

cine or to determine areas of weakness before instruction or practice.

The test may also be useful to show competency in evidence based medicine. A passing score could be defined by asking individuals who are agreed to meet or exceed minimum competence to take the test and setting a minimum proficiency score based on the range of these scores.

Validity and reliability

There are limitations to the validity, reliability, and general utility of the Fresno test. The groups we used to develop and validate the test probably represented the extremes of proficiency, leaving the middle ground relatively under-represented. The properties of the test may change when it used to assess groups of people that are more representative of the full range of proficiency in evidence based medicine.

The content of the test is based on the domains of evidence based medicine as promulgated by several widely read authors.⁴⁻⁹ Nonetheless, there may be disagreement about whether these are the most relevant areas or about whether the questions and grading rubric accurately represent ideal content. For example, on the test item about external validity (or relevance) the expert group did not score significantly higher than the novice group. We chose to retain this item because it examines the recently emphasised issue of clinical relevance,^{5,6} which we have found useful in our curriculum. As the evidence based medicine evolves, individual items may be more or less representative of current practice.

This test relies exclusively on the opinion of experts as the ultimate standard against which candidates are judged. Although expert opinion is the standard when developing tests, practising physicians are more concerned with improved patient outcomes. However, as no test exists that measures patient outcomes, the Fresno test is an improvement over current methods of assessing learning by self report.¹

The inter-rater reliability reported here is high despite the inherent subjectivity of a test of this nature. The two raters participated in the construction and revision of the rubrics and therefore knew them well when scoring these tests. This familiarity with the rubrics may have led to unrealistically consistent scoring.

Also, the test presently has only one set of clinical vignettes and one set of numeric examples for calculation questions. We have written, but not tested, new clinical vignettes. Other vignettes will probably be needed if the test is used in other clinical disciplines.

Conclusions

The Fresno test is the first standardised, objective measure of ability in evidence based medicine that requires learners to demonstrate knowledge and skill. It can assess the effectiveness of teaching in evidence based medicine and identify strengths and weaknesses of curriculums and individuals. Further investigation might examine whether reliability and validity extends to new sets of raters and learners in other clinical disciplines and to other clinical vignettes. Medical educators may be further challenged to develop tests that reliably assess use of evidence based medicine in real clinical circumstances, not simulated or prompted by vignettes.

We thank John Smucney, Upstate Medical University, New York, for providing the validation data set.

What is already known on this topic

Instruction in evidence based medicine is provided in many medical education settings, but its effectiveness is unknown

Existing measures to assess competence tend to be narrowly focused and of uncertain validity

What this study adds

The Fresno test measures a wide range of knowledge and skills necessary for evidence based practice

The standardised grading systems produced a high degree of consistency between graders

Experts scored significantly higher on the test than novices in evidence based medicine, showing that the test has construct validity

Contributors: KDR participated in constructing the initial test and grading rubric, grading of test, and subsequent regrading of revisions, was responsible for revisions of rubric, administered test to subjects, solicited volunteer experts to take test, and wrote the manuscript. SS participated in constructing the initial test and grading rubric, grading of test, subsequent regrading of revisions, and writing the manuscript. SMT participated in constructing the initial grading rubric, statistical analysis of data, and wrote part of the methods section of the manuscript. KDR and SS are guarantors

Funding: This research was funded as part of a Residency Training in Primary Care grant to teach evidence based medicine, from the Bureau of Health Profession, Health Resources and Services Administration.

Competing interests: None declared.

- Hatala R, Guyatt G. Evaluating the teaching of evidence-based medicine. *JAMA* 2002;288:1110-1.
- Grad R, Macaulay AC, Warner M. Teaching evidence-based medical care: description and evaluation. *Fam Med* 2001;33:602-6.
- Stern DT, Linzer M, O'Sullivan PS, Weld L. Evaluating medical residents' literature-appraisal skills. *Acad Med* 1995;70:152-4.
- Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based medicine: how to practice and teach EBM*. 2nd ed. London: Churchill Livingstone, 2000.
- Slawson DC, Shaughnessy AF, Barry H. Which should come first: rigor or relevance? *J Fam Pract* 2001;50:209-10.
- Slawson DC, Shaughnessy AF. Obtaining useful information from expert based sources. *BMJ* 1997;314:947-9.
- Bloom BS. *Taxonomy of educational objectives, handbook I: the cognitive domain*. New York: McKay, 1956.
- Wiersma W, Jurs S. *Educational measurement and testing*. 2nd ed. Boston: Allyn and Bacon, 1990.
- Guyatt G, Rennie D, eds. *Users' guides to the medical literature: a manual for evidence-based clinical practice*. Chicago: American Medical Association Press, 2002.

(Accepted 16 October 2002)

Corrections and clarifications

Testing new pharmaceutical products in children

We inadvertently omitted to publish the name and affiliation details of the second author of this editorial (11 January, pp 64-5). We published only the details of Alastair G Sutcliffe, implying that he was the sole author; his coauthor, however, was Vic Larcher, a consultant paediatrician and paediatric ethicist at the Royal London Hospital, London EC1 2DP. We apologise for this error.

Involving patients can work in home blood glucose testing

The author of this letter, David Kerr, has informed us of authorship errors in reference 5 (11 January, pp 103-4). The authors are Ingleby J, Trowbridge S, Kerr D, Cavan DA.