Teaching Residents Evidence-based Medicine Skills
A Controlled Trial of Effectiveness and Assessment of Durability
Christopher A. Smith, MD, Pamela S. Ganschow, MD, Brendan M. Reilly, MD, Arthur T. Evans, MD, MPH, Robert A. McNutt, MD, Albert Osu, MD, Muhammad Saquib, MD, Satish Surabhi, MD, Sunil Yadav, MD

OBJECTIVES: To measure the effectiveness of an educational intervention designed to teach residents four essential evidence-based medicine (EBM) skills: question formulation, literature searching, understanding quantitative outcomes, and critical appraisal.

DESIGN: Firm-based, controlled trial.

SETTING: Urban public hospital.

PARTICIPANTS: Fifty-five first-year internal medicine residents: 18 in the experimental group and 37 in the control group.

INTERVENTION: An EBM course, taught 2 hours per week for 7 consecutive weeks by senior faculty and chief residents focusing on the four essential EBM skills.

MEASUREMENTS AND MAIN RESULTS: The main outcome measure was performance on an EBM skills test that was administered four times over 11 months: at baseline and at three time points postcourse. Postcourse test 1 assessed the effectiveness of the intervention in the experimental group (primary outcome); postcourse test 2 assessed the control group after it crossed over to receive the intervention; and postcourse test 3 assessed durability. Baseline EBM skills were similar in the two groups. After receiving the EBM course, the experimental group achieved significantly higher postcourse test scores (adjusted mean difference, 21%; 95% confidence interval, 13% to 28%; P < .001). Postcourse improvements were noted in three of the four EBM skill domains (formulating questions, searching, and quantitative understanding [P < .005 for all], but not in critical appraisal skills [P = .4]). After crossing over to receive the educational intervention, the control group achieved similar improvements. Both groups sustained these improvements over 6 to 9 months of follow-up.

CONCLUSIONS: A brief structured educational intervention produced substantial and durable improvements in residents’ cognitive and technical EBM skills.

KEY WORDS: evidence-based medicine; clinical trial; graduate medical education; internship and residency.


Clinicians recognize the potential value of evidence-based medicine (EBM) but many feel ambivalent about its use in patient care. Major barriers to implementation include a perception that the appeal of EBM is more academic than practical, and clinicians cite their own lack of EBM skills as a contributing factor.1

Many postgraduate training programs currently attempt to teach EBM skills, primarily in journal clubs and through didactic lectures. However, published outcomes of these efforts have shown only limited effectiveness.2–8 Most of the studies have focused on critical appraisal while excluding or inadequately measuring other EBM skills.9 Consequently, there are no published studies that assess how best to teach more fundamental EBM skills: the ability to create searchable questions that arise from knowledge gaps in patient care and the ability to perform efficient literature searches to answer those questions.10 The acquisition of these skills, in addition to those of critical appraisal, is essential if EBM is to favorably influence patient outcomes.11–13

We conducted a trial to test whether an educational program could improve four essential EBM skills among first-year internal medicine residents: (1) posing well-structured, searchable questions arising from clinical cases; (2) performing efficient electronic literature searches to answer these questions; (3) understanding quantitative outcomes from published studies about diagnosis and treatment; and (4) evaluating the methodological quality of published studies and their clinical relevance to specific patients. We also carried out repeat assessments to see if these acquired skills were durable over time. These four skills are similar, but not identical, to those espoused by proponents of EBM—namely, question formulation, literature searching, critical appraisal skills, and application of the results to the patient.14

METHODS

Study Design

We performed a firm-based, controlled trial of an EBM educational course for first-year residents in the Cook County Hospital internal medicine training program. All residents and faculty in the program are assigned (without formal randomization) to one of three firms in the Department of Medicine.

There were 55 eligible first-year residents in our program at the start of the study. First-year residents in one firm (n = 18) were assigned to the experimental group;
those in the remaining two firms \((n = 37)\) were assigned to the control group. The experimental firm received a 7-week course taught twice weekly by three faculty members (BMR, ATE, and RAM), and six chief medical residents. During this time, the control firms participated in equally intense educational curricula of identical duration, usually case-based clinical presentations in small interactive groups emphasizing clinical practice, but without formal EBM teaching.

Evidence-based medicine skills were assessed before the intervention in January 1998 (precourse test); after the experimental firm completed the EBM course in March 1998 (postcourse test 1); after the control group received the EBM course in June 1998 (postcourse test 2); and a final time in November 1998 (postcourse test 3). The primary outcome was the difference in test scores at postcourse test 1. The study design is shown in Figure 1.

### Educational Intervention

The EBM course consisted of 1-hour noon conferences twice a week for 7 consecutive weeks and a 1.5-hour session in a computer lab. The course curriculum highlighted diagnosis and treatment issues—not risk, prognosis, or other content areas. The first session of each week was a didactic seminar which focused on selected components of the EBM curriculum. In the second session, these skills were practiced in an interactive format involving actual patient cases (Table 1).

In teaching question formulation, we focused on four components: (1) a defined patient population; (2) the intervention of interest (a test or a treatment); (3) clinically meaningful outcome measures; and (4) a comparison group. Literature searching was taught in 1 didactic and 1 computer lab session; the searching process was also reinforced weekly during follow-up discussions. Diagnosis seminars focused on understanding 2 × 2 tables and core criteria necessary to appraise studies about diagnostic tests rather than likelihood ratios and odds. Treatment seminars focused on the numerical measures of treatment benefits and harms (risk reductions and number needed to treat \([NNT]\)). Understanding confidence intervals \([CIs]\), and core criteria to appraise the validity of treatment trials.

### Outcome Measures

The primary outcome measure was a written test of EBM skills. We also surveyed residents’ self-reported attitudes and behaviors. The EBM skills test included 5 sets of questions; each set was prefaced by a clinical case scenario and designed to evaluate one of five content areas: (1) posing well-structured clinical questions; (2) performing computerized literature searches (using \(OVID\) MEDLINE—the only resource available at our institution) for both diagnostic and treatment questions; (3) understanding quantitative aspects of published diagnostic studies \((2 \times 2\) tables, predictive value, pretest and posttest probabilities, etc.); (4) understanding quantitative aspects of published treatment studies (relative risk reduction, absolute risk reduction, number needed to treat, confidence intervals); and (5) assessing the quality of treatment studies and their clinical relevance to particular patients. For purposes of scoring the tests, we collapsed content areas 3 and 4 above into one skill domain (quantitative understanding), yielding four tested skill domains. The maximum total score \((100\%)\) was a weighted average of scores in the four domains: formulating questions \((16\%)\), literature searches \((36\%)\), quantitative understanding \((24\%)\), and appraisal of study quality and clinical relevance \((24\%)\).

The overall structure of the test, the weighting of each skill domain, and the specific content areas remained unchanged for the four tests given between January and November. However, the clinical cases were changed and the wording of test questions varied somewhat to minimize recall bias on test performance. Two raters graded each exam independently; disagreements were resolved by consensus or after consultation with a third or fourth co-investigator; however, raters were not blinded to study group assignment.

In addition, we collected basic demographic data from the residents and a survey of self-reported EBM skills and behavior. This survey was repeated each time