

RESEARCH

Adaptation and Validation of the Fresno Test of Competence in Evidence-Based Medicine in Doctor of Pharmacy Students

Valerie Copenrath, PharmD, Lee Anne Filosa, PharmD, Elizabeth Akselrod, PharmD, Katherine M. Carey, PharmD

MCPHS University, School of Pharmacy Worcester/Manchester, Worcester, Massachusetts

Submitted May 20, 2016; accepted April 24, 2017; published August 2017.

Objective. To adapt and validate an instrument assessing competence in evidence-based medicine (EBM) in Doctor of Pharmacy students.

Methods. The Fresno test was validated in medical residents. We adapted it for use in pharmacy students. A total of 120 students and faculty comprised the validation set. Internal reliability, item difficulty, and item discrimination were assessed. Construct validity was assessed by comparing mean total scores of students to faculty, and comparing proportions of students and faculty who passed each item.

Results. Cronbach's alpha was acceptable, and no items had a low item-total correlation. All of the point-biserial correlations were acceptable. Item difficulty ranged from 0% to 60%. Faculty had higher total scores and also scored higher than students on most items, and 8 of 11 of these differences were statistically significant.

Conclusion. The Pharm Fresno is a reliable and valid instrument to assess competence in EBM in pharmacy students. Future research will focus on further refining the instrument.

Keywords: evidence-based medicine, psychometrics, reliability, validity

INTRODUCTION

Practicing evidence based medicine (EBM) is the purposeful search for and utilization of the best available evidence to make decisions about an individual's care.¹ The ability to identify, appraise, and apply biomedical research is an essential skill for pharmacists in their role as drug experts. As a result, the Accreditation Council for Pharmaceutical Education (ACPE) and the Center for the Advancement of Pharmaceutical Education (CAPE) require curricula in EBM as part of the education for the Doctor of Pharmacy degree (PharmD).^{2,3} Various strategies to measure knowledge and skills in EBM have been described.⁴⁻⁹ No published studies of an instrument validated to measure competence in EBM for pharmacy students were identified in a literature search conducted in March 2017.

In order to assess competency in EBM, the profession of pharmacy needs an instrument that is valid and reliable. Such an instrument will facilitate the design of pedagogical interventions to increase competency in EBM, which may contribute to improvements in patient care and how pharmacists are perceived by other health care providers.

Corresponding Author: Valerie Copenrath, MCPHS University, 19 Foster St., Worcester, MA 01608. Tel: 508-373-5692. Fax: 508-756-8715. E-mail: valerie.copenrath@mcphe.edu

The four steps of EBM include formulating an answerable question (Ask), finding the most relevant evidence (Acquire), appraising the evidence for validity and usefulness (Appraise), and applying these results to clinical practice (Apply).¹⁰ The clinical question is typically framed in the PICO format (Patient, Intervention, Comparator, Outcome). There are a multitude of instruments available to assess EBM competence, knowledge, attitudes, and behavior, but few have been validated in the medical profession that assess all steps of the EBM process.¹¹ The Fresno test,¹² the Berlin questionnaire,¹³ and the ACE tool¹⁴ each assess all four steps of the EBM process and have robust psychometric properties to support their validation. The Berlin questionnaire is a 15-item multiple choice test, and the ACE tool is a 15-item test that uses yes/no responses. We chose to use the Fresno tool as a starting point for our instrument because it requires open-ended responses rather than the identification of the correct answer from a list. The Fresno tool was developed and validated to assess medical residents' competence in the four steps of EBM¹² (the complete question and answer key is available at <http://uthscsa.edu/gme/documents/PD%20Handbook/EBM%20Fresno%20Test%20grading%20rubric.pdf>). However, the Fresno tool has not been validated in pharmacy students.

Additionally, because the original questions and answer key are widely available on the Internet, use of the Fresno in a PharmD curriculum would be problematic. To overcome these limitations, the original Fresno was adapted based on the expected skillset of pharmacists, and was sought to be validated for use in pharmacy students.

METHODS

The original Fresno test¹² is comprised of two clinical scenarios and contains 12 items: seven open response, three calculations, and two fill-in-the-blank questions. After obtaining permission from the authors of the original Fresno, the instrument was adapted for use in pharmacy students. These adaptations were made in order to be consistent with all four steps of the EBM process, the ACPE Standards 2016 (Standard 2.1),² and CAPE Outcomes 2013 (Subdomain 2.1).³ The most significant modification made was to include a published, randomized, controlled trial for participants to use when completing the instrument. This enabled direct assessment of participant ability to critically evaluate biomedical evidence (Appraise) and provide a specific recommendation based on trial results (Apply). Additional modifications included the creation of a new clinical scenario, revision of the calculation questions to include interpretation of number needed to treat (NNT) and confidence intervals (CI), and addition of a written summary of the included clinical trial with an answer to the clinical question.

The original Fresno contains three questions related to prognosis and diagnosis, which were omitted because these skills are less relevant to EBM application by pharmacists, and are not included in either the ACPE Standards 2016 or the CAPE Outcomes 2013.^{2,3} Other items related to treatment were kept the same as the original test. For any items that warranted development of new or modified “correct” answers for the scoring rubrics, two authors independently determined correct responses and then discussed any discrepancies to determine consensus. A scoring system similar to the original Fresno was used. The adapted instrument was titled the “Pharm Fresno test of competence in evidence-based medicine.” Table 1 describes the relationship between the Pharm Fresno items and the original Fresno items, as well as the relationship to each step of the EBM process.

A randomized clinical trial¹⁵ was selected by the investigators and several questions were edited in order to directly assess participants’ ability to complete EBM related skills.

A sample of convenience was recruited from students in the Doctor of Pharmacy classes of 2015 and 2016 and faculty members in the department of pharmacy practice. In the summer of 2014, the Class of 2016 completed the first professional year of three years in the

program, which included one stand-alone course on drug literature evaluation. Participants were recruited through a mandatory dean’s forum and by e-mail. Interested participants were asked to attend a face-to-face session in order to learn the study objectives, purpose, and methods. Students who agreed to participate were asked to provide informed consent and stay to complete the Pharm Fresno. The session was 60 minutes in length, which correlated with available class time. Participation in the study was voluntary, and the score of the Pharm Fresno was not linked to any grades. All participants in the School of Pharmacy Worcester/Manchester Class of 2015 who attended the sessions and provided consent completed the instrument. Additionally, students in the Class of 2015, who were assigned to the Principal Investigators’ Advanced Pharmacy Practice Experiences (APPE) during the 2014-2015 academic year, were recruited. These students had completed all didactic coursework, including the two courses described above. Further, they completed between 0-5 APPE rotations prior to the rotation of the investigators’, where they likely practiced EBM skills. All of the students on rotation with the investigators were required to complete the instrument as part of the APPE experience. However, students provided consent to include their data in the analysis. Investigators were unaware of which students in the Class of 2015 provided consent until after all APPEs for the year were completed. Similarly, scores on the Pharm Fresno were not linked to any grades. Faculty in the department of pharmacy practice were recruited at a departmental retreat. Those faculty who provided consent completed the instrument.

Responses to baseline demographic characteristic questions and the Pharm Fresno were collected electronically and anonymously using Qualtrics Online Survey Software [Insight Platform, licensed account under MCPHS University (formerly known as Massachusetts College of Pharmacy and Health Sciences)]. All participants who opened the Qualtrics link and responded to at least one item of the Pharm Fresno were included in the analysis. This protocol was submitted to the MCPHS University Institutional Review Board and was deemed “exempt.”

The responses collected with Qualtrics were scored by a team of evaluators, which included four faculty and one research fellow using standardized grading rubrics. The first 20 responses were scored as a team, with the differences discussed and the rubrics refined until consensus was reached. The remaining 100 responses were scored independently. Inter-rater reliability was assessed by determining the intra-class correlation of a convenience sample of 16 responses independently scored by two evaluators. As with the validation of the original

Table 1. Pharm Fresno and Map to EBM Process

| | Pharm Fresno Item | EBM Step | Comments |
|----|--|-----------------|---|
| 1 | Write a focused clinical question that will help you organize a search of the clinical literature and choose the best article. | Ask | Same question, same grading rubric. Different clinical scenario will make responses different than published key. |
| 2 | Where might clinicians go to find an answer to questions like these? Name as many categories as you can and describe advantages and disadvantages. | Acquire | Same question, same grading rubric. |
| 3 | Describe your strategy for searching MEDLINE for original research on this clinical question. Explain your rationale and how you would limit your search. | Acquire | Same question, same grading rubric. |
| 4 | What type of study design would best be able to address this question? Why? | Appraise | Same question, same grading rubric. |
| 5 | You conduct your search and identify a report of a well-designed study. What characteristics of the study will you consider in order to determine if it is RELEVANT to your clinical question? | Appraise | Similar question, adapted grading rubric to follow PICO format. |
| 6 | What study characteristics will you consider to evaluate the quality of the METHODS? | Appraise | Adapted question, maintained same grading rubric. |
| 7 | Use the attached article to answer the following question. Which of the outcomes (listed above in the actual test) demonstrated a STATISTICALLY SIGNIFICANT difference(s) between the intervention and control groups? | Appraise | Essentially split one item into two items. Adapted first of these two items to focus on statistical significance. Adapted a portion of one of the rubrics that focused on statistical significance. |
| 8 | Same wording as above, except “statistically significant” was replaced with “clinically significant.” | Appraise | Adapted second of these two items to focus on magnitude and clinical significance. Adapted a portion of one of the rubrics that focused on magnitude. |
| 9 | Using information provided, calculate the following: absolute risk reduction, relative risk reduction, and number needed to treat. Express the number needed to treat in your own words. | Appraise | Changed the provided information. Added a fourth component of expressing NNT in own words. Rubric adapted to include this fourth component. |
| 10 | Provided a confidence interval, 1. In your own words, what does this confidence interval tell you? 2. Does this confidence interval demonstrate a STATISTICALLY significant result? Why or why not? 3. Does this confidence interval demonstrate a CLINICALLY significant result? Why or why not? | Appraise | Adapted question. Created new rubric. |
| 11 | Rate the quality of the trial described in the provided article using a scale of 1-3, with 1 indicating the highest quality studies and 3 indicating the lowest quality studies. Justify your response using 1-2 sentences. | Appraise | Created new question related to appraisal. Created new rubric. |
| 12 | Use the provided article to provide a specific recommendation in the patient scenario, which has been copied below. Be sure to include a concise summary of the trial as the rationale for your recommendation. | Apply | Created new question related to apply step. Created new rubric. |

Fresno, the four properties used to test the Pharm Fresno were internal reliability, item difficulty, item discrimination, and construct validity.¹²

Cronbach’s alpha and item-total correlation were used to determine internal reliability. Cronbach’s alpha was used as the index of internal consistency of the test, with an acceptable range from .7 to .95.¹⁶⁻¹⁹ Item-total correlation was used to determine the reliability of a scale, with an acceptable value of 0.2 or higher.¹⁶

Item difficulty was defined as the proportion of candidates who achieve a passing score on each item.¹² In the Pharm Fresno, passing was defined as an average of “strong” scores for each item. For example, Item 1 has four domains. A score of “excellent” across each domain would result in a score of 24, which is the highest possible score. A score of “strong” across each domain would result in a score of 16. Thus, a score of 16 or higher was considered passing for Item 1.

A point-biserial correlation was used to determine if the highest 27% of performers were more likely to score higher on each item, with positive correlations desired. Correlations between 0.2 and 0.4 are considered acceptable.¹⁹

Two measures were used to test construct validity. First, the mean overall scores of students and faculty were compared using a *t*-test. A statistically significant difference between students and faculty was expected, with faculty scoring higher than students. Secondly, the proportion of participants passing each item in the student and faculty groups was compared using a χ^2 . Again, statistically significant differences were expected, with higher proportions of faculty passing.

Statistical tests were conducted using Microsoft Excel 2013 (Full Version 15) (Redmond, WA) and IBM SPSS Statistics (Version 22) (Armonk, NY). A *p* value of <.05 was considered statistically significant.

RESULTS

A total of 140 participants took the test. Of these, 120 completed at least one question and were included in the analysis. These participants included 45 Class of 2016 (P2) students, 46 Class of 2015 (P3) students, and 30 faculty. A majority of the participants were students, 75%, and 37% were male. A majority of the P2 and P3 students had a bachelor’s degree, 89% and 67% respectively. A majority of the faculty had a PharmD or other doctorate degree (93%). More than half of the P2 and P3 students had a self-reported GPA of 3.0 or higher, 60% and 52% respectively. However, 15% of P3 students did not provide a GPA value. When asked about the comfort level of using EBM, 74% of faculty responded that they were somewhat or very comfortable, as opposed to 64% of P2 students

and 58% of P3 students. Demographic data are reported in Tables 2 and 3.

The mean score for all participants on the 210 point test was 98 points [Standard deviation 25.918; Standard error of measurement 14; 95% Confidence Interval (70-122)]. The internal reliability, item difficulty, and item discrimination are reported in Table 4. The intra-class correlation for the sample of responses scored by two evaluators was .95 (95% Confidence Interval .85-.98), which indicates a high level of agreement between the two raters.

Cronbach’s alpha was .71, which is acceptable. Cronbach’s alpha was calculated with each item removed from the test, and the removal of any of the items would reduce the reliability of the test. The lowest item-total correlation was .25, which is above the accepted value of .2. Item difficulty in the Pharm Fresno ranged from 0% to 60%, indicating a high level of difficulty for the population tested. No participant passed Item 12, which was one of the items newly created in the Pharm Fresno. Excluding Item 12 resulted in a lower Cronbach’s alpha (.7, data not shown) and changed the range of difficulty from 9% to 60%. Being a high performer (the upper 27%) was positively correlated with passing each item. Among faculty in the upper 27% who self-reported advanced competence in EBM, the average score was 121 (n=3), which is 8 points higher than the average faculty score. All of the point-biserial correlations were greater than or equal to 0.2. Item 1 was the least difficult item, with 60% of participants achieving passing scores.

The construct validity is provided in Table 5. The mean total score on the 210 point test was 20 points higher in the faculty group compared with the student group

Table 2. Demographic Data – Age, Gender, and Race

| | Students (n=91) % | Faculty (n=30) % |
|------------------------|----------------------|---------------------|
| Age | | |
| 21-25 | 43 | 0 |
| 26-29 | 43 | 10 |
| 30+ | 14 | 87 |
| No response | 0 | 3 |
| Gender | | |
| Male | 40 | 27 |
| Female | 60 | 67 |
| Race/Ethnicity | | |
| White/Caucasian | 37 | 77 |
| Asian | 43 | 10 |
| Black/African American | 8 | 0 |
| Hispanic | 6 | 33 |
| Other | 7 | 3 |
| No response | 0 | 7 |

Table 3. Demographic Data – Education and EBM Comfort Level

| | % P2 Students (n=45) | % P3 Students (n=46) | % Faculty (n=30) |
|----------------------------|----------------------|----------------------|------------------|
| Prior Education | | | |
| Associate’s Degree | 0 | 4 | 3 |
| Bachelor’s Degree | 89 | 68 | 0 |
| Doctorate Degree | 4 | 2 | 93 |
| Other (incl. undergrad) | 7 | 26 | 0 |
| No response | 0 | 0 | 3 |
| GPA (self-reported) | | | |
| 3.30-4.0 | 33 | 33 | - |
| 3.0-3.29 | 27 | 20 | - |
| 2.30-2.99 | 36 | 28 | - |
| 2.0-2.29 | 2 | 4 | - |
| Can’t remember | 2 | 11 | - |
| No response | 0 | 4 | - |
| EBM Comfort Level | | | |
| Very comfortable | 4 | 4 | 17 |
| Somewhat comfortable | 60 | 54 | 57 |
| Neutral | 16 | 24 | 20 |
| Somewhat uncomfortable | 18 | 13 | 7 |
| Very uncomfortable | 2 | 2 | 0 |
| No response | 0 | 2 | 0 |

[(113 vs 93, $p < .005$ (95% CI 9.2-29.8)]. There were no statistically significant differences between faculty and student scores on Items 1, 5, 7, and 12. Faculty scored higher than students on the remaining eight items, and these differences were statistically significant.

DISCUSSION

Based on our search of the MEDLINE database in March 2017, this is the first documented validation of a test of competency in EBM in pharmacy students. It was modified based on the expected skill set of pharmacists. We replicated Ramos and colleagues’ measures for

validation.¹² For any instances in which we were unable to elicit the exact testing procedure utilized by Ramos and colleagues, we based our reference range for acceptable results on published literature. The Pharm Fresno is likely suited to small groups, as it was cumbersome to score the large number of responses even with a team of five individuals familiar with the rubric. The Pharm Fresno meets the criteria for a Level 1 Instrument to measure competency in EBM, as described by the 2006 systematic review.¹¹ Specifically, we established inter-rater reliability, used objective outcome measures, and used three or more types of established validity tests. The greatest

Table 4. Results of Reliability and Validation Studies

| | Cronbach’s alpha | Item-Total Correlation | Point Biserial ^a | Item Difficulty(%) | Not Evident, n (%) |
|-----------|------------------|------------------------|-----------------------------|--------------------|--------------------|
| Item 1 | - | 0.33 | 0.12 ^b | 59 | 1 (1) |
| Item 2 | - | 0.50 | 0.47 | 26 | 12 (10) |
| Item 3 | - | 0.41 | 0.32 | 60 | 2 (1.7) |
| Item 4 | - | 0.35 | 0.24 | 24 | 3 (2.5) |
| Item 5 | - | 0.40 | 0.51 | 9 | 9 (7.5) |
| Item 6 | - | 0.42 | 0.53 | 43 | 22 (18.3) |
| Item 7 | - | 0.38 | 0.48 | 54 | 22 (18.3) |
| Item 8 | - | 0.38 | 0.50 | 14 | 64 (53.3) |
| Item 9 | - | 0.38 | 0.33 | 46 | 28 (23.3) |
| Item 10 | - | 0.25 | 0.25 | 23 | 30 (25) |
| Item 11 | - | 0.28 | 0.27 | 17 | 40 (33.3) |
| Item 12 | - | 0.26 | 0.38 | 0 | 15 (12.5) |
| All Items | 0.71 | - | - | - | - |

^aAll correlations are significant at the 0.1 level (2-tailed) unless otherwise specified

^bCorrelation is significant at the 0.5 level (2-tailed)

Table 5. Construct Validity

| | Students | Faculty | <i>p</i> value |
|-------------------------|----------|---------|----------------|
| Mean score ^a | 93 | 113 | <.005 |
| % Passing ^b | | | |
| Item 1 | 57 | 67 | .33 |
| Item 2 | 20 | 43 | .01 |
| Item 3 | 56 | 77 | .04 |
| Item 4 | 19 | 40 | .02 |
| Item 5 | 9 | 10 | .86 |
| Item 6 | 38 | 63 | .01 |
| Item 7 | 56 | 50 | .59 |
| Item 8 | 9 | 30 | .004 |
| Item 9 | 36 | 80 | .00002 |
| Item 10 | 19 | 37 | .046 |
| Item 11 | 10 | 40 | .0002 |
| Item 12 | 0 | 0 | n/a |

^aCompared with *t* test

^bCompared with chi square

strengths are the inclusion of a published randomized controlled trial as a way of assessing actual skills in EBM, and utilizing open-ended test items. However, this proved to be very difficult in the population tested.

The Pharm Fresno was determined to be a reliable instrument based on Cronbach’s alpha and the item-total correlation measures. Inter-rater reliability, as measured by the intra-class coefficient, was high. Further, the researchers worked closely together on scoring a sample of the responses and refining the rubrics. It is reasonable to conclude the agreement between reviewers would be high among the rest of the sample. The Pharm Fresno in the current form was challenging for the participants we included in the validation set, including faculty. This is likely due to several factors. First, Ramos and colleagues validated the Fresno test in Fresno University medical residents and faculty members, and used a self-identified group of experts in evidence-based medicine. The original cohort likely was exposed to some of the basic precepts or application in earlier training and prior to taking the Fresno test. We validated the Pharm Fresno in pharmacy students and a sample of pharmacy practice faculty. As pharmacy practice faculty ourselves, we are most interested in the use of an instrument such as this in pharmacy students. This cohort of pharmacy practice faculty members included some who self-reported intermediate and advanced competence levels. Lower scores in our research may reflect a lower baseline knowledge of EBM of our participants compared to Ramos and colleagues. Second, we placed a time limit on completing the test. It is unclear if Ramos and colleagues limited time for participants to complete the instrument during its validation, although another publication states the original Fresno test was allotted 60 minutes.¹⁴

The time limit may partially explain why no participants earned passing scores on Item 12. It was the last item on our timed exam. This new item requested an answer to the clinical question, a concise summary of the clinical trial and included four domains: population, interventions/comparator, outcomes/results, and application. Only 60 minutes were allotted for the test in the present study. If the original Fresno was allotted 60 minutes and we allotted the same amount of time for a longer test, it is unsurprising that no participants were able pass this item. Fifty-three percent of participants earned a passing score on the single application domain, but this alone was not enough to pass Item 12 based on the criteria we set. Alternatively, the difficulty of Item 12 may indicate weakness in the application step of EBM in the participants we studied, and some participants may benefit from further training. It may be worth separating the two constructs of summary and application in Item 12 on future versions of this instrument, increasing the time allotment, or removing the trial summary.

There was no statistically significant difference between faculty’s and students’ scores on items 1, 5, 7, and 12. Item 1 was one of the easiest items, and Items 5 and 12 were among the most challenging, so these questions were not able to discern between faculty and students. Item 7 was only of moderate difficulty, so it is unclear why it was unable to distinguish between faculty and students. This item was related to the concept of statistical significance. Participants were able to identify statistically significant differences, but were unable to provide two examples of information they used to determine this. A *p* value of <.05 was one acceptable response, but participants rarely discussed the use of confidence intervals.

This research was subject to several limitations. First, we had chosen to adapt an existing tool rather than create a new tool suited for pharmacy students. We had previously considered this to be a strength. A systematic review of the literature identified it as being one of the many instruments available that assessed all four steps of the EBM process, and one that had psychometric properties to support its validation.¹¹ Further, the Fresno test has been used for pharmacy students but never validated.²⁰ However, conducting this research has revealed several drawbacks of the original instrument. To be considered “excellent,” responses must be several statements long. These open-ended questions take much longer to score, which makes the tool better suited for small groups. Items designed with short answer responses might be easier to score and better suited for medium to large groups. We had maintained a similar scoring structure as the original instrument. However, the difference between a “strong” response and a “limited” response is quite large. A competency level between these two, such as “average” may be warranted. Another drawback of the original tool

is the competency level of “not evident.” This means that a participant who did not respond at all and a participant who responded but did not meet criteria for “limited” earned the same score. While this may seem reasonable, it precluded the ability to determine how many participants completed the tests, which is a second limitation of our work. A third limitation of our research is that we did not utilize a group of self-identified experts as did the authors of the original Fresno. The issues with this are discussed above. Fourth, we only replicated the analysis of the original test rather than conducting exploratory and confirmatory factor analysis. Further, we used a sample of convenience rather than determining a target sample size to achieve statistically significant results.

Based on our literature search conducted in March 2017, Gardner and colleagues²⁰ are the only researchers to publish on use of the Fresno to measure skill in EBM in pharmacy students. To assess their elective course, the authors used seven of the original 12 Fresno test questions. They omitted mathematical calculations and diagnosis/prognosis questions, but kept the clinical scenario and scoring rubric the same. Using a pre/post-test study design, they found that mean composite scores on the Fresno increased significantly in both cohorts of their elective offering: pre=60.1 and post=95.9 for cohort 1; pre=93.7 and post=122.0 for cohort 2. They used the results of the Fresno for cohort 1 to make changes to their course and improve student learning during cohort 2. We feel that the validated Pharm Fresno could similarly be used by faculty at academic pharmacy institutions to assess changes in EBM skill and EBM-related teaching activities.

Several adaptations and validations of the original Fresno have been published for use in different subgroups within medicine, including pediatrics,²¹ psychiatry,²² and medical students in Malaysia.²³ The Fresno also has been translated into Spanish,²⁴ and an adapted version has been translated into Brazilian-Portuguese.²⁵ There have been several adaptations of the original Fresno to other health disciplines.²⁶⁻²⁹ We did not identify any adaptations of the Fresno to pharmacy.

McCluskey and Bishop²⁶ adapted the Fresno to occupational therapy (OT). Similar to our methods, they wrote new clinical scenarios, eliminated the diagnosis and prognosis calculations, and revised the rubric to include examples of what constituted an excellent, strong, limited, or not evident response for each item. Two faculty with expert EBM knowledge scored their responses. Their IRR was .91 (95% CI, .69-.98), consistent with our IRR of .95 (95% CI, .85-.98). Their internal consistency was Cronbach’s alpha=.74 (moderate), also consistent with our finding of Cronbach’s alpha of .71. They did not determine discriminative validity as in the current

study, but they did measure responsiveness to change. They found that participants’ scores increased by an average of 20.6 points after an EBM educational intervention. McCluskey and Bishop considered a 10% change on the adapted Fresno educationally important.²⁶

Tilson and colleagues²⁷ validated a modified Fresno for physical therapists. In their version, the calculation-type question point values were decreased, and the negative predictive value calculation was changed to *p* value/alpha interpretation. As in our study, the authors used a convenience sample made up of physical therapy (PT) students (novice to intermediate-level knowledge) and EBM-expert PT faculty. Two independent, blinded reviewers scored the test. Consistent with our results, their IRR was .91 (95% CI, .87-.94). Their internal consistency (Cronbach’s alpha) = .78. Their absolute difference in mean score between novice and expert was 56.2 points (24.2%). However, the weighting of items was different than in our study, so it is difficult to compare the absolute total score of their instrument to ours. Their item discrimination was acceptable (>.2) for all items except item 9. Like us, the author acknowledged that using the rubric required extensive training, and administering and scoring the exam was time-intensive.

Lewis and colleagues²⁸ focused on modifying the Fresno for health discipline students (novice level). They kept the original first seven items, and adapted items 8-12 to de-emphasize statistical calculations and focus more on interpretation of statistical results. They added knowledge of levels of evidence. Like the other adaptations, the authors rewrote scenarios to be applicable to the following health disciplines: physiotherapy, PT, OT, podiatry, and human movement. Because Lewis and colleagues wanted their adaptation to be relatively fast to take and easy to score, they made their revised questions short answer, true/false, and multiple choice. During development, their instrument was reviewed by three content experts and two novice students, with three rounds of review. Their instrument focused on Ask, Acquire, and Appraise, but not Apply. Unlike ours and other adaptations that had a completion time of around 60 minutes, their average completion time was 10 minutes. A total of seven out of 10 items showed statistically significant differences between students who had been exposed to EBM teaching vs those who had not been exposed. Difficulty ranged from 15% to 26% (question 8a/b) to 84%. They reported an overall IRR of .97. The use of short answer, true/false, and multiple choice questions by Lewis and colleagues made their instrument faster to complete and score, compared to other adaptations.

Lizorando and colleagues²⁹ validated the Adapted Fresno²⁶ to speech pathologists, social workers, and

dietitians/nutritionists. They used a panel of four experts of different disciplines to content-validate the instrument. New clinical scenarios were developed. Most of their questions were consistent with the original Fresno, but they revised three questions (1, 2, and 4) and revised the rubric. Four raters scored the tests of respondents. Similar to our scoring method, the raters first scored a sample of tests and discussed any differences, and then independently scored the remaining tests. They found an IRR of .93, .83, and .92 for speech pathology, social work, and dietitian/nutrition, respectively. Their Cronbach's alpha was .71, .68, and .74 for speech pathology, social work, and dietitian/nutrition, respectively. IRR and internal consistency were similar to our results.

Most of the adaptations of the Fresno to other health disciplines have used methods similar to our study. The Pharm Fresno is unique in that it incorporated a clinical trial into the instrument and added a summary of the clinical trial. A common barrier that we and other authors have reported is the time needed to administer and score the instrument when open-ended responses are required. One potential area for future research would be to use methods similar to Lewis and colleagues,²⁸ and make the questions short answer, true/false, and multiple choice. While this would make it difficult to use the same test multiple times, creation of a template that could be revised when new questions are needed is possible.

Future research will focus on refining the instrument to further improve the internal consistency, and validating another clinical scenario and related randomized controlled trial. Validating this instrument in pharmacy residents would be useful. Although at least one study compared self-assessment of skills in EBM to measurement with the Fresno²³ and found a poor correlation, our literature search did not identify any studies which looked at a comparison of two validated, objective measures of skills in applying the EBM process. A validation of the Berlin questionnaire¹³ and the ACE tool¹⁴ in pharmacy students or residents is warranted.

CONCLUSION

The Pharm Fresno is a reliable and valid test to assess the competence of EBM in pharmacy students. We hope this will be a catalyst for the creation and validation of other tools to evaluate competence in EBM in pharmacists.

ACKNOWLEDGMENTS

This research project was supported by a faculty development grant awarded by MCPHS University's faculty development committee. We thank Drs. Aimee Dietle and Amy LaMothe for their assistance scoring the responses, and Dr. Teresa Hedrick for her review and feedback of the

items and the scoring rubrics. Many thanks to Dr. Kathleen Ramos for granting us permission to adapt the Fresno test, and to Drs. Greta von der Luft and Matthew Silva for their assistance with this project.

REFERENCES

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence-based medicine: what it is and what it isn't. *BMJ* 1996;312(7023):71-72.
2. Accreditation Council for Pharmacy Education. Accreditation standards and key elements for the professional program in pharmacy leading to the doctor of pharmacy degree. Standards 2016. <https://www.acpe-accredit.org/pdf/Standards2016FINAL.pdf>. Accessed April 13, 2016.
3. Medina MS, Plaza CM, Stowe CD, et al. Center for the Advancement of Pharmacy Education (CAPE) Educational Outcomes 2013. *Am J Pharm Educ*. 2013;77(8):Article 162.
4. Bennett KJ, Sackett DL, Haynes RB, Neufeld VR, Tugwell P, Roberts R. A controlled trial of teaching critical appraisal of the clinical literature to medical students. *JAMA* 1987;257(18):2451-2454.
5. Schoenfeld P, Cruess D, Peterson W. Effect of an evidence-based medicine seminar on participants' interpretations of clinical trials: a pilot study. *Acad Med*. 2000;75(12):1212-1214.
6. Harewood GC, Hendrick LM. Prospective, controlled assessment of the impact of formal evidenced-based medicine teaching workshop on ability to appraise the medical literature. *Ir J Med Sci*. 2010; 179(1):91-94.
7. Gruppen LD, Rana GK, Arndt TS. A controlled comparison study of the efficacy of training medical students in evidenced-based medicine literature searching skills. *Acad Med*. 2005;80(10):940-944.
8. Ilic D, Tepper K, Misso M. Teaching evidence-based medicine literature searching skills to medical students during the clinical years: a randomized controlled trial. *J Med Lib Assoc*. 2012; 100(3):190-196.
9. Taylor R, Reeves B, Ewings P, Binns S, Keast J, Mears R. A systematic review of the effectiveness of critical appraisal skills training for clinicians. *Med Educ*. 2000;34(2):120-125.
10. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based Medicine: How to Practice and Teach EBM*. 2nd ed. London, UK: Churchill Livingstone; 2000.
11. Shaneyfelt T, Baum KD, Bell D, et al. Instruments for evaluating education in evidence-based practice: a systematic review. *JAMA*. 2006;296(9):1116-1127.
12. Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence-based medicine. *BMJ*. 2003; 326(7384):319-321.
13. Fritsche L, Greenhalgh T, Falck-Ytter Y, et al. 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):1338-1341.
14. Ilic D, Nordin RB, Glasziou P, Tilson JK, Villanueva E. Development and validation of the ACE tool: assessing medical trainee's competency in evidence based medicine. *BMC Med Educ*. 2014;14:114.
15. Fried LF, Emanuele N, Zhang JH, et al. Combined angiotensin inhibition for the treatment of diabetic nephropathy. *N Engl J Med*. 2013;369(20):1892-1903.
16. Everitt, BS. *The Cambridge Dictionary of Statistics*. Cambridge, UK: Cambridge University Press; 2002.

17. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53-55.
18. Peeters MJ, Beltyukova SA, Martin BA. Educational testing and validity of conclusions in the scholarship of teaching and learning. *Am J Pharm Educ.* 2013;77(9):Article 186.
19. Kelley T, Ebel R, Linacre JM. Item Discrimination Indices. *Rasch Measure Trans.* 2002;16(3):883-884. <http://www.rasch.org/mmt/mmt163a.htm>.
20. Gardner A, Lahoz MR, Bond I, Levin L. Assessing the effectiveness of an evidence-based practice pharmacology course using the Fresno test. *Am J Pharm Educ.* 2016;80(7):Article 123.
21. Dinkevich E, Markinson A, Ahsan S, Lawrence B. Effect of a brief intervention on evidence-based medicine skills of pediatric residents. *BMC Med Educ.* 2006;6:1.
22. Rothberg B, Feinstein R, Guiton G. Validation of the Colorado psychiatry evidence-based medicine test. *J Grad Med Educ.* 2013; 5(3):412-416.
23. Lai NM, Teng CL. Self-perceived competence correlates poorly with objectively measured competence in evidence based medicine among medical students. *BMJ Med Educ.* 2011;11:25.
24. Argimon-Pallas JM, Flores-Mateo G, Jimenez-Villa J, Pujol-Ribera E. Psychometric properties of a test in evidence-based practice: the Spanish version of the Fresno test. *BMC Med Educ.* 2010;10:45.
25. Silva AM, Costa LCM, Comper ML, Padula RS. Cross-cultural adaptation and reproducibility of the Brazilian-Portuguese version of the modified Fresno test to evaluate the competence in evidence based practice by physical therapists. *Braz J Phys Ther.* 2016; 20(1):26-47.
26. McCluskey A, Bishop B. The adapted Fresno test of competence in evidence-based practice. *J Contin Educ Health Prof.* 2009;29(2): 119-126.
27. Tilson JK. Validation of the modified Fresno test: assessing physical therapists' evidence-based practice knowledge and skills. *BMC Med Educ.* 2010;10:38.
28. Lewis LK, Williams MT, Olds TS. Development and psychometric testing of an instrument to evaluate cognitive skills of evidence based practice in student health professionals. *BMC Med Educ.* 2011;11:77.
29. Lizarondo L, Grimmer K, Kumar S. The adapted Fresno test for speech pathologists, social workers, and dietitians/nutritionists: validation and reliability testing. *J Multidiscip Healthc.* 2014;7:129-135.