Published in final edited form as:

J Manipulative Physiol Ther. 2012 ; 35(9): 692–700. doi:10.1016/j.jmpt.2012.10.011.

Development and Psychometric Evaluation of an Evidence-Based Practice Questionnaire for a Chiropractic Curriculum

Michael Leo, PhD^a, David Peterson, DC^b, Mitchell Haas, DC^c, Ron LeFebvre, DC^d, and Shireesh Bhalerao, DC^e

^aConsulting Statistician, Kaiser Permanente Center for Health Research, Portland, Oregon, USA

^bProfessor of Chiropractic Sciences, University of Western States, Portland, Oregon, USA

^cAssociate Vice President of Research, Center for Outcomes Studies, University of Western States, Portland, Oregon, USA

^dProfessor of Clinical Sciences, University of Western States, Portland, Oregon, USA

^eAssociate Professor of Chiropractic Sciences, University of Western States, Portland, Oregon, USA

Abstract

Objective—The purpose of this study was to describe the questionnaire development process for evaluating important elements of an evidence-based practice (EBP) curriculum and to report on initial reliability and validity testing for the primary component of the questionnaire, an EBP knowledge exam.

Methods—The EBP knowledge test was evaluated with students enrolled in a doctor of chiropractic program. The initial version was tested with a sample of 374 and a revised version with a sample of 196 students. Item performance and reliability were assessed using item difficulty, item discrimination, and internal consistency. An expert panel assessed face and content validity.

Results—The first version of the knowledge exam demonstrated a low internal consistency (KR20=0.55) and a few items had poor item difficulty and discrimination. This resulted in an expansion in the number of items from 20 to 40, as well as a revision of the poorly performing items from the initial version. The KR20 of the second version was 0.68; 32 items had item difficulties of between 0.20 and 0.80 and 26 items had item discrimination values of 0.20 or greater.

Conclusions—A questionnaire for evaluating a revised EBP integrated curriculum was developed and evaluated. Psychometric testing of the EBP knowledge component provided some initial evidence for acceptable reliability and validity.

Keywords

Evidence-Based Practice; Reproducibility of Results; Questionnaires; Chiropractic; Knowledge

INTRODUCTION

Adhering to the principles of evidence-based practice has become increasingly important in all aspects of health care, as a growing body of population-based outcome studies has documented that patients who receive evidence-based therapy have better outcomes than patients who do not.^{1–7} Evidence-based practices now help shape clinical education curricula as well as real-world clinical decisions and the issuing of guidelines and standards by the organizations that determine standards of care^{8–12} such as the Institute of Medicine¹³ and the American Nurses Credentialing Center.¹⁴

Chiropractic educators now also recognize the necessity of incorporating evidence-based practice principals into the clinical education of chiropractic students, interns, and doctors.¹⁵ Clinical decision-making should be informed by the best available evidence. For this to happen, practitioners must have the ability to access the best evidence, evaluate its quality, and use practice experience to determine its applicability to the individual patient.¹²

To address the need for chiropractors to expand their knowledge of the principles of evidence-based practice (EPB), the College of Chiropractic at the University of Western States (UWS) obtained an R25 grant from the National Center for Complementary and Alternative Medicine at the National Institutes of Health in August 2005. The primary goal of the grant was to develop an integrated EBP curriculum for chiropractic medicine and to train faculty to use EBP teaching methods in the classroom and clinics.

To determine whether the new curriculum had an impact on chiropractic student EBP attitudes, self-rated skills and behaviors, and knowledge, an evaluation committee composed of UWS faculty and an external consultant was formed in September 2005 to develop a questionnaire to measure these important outcomes. This paper describes the questionnaire development process for evaluating important domains of the EBP curriculum, with a specific focus on the development and psychometric testing of the knowledge component.

METHODS

Design and Protocol Overview

This was an instrument development study using methods and procedures consistent with best practices for developing psychological measures.^{16–19} Development protocols are presented in Figure 1. The EBP evaluation committee sought to identify qualitative and quantitative measures that would be valid, practical, and useful for EBP-related progress and conducted a search in the fall of 2005 for previously vetted questionnaires/exams from various health professions. The committee then agreed that our instrument should include the following components: a) a block of sociodemographic items b) a self-assessment of attitudes toward EBP behaviors and self-evaluation of his or her EBP skills, c) a self-assessment of the personality trait of “intellectual curiosity,” and d) an assessment of EBP knowledge. The committee, for reasons described in detail below, opted for a hybrid instrument incorporating available and developed components.

Subjects

Participants were enrolled students in the UWS chiropractic doctoral program. We administered questionnaires to all students who were present and willing during a prearranged regular scheduled class time. We assured students that their participation was completely voluntary and that all data collected was strictly confidential. Questionnaires were secured in the University’s Center for Outcomes Studies. The study was approved by the University of Western States, Institutional Review Board (FWA 851).

We administered version 1 of the questionnaire to students from three separate cohorts comprising the entering classes from fall 2005 through winter 2008. The knowledge exam component was completed by 263 students in the 9th and by 111 students in the 11th quarter. The number of 11th quarter students taking the exam is significantly decreased because many of the cohort 3 students had not yet reached 11th quarter. We administered version 2 to the fourth and fifth cohorts matriculating fall 2008 to winter 2010 that included 196 students in the 6th quarter. We present the baseline characteristics for all students who participated at initial data collection (N=606) in Table 1. Note that not all participants who were available at baseline (1st quarter) were present for the administration of the knowledge exam and skill self-assessment in future quarters.

Questionnaire Development – Version 1

We surveyed new students in the first week of classes to collect sociodemographic information, personality traits, and baseline attitudes toward EBP. We excluded the EBP knowledge exam and skills self-assessment because the majority of students entering chiropractic education have no formal training in EBP. We believed that scores would likely have been extremely low, resulting in little variability and thus non-informative item characteristic statistics. Study staff collected all questionnaires that were handed out during the class period. All data from quantitative assessment instruments were collected on paper and pencil forms. A research assistant verified all of the data using double entry.

Attitudes—The aim of this section of the questionnaire was to determine whether various students' attitudes toward EBP changed as a result of the revised EBP curriculum. This section consisted of 9 items that were measured on a 7-point, Likert-type scale. We constructed the items to measure attitudes that were likely to be present in clinicians who actively engage in EBP activities. For example, items focused on attitudes regarding the comparative weight of research evidence versus expert and clinical opinion, whether all types of evidence are equally important in making clinical decisions, the need to access and stay abreast of the most current information, and the need to critically review research literature.

Self-Rated Skills and Behavior—This section was devoted to self-appraisal of the respondent's ability to apply EBP knowledge. It consists of four questions asking the respondents to assess their understanding of basic biostatistical concepts and ability to find, critically appraise, and integrate clinical research into clinical practice (adapted from similar materials used by Northwestern Chiropractic College). The behavior self-appraisal includes three items asking respondents to evaluate the time spent reading original research, accessing PubMed, and applying EBP methods to patient care. The items were rated on a 7-point, Likert-type scale.

Personality—Intellectual curiosity was based on the construct of Openness to Experience from the Five Factor Model of personality traits. Openness was measured using the corresponding 10-item subscale from the Big Five Inventory (BFI).²⁰ For this questionnaire participants chose how much they agreed or disagreed with statements such as "I see myself as someone who is curious about many different things" on a five point, Likert-type scale. Previous research has presented adequate evidence for the validity and reliability of the BFI V44 questionnaire.^{21,22}

Knowledge—Our 2005 literature search for instruments measuring EBP knowledge found only measures that used presentation of a clinical scenario with most requiring a short-answer or essay response.^{23,24} For example, the Fresno Test²³ is an essay-based knowledge test that measures the ability to frame an answerable clinical question; choose appropriate

key words, databases, and search engines for a literature search; and evaluate the quality and applicability of research studies.

Given that a case-based approach can be impractical and unwieldy when large sample sizes need to be examined at multiple times points, we decided that a multiple choice format was the most feasible choice for objectively assessing EBP knowledge. One popular multiple choice option was the Berlin questionnaire,^{25,26} which consists of scenario-based application questions to assess EBP knowledge among postgraduate medical physicians. We also identified multiple choice knowledge tests for evaluating an EBP curriculum in a family practice residency,²⁷ medical students in a journal club,²⁸ and medical students during their medical rotation.²⁹ Unfortunately, the content of these exams were too context-specific for our needs and did not provide balanced coverage of the topics in our curriculum.

Therefore, we decided to develop our own multiple-choice measure of EBP knowledge. We tailored the exam to the topic areas presented in EBP reference textbooks,^{30–32} the university's EBP Standards document,³³ and topics covered in EBP faculty training workshops. An initial pool of 110 questions covering 10 domains was generated in a multiple choice format with four response alternatives (Table 2).

To achieve face validity, the EBP evaluation committee reviewed each item to ensure that the item was clearly worded and unambiguous. For content validity, the committee evaluated the relevance of each item and reduced the pool of questions in each of the 10 topic areas to 5 items each based on the degree of relevance. After numerous meetings and revisions, the committee pilot tested a 20-item draft exam for face and content validity and comprehensibility by a select group of upper quarter students, other grant personnel, and faculty and residents from other colleges. Based on the feedback from the pilot testing, the committee edited the 20 items to further improve clarity.

Data Analysis

We assessed each item of the knowledge instrument with two item-level psychometric indices, item difficulty, and item discrimination.^{16,34,35} We performed all analyses using Stata 11.2.

Item Difficulty—The difficulty index value of an item is measured by the proportion of respondents who answer the question correctly. Items that are answered correctly by more than 80% of the respondents (value >0.80) are considered “too easy” and items answered correctly by less than 20% of the respondents (value <0.20) are considered “too difficult.” Items with difficulty levels close to 0.50 maximize score variability,³⁵ which is highly desirable for distinguishing individuals of varying levels of ability. However, for criterion-referenced tests (evaluating absolute level of ability), higher difficulty index values (i.e., greater proportion correct) are desirable for individuals who are expected to learn the material.³⁶

Item Discrimination—Item discrimination refers to the extent to which an item distinguishes between those who score high and those who score low. It is computed by taking the difference between the item difficulty for those who score in the upper 33rd percentile and for those who score in the lower 33rd percentile. We considered item discrimination values of 0.35 and higher as excellent discriminators, values between 0.20 and 0.34 as acceptable, and values less than 0.20 as unsatisfactory.³⁵

Kuder-Richardson 20 (KR20)—This is a measure of the internal consistency across all of the scale items. It is a derivative of Cronbach's alpha for use in dichotomously scored

measures. We considered KR20 values of 0.70 and greater to indicate acceptable internal consistency.^{16,17}

Questionnaire Development - Version 2

We decided to revise the knowledge exam in September of 2009 based on the results of the item analysis of the version 1 and an examination of the test by new committee members and content experts. We dropped two items because of poor difficulty/discrimination ratings, overemphasis of a low curricular content area, and our judgment that an item was more value-based than factual. We retained the remaining 18 items, of which 15 were revised to improve clarity and/or make the item more clinically relevant.

We also decided to expand the number of items, tapping each of the EBP domains to ensure that we included a sufficient number of items to adequately measure the domain of interest and improve internal consistency, and domain coverage. We pilot-tested the new 45-item version of the exam on 6 individuals in December 2009 (2 faculty and 4 students). Both faculty members had received variable levels of EBP training and the students had varying levels of clinical ability. We asked participants to record how long it took to answer all of the items and write down any comments regarding test content clarity. Based on their feedback, we deleted 5 items and edited the remaining items for improved comprehensibility. Thus, version 2 of the knowledge exam consisted of 40 items.

We discussed the questions with high levels of difficulty or poor discrimination values. We determined that the response to four questions was likely a result of poor wording and/or incorrect information being presented to the students. Two of the questions were reworded. We also notified instructors and library personnel about content inaccuracies and corrected the relevant online learning modules for one question. Another question that had a high degree of difficulty represents an area that receives minimal coverage in our EBP curriculum and has been dropped and replaced with another question intending to measure the same domain of diagnosis.

RESULTS

Item difficulty and discrimination values are found in Table 3 for version 1 and in Table 4 for version 2 of the knowledge exam. Questions are presented by content domain only because of concern for the security of the content of specific items on the test. This information and a copy of the full knowledge exam are available upon request.

Version 1

Item difficulties ranged from 0.19 (hardest) to 0.76 (easiest). Item discrimination ranged from a low of 0.04 to a high of 0.49 (Table 3). Note that we also conducted the analyses for both follow-ups separately and the results were very similar to the combined results. Across all of the knowledge items, the KR20 was 0.55.

Version 2

Item difficulty ranged from 0.04 to 0.96 (Table 4). Thirty-two of the 40 items were in the acceptable range of 0.20 to 0.80. The most difficult items were in the content domains of "Asking answerable questions, finding evidence" (item difficulty=.04), "Critical evaluation of diagnostic studies" (item difficulty=.17), and "Critical evaluation of therapy articles" (item difficulty=.18). The easiest items were in the content domains of "EBP overview and clinical application" (item difficulty=.96), "Asking answerable questions, finding evidence" (item difficulty=.93), "EBP overview and clinical application" (item difficulty=.86),

“Asking answerable questions, finding evidence” (item difficulty=.84), and “Overview of clinical biostatistics (item difficulty=.81).

Item discrimination values ranged from a low of -0.10 to a high of 0.53. Fourteen items had discrimination values of less than 0.20. One item had a negative discrimination value, indicating that students in the lower 33rd percentile of knowledge scores were more likely to correctly answer this item than students in the upper 33rd percentile. The KR20 improved to 0.68.

DISCUSSION

In any research or program evaluation endeavor, it is important to ensure that the outcomes of interest are clearly defined and that the outcomes are evaluated using reliable and valid measures. The purpose of this paper was to describe the development of a questionnaire for measuring EBP-related attitudes, self-rated skills and behaviors, and knowledge. We specifically emphasized the knowledge component, given that it is the primary measure of curricular change.

The content validity and item analysis results for the knowledge exam provide a solid foundation of evidence for the reliability and validity for this component. Psychometric evaluation of the knowledge exam and the ensuing internal discussion lead to the development of a more comprehensive (items with a wide range of item difficulty values covering the content domain) and sensitive (majority of items with acceptable item discrimination values) version 2 of the exam. Though the current version of the exam had less than the stated acceptable level of .70 for internal consistency, some psychometricians (e.g., Devellis¹⁶) deem values between .65 and .70 as “minimally acceptable.”

To our knowledge, this section of the instrument is the first published measure of specific EBP knowledge tailored to a chiropractic setting. These questions will help determine whether our revised curriculum is related to improving EBP knowledge compared to the old curriculum. We believe that other initiatives can use this measure “off the shelf” with minimal modification, and thus not have to begin the development process from scratch. A noteworthy feature of our knowledge exam is that its domain extends beyond statistical knowledge to cover other important EBP dimensions, such as the ability to conduct a search of the literature and knowledge of research design.

A unique feature of the student assessment process was that the knowledge assessment tool was not tied to any specific EBP course. These assessments were additions to the traditional midterm or final examinations. As a result, some of the content tested was presented months and, in some cases, 2 years, before the assessment. The timing of our exam differed from the post-course assessments that are usually reported in the literature. Consequently, some of the results may be reflective of more long-term retention. On the other hand, any potentially lower scores resulting from this lag time may have been counterbalanced by additional EBP exposure in other traditional courses offered subsequently to the core EBP courses.

Since the initial development of our EBP questionnaire, there has been a comprehensive review by Shaneyfelt et al.³⁷ and Tilson et al.³⁸ of the available questionnaires and exams designed to assess EBP attitudes, knowledge, skills and behavior. We considered most of the knowledge-based exams discussed in that review, and the ones that were accessible at the time of our 2005 search did not have the content and structure we required. In fact, other disciplines have also had to modify and develop their own EBP measures in order to have a more applicable measure to the content of their programs (e.g., dentistry³⁹, occupational⁴⁰ and physical therapy⁴¹). Another instrument, the Knowledge of Research Evidence Competencies (K-REC),⁴² was designed to be applicable across multiple health disciplines.

However, this instrument assesses knowledge expected at the undergraduate level, whereas our instrument was designed to cover a wider range of knowledge (i.e., from novice to expert).

For assessing EBP knowledge, we believe our measure is a comprehensive (40 items covering a wide gamut of EBP concepts), non-burdensome (20 to 30-minute completion time), and pragmatic tool with sufficient level of precision. The majority of identified existing tools was either too complex or onerous to administer to a large number of students or too focused to satisfactorily assess the full range of knowledge of an EBP-integrated curriculum. Although our knowledge measure was developed with our specific curriculum goals in mind, we believe that many, if not all, of the items can be easily modified for use in other contexts.

Lessons Learned

In the process of evaluating EBP curricular changes, we discovered an inherent conflict between providing summative feedback about student knowledge and providing formative feedback to instructors and administrators to aid in program development and refinement. The revision process highlighted this dynamic tension. On the one hand, we wanted a stable instrument to measure intern knowledge, compare cohorts, and to function as one data stream measuring the program's success. On the other hand, we wanted a flexible instrument that would evolve as our learning objectives evolved and matured with time and experience. Modifications were deemed necessary to improve performance of the instrument and, in part, to reflect changes that were occurring relative to the depth and focus of course content. This required revising questions, but at the same time improving the reliability and validity of the test. The goal was to maintain an exam in which the questions were distributed by domain, and to avoid making changes that were so substantial that we would no longer be able to compare cohorts.

The process of revising test questions presented an additional challenge. It was important to provide feedback to core instructors and to solicit input from them concerning course content without creating a situation where they taught to the exact questions on the exam. The original exam was written by the evaluation committee in which only one member had actually taught in the core EBP series. The lag time between test design and application was judged to be long enough to mitigate the likelihood that this member of the faculty would "teach to the test" rather than to the learning objectives. However, because one aim of the knowledge test was to provide formative feedback to the EBP curricular designers, the results were circulated throughout that committee. It was the EBP Curriculum Committee that detected the gaps between certain test questions and the revised learning objectives. In some cases the content was no longer even being taught. The EBP Curriculum Committee has become instrumental in revising test questions. However, the Committee needs to be mindful of potential conflict of interest because it will evolve to include more members who are directly involved in teaching.

Limitations

The process of devising and testing our exam was subject to uncertainties inherent in this type of research. With regard to measures of objective knowledge, there is a potential threat to validity regarding testing effects.⁴³ That is, repeated administrations of the same measure may result in an inflation of scores, because either priming (exposure at the pre-test cues the material to focus on for post-test) or from cheating. These effects were minimized because students were not given feedback on their performance, were not allowed to retain copies of the survey, had little motivation to cheat as performance was not linked to grades, and were limited to widely spaced testing occasions. However, without significant stakes tied to test

performance, students may not have been highly motivated to perform well on the EBP knowledge exam (this was occasionally reflected in written comments on the test). These circumstances pertained equally to the two administrations of the exam and so the issue of motivation should not be a significant confounder when comparing the results for the same cohort at two points in time. Nevertheless, the absolute overall scores are lower than we desired.

Another potential limitation may be selection bias with regard to participation, limiting our ability to generalize the item characteristic results. A disparity in number of students completing the questionnaire in the 9th term (263) and 11th term (111) could have introduced bias to outcomes. Potential bias could have been introduced if those not completing the 11th quarter questionnaire had characteristics that influenced their responses. For example, those who chose to participate may be more motivated or more confident in their EBP knowledge than those who did not choose to participate, artificially inflating the scores. However, it is important to note that the disparity between sample sizes can be partially accounted for by a considerable percentage of students that took time off between the 9th and 11th quarter. The institution has also implemented a policy that should increase 11th quarter response rates by requiring students to complete the survey to pass their 11th quarter EBP journal club class.

Our sample size was not large enough to engage in more sophisticated psychometric analyses. Measures of objective knowledge are well suited to techniques that are associated with item response theory, which go beyond simple metrics such as classical item difficulty, item discrimination, and internal consistency. However, with the addition of future cohorts to our database, we expect to have sufficient sample size to further examine the properties of this questionnaire. Finally, we must acknowledge that while the multiple choice format may be useful to measure the knowledge base that comprises EBP; the composite skill of finding, assessing, and applying research evidence is not directly measured by this method.

CONCLUSIONS

This research resulted in the development of a questionnaire that measures important outcomes related to a revised EBP integrated curriculum. The primary outcome of interest was EBP knowledge, and the results of psychometric testing of this component provide some initial evidence for its reliability and validity. Therefore, it may be a useful measure for other studies, in the classroom, or as a basis for personalization in other healthcare educational contexts.

Acknowledgments

This study was supported by the National Center for Complementary and Alternative Medicine, Department of Health and Human Services (grant no. R25 AT002880). The study was approved by the University of Western States Institutional Review Board (FWA 851).

REFERENCE LIST

1. Krumholz HM, Radford MJ, Ellerbeck EF, et al. Aspirin for secondary prevention after acute myocardial infarction in the elderly: prescribed use and outcomes. *Annals of Internal Medicine*. 1996 Feb 1; 124(3):292–8. [PubMed: 8554223]
2. Krumholz HM, Radford MJ, Wang Y, et al. National use and effectiveness of beta-blockers for the treatment of elderly patients after acute myocardial infarction: National Cooperative Cardiovascular Project. *JAMA*. 1998 Aug 19; 280(7):623–9. [PubMed: 9718054]
3. Mitchell JB, Ballard DJ, Whisnant JP, et al. What role do neurologists play in determining the costs and outcomes of stroke patients? *Stroke*. 1996 Nov; 27(11):1937–43. [PubMed: 8898795]

4. Wong JH, Findlay JM, Suarez-Almazor ME. Regional performance of carotid endarterectomy. Appropriateness, outcomes, and risk factors for complications. *Stroke*. 1997 May; 28(5):891–8. [PubMed: 9158621]
5. Sackett, D.; Straus, SE.; Richardson, W.; Rosenberg, WM.; Haynes, RB. Evidence-based medicine: How to practice and teach EBM. Edinburgh: 2000.
6. Swan BA, Boruch RF. Quality of evidence: usefulness in measuring the quality of health care. *Medical Care*. 2004 Feb.42(2:Suppl):Suppl-20.
7. Joint Commission on Accreditation of Healthcare Organizations. Patient Safety Essentials for Health Care. Joint Commission on Accreditation of Healthcare Organizations; 2006.
8. Hatala R, Guyatt G. Evaluating the teaching of evidence-based medicine. *JAMA*. 2002 Sep 4; 288(9):1110–2. [PubMed: 12204080]
9. Straus SE, Green ML, Bell DS, et al. Evaluating the teaching of evidence based medicine: conceptual framework. *BMJ*. 2004 Oct 30; 329(7473):1029–32. [PubMed: 15514352]
10. Davidson RA, Duerson M, Romrell L, et al. Evaluating evidence-based medicine skills during a performance-based examination. *Academic Medicine*. 2004 Mar; 79(3):272–5. [PubMed: 14985203]
11. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992 Nov 4; 268(17):2420–5. [PubMed: 1404801]
12. Sackett DL, Rosenberg WM. The need for evidence-based medicine. *Journal of the Royal Society of Medicine*. 1995 Nov; 88(11):620–4. [PubMed: 8544145]
13. Greiner, A.; Knebel, E. Health professions education: a bridge to quality. National Academies Press; 2003.
14. McClure, ML.; Hinshaw, AS. Magnet hospitals revisited: attraction and retention of professional nurses. American Nurses Association; 2002.
15. Delaney PM, Fernandez CE. Toward an evidence-based model for chiropractic education and practice. *Journal of Manipulative & Physiological Therapeutics*. 1999 Feb; 22(2):114–8. [PubMed: 10073628]
16. DeVellis, RF. Scale development: theory and applications. Sage Publications, Inc; 2003.
17. Nunnally, JC.; Bernstein, IH. Psychometric theory. McGraw-Hill; 1994.
18. Crocker, LM.; Algina, J. Introduction to classical and modern test theory. Holt, Rinehart, and Winston; 1986.
19. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing. Standards for educational and psychological testing. American Educational Research Association; 1999.
20. John, O.; Donohue, E.; Kentle, R. The Big Five Inventory—Versions 4a and 54 (Tech Rep). Berkeley, CA: Institute of Personality and Social Research, University of California; 1991.
21. Benet-Martinez V, John OP. Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*. 1998; 75(3):729–50.10.1037/0022-3514.75.3.729 [PubMed: 9781409]
22. John, O.; Srivastava, S. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In: Pervin, L.; John, O., editors. *Handbook of Personality: Theory and Research*. Guilford Press; 1999. p. 102-38.
23. Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. *BMJ*. 2003 Feb 8; 326(7384):319–21. [PubMed: 12574047]
24. Green ML, Ellis PJ. Impact of an evidence-based medicine curriculum based on adult learning theory. *Journal of General Internal Medicine*. 1997 Dec; 12(12):742–50. [PubMed: 9436893]
25. Akl EA, Izuchukwu IS, El-Dika S, et al. Integrating an evidence-based medicine rotation into an internal medicine residency program. *Academic Medicine*. 2004 Sep; 79(9):897–904. [PubMed: 15326018]
26. 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 Dec 7; 325(7376):1338–41. [PubMed: 12468485]

27. Ross R, Verdick A. Introducing an evidence-based medicine curriculum into a family practice residency--is it effective? *Academic Medicine*. 2003 Apr; 78(4):412–7. [PubMed: 12691976]
28. Linzer M, Brown JT, Frazier LM, et al. Impact of a medical journal club on house-staff reading habits, knowledge, and critical appraisal skills. A randomized control trial. *JAMA*. 1988 Nov 4; 260(17):2537–41. [PubMed: 3050179]
29. Landry FJ, Pangaro L, Kroenke K, et al. A controlled trial of a seminar to improve medical student attitudes toward, knowledge about, and use of the medical literature. *Journal of General Internal Medicine*. 1994 Aug; 9(8):436–9. [PubMed: 7965237]
30. Guyatt, G.; Deditors, Rennie. Users' guides to the medical literature: a manual for evidence-based clinical practice. Chicago: American Medical Association; 2002.
31. Straus, SE. Evidence-based medicine: how to practice and teach EBM. Elsevier/Churchill Livingstone; 2005.
32. Dawes M, Summerskill W, Glasziou P, et al. Sicily statement on evidence-based practice. *BMC Medical Education*. 2005 Jan 5.5(1):1. [PubMed: 15634359]
33. Lefebvre RP, Peterson DH, Haas M, et al. Training the evidence-based practitioner: university of Western States document on standards and competencies. *Journal of Chiropractic Education*. 2011; 25(1):30–7. [PubMed: 21677870]
34. Haladyna, TM. Developing and validating multiple-choice test items. L. Erlbaum Associates; 1999.
35. Thorndike, RM. Measurement and evaluation in psychology and education. Pearson Merrill Prentice Hall; 2005.
36. McCowan, RJ.; McCowan, SC. Item Analysis for Criterion-Referenced Tests. Center for Development of Human Services; State University of New York (SUNY), Research Foundation; 1999.
37. Shaneyfelt T, Baum KD, Bell D, et al. Instruments for evaluating education in evidence-based practice: a systematic review. *JAMA*. 2006 Sep 6; 296(9):1116–27. [PubMed: 16954491]
38. Tilson JK, Kaplan SL, Harris JL, et al. Sicily statement on classification and development of evidence-based practice learning assessment tools. *BMC Medical Education*. 2011; 11:78. [PubMed: 21970731]
39. Hendricson WD, Rugh JD, Hatch JP, et al. Validation of an instrument to assess evidence-based practice knowledge, attitudes, access, and confidence in the dental environment. *Journal of Dental Education*. 2011 Feb; 75(2):131–44. [PubMed: 21293036]
40. McCluskey A, Bishop B. The Adapted Fresno Test of competence in evidence-based practice. *Journal of Continuing Education in the Health Professions*. 2009; 29(2):119–26. [PubMed: 19530195]
41. Tilson JK. Validation of the modified Fresno test: assessing physical therapists' evidence based practice knowledge and skills. *BMC Medical Education*. 2010; 10:38. [PubMed: 20500871]
42. 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 Medical Education*. 2011; 11:77. [PubMed: 21967728]
43. Shadish, WR.; Cook, TD.; Campbell, DT. Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin; 2002.

Key Learning Points

- Description of the questionnaire development process for evaluating an evidence-based practice (EBP) curriculum.
- The initial reliability and validity evidence for the knowledge component of the questionnaire is promising.
- Though further data regarding reliability and validity for the knowledge component is required, the measure can be used by other programs for evaluating EBP knowledge.

Highlights

This research resulted in the development of a questionnaire to measure the effectiveness of a revised EBP integrated curriculum. The primary outcome was EBP knowledge, and the results of psychometric testing provide some initial evidence for its reliability and validity.

Fig 1.
Overview of study protocol

Table 1

Baseline characteristics of first-year students

	Cohorts 1 to 3 (2005 – 2008) (n = 339)	Cohort 4 to 5 (2008 – 2010) (n = 267)
Age	27.4 (5.0)	26.3 (4.7)
Female (%)	36	34
White non-Hispanic (%)	83	81
Bachelor's degree or higher (%)	71	74
# prior research methods courses	2.5 (2.7)	2.8 (4.1)
# prior probability & statistics courses	1.7 (1.3)	1.6 (1.0)
Prior college within 2 years (%)	93	96
Prior job in health care (%)	36	41

Mean (SD) or %.

Table 2

Content domains represented on the EBP knowledge exam

Domains	
1	EBP overview and clinical application
2	Asking answerable questions, finding evidence
3	Overview of clinical biostatistics
4	Research study design and validity overview
5	Critical evaluation of therapy articles
6	Critical evaluation of diagnostic studies
7	Critical evaluation of preventive studies
8	Critical evaluation of harm studies
9	Critical evaluation of prognosis studies
10	Critical evaluation of systematic reviews/guidelines

Table 3

Version 1 knowledge exam item difficulty and discrimination

Item #	Content domain	Item Difficulty	Item Discrimination
1	EBP overview and clinical application *	0.27	0.10
2	Asking answerable questions, finding evidence	0.72	0.34
3	Critical evaluation of therapy articles	0.24	0.08
4	Research study design and validity overview	0.76	0.32
5	Asking answerable questions, finding evidence	0.74	0.41
6	Research study design and validity overview	0.28	0.26
7	Asking answerable questions, finding evidence	0.41	0.40
8	Critical evaluation of diagnostic studies	0.69	0.38
9	Critical evaluation of diagnostic studies	0.19	0.04
10	Overview of clinical biostatistics	0.53	0.32
11	Critical evaluation of diagnostic studies	0.25	0.17
12	Overview of clinical biostatistics	0.60	0.38
13	Research study design and validity overview	0.39	0.34
14	Critical evaluation of systematic reviews/guidelines	0.48	0.36
15	Critical evaluation of systematic reviews/guidelines	0.75	0.33
16	Critical evaluation of diagnostic studies	0.34	0.31
17	Critical evaluation of diagnostic studies	0.50	0.42
18	EBP overview and clinical application	0.73	0.45
19	Critical evaluation of therapy articles	0.24	0.32
20	Introduce practice guidelines into your clinic *	0.76	0.49

* items dropped in next revision

Table 4

Version 2 knowledge exam item difficulty and discrimination

Item #	Content domain	Item Difficulty	Item Discrimination
1.	Critical evaluation of harm studies	0.50	0.25
2.	Critical evaluation of systematic reviews/guidelines	0.36	0.16
3.	Overview of clinical biostatistics	0.81	0.30
4.	Critical evaluation of therapy articles	0.80	0.35
5.	Critical evaluation of systematic reviews/guidelines	0.58	0.46
6.	Critical evaluation of systematic reviews/guidelines	0.79	0.07
7.	Critical evaluation of diagnostic studies (Rev of Q17)	0.43	0.18
8.	Overview of clinical biostatistics	0.41	0.24
9.	Critical evaluation of prognosis studies	0.66	0.31
10.	Critical evaluation of systematic reviews/guidelines (Rev of Q14)	0.50	0.29
11.	Critical evaluation of harm studies	0.29	-0.10
12.	Critical evaluation of therapy articles	0.54	0.37
13.	Critical evaluation of diagnostic studies (Rev of Q16)	0.24	0.36
14.	Critical evaluation of diagnostic studies	0.64	0.19
15.	Critical evaluation of therapy articles (Rev of Q19)	0.51	0.52
16.	Critical evaluation of therapy articles	0.56	0.53
17.	Overview of clinical biostatistics	0.43	0.51
18.	Overview of clinical biostatistics	0.58	0.22
19.	Critical evaluation of preventive studies	0.66	0.10
20.	Critical evaluation of preventive studies	0.38	0.41
21.	Research study design and validity overview	0.70	0.24
22.	Critical evaluation of diagnostic studies (v1 Q8)	0.54	0.23
23.	Critical evaluation of diagnostic studies (Rev of Q9)	0.26	0.19
24.	Overview of clinical biostatistics (Rev of Q10)	0.61	0.32
25.	Critical evaluation of therapy articles (Rev of Q3)	0.18	0.07
26.	Research study design and validity overview (Rev of Q6)	0.41	0.23
27.	Research study design and validity overview	0.66	0.44
28.	Critical evaluation of diagnostic studies	0.54	0.40
29.	Research study design and validity overview (v1 Q4)	0.79	0.27
30.	Critical evaluation of systematic reviews/guidelines (Rev of Q15)	0.54	0.38
31.	EBP overview and clinical application	0.96	0.11
32.	Asking answerable questions, finding evidence (Rev of Q5)	0.84	0.28
33.	EBP overview and clinical application (Rev of Q18)	0.86	0.21
34.	Critical evaluation of diagnostic studies (Rev of Q11)	0.17	0.24
35.	Asking answerable questions, finding evidence (v1 Q2)	0.93	0.12
36.	Research study design and validity overview (Rev of Q13)	0.58	0.12
37.	Asking answerable questions, finding evidence (Rev of Q7)	0.63	0.38
38.	Asking answerable questions, finding evidence	0.04	0.06
39.	Asking answerable questions, finding evidence	0.32	0.18

Item #	Content domain	Item Difficulty	Item Discrimination
40.	Overview of clinical biostatistics (Rev of Q12)	0.37	0.18

Note: Parentheticals indicates the item number from an unchanged item of Version 1 (v1) or revised item from Version 1 (Rev of). Items without parentheticals are new items not present in the Version 1.