Avorn J, Porsius AJ, de Boer A. Reducing prescribing of highly anticholinergic antidepressants for elderly people: Randomised trial of group versus individual academic detailing. BMJ. 2001;322 (7287):654–657.
51. Fender GR, Prentice A, Gorst T, Nixon RM, Duffy SW, Day NE, et al. Randomised controlled trial of educational package on management of menorrhagia in primary care: The Anglia menorrhagia education study. BMJ. 1999;318(7193):1246–1250.
52. White C, Albanese M, Brown D, Caplan R. The effectiveness of continuing medical education in changing the behavior of physicians caring for patients with acute myocardial infarction: A controlled randomized trial. Ann Intern Med. 1985;102(5):686–692.
53. Clark NM, Gong M, Schork MA, Evans D, Roloff D, Hurwitz M, et al. Impact of education for physicians on patient outcomes. Pediatrics 1998;101(5):831–836.
54. Majumdar SR, Guirguis LM, Toth EL, Lewanczuk RZ, Lee TK, Johnson JA. Controlled trial of a multifaceted intervention for improving quality of care for rural patients with type 2 diabetes. Diabetes Care. 2003;26:3061–3066.
55. Pazirandeh M. Measuring continuing medical education effectiveness and its ramification in a community hospital. J Contin Educ Health Prof. 2000;20(3):176–180.
56. Cantillon P, Jones R. Does continuing medical education in general practice make a difference? BMJ. 1999;318(7193):1276–1279.
57. Hodges B, Inch C, Silver I. Improving the psychiatric knowledge, skills, and attitudes of primary care physicians, 1950–2000: A review. Am J Psychiatry. 2001;158:1579–1586.
58. Kerwick S, Jones R. Educational interventions in primary care psychiatry: A review. Prim Care Psychiatry. 1996;2:107–117.
59. Wensing M, van der Weijden T, Grol R. Implementing guidelines and innovations in general practice: Which interventions are effective? Br J Gen Pract. 1998;48(427):991–997.
60. Allery LA, Owen PA, Robling MR. Why general practitioners and consultants change their clinical practice: A critical incident study. BMJ. 1997;314(7084):870–874.
61. Knowles M, Holton E, Swanson R, Holton E. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development:* Houston, Tex. Gulf Professional Publishing, 1998.
62. Slotnick H. How doctors learn: The role of clinical problems across the medical school-to-practice continuum. Acad Med. 1996;71(1):28–34.

63. Beer M, Eisenstat R, Spector B. The Critical Path to Corporate Renewal. Boston: Harvard Business School Press; 1990.
64. Weick KE, Quinn RE. Organizational change and development. Annu Rev Psychol. 1999;50:361–386.
65. Smith G. *Meta-Analysis in Social Research*. Beverly Hills, California: Sage, 1981.
66. Jackson G. Methods for integrative reviews. Rev Educ Res. 1980;50(3): 438–460.
67. American Educational Research Association. Standards for Reporting on Empirical Social Science Research in AERA Publications. Washington, DC: American Educational Research Association, 2006.

DOI: 10.1002/chp