Development and Validation of a Test for Competence in Evidence-Based Medicine

Rushad Patell, MD[1], Paola Raska, PhD[2], Natalie Lee, MD, MPH[3], Gina Luciano, MD[4], Deborah J. DiNardo, MD, MS[5], Amiran Baduashvili, MD[6], Mel L. Anderson, MD[7,8], Frank Merritt, MD[7], and Michael B. Rothberg, MD, MPH[9]

[1]Department of Hematology-Oncology, Beth Israel Deaconess Medical Center, Boston, MA, USA; [2]Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH, USA; [3]National Clinician Scholar, Medicine, Perelman School of Medicine, Pennsylvania, PA, USA; [4]Division of General Medicine and Geriatrics, Baystate Health, Springfield, MA, USA; [5]Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; [6]Section of Hospital Medicine, Department of General Internal Medicine, Weill Cornell Medical College, New York, NY, USA; [7]Department of Medicine, University of Colorado Denver School of Medicine, Aurora, CO, USA; [8]Medicine Service, Rocky Mountain VA Regional Medical Center, Aurora, CO, USA; [9]Department of Internal Medicine , Cleveland Clinic, Cleveland, OH, USA.

BACKGROUND: Medical educators need valid, reliable, and efficient tools to assess evidence-based medicine (EBM) knowledge and skills. Available EBM assessment tools either do not assess skills or are laborious to grade.

OBJECTIVE: To validate a multiple-choice–based EBM test—the Resident EBM Skills Evaluation Tool (RESET).

DESIGN: Cross-sectional study.

PARTICIPANTS: A total of 304 medicine residents from five training programs and 33 EBM experts comprised the validation cohort.

MAIN MEASURES: Internal reliability, item difficulty, and item discrimination were assessed. Construct validity was assessed by comparing mean total scores of trainees to experts. Experts were also asked to rate importance of each test item to assess content validity.

KEY RESULTS: Experts had higher total scores than trainees (35.6 vs. 29.4, $P < 0.001$) and also scored significantly higher than residents on 11/18 items. Cronbach's alpha was 0.6 (acceptable), and no items had a low item-total correlation. Item difficulty ranged from 7 to 86%. All items were deemed "important" by > 50% of experts.

CONCLUSIONS: The proposed EBM assessment tool is a reliable and valid instrument to assess competence in EBM. It is easy to administer and grade and could be used to guide and assess interventions in EBM education.

KEY WORDS: evidence-based medicine; medical education; test.

BACKGROUND

Evidence-based medicine (EBM) is a method of medical practice that entails utilizing the most relevant and valid evi-

dence available to make clinical decisions.[1, 2] Incorporating formal EBM education into residency training is necessary to develop EBM skills and habits.[3] Many institutions attempt to teach EBM through journal clubs, seminars, or educational sessions that blend EBM instruction with patient care.[4, 6] However, assessments of these programs are limited.

In order to assess curricula, educators need reliable, validated, and easy-to-use tools to test gained knowledge and skills.[7] Ideally, these tests accurately measure applied knowledge and skills, can be easily graded, and are applicable to practicing physicians. Many available tools may measure attitudes and self-efficacy rather than knowledge and skills; recent studies have shown these are poorly correlated.[8] Few tools that assess knowledge have been tested for psychometric properties and the ability to discriminate between different expertise levels in EBM.[9, 11] The available tools also evaluate limited constructs or use an open-ended question design, which make them impractical for training programs given time constraints and the burden of grading such tests. Performance of these tests has not been shown to correlate well with others, even for the same constructs. [12]

Due to the limitations of current EBM assessment tools, we developed an EBM examination, the Resident EBM Skills Evaluation Tool (RESET), at Baystate Health, in Springfield, MA, to evaluate an EBM Ambulatory Morning Report Curriculum.[13] This instrument evaluates various aspects of the EBM process and, if validated, would provide an easy-to-grade assessment of EBM curricula. The goal of this study was to validate the RESET by administering it to residents at 5 US training programs and to EBM experts.

MATERIAL AND METHODS

EBM Tool and Development

Based on the approach outlined in the JAMA User's Guide to the Medical Literature[14] we developed questions to cover selected representative EBM content and using day-to-day clinical scenarios connected to published studies. We began with a true

statement, then created additional false answers that sounded plausible. Incorrect answers often paraphrased erroneous statements we heard from residents in class. We reviewed and edited the preliminary test based on the ABIM criteria for test writing[15] and for accuracy, and discarded any duplicative questions.

We then distributed the test, without answers, to EBM colleagues nationally, who completed the test and also commented on the content and wording of the questions. We made additional edits based on this feedback. We discarded additional questions if another question asked for the examinee to perform a similar skill (e.g., a question on sensitivity was discarded as a question on specificity already existed) or if the question addressed knowledge rather than applied skills. We added questions based on content that the EBM experts felt should be represented on the test. Finally, if expert answers differed on a question, we either edited the question or its corresponding set of answers until the experts all agreed on the answer, or we discarded the question completely.

We distributed the test again, without answers, to a new set of education and EBM colleagues nationally, and clarified questions further based on their feedback. One education expert commented that he was able to answer a few questions based solely on test-taking skills, and these questions were edited to avoid that pitfall.

The RESET consists of eighteen multiple-choice questions assessing EBM applied skills. The test employs one short answer and seventeen multiple-choice questions. The first item asks the respondent to write a focused clinical question. Responses are given one point for each of the following included components identified: a patient population, an intervention, a comparison, and an outcome, for a total of 4 possible points. The remaining items are given one point if the correct option was selected and no points otherwise. Questions represented four EBM constructs: ask (Q1), appraise (Q2–Q7), apply (Q9, Q13, Q15–Q18), and interpreting results (Q8, Q10–Q12, Q14). The questions were weighted by multiplying by the following scalar: ask (3) appraise (2) apply (2) and interpreting results (2.4) and then summed to obtain the overall score with equal weightings for each construct, ranging from 0 to 48 points. Thus, questions were weighted to produce an overall score with equal weightings for each construct, ranging from 0 to 48 points. Questions that were left unanswered were marked incorrect.

Study Population and Design

In order to validate this EBM test, we administered it to internal medicine residents in multiple centers across the USA and to a new cohort of EBM experts between August 2016 and July 2017.

Residents

IRB approval was obtained at each site prior to administration. The test was administered to all available internal medicine residents at the following institutions: Cleveland Clinic (Cleveland, OH), New York Presbyterian – Weill Cornell Medicine (New York, NY), Baystate Health (Springfield, MA), University of Colorado (Aurora, CO), and University of Pittsburgh Medical Center (Pittsburgh, PA). Tests were administered on paper, collected by the site director, and mailed in a sealed envelope to Dr. Rothberg. Resident data was de-identified prior to the analysis. For residents at institutes other than CCF, participation was voluntary. An information sheet accompanying the test emphasized that participation in the study was voluntary and further specified how identity and test results would remain confidential. Test instructions specified that no resources apart from a calculator were to be used to answer the questions. In addition to the test questionnaire itself, residents were asked to specify their training level, to rate their global self-perceived proficiency in using EBM clinically (as either novice, competent, or expert), and to recall the number of times they had used a primary article to answer a clinical question in the past week.

Experts

A list of self-identified experts was contacted via email and asked to complete the test and rate the importance of each question with respect to its importance to EBM competency for a graduating Internal Medicine resident. Invitations were sent to members of the following groups: Society of General Internal Medicine EBM task force and EBM interest group, AHRQ Evidence-Based Practice Centers, and the NY Academy of Medicine EBM task force. Two follow-up emails were sent at weekly intervals to ensure maximum participation. Experts completed the test online, but the questions were in every other way identical to the paper version. Participation was voluntary and no incentive was provided. In addition, respondents were asked to suggest names of other EBM experts (snowball sampling).

Study data were collected and managed using REDCap Survey (Research Electronic Data Capture) a secure, web-based application designed to collect survey data for research. Resident responses were entered into the database by a research assistant, whereas expert responses were entered directly into REDCap via the web interface. The database is hosted at the Cleveland Clinic Datacenter behind a firewall. Only the research team had access to the responses to the survey.

Analysis

Test Properties. To establish content validity, experts in the validation cohort scored each question for importance in testing EBM proficiency, using a 5-point Likert scale ranging from "Not at all important" to Very important." Inter-rater reliability was assessed for the short answer question (Q1); all other questions were multiple-choice. Two raters (RP and MR) scored Q1 for a subset of 75 tests, and subsequent tests were scored by RP. Internal reliability and individual item discrimination were computed for experts and novices separately. Cronbach's alpha and item-total correlation were used

to measure internal reliability and the item discrimination index for individual item discrimination. Individual item difficulty was defined as the inverse percentage answered correctly by residents and experts. To assess construct validity, mean scores of experts and novices were compared by two-tailed *t* test. Significance was assessed by a *p* value of < 0.05.

Sample Size. A power calculation was conducted to identify the required numbers of "experts" and "novices" to demonstrate construct validity (mean scores of experts and novices compared by *t* test). Based on test scores of a previous cohort of novices, the variance in the overall score was 57. Assuming that the variance in overall test scores for the experts would be half that of novices, having at least 300 novices and 30 experts would provide an 80% power to detect a mean difference of at least 3 points at an alpha of 0.05.

Regression Analysis. As an exploratory analysis, univariable and multivariable regression were performed in the resident cohort to assess associations between total test scores and level of training (PGY1, PGY2, PGY3/4), self-perceived proficiency, number of research articles used in the past week, and training site. Analyses were conducted using R Statistical Software.

RESULTS

A total of 304 residents completed the test. These included 106/175 (60.5%) residents from program A, 58/130 (44.6%) from program B, 21/60 (35%) from program C, 61/155 (39.3%) from program D, and 58/110 (52.7%) from program E. Residents were mostly PGY1 (*n* = 163, 54%) followed by PGY2 (*n* = 67, 22%) and PGY3 or higher (*n* = 74, 24%). Most residents rated their current proficiency at using EBM clinically at "novice" (*n* = 183, 60%) or competent (*n* = 114. 38%). Only one resident (0.3%) rated their proficiency as expert and 6 residents did not respond to this question. The median number of times they used a primary article to answer a clinical question in the past week was 1 (IQR 0, 2]. Both self-perceived proficiency and number of articles used were strongly associated with PGY level (*p* < 0.001 for both). Of the 55 EBM experts who were invited via email to participate, 34 (62%) completed the survey (Table 1). Ninety-five percent of the residents and 97% of the experts made it to the last question.

The mean total score among experts was significantly higher than that of the residents (35.6 vs. 29.4, *p* < 0.001). The distribution of the total scores for residents and experts is shown in Fig. 1. The percentage of correct answers for individual questions ranged from 7 to 86% for trainee respondents, and from 30 to 100% for experts. The experts significantly outperformed the residents on 11 items (Fig. 2). The largest differences in correct answers appeared on questions 4, 16, and 18, which tested understanding of confounding in observational studies, interpretation of likelihood ratios, and understanding of odds ratios, respectively. Both groups performed poorly on question 17, which assessed understanding of survival curves. With regard to individual constructs, the experts significantly outperformed the novices in appraise (4.5 vs. 3.5, *p* < 0.001), apply (3.4 vs. 2.3, *p* < 0.001), and interpret (4.5

Table 1 Scores by Level of Training, Self-Perceived Competence, and Institution

	Univariable analysis			Multivariable analysis	
	N	Score	*p* value†	Adjusted score*	*p* value†
Cohort			< 0.001		
Residents	304	29.4		–	
Experts	33	35.6		–	
Residents					
Level Of Training			0.08		0.01
PGY1	167	28.9		26.6	
PGY2	67	29.1		28.5	
PGY3/4	74	30.9		29.5	
Self-Perceived Competence			0.21		0.14
Novice	183	29.0		26.6	
Competent	114	30.3		27.4	
Expert	1	31.6		29.6	
Number of articles used/week			0.11		0.15
0	89	30.0		26.6	
1	78	30.5		27.7	
>1	129	28.7		25.9	
Institution of training			0.005		0.01
A	106	28.0		26.6	
B	58	31.0		30.0	
C	21	27.2		25.8	
D	58	30.0		29.4	
E	61	30.7		29.1	

*Adjusted for level of training, self-perceived competence, number of articles used per week, and institution
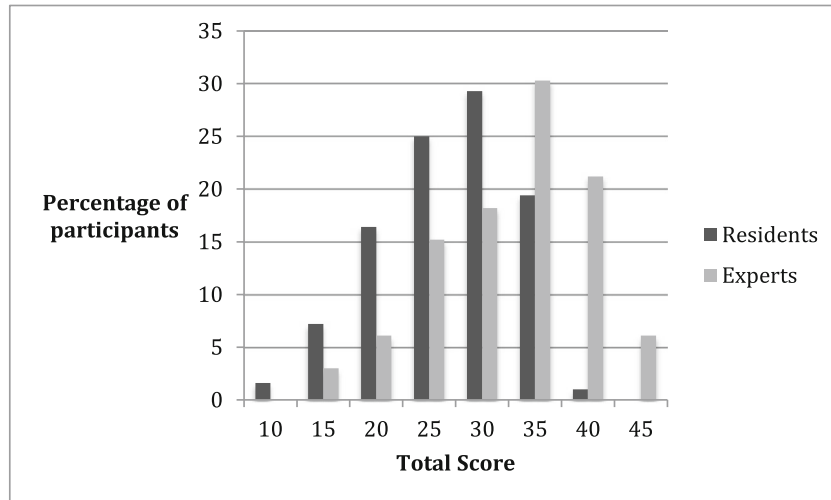†p values compare scores across all levels of each category

Figure 1 Distribution of total scores for residents and experts.

vs. 3.4, $p < 0.001$) but not ask (3.0 vs. 3.3, $p = 0.83$). Every question was deemed important (≥ 3 on the Likert scale) by more than 50% of the experts, including three questions that more than half of the experts answered incorrectly (Fig. 3).]–>]–>]–>

Internal consistency was higher for experts than residents, but the difference was not significant (0.6 vs. 0.4, $p = 0.14$). As an additional measure of internal reliability, per-item correlations to the total score for the construct were computed and averaged (Fig. 4). The mean individual item correlation for residents was 0.3 and 0.37 for the experts. Inter-rater reliability for the single short answer question was 0.9; this measure was not applicable to the other multiple-choice questions.]–>

Univariable regression analysis is presented in Table 1. On multivariable analysis, there was a statistically significant difference in performance by site, with three sites outperforming the other two. Higher levels of training were associated

with higher scores, whereas self-perceived competency and number of articles used in the previous week were not.

DISCUSSION

In this study, we describe a multiple-choice EBM test developed to assess the EBM skills of trainees. When test results from trainees of different levels at 5 different institutions were compared with results from EBM experts, the test demonstrated good discriminative ability, internal validity, and construct validity. The one open-ended question assessing the ability to ask an answerable clinical question displayed excellent inter-rater reliability between graders, suggesting that the test should be highly reproducible across institutions.

The RESET was designed to address multiple components in the EBM process: ask, appraise, and apply, with appraise
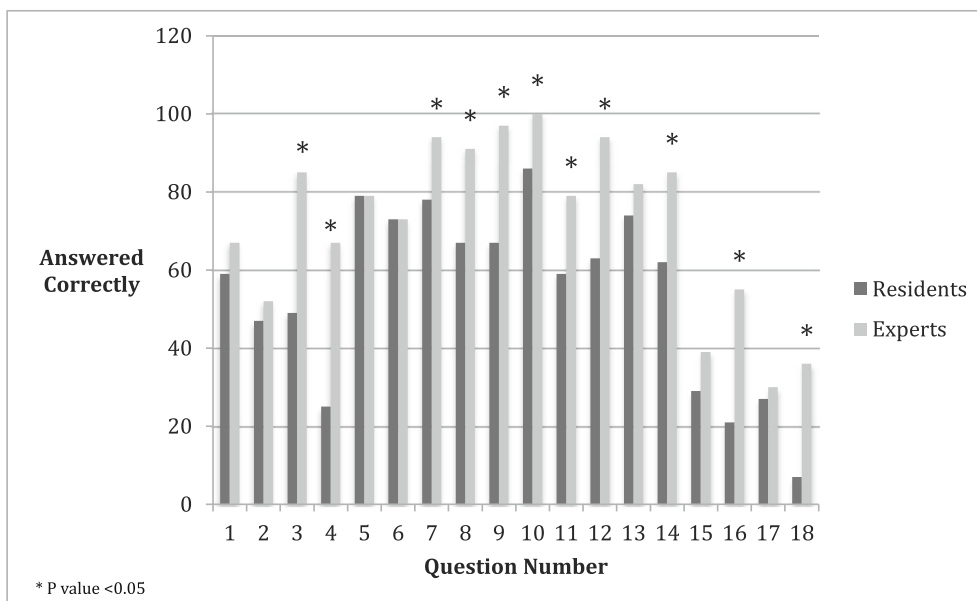

Figure 2 Percent correct for each item in the 18-item questionnaire, item difficulty.
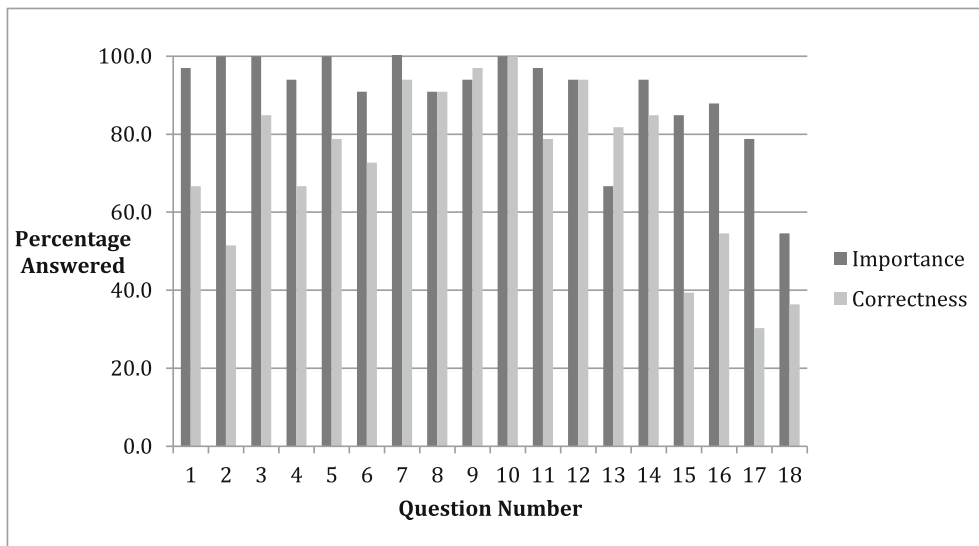
Figure 3 Percent of experts that deemed each question important (>=3 in Likert scale) and percent item was answered correctly.

further divided into "validity" and "interpretation of results." Moreover, the results were weighted over the four constructs to create a balanced total score. Although open-ended questions may assess higher order thinking in response to an authentic task, they require more time and expertise to grade and such tests are impractical for higher throughput testing.[15] Accordingly, the proposed tool was designed in a multiple-choice format, using scenarios and examples that would be relevant to internal medicine practice to test knowledge and skills.

Assessing an EBM educational program requires a validated, objective tool, as self-perception of ability corresponds poorly to objective assessment of skills.[16] We also found that residents' self-perception did not correlate with their test scores. In contrast, the RESET successfully differentiated self-identified EBM experts from residents at five different training programs across the country, demonstrating robust discriminative capacity and construct validity. Further, senior residents performed better than interns, adding to construct validity and suggesting that residents acquire some EBM skills during training, although the increase in scores was modest, suggesting that our EBM curricula may need to be improved. Residents also varied across programs, implying that some programs were more effective than others. The greatest variation, however, was among residents within each program. This could reflect varying uptake of EBM material by individuals or simply better test-taking skills. We tried to make the questions clinically relevant and to avoid answers that could be discerned without a thorough knowledge of the material. Content validity was also confirmed by the experts, who rated every question as "important."

A recent systematic review of EBM educational practices identified that of 24 instruments used, only 6 were considered high-quality (achieved ≥ 3 types of established validity
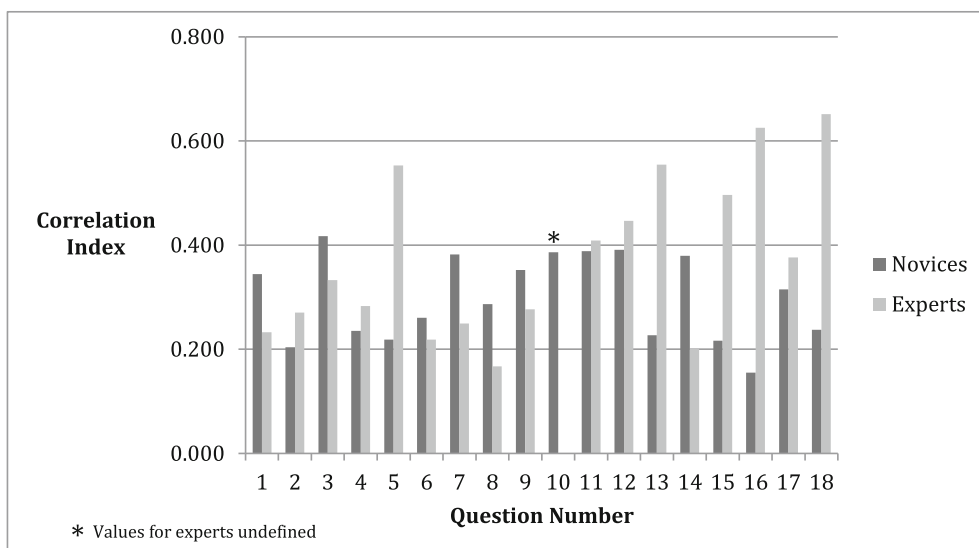

∗ Values for experts undefined

Figure 4 Correlations between each item and the total score.

Table 2 Psychometric Properties of Study Tool Compared With Other Available EBM Competence Tests

Test property	Measure to be used	Acceptable results	RESET tool	ACE tool	Berlin (two different sets)	Fresno
Content validity	Expert opinion	Test covers EBM topics	All items were scored at least "important" by > 50% of experts	Acceptable	Not reported	"revisions based on experts' suggestions"
Inter-rater reliability	Inter-rater correlation	Expected to be high (> 0.6)	0.9	N/A	N/A	0.76–0.98
Internal reliability	Cronbach's alpha [23] Item-total correlation	0.6–0.7 = questionable 0.7–0.8 = acceptable 0.8–0.9 = good > 0.9 = excellent	0.6	0.69	0.75, 0.82	0.88
		> 0.30 (per Fresno)	0.37	0.14 to 0.20, apart from three items (0.03, 0.04, and 0.06)		0.47 to 0.75
Item difficulty	Percentage of candidates who correctly answer each item	Wide range of results allows test to be used in expert and novice groups	7–86% (trainees), 30–100% (experts)	36–84%	Not reported	24 to 73%
Item discrimination	Item discrimination index (ranges from − 1.0 to 1.0)	All items should be positively indexed, > = 0.2 is considered acceptable	0.17–0.65	0.37–0.84	Not reported	0.41–0.86
Construct validity	Mean scores of experts and novices compared by t test % passing for expert and novice groups compared by χ^2 test		48-point test: novice 29.4, expert 35.4 ($p < 0.001$)	15-point test: novice 8.6, intermediate 9.5, advanced 10.4 ($p < 0.001$)	ANOVA: 4.2 control, 6.3 course participants, 11.9 expert ($p < 0.001$)_	212-point test: novice 95.6, expert 147.5 ($p < 0.001$)

N/A, not assessed

evidence), and only 4 measured both knowledge and skills.[17] Of these, the Berlin and Fresno tests had robust psychometric properties and the ability to discriminate among different expertise levels in EBM.[9][11] The Berlin assessment tool consists of 15 multiple-choice questions and largely assesses the "Appraise" step.[17] The Fresno assessment tool evaluates four steps—ask, acquire, appraise, and apply—using clinical scenarios and open-ended questions.[17] The ACE tool was more recently developed and was not available when we developed the RESET. This tool may more comprehensively address EBM steps than our proposed tool, which does not assess the "acquire" step in EBM. However, internal reliability (which they defined as the item-total correlation using Pearson product-moment correlation coefficients) was poor for three of the fifteen items in the ACE tool, all three of which were related to critical appraisal. Moreover, a test format limiting choices to yes/no answers may fail to capture the full spectrum of thought processes. Finally, the ACE tool has been validated only for medical undergraduates and not other medical trainees.[18] We used standard psychometric measures to compare our test against other available tools for EBM competence assessment (Table 2). Cronbach's alpha for the RESET was lower than other instruments but within the acceptable range.

Our study has limitations. Responsive validity refers to the ability of a tool to assess the impact of an EBM educational intervention and requires statistical comparison of the same participant's scores before and after an EBM educational intervention.[9] The RESET has not yet been assessed to evaluate programmatic rather than individual impact of EBM interventions, although we plan to do so. Using the same tool longitudinally may not be practical to gauge the impact of an EBM educational intervention or a change in learners' skill because of familiarity with the questionnaire. Multiple versions of the test could be developed to overcome this potential shortcoming; however, the versions would have to correlate with each other to ensure comparability. In our experience, residents do not learn from the test and often get the same questions wrong the following year. Further, this tool does not explicitly test the learner's ability to "acquire" evidence to answer their question nor is it designed to assess meaningful changes in behavior.[17] Given these limitations, it would be premature to use the test for high-stakes evaluation purposes. Our "experts" were self-identified and their level of EBM skill may not have been uniform. Nevertheless, the tool broadly distinguished them from trainees. Like the Fresno and Berlin tools, the RESET was tested in a physician population, so results may not apply more broadly. Non-physicians were rarely evaluated in the development of other instruments[9] and yet EBM education is part of curricula in allied medical fields such as pharmacy, nursing, physical therapy, and others.[19][20] The Fresno tool has been successfully adapted for doctor of pharmacy and physical therapy students.[21][22]

Future studies can likewise examine the performance of this test in other professions. Lastly, the RESET takes at least 30 min to administer, which could be burdensome. Most residents completed it, but there was some drop-off as the questions became increasingly difficult. We did not formally assess the average time taken to complete the test, and the 30-min estimate is based on our experience across the centers. A formal time limit was not enforced. Only 11 of the 18 questions clearly differentiated residents from experts. Future studies could attempt to reduce the number of questions and still maintain appropriate discrimination. For purposes of identifying weaknesses for individual residents, however, the complete test would still be required.

In conclusion, we present a validated, practical questionnaire that can used to assess knowledge and skills in the recognized key constructs of EBM. With this practical and accurate means to test trainee's knowledge, educators can identify gaps, design interventions, and assess effectiveness of these interventions in a longitudinal manner.

Corresponding Author: *Michael B. Rothberg, MD, MPH; Department of Internal Medicine Cleveland Clinic, 9500 Euclid Avenue, Mail Code G10, Cleveland, OH 44195, USA (e-mail: rothbem@ccf.org).*

REFERENCES

1. **Sackett DL SS**, **Richardson WS**, **Rosenberg W**, **Haynes RB**. Evidence-Based Medicine: How to Practice and Teach EBM. 2 Edinburgh: Churchill Livingstone; 2000.
2. **Ilic D**, **Maloney S**. Methods of teaching medical trainees evidence-based medicine: a systematic review. Med Educ. 2014;48(2):124–35.
3. **Paulsen J**, **Al Achkar M**. Factors associated with practicing evidence-based medicine: a study of family medicine residents. Adv Med Educ Pract. 2018;9:287–93.
4. **Coomarasamy A**, **Taylor R**, **Khan KS**. A systematic review of postgraduate teaching in evidence-based medicine and critical appraisal. Med Teach. 2003;25(1):77–81.
5. **Hatala R**, **Keitz SA**, **Wilson MC**, **Guyatt G**. Beyond journal clubs. Moving toward an integrated evidence-based medicine curriculum. J Gen Intern Med. 2006;21(5):538–41.
6. **Green ML**. Evidence-based medicine training in internal medicine residency programs a national survey. J Gen Intern Med. 2000;15(2):129–33.
7. **Albarqouni L**, **Hoffmann T**, **Glasziou P**. Evidence-based practice educational intervention studies: a systematic review of what is taught and how it is measured. BMC Med Educ. 2018;18(1):177.
8. **Snibsoer AK**, **Ciliska D**, **Yost J**, **Graverholt B**, **Nortvedt MW**, **Riise T**, et al. Self-reported and objectively assessed knowledge of evidence-based practice terminology among healthcare students: a cross-sectional study. PLoS One. 2018;13(7):e0200313.
9. **Shaneyfelt T**, **Baum KD**, **Bell D**, **Feldstein D**, **Houston TK**, **Kaatz S**, et al. Instruments for evaluating education in evidence-based practice: a systematic review. JAMA. 2006;296(9):1116–27.
10. **Ramos KD**, **Schafer S**, **Tracz SM**. Validation of the Fresno test of competence in evidence based medicine. BMJ. 2003;326(7384):319–21.
11. **Fritsche L**, **Greenhalgh T**, **Falck-Ytter Y**, **Neumayer HH**, **Kunz R**. Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. BMJ. 2002;325(7376):1338–41.
12. **Lai NM**, **Teng CL**, **Nalliah S**. Assessing undergraduate competence in evidence based medicine: a preliminary study on the correlation between two objective instruments. Educ Health (Abingdon). 2012;25(1):33–9.
13. **Luciano GL**, **Visintainer PF**, **Kleppel R**, **Rothberg MB**. Ambulatory morning report: a case-based method of teaching EBM through experiential learning. South Med J. 2016;109(2):108–11.
14. **Gordon Guyatt DR**, **Maureen O. Meade**, **Deborah J. Cook**. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3: McGraw-Hill Medical, 2008.
15. **Melovitz Vasan CA**, **DeFouw DO**, **Holland BK**, **Vasan NS**. Analysis of testing with multiple choice versus open-ended questions: outcome-based observations in an anatomy course. Anat Sci Educ. 2018;11(3):254–61.
16. **Khan KS**, **Awonuga AO**, **Dwarakanath LS**, **Taylor R**. Assessments in evidence-based medicine workshops: loose connection between perception of knowledge and its objective assessment. Med Teach. 2001;23(1):92–4.
17. **Ilic D**. Assessing competency in evidence based practice: strengths and limitations of current tools in practice. BMC Med Educ. 2009;9:53.
18. **Thomas RE**, **Kreptul D**. Systematic review of evidence-based medicine tests for family physician residents. Fam Med. 2015;47(2):101–17.
19. **Horntvedt MT**, **Nordsteien A**, **Fermann T**, **Severinsson E**. Strategies for teaching evidence-based practice in nursing education: a thematic literature review. BMC Med Educ. 2018;18(1):172.
20. **Manske RC**, **Lehecka BJ**. Evidence - based medicine/practice in sports physical therapy. Int J Sports Phys Ther. 2012;7(5):461–73.
21. **Coppenrath V**, **Filosa LA**, **Akselrod E**, **Carey KM**. Adaptation and validation of the Fresno test of competence in evidence-based medicine in doctor of pharmacy students. Am J Pharm Educ. 2017;81(6):106.
22. **Tilson JK**. Validation of the modified Fresno test: assessing physical therapists' evidence based practice knowledge and skills. BMC Med Educ. 2010;10:38.
23. **Bland JM**, **Altman DG**. Statistics notes: Cronbach's alpha. BMJ. 1997;314(7080):572.