IJAHSP
The Internet Journal of Allied Health Sciences and Practice
http://ijahsp.nova.edu

A Peer Reviewed Publication of the College of Health Care Sciences at Nova Southeastern University
Dedicated to allied health professional practice and education
http://ijahsp.nova.edu Vol. 10 No. 2 ISSN 1540-580X

Question the Experts: An Adaptable Model to Increase Understanding of Research Methodology and Improve Evaluation Skills

Marianne Hollis, PhD, RN[1]
Ann Hallyburton, MSLS, AHIP[2]
Michael Hubble, PhD[3]

1. Associate Professor, School of Health Sciences, Western Carolina University, Cullowhee, North Carolina
2. Research and Instruction Librarian, Liaison to the Health Professions, Hunter Library, Western Carolina University, Cullowhee, North Carolina
3. Professor, School of Health Sciences, Western Carolina University, Cullowhee, North Carolina

United States

CITATION: Hollis M, Hallyburton A, Hubble M. Question the Experts: An Adaptable Model to Increase Understanding of Research Methodology and Improve Evaluation Skills. *The Internet Journal of Allied Health Sciences and Practice.* April 2012. Volume 10 Number 2.

ABSTRACT

Students new to research methods may fail to acknowledge that research findings are open to interpretation. They may see the sole purpose of research as proving or disproving a hypothesis. **Purpose:** The study aims to determine whether an exercise asking research methods students to compare studies using preset evaluation criteria helps those students to better grasp the intricacies of research. The goal of the exercise is to better inform future and current practitioners' professional practice and research. **Methods:** Students in a multidisciplinary health sciences graduate program were asked to compare the methods and findings of two large-scale research studies dealing with a particular topic, the effect of aspirin on the cardiovascular health of women. **Results:** Student participants identified basic components of clinical research: population, intervention, sampling, and screening criteria. These components proved similar across the arguably landmark Nurses' Health Study and Women's Health Study. Students had to move past acceptance of what researchers from prestigious studies asserted and deal with dichotomies arising from their disparate conclusions. **Conclusions:** By examining and comparing methods used in research studies with differing conclusions, future and current health care providers can better appreciate the dichotomy of conflicting findings from similar studies, the utility of applying different methodologies when studying a single issue, and the importance of evidence-based practice appropriate to their discipline.

INTRODUCTION

Advanced students in the health professions may lack skills to critically evaluate the evidence that informs their practice.[1-3] While health care students and professionals receive the exhortation to keep up with "the literature" and apply evidence-based practice principles, even clinical educators have great difficulties in grasping these concepts.[4] Research methods courses offer instruction on study design, research ethics, information-finding practices, proposal writing, and encourage the beginnings of original research. However, this schedule leaves little time for individuals to process the implications of what they learn and apply their new knowledge to the literature of their profession.

The topic of how health care consumers find and evaluate health information has received extensive coverage; however, practitioners' health information literacy, defined here as how health professionals find, evaluate, and use health information, has received little attention. The practice goes beyond whether research is published in a peer reviewed journal, a factor not

automatically conferring quality.[5,6] An exercise meant to evaluate practitioners' health information literacy was tested in one allied health-focused graduate program's research methods and biostatistics courses. The method used involved a repeated exercise in which students, at the end of specific courses, used a research evaluation tool to assess similar, large scale studies with divergent findings. This exercise is easily adaptable in research courses and other settings where increased awareness of research quality indicators and limitations is desired.

Difficulties arise in effectively evaluating research literature for a number of reasons. Individuals evaluating research quality use diverse methods, even during the peer review process.[7,8] Evaluators judge a study by its own merits, compare it to a like piece of research, or use established criteria to score the quality of a given piece of research.[9] A journal's reputation, selection standards, and impact within its field (impact factor) may offer good measures of the quality of the information it publishes.[10] In the clinic, however, a journal's impact factor often has little to do with how practitioners evaluate and apply information.[11] In deciding whether to apply new information to their practice, clinicians evaluate research according to relevance, perceived quality, practicality, effectiveness, external regulation, and how well new practices will work within current clinical operations.[12,13]

Many science researchers place a higher premium on easy access to research than on its timeliness or relevance.[14,15] This predilection conflicts with the principle that "best practice health care depends on clinicians understanding and applying the results of trials in clinical practice."[16] Health care providers' grasp of current best practices is crucial as these professionals provide patient care.[17]

In order for practitioners to effectively find, evaluate, and apply research information, they must receive instruction.[18] Research methods courses are common in upper level and continuing education curricula. Students in research methods courses cover a wide spectrum of ages, professional experiences, and lifestyles.[19] Courses encompass experiment design, problem-based learning, clinic-based research, information evaluation, and evidence-based practice principles while delivery methods range from in-person, to videoconference, to online, to a melding of approaches.[20] Independent of delivery mode, many courses focus on teaching students to conduct qualitative or quantitative research with little crossover between the methods.[21] Students often lack a sufficient basis in cross-comparison of the results of these methods.[22] One approach to addressing this disconnect involves focusing on the related concepts of practitioners' health information literacy and evidence-based practice. Lacking firm footing in basic research principles, students (and the health professionals they become) can go astray by following flawed research or biased interpretations. Depending upon area of practice, this misdirection may negatively affect one person or countless individuals.[23]

No matter the expertise, health professionals are unlikely to have the time or ability to evaluate the quality of every piece of published research on a topic. Even when research conduct is beyond reproach, findings on an issue may differ from one study to the next. At this juncture, the principles of evidence-based practice assume vital importance.[24] Basing practice on evidence entails the "conscientious, explicit, and judicious use of current best evidence in making decisions . . . integrating individual clinical expertise with the best available external clinical evidence from systematic research."[25] Basic principles of evidence-based practice apply across health care, education, librarianship, the social sciences, and many other disciplines. While the concept of evidence-based practice does have detractors, the value of a system in which a preponderance of research findings is used to better inform interventions is widely recognized.[26]

A key element in sorting through evidence-based practice information is the application of a hierarchical structure, often depicted as a pyramid, where certain types of evidence wield greater influence than others.[27] Weight given to different evidence types depends upon the discipline to which they apply, with systematic reviews, meta-analyses, and practice guidelines consistently residing at the highest levels.[28,29] Systematic reviews, meta-analyses, and practice guidelines provide critical analyses of evidence from multiple studies, enabling subject area experts conducting those reviews to tease out recommendations on practice. If available evidence lacks sufficient depth or quality, reviewing teams may not offer overall recommendations; this reticence strengthens rather than diminishes review integrity. Systematic reviews from the international Cochrane Collaboration (health care) and Campbell Collaboration (social sciences) carry particular authority due to the meticulous protocols set forth for the conduct of their reviews.

Evidence-based practice goes beyond the classification of publication types, though. This paper explores how health care students and professionals can improve their research assessment skills using structured evaluation criteria and the tenets of evidence-based practice. The use of a research analysis tool may better enable students to weed through and apply the best evidence, and if those students become researchers, foster improved study design and analysis. Foremost, students should begin to recognize the frequently subjective nature of research and go beyond the threshold of whether or not a report has undergone peer review.

METHODS

To test these assertions, the authors adapted study quality questions drawn from a standardized abstraction form created by the Task Force on Community Preventive Services, a group of public health and prevention experts appointed by the Director of the US Centers for Disease Control and Prevention (CDC). The Task Force based their criteria on systematic review methodologies, social and health science journal standards, meta-analytic and statistical publications, expert review, and pilot testing.[30]

The authors then identified research articles detailing the findings of two influential studies having disparate findings on similar issues for comparative evaluation. The first article, by Ridker et al, details the findings of the US National Heart, Lung and Blood Institute (NHLBI)-sponsored Women's Health Study, a ten-year, double blind, randomized placebo-controlled study of approximately 40,000 healthy women aged 45 and older.[31] Findings from this study suggest low-dose aspirin intake does not significantly lower women's incidence of major cardiac events and overall death rates from cardiovascular causes. For clarity, the article is identified as "Ridker-Women's Health Study" for the remainder of the paper. The second article, by Chan et al., reports the results of a nested, prospective, case-control study of nearly 80,000 women with no cardiovascular disease or cancer history.[32] The study group is part of the expansive Nurses' Health Study, a decades-long, multi-faceted study funded by the National Institutes of Health (NIH) and conducted under the auspices of multiple academic and research institutions. Results from this study group indicate a significant association between low-to-moderate dose aspirin use and reduced cardiovascular events and cardiovascular system-related mortality in women. For clarity, this article is listed as "Chan-Nurses' Health Study" for the remainder of the paper.

The authors conducted a pilot test of the evaluation exercise using a convenience sampling of health sciences graduate students taking part in research methods and biostatistics courses. Though this study population is comprised of upper level health sciences students at a US university, its format is applicable in settings outside academia (see "Recommendations for Application"). Two nearly identical surveys were created using Task Force adapted questions. The second survey (covering the Chan-Nurses' Health Study article) differs from the first in that it asks respondents to choose which article they find more credible, and what recommendations they would make to others as a result their evaluation of the article. The Appendix provides a list of these adapted questions, plus demographic queries.

This research received approval through the academic institution's Institutional Review Board (ethics committee), and students had the option of choosing not to have responses used in the study. Students received identification numbers drawn from a random number generation tool; this method enabled limiting of investigator bias while still tracking how many students took the survey in more than one course. The first administration of the survey took place at the end of a summer term research methods seminar. This elective seminar followed the previous spring semester's introductory research methods course and included 13 students. The second administration took place in a fall-term biostatistics course containing 24 students who had received instruction on all course content, with the exception of regression. Respondents provided information on age, gender, program concentration, stage within the program, and whether they had previously taken research methods or biostatistics courses.

The separate biostatistics course, with the purpose of increasing students' grasp of research findings and methods, was a new addition to the graduate curriculum. Comparing the research evaluation skills of students who had and had not been exposed to biostatistics instruction within the academic program provided an additional, institution-specific goal for the assessment. To that end, descriptive and comparative analyses were performed for each group of students. Contingency coefficient analyses were performed on responses to each of the evaluation questions.

RESULTS

Over the course of the two-semester study period, 28 evaluations were completed: 15 during the summer research methods course (eight students provided responses for the Ridker-Women's Health Study article and seven of those same students provided responses for the Chan-Nurses' Health Study article), 13 in the fall biostatics course (seven students responded to the Ridker-Women's Health Study evaluation and six of those same students provided responses to the Chan-Nurses' Health Study evaluation). These evaluations represent 14 students, with one student completing evaluations in both courses. Course members fell in different stages of the program. Most students had progressed from one-third to halfway through the program at the time of the research. Only three students were in their first semester. Respondents provided information on age, gender, and graduate program concentration (see Table 1: Summary of Participants by Course).

Table 1. Summary of Participants by Course

	Research Methods	Biostatistics
Age Range		
20-29	5	1
30-39	1	5
40-49	2	0
50-59	0	1
Gender		
Female	7	2
Male	1	5
Program Concentration		
Emergency Medical Care	0	3
Health Education	3	1
Health Management	2	3
Nutrition	3	0

In the summer research methods course, five of the eight students had taken a preliminary research methods course within the program (see Table 2: Research Course Exposure). Two respondents had taken undergraduate research methods courses. None had taken a graduate level research course outside of the program. None had taken a biostatistics course at any level previously.

In the fall biostatistics course, three respondents reported having completed a prior research methods course within the program. One had taken a spring research methods course while another had taken both a spring research methods and the summer elective research methods course (the other respondent group in this research). The third had taken only the summer elective research methods course. One of the two respondents who had taken the research methods courses completed both the Ridker-Women's Health Study and the Chan-Nurses' Health Study article evaluations. No respondents had taken research methods courses at the undergraduate level.

Table 2. Research Course Exposure

	Research Methods		Biostatistics	
Exposure Status	**Yes**	**No**	**Yes**	**No**
Any other research course prior to assessment	7	1	7	0
Spring research methods course	5	3	2	5
Summer research methods course (*research participants*)	0	8	2	5
Fall biostatistics course (*research participants*)	0	8	0	7
Undergraduate research course	2	6	0	7
Graduate research course outside program	0	8	0	7

Agency for Healthcare Research and Quality (AHRQ) Evaluation Criteria

The AHRQ criteria used to evaluate each of the chosen articles consist of 21 questions related to the research design, potential biases, validity, reliability, and analysis of study findings. Twelve additional qualitative questions were triggered based on a given response by the student. These follow-up questions asked respondents to give the reasoning behind their initial responses. These qualitative responses provide insight into the students' levels of understanding of the findings presented in the articles.

Responses to three questions revealed significant differences between evaluations of the articles by the two classes. While a question on how well students felt each study population was described (Appendix, question 1) elicited no difference from the assessment of the Ridker-Women's Health Study article ($r=.344$, $p=.155$), there was a significant difference in respondents' evaluations of the Chan-Nurses' Health Study article ($r=.509$, $p=.033$). All students taking part in the summer research methods course responded "yes" when queried on whether the study population in the Chan-Nurses' Health Study article was well described; students in the fall biostatistics course split their assessments between "yes" and "no."

Students were then asked to determine if the authors had properly controlled for study design effects in their analyses (Appendix, question 17). In this case, there were no differences in students' assessments of the Chan-Nurses' Health Study article ($r=.258$, $p=.629$). However, students in the fall biostatistics course unanimously responded "yes" to this question for the Ridker-Women's Health Study article, differing statistically from their summer research methods classmates ($r=.607$, $p=.013$).

The final AHRQ-based question (Appendix, question 24) asked students whether the authors had corrected for controllable variables in order to limit the effect of bias. Differences approaching significance were found in two assessments of the Ridker-Women's Health Study article (r=.513, p=.069), but not found in assessments of the Chan-Nurses' Health Study article (r=.315, p=.489).

No other significant differences existed between students' evaluations of the articles. Marginal differences existed related to intervention detail (Appendix, questions 3-4), sampling (Appendix, question 5), and screening criteria (Appendix, questions 6-7) with the Ridker-Women's Health Study article scoring slightly higher each time. Students from both courses expressed concerns over selection bias and whether entire study populations or eligibility samples were used in both articles (Appendix, questions 8-9). These results suggest uncertainty about how study samples for each article were determined. Selection bias issues in student comments centered on the sample populations' overly homogenous makeup.

Five questions (Appendix, questions 10-14) were related to exposure variables and the validity and reliability of exposure and outcome variables. The initial query in this group questioned whether there was an attempt to measure exposure to the intervention in the study. Aside from possible answers of "yes" and "no," this question gave students a third choice of "not applicable." Follow-up questions dealing with validity and reliability for exposure and outcome variables were triggered automatically. Student responses showed no clear consensus on these issues for either article.

Questions on analysis procedures for each of the studies (Appendix, questions 15-20) produced stronger levels of agreement between respondents in both summer and fall evaluations. All but one student agreed that the articles' authors had used appropriate data analysis procedures. The one dissenting student, enrolled in the summer course, marked "undecided." Students found that the authors in both studies adequately reported the procedures they (the authors) used in performing statistical analyses. Students also stated that the statistical models applied in each article were designed to accurately report both individual and group-level data. From the students' perspectives, few issues based solely on reading each article limited interpretation of the studies' results. Several of the students' follow-up comments, though, dealt with perceived limits of self-reported data, the authors' lack of a minimum timeline to establish aspirin benefit in either study, and the authors' inability to establish causality in either study.

Study Deemed Most Credible

After reviewing each article and completing separate surveys (Ridker-Women's Health Study first, then Chan-Nurses' Health Study), the survey for the Chan-Nurses' Health Study article asked students to choose which article they found more credible (Appendix, question 32). Answer options were "Chan," "Ridker," "equally credible," "neither one credible," and "undecided." In the summer research methods course, decisions on credibility spread across available answer choices, with "Chan" (n=3), "Ridker" (n=1), "equally credible" (n=1), and "undecided" (n=2). In the fall biostatistics administration, students chose "Chan" (n=1), "Ridker" (n=4) and "equally credible" (n=1), with none selecting "neither" or "undecided." The number of total student responses to this question (Table 3: Most Credible Study by Course) may at first appear inconsistent with the number of overall responses. This difference arises from the inclusion of the "study deemed most credible" question only in the survey for the Chan-Nurses' Health Study article. Not all students who initially completed the Ridker-Women's Health Study article responded to the Chan-Nurses' Health Study survey question about credible findings.

Students' choices of which study they considered more credible (see Table 3: Most Credible Study by Course) showed a decided shift between research methods and biostatistics course respondents. The majority of research methods course participants chose the Chan-Nurses' Health Study article; the remaining participants found both studies equally credible or were undecided. The majority of fall biostatistics course participants identified the Ridker-Women's Health Study article as more credible, with none choosing "undecided." No significant difference existed between the groups (*r*=.517, p=.191), however.

Table 3. Most Credible Study by Course

Course	Chan	Ridker	Equally	Neither	Undecided	Total
Research Methods	3	1	1	0	2	***7***
Biostatistics	1	4	1	0	0	***6***
Total	4	5	2	0	2	***13***

Personal Aspirin Use, Actions, and Recommendations to Others

Students also identified their personal aspirin consumption (Appendix, question 31). They then commented on actions they might take in regard to aspirin consumption and what recommendations they would consider making to others due to their evaluation of the articles. Regardless of gender or age range, no statistical differences existed in respondents' use of aspirin. Four students in

the summer research methods course recorded no aspirin use; three listed that they consumed aspirin only for minor aches and pains. Regarding personal aspirin intake, students chose from preset options and stated they may "take aspirin when I get older," "talk to physician about taking aspirin," "only take if necessary," and "take no action." Several students in this group stated they would likely offer recommendations to others on aspirin use, particularly in relation to cardiovascular risk. One student commented that they would "talk with physician about taking low dose aspirin if prone to cardiovascular disease (CVD)."

In the fall biostatistics survey administration, one student identified as a daily user of low dose aspirin, one took aspirin only for minor aches and pains, and four took no aspirin at all. In the comments section, the current "low dose" user stated they would continue their aspirin regimen. In addition to minor aches and pains, another student stated they would consider taking low dose aspirin after consulting with a personal physician. Two of the four students who took no aspirin stated they would consider taking low dose aspirin after reviewing the study articles. The other two students not taking aspirin did not intend to start nor would they recommend it for others. The same students already taking low dose aspirin or considering beginning a regimen stated they would recommend to others that they consult a physician about taking aspirin on a regular basis.

DISCUSSION

Through the exercise, students identified basic study components: population, intervention, and sampling and screening criteria. These components proved quite similar across the Nurses' Health Study and the Women's Health Study, two arguably landmark studies. Students had to move past acceptance of what prestigious researchers asserted and deal with the dichotomy arising from disparate conclusions. Students received instruction on the importance of using a wide range of evidence (particularly the rigorous critical appraisals of multiple studies provided in systematic reviews and meta-analyses) and found that a preponderance of evidence was not always clear or available.

Limitations

The influence of the biostatistics content covered during the fall course presents a possible factor leading to differing responses between students in the two courses. In addition, several students from the research methods course also took part in the fall biostatistics course. The influence of repeated exposure to the articles and discussion among classmates may have promoted a level of consensus in those respondents. The phenomenon of regression to the mean may also have played a role. In the research course, decisions on credibility varied.

Differences in demographic composition between the summer research methods and fall biostatistics courses may also have affected results. Students in the summer research methods course were more likely to have taken a prior research methods course within the program, though these students would have received much less instruction in biostatistics than students taking the biostatistics course. Most respondents in the summer research methods administration were female, younger, and less likely to have considered the subject of the two articles (aspirin's effect on cardiovascular illness). Most respondents in the fall biostatistics course were male and older. This group included three students working in emergency medical services, a discipline that operates in emergent situations frequently of cardiovascular origin and in which aspirin is often considered an early therapy.

A comparatively small sample size (n=14) limited conclusions. This limitation was further complicated by having only one student complete the review of both articles in both classes. A revised design would schedule the review of articles as a regular assignment within the courses. Students would complete the evaluation, then engage in discussion relative to the merits of the pair.

Another possible limitation may have been the evaluation instrument itself. While well validated in the literature, its utility for graduate students with varying exposure to research content or specific terminology may have been problematic. Assessment of student exposure to prior research content did not include whether or not they had completed a graduate level epidemiology course. Information focusing on the identification of exposure and outcome variables is a key feature of epidemiology content. Lack of exposure to key epidemiology content may have contributed to mixed results in those evaluation items.

Recommendations for Application

Principles of this exercise are adaptable in multiple areas. The first step, no matter the setting, requires establishing research evaluation criteria. Educators at the graduate, undergraduate, and associates degree levels, continuing education professionals, and individuals wishing to sharpen their research evaluation skills may adapt questions from the Task Force-based questionnaire (Appendix). Individuals in fields outside the health professions will likely find these questions sufficiently general in scope, but specific enough to research methods, to be applicable. The UK's National Institute for Health and Clinical Excellence (NICE) provides another valuable resource from which research evaluation guidelines can be drawn and questions formulated. In particular, the chapter on evidence review from NICE's *Guidelines Manual* offers great detail on research evaluation.[33] The

international AGREE Collaboration offers succinct queries for the evaluation of clinical guidelines.[34] For increased field specificity, researchers may wish to consult guidance provided through their professional organizations or weighty publications (e.g., recognized handbooks, manuals, etc.).

A librarian may prove helpful in finding research publications to compare, the second step in preparing the exercise. Researchers may use free tools such as the Agency for Healthcare Research and Quality's (AHRQ) National Guideline Clearinghouse (http://guideline.gov) or NICE's Guidance by Topic (http://guidance.nice.org.uk/Topic) to locate guidelines on an issue and evaluate differences between recommendations. While similar studies with conflicting findings provided the basis for the authors' pilot test, instructors wishing to expose students to different types of research may choose studies on similar topics employing dissimilar or combined research methods. For instance, studies on diabetes patient education using qualitative, quantitative, and mixed methods formats could be explored.

Instruction on how and why to seek out the preponderance of evidence on a topic should follow the exercise. Individuals should also be introduced to reviews available through agencies like Cochrane Collaboration (http://cochrane.org), Campbell Collaboration (http://campbellcollaboration.org), NICE (http://nice.org.uk), and AHRQ (http://ahrq.gov). These organizations provide free access to their complete or summarized reviews.

CONCLUSION

While students may possess prior experience in their chosen disciplines (emergency medical care, nursing, etc.), the purpose of graduate education is to move individuals beyond skill level proficiency toward higher level thinking and decision making.[35] A successful student and advanced practitioner must know how to evaluate research in order to improve professional practice. These individuals' instructors should continuously assess research content, align learning experiences with desired outcomes, and develop successful competencies in this arena.

Inherent in any research review is that conclusions drawn from data are open to interpretation.[36] This lack of concreteness may prove difficult for students to acknowledge, much less assimilate into their way of thinking; they may view the purpose of research as proving something definitively. Contradictory research findings provide the basis of teachable moments in working through differences in design and interpretation.

REFERENCES

1. Hadley J, Wall D, Khan K. Learning needs analysis to guide teaching evidence-based medicine: knowledge and beliefs amongst trainees from various specialities. *BMC Medical Education*. 2007;7:11.
2. Harewood G, Hendrick L. Prospective, controlled assessment of the impact of formal evidence-based medicine teaching workshop on ability to appraise the medical literature. *Irish Journal of Medical Science*. 2010;179(1):91-4.
3. Windish D, Huot S, Green M. Medicine residents' understanding of the biostatistics and results in the medical literature. *JAMA: Journal of the American Medical Association*. 2007;298(9):1010-22.
4. Nicholson L, Warde C, Boker J. Faculty training in evidence-based medicine: improving evidence acquisition and critical appraisal. *Journal of Continuing Education in the Health Professions*. 2007;27(1):28-33.
5. Callaham M, McCulloch C. Longitudinal trends in the performance of scientific peer reviewers. *Annals of Emergency Medicine*. 2011;57(2):141-8.
6. Gray M. Journals and the quality of peer reviewing -- a customer's perspective. *Serials*. 2008;21(2):75-7.
7. West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence. *Evidence report/technology assessment* (Summary). 2002;47:1-11.
8. Lipworth W, Kerridge I, Carter S, Little M. Journal peer review in context: A qualitative study of the social and subjective dimensions of manuscript review in biomedical publishing. *Social Science and Medicine*. 2011;72(7):1056-63.
9. Clyde, L. Evaluating the quality of research publications: A pilot study of school librarianship. *Journal of the American Society for Information Science and Technology*. 2004;55(13):1119-30.
10. Lee K, Schotland M, Bacchetti P, Bero L. Association of journal quality indicators with methodological quality of clinical research articles. *JAMA: Journal of the American Medical Association*. 2002;287(21):2805-8.
11. Dougherty M, Freda M, Kearney M, Baggs J, Broome M. Online Survey of Nursing Journal Peer Reviewers: Indicators of Quality in Manuscripts. *Western Journal of Nursing Research*. 2011;33(4):506-21.
12. French B. Evaluating research for use in practice: What criteria do specialist nurses use? *Journal of Advanced Nursing*. 2005;50(3):235-43.
13. Genuis SK, Genuis SJ. Exploring the continuum: Medical information to effective clinical practice. Paper I: The translation of knowledge into clinical practice. *Journal of Evaluation in Clinical Practice*. 2006;12(1):49-62.

14. Grefsheim SF, Rankin JA. Information needs and information seeking in a biomedical research setting: A study of scientists and science administrators. *Journal of the Medical Library Association*. 2007;95(4):426-34.
15. Hemminger BM, Lu D, Vaughan KTL, Adams SJ. Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*. 2007;58(14):2205-25.
16. Cockburn DM. How to make clinical decisions from statistics. *Clinical and Experimental Optometry: Journal of the Australian Optometrical Association*. 2006;89(3):176-83.
17. Hurwitz S, Slawson D. Should we be teaching information management instead of evidence-based medicine? *Clinical Orthopaedics and Related Research*. 2010;468(10):2633-9.
18. Jones S, Crookes P, Johnson K. Teaching critical appraisal skills for nursing research. *Nurse Education in Practice*. 2011;11(5):327-32.
19. El Ansari W. Appraisal skills as a public health competency for evidence-based care: Students' satisfaction with their research education. *Journal of Public Health Management and Practice*. 2004;10(4):354-65.
20. Greenwald ML. Teaching research methods in communication disorders: A problem-based learning approach. *Communication Disorders Quarterly*. 2006;27(3):173-9.
21. Tashakkori A, Teddlie C. Issues and dilemmas in teaching research methods courses in social and behavioural sciences: US perspective. *International Journal of Social Research Methodology*. 2003;6(1):61-77.
22. Wiles R, Durrant G, De Broe S, Powell J. Methodological approaches at PhD and skills sought for research posts in academia: A mismatch? *International Journal of Social Research Methodology*. 2009;12(3):257-69.
23. Northam S, Lakomy J. Promoting advanced practice APRNS' buy-in for evidence based practice. *Internet Journal of Advanced Nursing Practice*. 2009;10(1):1-14.
24. Chou R. Using evidence in pain practice: Part I: Assessing quality of systematic reviews and clinical practice guidelines. *Pain Medicine*. 2008;9(5):518-30.
25. Sackett D, Rosenberg W, Gray J, Haynes R, Richardson W. Evidence based medicine: What it is and what it isn't. *BMJ*. 1996;312(7023):71-2.
26. Porter S, O'Halloran P. The postmodernist war on evidence-based practice. *International Journal of Nursing Studies*. 2009;46(5):740-8.
27. Paynter RA. Evidence-based research in the applied social sciences. *Reference Services Review*. 2009;37(4):435-50.
28. Morago P. Evidence-based practice: from medicine to social work. *European Journal of Social Work*. 2006;9(4):461-77.
29. HealthLinks, University of Washington. *Evidence-Based practice tools summary*. 2007. (Online). http://healthlinks.washington.edu/ebp/ebptools.html.
30. Zaza S, Wright-De Aguero LK, Briss PA, et al. Task Force on Community Preventive Services. Data collection instrument and procedure for systematic reviews in the guide to community preventive services. *American Journal of Preventive Medicine*. 2000;18(1S1):44-74.
31. Ridker P, Cook N, Lee I, et al. A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women. *New England Journal of Medicine*. 2005;352(13):1293-1304.
32. Chan A, Manson J, Feskanich D, Stampfer M, Colditz G, Fuchs C. Long-term aspirin use and mortality in women. *Archives of Internal Medicine*. 2007;167(6):562-572.
33. National Institute for Health and Clinical Excellence (NICE). Reviewing the evidence. *Guidelines Manual*. 2009:69-80. (Online). http://www.nice.org.uk/media/5F5/22/The_guidelines_manual_2009_-_Chapter_6_Reviewing_the_evidence.pdf.
34. AGREE Collaboration. (2001). Appraisal of guidelines for research & evaluation (AGREE) instrument. (Online). http://www.agreecollaboration.org/pdf/agreeinstrumentfinal.pdf
35. Seldomridge L, Walsh C. Measuring critical thinking in graduate education: What do we know? *Nurse Educator*. 2006;31(3):132-137.
36. Chou R. Using evidence in pain practice: Part II: Interpreting and applying systematic reviews and clinical practice guidelines. *Pain Medicine*. 2008;9(5):531-541.

KEY TERMS
Evaluation, Graduate Education, Information Literacy, Comparative Research

APPENDIX A
Research Evaluation Questions

Adapted from Task Force on Community Preventive Services. (2000). Data collection instrument and procedure for systematic reviews in the guide to community preventive services. *American Journal of Preventive Medicine*. 2000;18(1S1):44-74.

1. Was the study population (i.e., the intervention and comparison population) in the X article well described?
2. Please explain why the study population in the X article was or was not well described.
3. Was the intervention well described in the X article?
4. Please explain why the intervention in the X article was or was not well described.
5. Did the authors of the X article specify (i.e., describe characteristics and size of) the sampling frame or universe of selection for the study population?
6. Did the authors of the X article specify the screening criteria for study eligibility (if applicable)?
7. If you responded YES to the previous question, please describe the screening criteria in the X article.
8. Was the population that served as the unit of analysis the entire eligible population or a probability sample at the point of observation in the X article?
9. Are there other selection bias issues in the X article not identified above?
10. Was there an attempt to measure exposure to the intervention in the X article?
11. Were the exposure variables valid measures of the intervention under study in the X article? Please explain your answer.
12. Were the exposure variables reliable (consistent and reproducible) measures of the intervention under study in the X article? Please explain your answer.
13. Were the outcome and other independent (or predictor) variables valid measures of the outcome of interest in the X article? Please explain your answer.
14. Were the outcome and other independent (or predictor) variables reliable (consistent and reproducible) measures of the outcome of interest in the X article? Please explain.
15. Did the authors of the X article conduct appropriate analysis by conducting statistics testing (when appropriate)?
16. Did the authors of the X article conduct appropriate analysis by reporting which statistical tests were used?
17. Did the authors of the X article conduct appropriate analysis by controlling for design effects in the statistical model?
18. Did the authors of the X article conduct appropriate analysis by controlling for repeated measures in the analysis, for study designs in which the same population was followed with repeated measurements over time?
19. Did the authors of the X article conduct appropriate analysis by accounting for different levels of exposure in segments of the study population in the analysis?
20. If the authors of the X article analyzed group-level and individual-level covariates in the same statistical model, was the model designed to handle multi-level data?
21. Were there other problems with data analysis that limit interpretation of the results of the study in the X article?
22. Did at least 80% of enrolled participants complete the study in the X article?
23. Did the authors of the X article assess whether the units of analyses were comparable prior to exposure to the intervention?
24. Did the authors of the X article correct for controllable variables or institute study procedures to limit bias appropriately (e.g., randomization, restriction, matching, stratification, or statistical adjustment)?
25. Please describe all potential biases or unmeasured contextual confounders described by the authors of the X article if there are any. For all responses, indicate the likely direction of effect on the results, if possible.
26. Please describe other potential biases or unmeasured contextual confounders NOT identified by the authors of the X article if there are any. For all responses, indicate the likely direction of effect on the results, if possible.
27. Which age range matches your current age in years?
28. What is your gender?
29. What is your Program concentration area?
30. Which of the following research courses have you completed?
31. Which of the following best describes your personal aspirin consumption?
32. Which study do you find most credible?
33. What action in regard to your personal health would you consider taking in light of your evaluation of these studies?
34. What recommendation in regard to personal health habits would you make to others in light of your evaluation of these studies?