

Evaluating evidence-based practice performance

THOUGHT IS THE BLOSSOM; LANGUAGE THE
BUD; ACTION THE FRUIT BEHIND IT.

RALPH WALDO EMERSON

Evidence-based practice (EBP) educators need valid and feasible approaches to evaluate the effect of new curricula and to document the competence of individual trainees. As of 1999, however, the published evaluation instruments often lacked established validity and reliability, focused on critical appraisal to the exclusion of other EBP steps, and measured knowledge and skills but not behaviors in actual practice (1). Editorialists at the time lamented, "Ironically, if one were to develop guidelines for how to teach evidence-based medicine (EBM) based on these results, they would be based on the lowest level of evidence" (2).

In the ensuing years, educators have responded to the challenge. We now have several instruments, supported by multiple types of evidence for validity, to evaluate EBP knowledge and skills. In the parlance of Miller's pyramid (3), these instruments can document that a trainee "knows" and "knows how" to practice EBM. And promising EBP Objective Structured Clinical Examinations (OSCE), although supported by limited psychometric testing, allow trainees to "show how" in realistic clinical settings.

To "bear fruit from the blossom," we must ensure that trainees implement their EBP skills as we move from competence (or skills) to performance (or behaviors). We can consider EBP performance at 2 levels. First, we can ask, *does a trainee perform the 5 EBP steps (Ask, Acquire, Appraise, Apply, and Assess) in the course of patient care activities?* Alternatively, we can look further downstream and examine her clinical practice data directly, asking *does she perform evidence-based clinical maneuvers and effect desirable patient outcomes?*

HOW CAN WE EVALUATE THE PERFORMANCE OF EBP STEPS IN PRACTICE?

We can simply ask a trainee if, for example, she consistently searches for the evidence to answer her clinical questions. However, retrospective self-reports of EBP behaviors remain extremely biased, as physicians tend to underestimate their information needs and overestimate their pursuit of them (4). At the opposite extreme of rigor, we can shadow a trainee in the course of her patient encounters, document her emerging clinical questions, and follow-up later to see if she has acquired, appraised, and applied the evidence. This question collection might involve passive anthropological observation (5) or active debriefing (4, 6). While this direct observation yields more valid data, it is not feasible outside of the research setting. Thus, educators have looked for intermediate approaches to document the performance of EBP steps.

Some investigators have analyzed audiotapes of resident-faculty interactions, looking for phrases related to literature searching, clinical epidemiology, or critical appraisal (7). Family practice residents' "EBM utterances" increased from 0.21 per hour to 2.9 per hour after an educational intervention. However, in my view, this outcome lacks face validity as a suitable surrogate for EBM performance. Another group questioned residents about their awareness of findings in recent journal articles deemed relevant to primary care practice (8). In this pre-post randomized controlled trial (RCT), residents exposed to academic detailing recalled more articles and correctly answered more questions about them.

Educators can also electronically capture trainees' searching behaviors, including number of log-ons, articles viewed, and time spent searching. In an RCT, Cabell demonstrated that these measures were responsive to an intervention that included the use of well-built clinical question cards and practical sessions in clinical question building (9). While this approach is quite feasible, the crude measure of searching volume fails to capture the pursuit and application of information in response to particular clinical questions.

Another approach is to have trainees catalogue their EBP learning activities in learning portfolios, which represent a purposeful collection of student work that exhibits to the student (and/or others) the student's efforts, progress, or achievement in a given area (10). EBP portfolios might include educational prescriptions, which faculty dispense (or trainees self-prescribe) when a moment of uncertainty arises in the course of patient care (11-13). A typical educational prescription describes the clinical problem, states the question, specifies who is responsible for answering it, and reminds the trainee and faculty of a follow-up time. Variations have the trainee articulate foreground questions in the Participant-Intervention-Comparison-Outcome (PICO) format, document the information resources searched, grade the level of evidence, state how her practice will change, or reflect on what she learned. EBP learning portfolios can be maintained in sophisticated Internet-based databases (14, 15). Educators implemented a Computerized Obstetrics and Gynecology Automated Learning Analysis (KOALA) at several residency programs (15). This portfolio allowed residents to record their clinical encounters, directly link to information resources, and document critical learning incidents. During a 4-month pilot at 4 programs, 41 residents recorded 7049 patient encounters and 1460 critical learning incidents. Residents at 1 of the programs, which had 1 year of prior experience with KOALA, demonstrated higher self-directed learning readiness. In another program, residents entered their clinical questions, accompanied by MEDLINE links and article summaries, into a similar Internet-based compendium (14). The EBP exercises produced useful information for 82% and altered patient management for 39% of 625 clinical questions over 10 months.

HOW CAN WE EVALUATE PERFORMANCE OF EVIDENCE-BASED CLINICAL ACTIONS?

Ellis devised a reliable method for determining the primary therapeutic intervention chosen by a practitioner and classifying the quality of evidence supporting it as 1) supported by individual or systematic reviews of RCTs, 2) supported by convincing nonexperimental evidence, or 3) lacking substantial evidence (16). This method has been employed in descriptive studies in inpatient medicine (16, 17), general outpatient practice (18), emergency ophthalmology (19), dermatology (20), anesthesiology (21), general surgery (22), pediatric surgery (23), and inpatient psychiatry (24) settings. In Straus' study, patients admitted after a multifaceted educational intervention were more likely to receive therapies proven to be beneficial in RCTs, providing initial evidence of the responsive validity of this evaluation strategy (25). The Ellis protocol seems most suited to evaluate changes in EBP performance after an educational intervention or simply over time. To use it to document some absolute threshold of performance, one would

Resident(s): _____ Faculty: _____ "Rx" date _____
 "fill" date _____

1. Rotation:	Wards	Elective	Geriatrics	Continuity Clinic	Ambulatory office	Other:
	ICU	Emergency	Night float		Ambulatory specialty	

2. Context	Question conference	EBM curriculum	Resident report	Gen. patient care	Other:
------------	---------------------	----------------	-----------------	-------------------	--------

3. Clinical scenario:

4. Question			
<u>patient</u>	<u>intervention</u>	<u>comparison</u>	<u>outcome</u>

5. Clinical task:	clinical findings	differential diagnosis	prognosis	prevention
	etiology / harm	diagnostic testing	therapy	manifestations of disease

6. Information source that provided answer:	Clinical Evidence	EBM On Call	MEDLINE original study
	Cochrane Library /DARE	UpToDate	MEDLINE meta-analysis
	ACPJC On line	PIER	MEDLINE narrative review
	Rational Clinical Exam	Textbook	Consultant/colleague
	Info Retriever (POEM or other source: _____)		Other: _____

7. The evidence (reference)

8. Your answer / Results:

9. What did you learn from this experience? How will it change your clinical practice, if at all?

Faculty _____

Date returned: _____

Describe the main teaching point (EBM slice) you made for this exercise:

<input type="checkbox"/> None	
<input type="checkbox"/> Ask	
<input type="checkbox"/> Acquire	
<input type="checkbox"/> Appraise	
<input type="checkbox"/> Apply	

Figure 1. Clinical question form.

have to know, for every trainee's set of patients, the denominator of evidence-based therapeutic options, making it impractical on a programmatic scale.

We can also document EBP performance by auditing records for adherence to evidence-based guidelines or quality indicators. Hardly a new development, this type of audit is commonly performed as part of internal quality initiatives or external reviews. Langham used a quality audit to evaluate the effect of an EBP curriculum, documenting improvements in practicing physicians' documentation, clinical interventions, and patient outcomes relating to cardiovascular risk factors (26). Finally, clinical vignettes may represent a more feasible, yet valid, alternative for estimating the provision of evidence-based care (27).

WHICH SHOULD WE MEASURE: STEPS OR ACTIONS?

One could argue that trainees' enactment of EBM steps, however measured, represents an intermediate behavioral outcome. That is, we assume that physicians who consistently perform EBP steps will provide more evidence-based care, which, in turn, will lead to more EBP actions and better patient outcomes. But our clinical experience reminds us that intermediate outcomes may fail to guarantee the ultimate outcomes of interest. Nonetheless, I believe we should document *both* types of EBP performance levels. While practice performance measures represent the ultimate outcome, they remain, by virtue of their downstream vantage, blunt instruments. A physician's performance, for instance, in screening her older male patients for abdominal aortic aneurysm represents the end result of myriad inputs, some of which remain outside of her control. Would a record audit detect that the patient did not adhere with her recommendation to undergo screening because of denied insurance coverage? Or, perhaps the physician did *not* recommend screening but her decision reflected a careful consideration of the patient's particular clinical circumstances and preferences (28, 29), rather than a failure to consider the new guidelines supported by a systematic review of the evidence. Perhaps, for example, a chest radiograph revealed a pulmonary nodule, and she deferred screening until lung cancer was excluded. And finally, we should ensure trainees' inclination to consistently perform EBP steps in clinical practice in anticipation that they will direct this behavior to the unforeseeable (and thus unauditable) clinical problems they will encounter in the future.

HOW SHOULD WE EVALUATE EBP PERFORMANCE IN OUR EDUCATIONAL SETTINGS?

With resources and time limited in educational settings, educators would be wise to start small, select feasible strategies off-the-shelf, and adapt the process to local contextual variables. A portfolio of educational prescriptions represents the most promising technology to document the performance of EBP steps in clinical practice. With a simple system to dispense and collect the forms in place, the trainees can do most of the data entry. In addition, unlike many other approaches, the prescription serves as an education intervention as well, particularly if the trainee reflects upon the EBP moment and reviews it with a faculty member. At a recent accreditation visit to our program, the Residency Review

Committee representative found our completed clinical question forms (Figure) to be more than satisfactory evidence of our residents' practice-based learning and improvement performance.

To document EBP performance, educators can borrow the quality data often already collected by health care organizations or team up with institutional officials to leverage resources for nascent efforts. Also, the American Board of Internal Medicine now offers Practice Improvement Modules for residency programs and Maintenance of Certification credit for faculty who facilitate them (30). Finally, as above, we can put the resident to work collecting her own practice performance data (often causing a fair measure of chagrin) in the context of a quality curriculum (31).

Michael L. Green, MD, MSc
Yale University School of Medicine
New Haven, Connecticut, USA

References

- Green ML. *Acad Med.* 1999;74:686-94.
- Hatala R, Guyatt G. *JAMA* 2002;288:1110-2.
- Miller GE. *Acad Med.* 1990;65:S63-7.
- Covell DG, Uman GC, Manning PR. *Ann Intern Med.* 1985;103:596-9.
- Osheroff JA, Forsythe DE, Buchanan BG, et al. *Ann Intern Med.* 1991; 114:576-81.
- Green ML, Ciampi MA, Ellis PJ. *Am J Med.* 2000;109:218-23.
- Ross R, Verdick A. *Acad Med.* 2003;78:412-7.
- Stevermer JJ, Chambliss ML, Hoekzema GS. *Acad Med.* 1999;74:70-2.
- Cabell CH, Schardt C, Sanders L, Corey GR, Keitz SA. *J Gen Intern Med.* 2001;16:838-44.
- Reckase MD. *Educational Measurement: Issues and Practice.* 1995;14: 12-4, 31.
- Rucker L, Morrison E. *Acad Med.* 2000;75:527-8.
- Oxford Centre for Evidence-Based Medicine. *Educational Prescription.* www.cebm.net/downloads/educational_prescription.rtf (accessed November 2005).
- Toronto Centre for Evidence-Based Medicine. *Educational Prescription.* www.cebm.utoronto.ca/practise/formulate/eduprescript.htm (accessed November 2005).
- Crowley SD, Owens TA, Schardt CM, et al. *Acad Med.* 2003;78:270-4.
- Fung MF, Walker M, Fung KF, et al. *Med Educ.* 2000;34:474-9.
- Ellis J, Mulligan I, Rowe J, Sackett DL. *Lancet.* 1995;346:407-10.
- Michaud G, McGowan JL, van der Jagt R, Wells G, Tugwell P. *Arch Intern Med.* 1998;158:1665-8.
- Gill P, Dowell AC, Neal RD, et al. *BMJ.* 1996;312:819-21.
- Lai TY, Wong VW, Leung GM. *Br J Ophthalmol.* 2003;87:385-90.
- Jemec GB, Thorsteinsdottir H, Wulf HC. *Int J Dermatol.* 1998;37:850-4.
- Myles PS, Bain DL, Johnson F, McMahon R. *Br J Anaesth.* 1999;82: 591-5.
- Kingston R, Barry M, Tierney S, Drumm J, Grace P. *Eur J Surg.* 2001;167:324-30.
- Kenny SE, Shankar KR, Rintala R, Lamont GL, Lloyd DA. *Arch Dis Child.* 1997;76:50-3.
- Geddes JR, Game D, Jenkins NE, et al. *Qual Health Care.* 1996;5:215-7.
- Straus SE, Ball C, Balcombe N, Sheldon J, McAlister FA. *J Gen Intern Med.* 2005;20:340-3.
- Langham J, Tucker H, Sloan D, et al. *Br J Gen Pract.* 2002;52:818-24.
- Peabody JW, Luck J, Glassman P, et al. *Ann Intern Med.* 2004;141:771-80.
- Oswald N, Bateman H. *J Eval Clin Pract.* 2000;6:139-48.
- Haynes RB, Devereaux PJ, Guyatt GH. *ACP J Club.* 2002;136:A11-4.
- American Board of Internal Medicine. *Practice Improvement Modules for Residency Training.* www.abim.org/cert/tet_pims.shtm (accessed 10 May 2006).
- Holmboe ES, Prince L, Green M. *Acad Med.* 2005;80:571-7.