

# Teaching Emergency Medicine Residents Evidence-Based Critical Appraisal Skills: A Controlled Trial

From the Department of Emergency Medicine, University of Rochester Medical Center, Rochester, NY,\* and the Department of Emergency Medicine, Allegheny University Health Sciences—MCP Division, Philadelphia, PA.†

Received for publication August 6, 1998. Revisions received November 17, 1998, and December 28, 1998. Accepted for publication January 18, 1999.

Presented as an oral platform presentation ("Is There Evidence for Evidence-Based Medicine?") at the annual meeting of the Society for Academic Emergency Medicine, Washington DC, May 1997.

Supported in part by an Innovations In Medical Education Grant from the Emergency Medicine Foundation.

**Address for reprints:** Jeffrey J Bazarian, MD, Department of Emergency Medicine, Box 4-9200, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, NY 14642; 716-273-4112, fax 716-473-3516; E-mail jeff\_bazarian@urmc.rochester.edu.

Copyright © 1999 by the American College of Emergency Physicians.

0196-0644/99/\$8.00 + 0  
47/1/97457

**Jeffrey J Bazarian, MD\***  
**Colleen O Davis, MD**  
**Linda L Spillane, MD\***  
**Howard Blumstein, MD†**  
**Sandra M Schneider, MD\***

**See editorial, p. 226.**

**Study objectives:** To compare the performance of an evidence-based medicine (EBM) approach and a traditional approach to teaching critical appraisal skills to emergency medicine residents.

**Methods:** This was a prospective, case-controlled trial of 32 emergency medicine residents (16 control and 16 intervention). Intervention residents were exposed to a monthly, 1-hour journal club using an EBM approach to critical appraisal over the course of 1 year. Control residents were exposed to a traditional, unstructured journal club, also monthly. Both groups were given a factitious article to evaluate in an essay format before and after the 12-month study period. The Wilcoxon rank sum test was used to compare mean improvement in test scores for each group.

**Results:** The mean improvement in test scores was 1.80 for the control group and 1.53 for the intervention group; these values were not significantly different ( $P=.90$ ). The difference in mean change in test score between the 2 groups was .27 points.

**Conclusion:** Compared with a traditional approach, an EBM approach to teaching critical appraisal did not appear to improve the critical appraisal skills of emergency medicine residents. However, because of the small number of subjects studied, small differences in critical appraisal skill improvement cannot be ruled out.

[Bazarian JJ, Davis CO, Spillane LL, Blumstein H, Schneider SM: Teaching emergency medicine residents evidence-based critical appraisal skills: A controlled trial. *Ann Emerg Med* August 1999;34:148-154.]

---

## INTRODUCTION

Evidence-based medicine (EBM) has emerged as a popular solution to the problem of harnessing data from published research to make patient care decisions. This approach relies heavily on a systematic, critical appraisal of the literature using “rules of evidence” that are derived from the fields of biostatistics and epidemiology.

Traditionally, the journal club has served as the forum for teaching critical appraisal skills. In the typical journal club, critical appraisal is taught by analysis of several journal articles, usually in a nonsystematic way. The group usually focuses on understanding how to integrate the conclusions into clinical practice, rather than on whether those conclusions are valid. Discussion of the quality of the evidence presented can often reflect the opinions of 1 or more dominant members of the group. In short, typical journal clubs often lack a process by which the evidence presented can be systematically scrutinized and understood. More recently, journal clubs using EBM, with its step-by-step approach to the interpretation of evidence, have emerged. Although many educators have accepted EBM as the de facto gold standard, it is unclear whether this approach to teaching critical appraisal skills is superior to traditional, less systematic methods. Several educators have attempted to measure the effect of loosely structured, EBM journal clubs on critical appraisal abilities without showing a positive effect.<sup>1-5</sup> Others showed an effect but did not use controls.<sup>6</sup> No researchers have studied emergency medicine residents.

Our objective was to determine whether a highly structured EBM journal club was better than a traditional, less systematic journal club in teaching critical appraisal skills to emergency medicine residents.

---

## MATERIALS AND METHODS

We performed a prospective, case-controlled trial comparing 2 journal club formats for teaching critical appraisal skills to emergency medicine residents. The duration of the study was 12 months. Requirement for informed consent was waived by the University of Rochester Research Subjects Review Board.

Subjects were 16 intervention residents and 16 control residents. The 16 intervention residents included 8 interns (EMR-1) and 8 second-year residents (EMR-2) from an academic, urban emergency medicine residency program. The control group consisted of 8 EMR-1 and 8 EMR-2 residents from a different academic emergency medicine residency program at a geographically distinct

site. This particular program was chosen as the site for the control group because it is a more established program (founded in 1971), it has had a set method for carrying out its journal club for 8 years, and its residency director was interested in EBM and was willing to abide by the study protocol. The emergency medicine program chosen as the study site was picked because it was first established in 1993, it did not yet have an established journal club, and the residency director was willing to experiment with new approaches. Both programs perform equally well in the match each year, getting their residents from the top 25% to 30% of their respective match lists, although the study program did not participate in the match in the first year it accepted residents. Both programs are 3 years in length.

Intervention residents participated in a highly structured, monthly EBM journal club containing 3 primary elements. The first element was a *case-based presentation format*. The presenting resident identified a clinical question derived from an actual patient case and found 1 journal article that best addressed the question.

The second element was *structured worksheets*. The presenting resident and all conference attendees used structured worksheets to guide their critical appraisal of the article. Two worksheets were created, 1 on diagnosis and 1 on therapy. These worksheets were derived from the textbook, *Clinical Epidemiology: A Basic Science for Clinical Medicine*,<sup>7</sup> by Sackett et al, and from the 5-part series in the *Journal of the American Medical Association*, “User’s Guide to the Medical Literature.”<sup>8-12</sup> As an example, the diagnosis worksheet asked residents to answer the following questions (instructions on how to calculate the parameters in questions 7 through 10 were included):

1. Is there a gold standard for this test?
2. Is the new test reported to be more accurate than the gold standard?
3. If not, what is its advantage?
4. Was there an independent blind comparison with the accepted gold standard?
5. Did the patient sample include an appropriate spectrum of patients?
6. Did the results of the test influence the decision to perform the gold standard?
7. What is the sensitivity and specificity of the test?
8. What are the negative and positive predictive values of the test?
9. Is the pretest probability (prevalence) of the target disorder known?
10. What is the likelihood ratio for a positive and negative test?

11. Has the reproducibility (precision) of the test been established?

12. Have intraobserver and interobserver variation been established?

13. Do these results help to answer your specific journal club question?

The third element of the evidence-based journal club was *close faculty supervision*. The presenting resident received 60 to 90 minutes of preparation from 1 of the investigators, who closely supervised the actual journal club session. Attendance at this conference was mandatory for all on-service emergency medicine residents.

Control residents attended a monthly journal club in which they reviewed a topic rather than a patient case and used multiple articles rather than a single journal article. No systematic, evidence-based analysis was applied to the articles reviewed. Consensus regarding the evidence was reached by group opinion. There was faculty supervision at each journal club, but there was no faculty preparation of the presenting resident before the presentation. Resident attendance at this conference was not mandatory.

A pretest-posttest format was used to measure key outcomes. Both groups were given a pretest consisting of a factitious article entitled, "Intravenous Propranolol for the Emergency Management of Acute Migraine Headache." This article was used with the author's permission. It was written originally to test the editorial skills of reviewers of the *Annals of Emergency Medicine* and was purposely created with 17 methodologic flaws.<sup>13</sup> Of the 203 referees who reviewed the article, 117 (59%) recommended rejection, 67 (33%) advised extensive revisions, and 15 (7%) accepted it. The errors in the factitious article were first revealed by the study's author during a lecture given at the 1995 ACEP Scientific Assembly, and subsequently in a published manuscript.<sup>13</sup> The reviewers who recommended rejection identified 7.1 errors per review; the 67 who advised revisions identified 5.8 errors per review; and the 15 who accepted the work identified 3.3 errors per review.<sup>13</sup>

Residents were asked to answer 1 question in an essay format: "After critically appraising this article, would you use intravenous propranolol for the treatment of migraine headache? Give at least 5 reasons to support this decision." At the end of the 12-month study interval, the same article and test were given to both groups of residents. Both pretest and posttest were given in a closed conference room and proctored by 1 of the investigators. No books or educational materials

were allowed. Residents had 1 hour to complete each test. The test date was not announced in advance, to prevent residents from reading up on the topic. Residents were not given any feedback on their answers to the pretest. Although residents were not told that the same test would be administered as a posttest, we could not control for the possibility that residents would read about intravenous propranolol for treatment of migraines between the pretest and posttest.

Scoring of tests was done by a weighted counting system devised with the aid of an educational consultant. Of the 17 flaws in the study, we considered 7 to be of major importance and 10 to be of minor importance for the critical appraisal of the evidence presented (Table 1). Major flaws consisted of methodologic errors that the EBM approach to critical appraisal emphasized (eg, improper randomization, missing patients). Minor flaws consisted of methodologic elements that were not emphasized, such as the use of histograms and visual pain scales (Table 1).

For each resident, article flaws were identified from the free text list of 5 reasons used to support the answer to the main test question. We assigned 2 points for each major flaw and 1 point for each minor flaw identified; an additional 3 points were given for answering the main test question correctly ("No"). For example, if the main test question was answered correctly and 5 major flaws were identified, the score was 13. The theoretic maximum score was 27. Tests were scored by 3 independent evaluators (JB, LS, CD), who were blinded to

**Table 1.**

*Methodologic errors contained in the factitious journal article used for the pretest and posttest.*

Major Errors	Minor Errors
No inclusion/exclusion criteria	Inappropriate pain scale
Improper randomization	Incorrect statistics
Missing patients	Ten eliminated patients not discussed
Not blinded	No monitoring of vital signs
Side effects not discussed	No institutional review board approval
Baseline characteristics of study and control groups not compared	Tables not discussed in text
Overstated conclusions	Inappropriate use of histograms
	No comparison to known agent
	Mean time in the ED not measured
	No references newer than 1989

Adapted from Baxt WG, Waeckerle JF, Berlin JA, et al: Who reviews the reviewers? Feasibility of using a factitious manuscript to evaluate peer reviewer performance. *Ann Emerg Med* 1998;32:310-317.

the identity of the test takers and to whether the answers were from the pretest or the posttest. Each evaluator scored all tests. Each of the 3 test graders used personal judgment to determine whether a free text answer came close enough to the predetermined answer to be counted as correct; exact phrasing in the resident's response was not necessary. Independent agreement by all 3 test graders was necessary for the answer to be counted as correct or incorrect. In cases in which 1 grader disagreed with the other 2, the resident's answer was reviewed by the 3 graders as a group until consensus was reached.

Residents in both groups were given a questionnaire asking about postcollege training, previous research experience, and epidemiology training. Residents were also asked to rate their journal reading habits at the beginning and at the end of the study interval.

Mean and median pretest and posttest scores were calculated for the intervention and control groups. The effect of the intervention was measured by calculating the change in score (posttest score minus pretest score)

**Table 2.**  
*Comparison of baseline characteristics.*

Characteristic	Intervention Group	Control Group
Age in years (mean [95% CI])	30.9 (28.9 to 33.0)	28.2 (27.0 to 29.4)
Female (%)	20	53
Previous training in statistics or epidemiology (%)	13	27
Research experience (%)	93	87
No. of journal club conferences attended (median)	6	2

**Table 3.**  
*Summary of test scores.*

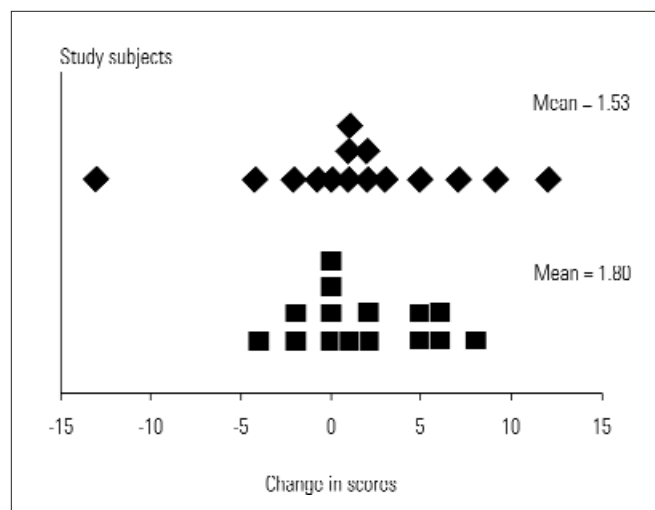
Item	Mean Score (95% CI)	Median Score
<b>Control group</b>		
Pretest	6.87 (5.1 to 8.6)	7
Posttest	8.67 (7.5 to 9.8)	8
Change in score	1.80 (.034 to 3.57)	1.0
<b>Intervention group</b>		
Pretest	6.33 (4.3 to 8.4)	7
Posttest	7.87 (5.6 to 10.1)	9
Change in score	1.53 (-1.41 to 4.48)	1.0

for each resident and then calculating the mean change in score for each group. Because the number of data points was relatively small, we assumed that they were not normally distributed. We therefore used the Wilcoxon rank sum test to compare the mean changes in score between the 2 groups.

Interrater reliability for the 3 test graders was analyzed by the intraclass correlation coefficient, rather than  $\kappa$ , because the measured variable (mean change in essay test scores) was considered to be a continuous, not nominal, variable. The Wilcoxon rank sum test was used to compare the number of journal clubs attended by each group. The median test score in the group of residents who answered the main test question incorrectly was compared with the median score of residents who answered it correctly (after subtracting the extra 3 points that were awarded for getting the main test question correctly), using the Wilcoxon rank sum test.

The ability to detect a significant difference in change in test score between the 2 groups was limited by the small number of residents in the 2 residency programs. We observed a difference of .27 in mean test score change between the groups, and estimate the standard deviation on this difference to be 4.71. Using this standard deviation, a sample size of 14 residents in each group has the ability to detect a difference in mean test score of 5.0 with a power of .80, assuming a 2-sided *t* test with a type I error ( $\alpha$ -value) of .05. Eighteen resi-

**Figure.**  
*Change in test score (posttest score minus pretest score) for intervention residents (diamonds) and control residents (squares).*



dents in each group would have been needed to detect a difference of 4.5 with similar power.

## RESULTS

A comparison of the baseline characteristics of the residents in the 2 groups is shown in Table 2. Over the 12-month study interval, 2 residents dropped out (1 from each group). One resident switched residency programs, and 1 did not complete the pretest despite several requests. The intervention group attended significantly more journal club conferences than the control group did (median, 6 versus 2 sessions [mean, 5.93 versus 1.67 sessions],  $P < .01$ ).

A summary of test score results is given in Table 3. Raw test scores ranged from 0 to 14 for the pretest and 0 to 17 for the posttest. The change in critical appraisal test score over the 12-month period for all participants ranged from a decline of 13 points to an improvement of 12 points (Figure). Ten intervention residents showed improvement in test scores, as did 8 controls. The mean change in test score was 1.80 (95% confidence interval [CI], .034 to 3.57) for the controls and 1.53 (95% CI, -1.41 to 4.48) for the intervention group. The difference between the mean changes in test score for the 2 groups was .27 (95% CI, -3.9 to 2.9). According to the Wilcoxon rank sum test, this difference was not significant ( $z = .125$ ,  $P = .90$ ). The median change in test scores for both groups was +1.0.

The 3 independent graders agreed fairly well; the intra-class correlation coefficient was .94.

At the start of the 12-month study, 4 intervention and 3 control residents answered the main test question incorrectly by responding "Yes." On the posttest, 3 intervention residents and no control residents answered "Yes." Although this difference in improvement is statistically insignificant, 2 of the 3 intervention residents answering the main test question incorrectly at the end of the study had answered it correctly at the start of the study. The median test score

was significantly higher among those who answered the main test question correctly, both before (5 versus 0,  $P = .002$ ) and after (6 versus 0,  $P = .04$ ) the intervention, even after subtracting the extra 3 points that were awarded for answering the main test question correctly. This effect was seen in both intervention and control groups.

There was no statistically significant difference in the change in journal reading habits between the groups (Table 4).

## DISCUSSION

EBM provides a framework not only for the critical appraisal of published trials, which was the focus of this investigation, but also for the conversion of information needs into answerable questions, for the application of critically appraised evidence to individual patients, and for the efficient retrieval of the best evidence from the ever-expanding pool of published trials. Although its philosophical origins date back to the mid-1800s,<sup>14</sup> it was not until 1985 that EBM became popularized with the publication of a text by DL Sackett et al entitled, *Clinical Epidemiology: A Basic Science for Clinical Medicine*,<sup>7</sup> in which the connections between real-life, everyday clinical practice and the principles of epidemiology and biostatistics were clarified. Beginning in 1993, Sackett and the other members of the Evidence-Based Medicine Working Group followed up this work with a series of articles published in the *Journal of the American Medical Association* entitled, "User's Guides to the Medical Literature."<sup>8-12</sup> These articles further clarified the concepts of EBM and reached an international audience. According to Sackett, gone are the days when the seasoned, elder authoritarian clinician was the only one able to make complex decisions. Now, with the rules of evidence in hand, even the most novice clinician can enter into complex decisionmaking processes. Since 1985, EBM has grown immensely in popularity and has

**Table 4.**

*Journal reading habits of residents before and after intervention.*

Habit	Intervention Group			Control Group		
	Before	After	Δ	Before	After	Δ
Read >2 journal articles per week	4	8	4	2	4	2
Read journal articles primarily for patient care	3	0	-3	1	3	2
Find journal articles very useful in clinical practice	3	4	1	1	0	-1
Spend >30 min/wk reading journal articles	6	11	5	7	9	2

---

attracted many advocates.<sup>14</sup> Many resident educators have formally incorporated the EBM approach to teaching critical appraisal into their curricula, although it has never been validated as a superior approach.<sup>14</sup>

Several attempts to prove the value of EBM have been made. In 1992, Langkamp et al<sup>1</sup> gave 2 introductory lectures in epidemiology and biostatistics followed by 8 unstructured journal clubs to pediatric residents but found no significant improvement in knowledge of these fields, compared with controls. In 1991, Seelig<sup>6</sup> gave internal medicine residents a 1-hour seminar on the basics of critical appraisal and found that their critical appraisal test scores did improve, although there was no comparison with a control group. Both Kitchens and Pfeifer<sup>2</sup> and Linzer et al<sup>3</sup> used loosely structured journal clubs for internal medicine residents and found no significant improvement in their knowledge base or critical appraisal scores. Norman et al<sup>15</sup> performed a systematic review of the 7 methodologically acceptable studies published between 1966 and 1995 that attempted to measure the impact of an EBM approach to critical appraisal in residents and medical students; they found consistent improvements in critical appraisal skills among medical students but not among residents.

In a survey of the directors of the 86 emergency medicine residency programs approved by the Accreditation Council for Graduated Medical Education in 1991–1992, residents were more likely to feel that the journal club was educationally successful in the 30% of programs that used structured guidelines for the critical appraisal of selected journal club articles.<sup>16</sup> We interpreted these trends in the data to imply that a more rigorous and structured attempt at teaching an EBM approach to critical appraisal would be needed to improve the critical appraisal skills of residents.

However, after applying what we believe to be a fairly rigorous and structured approach, we were unable to demonstrate any significant difference in critical appraisal scores between the intervention and control groups. Although this study was not designed prospectively for power, it in fact had good ability to detect educationally significant differences in mean scores (ie, differences of 5 points or greater). Although some may argue that the small number of residents studied makes it difficult to rule out smaller differences in score improvement that may be important, the data suggest that the performance of the residents exposed to EBM was not any better, and may actually have been worse, than that of the controls.

Not only did the critical appraisal test scores of 4 residents exposed to EBM decline, but 2 intervention resi-

---

dents who got the correct clinical bottom line at the start of the study got it wrong at the end. These results occurred despite the fact that intervention residents attended significantly more journal club conferences than did controls (median, 6 versus 2). Did our rigidly structured and time-consuming intervention actually confuse some residents? Or, are 6 structured lessons on evidence-based critical appraisal enough to confuse but not enlighten residents who may have a greater “innate” sense of critical appraisal than we would have anticipated? The important variable may be not the degree of structure in the teaching method but the total exposure time to the teaching method. Because EBM represents a new way of thinking for most residents, a certain degree of unlearning of less systematic approaches to critical appraisal may be necessary before this new approach can be accepted and understood. As the Figure demonstrates, some intervention residents demonstrated remarkable improvements in critical appraisal abilities, whereas others showed regression in abilities. This wide variation in test results was not seen among controls, who tended to cluster around the mean.

In regard to change in journal reading habits, there was no statistically significant difference between the groups, although some trends in the data are worth mentioning. At the end of the study, both control and intervention groups showed an increase in the number of residents reading more than 3 journal articles per week and in the number reading for longer than 30 min/wk. This increase was greatest for the intervention group. The number of intervention residents reading “primarily for patient care reasons” declined over the study interval, whereas at the close of the study more residents in the control group were reading to help with patient care (Table 3). Again, did our structured EBM approach to critical appraisal interfere with the natural tendency for residents to rely more on journal articles to make patient care decisions as they progress through their residency? Rather than fostering the use of published evidence, we appear to have accomplished the opposite. A subset analysis of this data revealed no significant difference in journal reading habits between EMR-1 and EMR-2 participants.

The effects of the 12-month journal club may go beyond what we were able to measure directly. The intervention residents seem to be talking about and actively using the concepts of EBM more now than before the start of this study. In addition, we believe that understanding of the concepts of EBM has helped our residents design and carry out their own research projects, which is now an Emergency Medicine Residency Review Committee requirement.

Although these are somewhat intangible results, we feel encouraged to continue to pursue teaching this approach to critical appraisal despite not being able to demonstrate improved test scores.

Our study was limited in several ways. We were limited to 16 emergency medicine residents in each group by the size of the 2 residency programs. Expansion of the study to multiple sites so as to include more residents would have introduced unwanted variation in teaching styles and exposure to EBM outside the journal club. However, our small sample size may have made it difficult to detect a small difference in critical appraisal scores between the groups, had such a difference existed.

This study was further limited by the validity of the essay test used to probe critical appraisal skills, which was not prospectively validated. Although demonstrating the predictive validity of this essay is not possible because there are no gold standards for testing of critical appraisal skills, the essay test does have some degree of internal consistency, since those who answered the main test question correctly also identified more flaws in the study. That the range of raw test scores of the study subjects was within the range of scores of the *Annals of Emergency Medicine* referee group provides some measure of reproducibility. Content validity is probably poor because the essay test involved only 1 factitious study, although the diversity of errors placed in the study mitigates against this effect somewhat. Therefore, although it may be an authentic measure of what we are testing (ie, critical appraisal skills), this essay test probably has face validity only.

Another limitation of this study involves our inability to control for residents' exposure to EBM outside the journal club. In addition, the extent to which the control group used EBM in an informal way to learn critical appraisal skills would also be difficult to estimate, and even more difficult to eliminate. However, problems such as these plague all educational studies, not just ones evaluating EBM.

Finally, EBM was meant to be a tool both to critically appraise articles and to enhance medical decisionmaking. The effect of our EBM journal club intervention on patient care was not evaluated, but this certainly could be an area of future investigation.

In summary, 1 year of highly structured, monthly, EBM journal clubs was not associated with significantly improved critical appraisal test scores or journal reading habits among emergency medicine residents, although, with the small number of subjects studied, small differences in these outcomes cannot be ruled out. More pro-

longed exposure to EBM may be necessary to demonstrate widespread improvement in critical appraisal skills.

We wish to acknowledge Randall Curren for expertise in designing and scoring the essay test, Mike McDermott and Hongwei Zhao for biostatistical guidance, Donald Bordley, MD, for assistance with manuscript preparation, Jennifer Pulver for technical support, and the emergency medicine residents of the University of Rochester and Allegheny University Health Sciences—MCP Division.

## REFERENCES

- Langkamp DL, Pascoe JM, Nelson DB: The effect of medical journal club on residents' knowledge of clinical epidemiology and biostatistics. *Fam Med* 1992;24:528-530.
- Kitchens JM, Pfeifer MP: Teaching residents to read the medical literature: A controlled trial of a curriculum in critical appraisal/clinical epidemiology. *J Gen Intern Med* 1989;4:384-387.
- Linzer M, Brown T, Frazier LM, et al: Impact of a medical journal club on house-staff reading habits, knowledge, and critical appraisal skills: A randomized control trial. *JAMA* 1988;260:2537-2541.
- Bennett KJ, Sackett DL, Haynes RB, et al: A controlled trial of teaching critical appraisal skill of the clinical literature to medical students. *JAMA* 1978;257:2451-2454.
- Linzer M, DeLong ER, Hupart KH: A comparison of two formats for teaching critical reading skills in a medical journal club. *J Med Educ* 1987;62:690-693.
- Seelig CB: Affecting residents' literature reading attitudes, behaviors, and knowledge through a journal club intervention. *J Gen Intern Med* 1991;6:330-334.
- Sackett DL, Haynes RB, Guyatt GH, et al: *Clinical Epidemiology: A Basic Science for Clinical Medicine*, ed 2. Boston: Little, Brown, 1991.
- Oxman AD, Sackett DL, Guyatt GH, et al: I. How to get started. *JAMA* 1993;270:2093-2095.
- Guyatt GH, Sackett DL, Cook DJ, et al: II. How to use an article about therapy or prevention: A. Are the results of the study valid? *JAMA* 1993;270:2598-2601.
- Guyatt GH, Sackett DL, Cook DJ, et al: II. How to use an article about therapy or prevention: B. What are the results and will they help me in caring for my patients? *JAMA* 1994;271:59-63.
- Jaeschke R, Guyatt GH, Sackett DL, et al: III. How to use an article about a diagnostic test: A. Are the results of the study valid? *JAMA* 1994;271:389-391.
- Jaeschke R, Guyatt GH, Sackett DL, et al: III. How to use an article about a diagnostic test: B. What are the results and will they help me in caring for my patients? *JAMA* 1994;271:703-707.
- Baxt WG, Waeckerle JF, Berlin JA, et al: Who reviews the reviewers? Feasibility of using a factitious manuscript to evaluate peer reviewer performance. *Ann Emerg Med* 1998;32:310-317.
- Sackett DL, Richardson WS, Rosenberg W, et al: *Evidence-Based Medicine: How to Practice and Teach EBM*. Edinburgh: Church Livingstone, 1997.
- Norman GR, Shannon SI: Effectiveness of instruction in critical appraisal (evidence-based medicine) skills: A critical appraisal [see comments]. *CMAJ* 1998;158:177-181.
- Singletary EM, Dorr CE: Journal club [letter]. *Acad Emerg Med* 1995;2:570-573.