Western States Chiropractic College

Annotated EBP Standards, Learning Objectives & Competencies

6/11/13

The standards and learning objectives are divided into separate chapters. Each chapter is punctuated by blocks of annotation. These annotations fall into 3 broad categories: teaching tips for individual instructors, curriculuar suggestions relative to implementation of the EBP program, and commentary offering editoral input to reflect the thinking of members of the committee.

Symbol for teaching tips 

Symbol for curriculum suggestions 

Commentary 

Table of Contents

Standards and Main Learning Objectives (outline)

[To jump to an item, scroll over an item, press control ("Ctrl") and left click the mouse.]

1. The EBP competent practitioner can present a general overview of the characteristics and principles of EBP.

1.1. Can describe EBP. (1 cn, 1 rl, 1 rg, 1 mh, 1 dp)
1.2. Appreciates the difference between scientific evidence and other forms of knowledge and opinion. (1 dp, 1 rl, 1 jt, 1 mh, 1 cn, 1 rg)
1.3. Appreciates the necessary balance between *patient-oriented evidence* and *disease/pathomechanical-oriented evidence.* (1 jt, 1 mh, 1 rl , 1 cn 1, 1 dp, 1 rg)
1.4. Can explain the steps involved in performing both rapid and in-depth acquisition and assessment of clinical evidence. (1 cn, 1 rg, 1 rl, 1 mh, 1 dp)
1.5. Can articulate the advantages of EBP.
1.6. Can address controversial issues regarding EBP. (1 dp, 2 rl, 1 rg, 1 mh, 1 cn)

2. The EBP competent practitioner can translate an issue of clinical uncertainty into an answerable question.

2.1. The practitioner understands the issues relating to clinical ambiguity and uncertainty. (1 dp, 1 rl, 1 mh, 1 jt, 1 rg, 1 cn)
2.2. The practitioner can translate uncertainty or knowledge gaps into a question that is searchable. (1 dp, 1 rl, 1 jt, 1 mh, 1 rg, 1 cn)

3. The EBP competent practitioner can effectively and efficiently access, retrieve and manage useful, up-to-date health care information and evidence.

3.1. Can design and conduct an effective and efficient literature/information search. (1 dp, 1 rl, 1 jt, 1 mh, 1 rg, 1 cn)
3.2. Is familiar with recommended "best" resources for finding evidence. (1 jt, 1 mh, 1 dp, 1 rg, 1 cn, 1 rl)
3.3. Has the knowledge and skills necessary to coalesce, organize, store and retrieve previously searched health care information. (1 dp, 1 rg, 1 rl, 1 mh, 1 jt, 1 cn)

4. The EBP competent practitioner can critically appraise the validity and clinical significance of relevant evidence.

4.1. Understands the inherent strengths and weaknesses of different levels of evidence and can rate their quality. (1 dp, 1 rg, 1 mh, 1 rl, 1 cn)
4.2. Can demonstrate a basic conceptual understanding of biostatistics. (1 dp, 1 cn, 1 mh, 1 rg, 1 rl)
4.3. Understands the design and hierarchy of different types of primary studies along with their inherent strengths and weaknesses. (1 mh, 1 jt, 1 cn, 1 dp, 1 rg, 1 rl)
4.4. Can describe the basic characteristics that determine the quality of research studies. (1 rg, 1 dp, 1 cn, 1 mh, 1 rl)
4.5. Can demonstrate an understanding of the role and basic characteristics of DIAGNOSTIC tests. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh)
4.6. Can appraise the validity and usefulness of a primary study of DIAGNOSTIC tests. (1 dp, 1 rl, 1 cn, 1 rg)
4.7. Can appraise the validity and usefulness of research on the process of DIFFERENTIAL DIAGNOSIS. (1 rg, 1 dp, 1 cn, 1 mh, 1 rl)
4.8. Can appraise the validity and usefulness of a primary study on THERAPY (e.g., an RCT). (1 mh, 1 dp, 1 cn, 1 rg, 1 rl)
4.9. Can appraise the validity and usefulness of a study on PROGNOSIS. (1 dp, 1 rg, 1 mh, 1 rl)
4.10. Can appraise the validity and usefulness of a study on HARM (prevention and side-effects). (1 dp, 1 cn, 1 rg, 1 mh, 1 rl)
4.11. Can appraise the validity and usefulness of a study on COST EFFECTIVENESS. (2 rl, 1 rg)

5. The EBP competent practitioner applies the relevant evidence to practice.

5.1. Assesses the relevance of the appraised evidence to the clinical problem at hand (*clinical applicability*). (1 dp, 1 mh, 1 cn, 1 rl, 1 rg)
5.2. Can select and interpret diagnostic tests appropriate to a particular patient's problem. (1 dp, 1 rg, 1 mh, 1 cn, 1 rl)
5.3. Understands how to decide if a potential therapy is likely to be appropriate and effective for a particular patient. (2 dp, 2 rg, 1 mh, 1 cn, 1 rl)
5.4. Can apply pertinent evidence to a particular patient situation when estimating potential harm from health care decisions (diagnostic test, treatments, lifestyle choices, etc.). (1 dp, 1 mh, 1 cn, 1 rg, 1 rl)
5.5. Understands and applies prognostic indicators to help predict a patient's outcome. (1 cn, 1 rg, 1 rl, 1 mh, 1 dp)
5.6. Understands how to select appropriate outcome measures. (1 dp, 1 rg, 1 cn, 1 rl)
5.7. Can develop and employ a plan to apply new evidence to the patient's situation. (1 dp, 1 rg, 1 mh, 1 cn, 1 rl)

6. The EBP competent practitioner engages in self evaluation of his/her process for accessing, appraising and incorporating new evidence into practice.

6.1. Demonstrates the behavior necessary to maintain and improve EBP skills. (1 dp, 1 rl, 1 mh, 1 rg, 1 jt, 1 cn)
6.2. Reflects on how well these activities are performed and continues to improve them. (1 jt, 1 cn, 1 rg, 1 rl, 1 mh)

OVERVIEW

STANDARD 1

The EBP competent practitioner can present a general overview of the characteristics and principles of evidence-based practice.

Standards, Main Learning Objectives, and Specific Competencies 5/2/08

1. The EBP competent practitioner can present a general overview of the characteristics and principles of EBP.

1.1. Can describe EBP. (1 cn, 1 rl, 1 mh, 1 dp, 1 rg, 1 jt) (1.0)
1. Can define EBP. (1 cn, 1 dp, 1 mh, 1 rl, 1 rg, 1 jt) (1.0)
2. Can explain what is meant by *best evidence.* (1 cn, 1 mh, 1 dp, 1 rl, 1 rg, 1 jt) (1.0)
3. Can explain what is meant by *clinical expertise.* (1 cn, 1 mh, 1 dp, 2 rl, 1 jt, 1 rg) (1.2)
4. Can explain what is meant by *patient values* and *circumstances.* (1 cn, 1 mh, 1 dp, 1 rl, 1 jt, 1 rg) (1.0)
5. Can outline the 5 classic steps in the application of EBP. (2 cn, 1 mh, 1 rl, 1 jt, 1dp, 1 rg) (1.2)

Teaching Tips: A quick and simple way to drill students in applying the steps is with the alliteration: ask, access, assess, apply, self-assess. The more often and the more places students hear this approach and are expected to follow it, the greater the likelihood that the process will be internalized [RL].

1.2. Appreciates the difference between scientific evidence and other forms of knowledge and opinion. (1 dp, 1 rl, 1 jt, 1 mh, 1 cn, 1 rg) (1.0)
1. Can differentiate data from assertions and opinions. (1 dp, 1 rl, 1 jt, 1 cn, 1 mh, 1 rg) (1.0)
2. Can differentiate a balanced, systematic consideration of the evidence from a selective data presentation ("cherry picking" data). (1 dp, 1 rl, 1 jt, 1 cn, 1 mh, 1 rg) (1.0)
3. Can differentiate among rational hypotheses, empirically-based hypotheses, and apriori beliefs. (1 dp, 1 rl, 1 jt, 1 cn, 1 mh , 1rg) (1.0)

Commentary: Below is a table which my good to trigger a student discussion relative to unscientific claims for some new medical or chiropractic technology and how one might respond.

Seller Assertions	Buyer Responses
"Help more patients"	Indications and contraindications, please
"Better than last year's game"	How much better – earlier discharge, less disability, and sooner back to work?
"Has more frequencies and amplitudes"	And so – what?
"Our research shows…"	How surprising is that?
"Good for everything"	But no thing is ever good for everything – except maybe nothing
"All patients better after six weeks"	But who isn't better after six weeks?
"Makes you more money"	But must I "sell" it?
"Developed by a medical doctor"	If it is so good, why aren't they using it?
"Developed at NASA"	Not much good news coming out of NASA these days; glad they are finally focusing on chiropractic
"Justifies care"	So, I can release patients sooner?
"Published in a major medical journal"	The Uzbekistan Medical Journal of Applied Aura Reading and Astrology?
"Clinical certainty"	Truly remarkable – the first time in the history of medicine and science. And the Nobel prize goes to..?
"Just look at all these references"	But they are on your Web site, not clinical studies, all in Japanese, etc.
"Too good to be true"	I totally agree

Dynamic Chiropractor, January 1, 2007

1.3. Appreciates the necessary balance between *patient-oriented evidence and disease/ pathomechanical-oriented evidence.* (1 jt, 1 mh, 1 rl , 1 cn 1, dp, 1 rg) (1.0)
1. Can articulate the difference between patient-oriented evidence and disease or pathomechanical evidence. (1 dp, 1 rl, 1 jt, 1 mh, 1 cn, 1 rg) (1.0)

a. Can define the characteristics of patient-oriented evidence (e.g., based on mortality, morbidity, pain status, functional capacity, and quality of life). (1 rl, 1 cn, 1 dp, 1 jt, 1 mh, 1 rg) (1.0)
b. Can define the characteristics of disease-oriented and pathomechanical evidence (i.e., based on understanding etiology and mechanisms, or measuring pathophysiologic, neurologic, or biomechanical changes). (1 rl, 1 cn, 1 dp, 1 mh, 1 rg, 1 jt) (1.0)
c. Can distinguish patient-oriented outcomes from changes in physical examination findings (e.g., palpatory tenderness, spinal motion, muscle tests, leg alignment). (1 rl, 1 cn, 1 dp, 1 jt, 1 mh, 1 rg) (1.0)

2. Can articulate the strengths and weakness of evidence which is based on pathophysiology (i.e., disease) and pathomechanical research. (1 jt, 2 mh, 1 rl, 1 cn, 1 dp, 1 rg) (1.2)
   a. Can cite the benefits of clinically oriented basic science research (e.g., best causal evidence) when compared to informal clinical experience or speculation based solely on extrapolation of basic science or biomechanical principles. (2 dp, 1 cn, 2 mh, 2 rl, 2 rg, 1 jt) (1.7) [explain in wiki]
   b. Can identify the limitations of pathophysiological and pathomechanical evidence compared to EBP/outcome-oriented evidence. (1 rl , 1 cn, 1 dp, 1 jt, 1 mh, 1 rg) (1.0)

Commentary: There are many examples of why there is often a shift from measuring treatment effects based on physiological changes vs. patient based clinical outcomes as illustrated in this alert from Physician's First Watch for August 9, 2007 regarding a drug used to treat diabetes.

"Writing online for the New England Journal of Medicine, he says the committee sought to evaluate the evidence about rosiglitazone, "a new 'wonder drug,' approved prematurely and for the wrong reasons by a weakened and underfunded government agency subjected to pressure from industry, [that] had caused undue harm to patients." The advisory committee concluded that use of rosiglitazone carried risks for myocardial ischemia, and recommended, not removal from the market, but label warnings and "extensive educational efforts. He says that among the studies evaluated, two of the largest "failed to find a significant reduction in cardiovascular events even with excellent glucose control. Rosen recommends that the FDA shift its primary efficacy end point away from surrogates, like glycated hemoglobin levels, to clinical outcomes. He says the agency took a similar step a generation ago when it shifted its end point for osteoporosis drugs from bone mineral density to fractures." [RL 8/9/07]

3. Can put into perspective the role of pathophysiologic and pathomechanical evidence in making clinical decisions. (1 jt, 1 mh, 1 rl , 1 cn, 1 dp , 1 rg) (1.0) [check this one]
   a. Can access meaningful evidence in these realms of knowledge. (1 rl , 1 cn, 1 dp, 1 jt, 1 mh 1, 1 rg) (1.0)
   b. Can appraise the quality and relevance/applicability of this type of evidence (1 rl , 1 cn, 2 dp, 1 jt, 1 mh, 2 rg) (1.3)

1.4. Can explain the steps involved in performing both rapid and in-depth acquisition and assessment of clinical evidence. (1 cn, 1 dp, 1 mh, 1 jt, 1 rg, 1 rl) (1.0)
   1. Can perform a comprehensive literature search and an in-depth critical analysis of the quality of individual primary studies, applying the classic steps of EBP ("doing mode"). (1 cn, 1 dp, 1 mh, 1 rl, 1 jt, 1 rg) (1.0)
      a. Understands that this process is most commonly applied to those conditions encountered routinely. (1 mh, 3 rl, 2 cn, 3 jt, 1 dp, 2 rg) (2.0)
   2. Can rapidly access dependable sources of pre-appraised evidence and judge its quality and applicability (skipping the critical appraisal of primary sources) ("using mode"). (1 cn, 1 dp, 1 rl, 1 jt, 1 rg, 1 mh) (1.0)
      a. Understands that this process is the most commonly applied to conditions encountered less frequently. (2 cn, 1 mh, 1 dp, 3 rl, 2 jt, 2 rg) (1.8)
      b. Understands that this process is most practical for addressing questions during clinical practice. (1 cn, 1 mh, 1 dp, 2 rl, 2 jt, 1 rg) (1.3)
   3. Can access and identify quality clinical guidelines and decision-making rules relevant to his/her patient ("replicating mode"). (1 cn, 1 dp, 1 mh, 1 rl, 1 jt, 1 rg) (1.0)
      a. Understands that this process is more commonly applied to conditions encountered very infrequently. (1 dp, 1 mh, 3 rl, 2 cn, 2 jt, 2 rg) (1.8)

1.5. Can articulate the advantages of EBP.

1. Understands that evidence-based care and best practice recommendations may lead to better patient outcomes. (1 dp, 1 rl, 1 mh, 1 rg, 1 cn, 1 jt) (1.0)
    a. Understands that clinical experience alone is not enough to provide the best possible care. (1 cn, 1 dp, 1 mh, 1 rl, 1 jt, 1 rg) (1.0)

Commentary: The following quote offers useful background. "How do experienced clinicians 'know what they know?' The inherent knowledge of clinical experience has been called 'knowing in practice.' (Hogarth 1987) Experience is what allows seasoned clinicians to come up with a diagnosis after only spending a moment with the patient. This ability can be developed only with years of seeing the same patterns emerge in patients with similar problems, allowing clinicians to gain the insight of "hearing between the lines." Performing countless physical examinations results in clinicians who "know" when an ovary is enlarged or how to maneuver an endoscope around the splenic flexure. Continuous experience, and learning from this experience, is how knowing in practice occurs.

In contrast, EBM knowledge comes from evaluating scientific research. This way of knowing requires the critical appraisal of the study's methods and an interpretation of the numeric results. Information Mastery furthers this way of knowing by preferentially relying on final outcomes, patient-oriented evidence that matters (POEM)."

Clinical guidelines can also be helpful. Farabrugh, writing about the benefits of the CCGPP's Best Practice Initiative, states "In his 1997 North American Spine Society Presidential address, Dr. Saul stated: '…physicians often prescribe treatment for their patients based upon their most recent success or failure. We skim our journals for articles that appeal to us and sort out information that does not support our frame of reference. Even learned people will tend to gather and synthesize information preferentially as it supports and relates to their own opinions and objectives. 'Sort out the information'"(Farabraugh 2006)

References

Hogarth R. Judgment and choice. 2nd ed. New York: John Wiley & Sons; 1987.

Ellis J, Mulligan I, Rowe J, Sackett DL. Impatient general medicine is evidence-based. Lancet 1995;346:407-10.

Farabrugh RJ, The Week in Chiropractic, Foundation for Chiropractic Education and Research. Vol. 12, No. 49. September 20, 2006

Gross CP, Anderson GF, Powe NR. The relation between funding by the National Institutes of Health and the burden of disease. N Engl J Med 1999;340:1881-7.

    b. Understands that knowledge of disease processes is not enough for effective patient management. (1 cn, 1 mh, 1 dp, 1 rl, 1 jt, 1 rg) (1.0)
2. Understands that EBP provides a method to maintain and update clinical skills. (1 dp, 1 rl, 1 mh, 1 rg, 1 cn, 1 jt) (1.0)
    a. Understands that there is a need to remain up-to-date in an environment of continuous and rapidly expanding health care information. (1 cn, 1 dp, mh 1, 1 rg, 1 rl, 1 jt) (1.0)
    b. Understands that finding up-to-date evidence on a particular clinical question may be more useful than depending solely on postgraduate education programs. (1 dp, 1 rl, 1 mh, 1 rg, 1 cn, 1 jt) (1.0)
3. Understands the role that EBP can play in furthering the goals of the profession. (1 cn, 1 dp, 1 mh, 2 rl, 1 rg 1 jt) (1.2)
    a. Understands the role that EBP can serve to improve professional credibility and recognition. (1 dp, 1 mh, 2 rl, 1 cn, 1 jt, 1 rg) (1.2)
    b. Understands the role that EBP can serve to improve chiropractic's positioning in societal, political, and the insurance environments. (1 dp, 1 mh, 2 rl, 1 cn, 1 jt, 1 rg) (1.2)
    c. Understands the role that EBP can serve to help establish chiropractic care in integrative health care. (1 dp, 1 mh, 2 rl, 1 cn, 1 jt, 1 rg) (1.2)
    d. Understands the potential role that EBP can play in expanding scope of practice. (1 dp, 1 mh, 3 rl, 2 cn, 2 jt, 2 rg) (1.8)

1.6. Can address controversial issues regarding EBP**.** (1 dp, 1 mh, 2 rl, 1 cn, 1 jt, 1 rg) (1.2)

1. Can articulate potential barriers to EBP. (1 dp, 2 rl, 1 mh, 1 jt, 1 cn, 1 rg) (1.2)
    a. Understands the natural apprehension that one can have of the new subject material that EBP represents (especially concern about level of biostatistics expertise needed). (1 dp, 2 rl, 1 cn, 1 mh, 2 jt, 1 rg) (1.3)

b. Understands the role of inherent human skepticism and resistance to change. (2 dp, 2 rl, 1 cn, 1 mh, 2 jt, 2 rg) (1.7)
c. Understands the challenge that there are large amounts of information to manage ("information overload"). (1 dp, 2 rl, 1 cn, 1 mh, 1 jt, 1 rg) (1.2)
d. Understands that there can be peer bias against EBP. (2 dp, 2 rl, 1 cn, 1 mh, 2 jt, 2 rg) (1.7)
e. Understands there can be a lack of professional support and encouragement in developing EBP skills out in practice. (2 dp, 3 rl, 1 cn, 1 mh, 2 jt, 2 rg) (1.8)
f. Understands there are limited mentors for role modeling. (2 dp, 3 rl, 1 cn, 1 mh, 2 jt. 2 rg) (1.8)

 **Commentary: Practitioner obstacles—**"In particular, the unrealistic expectation that evidence should be tracked down and critically appraised for all knowledge gaps led to early recognition of practical limitations and disenfranchisement amongst some practitioners. (McAllister)"

McAllister FA, Graham I, Karr GW, Laupacis A: Evidenced-based medicine and the practicing clinician. J Gen Intern Med 1999, 14:236-242.

Top 10 Pearls for Translating Knowledge to Practice

1. Do not make the assumption that knowledge equals behavior change. Interventions for change need to include both behavioral and knowledge strategies.
2. Try your ideas in the clinic sooner than later. Do not wait until you have a perfect product.
3. Spend at least as much or more time on determining your barriers to change as on analyzing the evidence.
4. Involve members of your office staff. This is especially true regarding tracking initiatives.
5. "Cherry-pick" and hunt for solutions. Do not reinvent the wheel.
6. Keep it simple but multifaceted. There is no magic bullet, but simplicity combined with a few different lines of attack seems to be most effective.
7. Befriend an expert in marketing or design. In the end, you are "selling" something.
8. Reduce the number of steps or people involved. For example, many knowledge products do better if they target the consumer rather than the physician, who must then translate to them to the consumer.
9. If you hope to get other colleagues in your clinic to implement your evidence-based intervention, do not assume that they care or will give their time freely.
10. Build in a simple evaluation system. This will be rewarding for everybody when you can see a change. p. 30

2. Understands the criticisms and misperceptions surrounding EBP. (1 jt, 1 dp, 1 cn, 1 rg, 1 mh, rl 1) (1.0)
   a. Understands the perception that EBP might be used to define evidence too narrowly, focusing too much on controlled studies (e.g., double-blind random controlled studies), minimizing the contribution of other study designs (e.g., observational studies). (1 rl, 1 jt, 1 dp, 1 rg, 1 mh, 1 cn) (1.0)
   b. Understands the perception that EBP might overemphasize evidence based on patient-centered outcomes while under valuing to an inappropriate degree evidence derived from pathophysiological and pathomechanical investigation. (1 mh, 1 rl, 1 jt, 1 cn, 2 dp, 2 rg) (1.3)
   c. Understands the perception that EBP may devalue clinical experience. (1 rl, 1 mh, 1 jt, 1 rg, 1 cn, 1 dp) (1.0)
   d. Understands the fear that EBP might minimize the role of patient values. (1 cn, 1 mh, 1 dp, 2 rl, 1 jt, 1 rg) (1.2)
   e. Understands the perception that EBP can promote a "cookbook" health care approach. (1 cn, 1 mh, 1 dp, 1 rl, 1 jt, 1 rg) (1.0)
   f. Understands the perception that EBP might threaten the autonomy of the doctor-patient relationship. (2 rl, 2 dp,1 mh, 2 rg, 2 cn, 1 jt) (1.7)
   g. Understands the fear that that EBP can be used inappropriately to promote cost cutting, poorer quality of care, and 3rd party payment denial. (1 cn, 2 dp, 1 mh, 1 rl, 1 jt, 2 rg) (1.3)
   h. Understands the concern that EBP may be too time intensive to be practical in busy clinical practice. (2 cn, 2 dp, 1 mh, 1 rl, 1 jt, 2 rg) (1.5)
3. Understands that EBP has limits in contributing to diagnostic or therapeutic certainty. (1 dp, 1 rl, 1 rg, 1 mh, 1 cn, 1 jt) (1.0)
   a. Understands that not all issues can be formulated into questions that will yield evidence-based answers. (1 cn, 1 dp, 1 rl, 1 mh, 1 jt, 1 rg) (1.0)
   b. Understands that often there is insufficient quantity and quality of evidence to make an evidence-based clinical decision. (2 rl, 1 mh, 2 jt, 2 rg, 2 cn, 2 dp) (1.8)

Commentary: "Compared with the breadth of clinical questions, the pool of research-supported clinical answers is small. In a study conducted in England, only about half (53%) of inpatient general medical services were evidence based (Ellis 1995); that figure dropped to 31% in ambulatory practice. (Gross 1999) p. 61

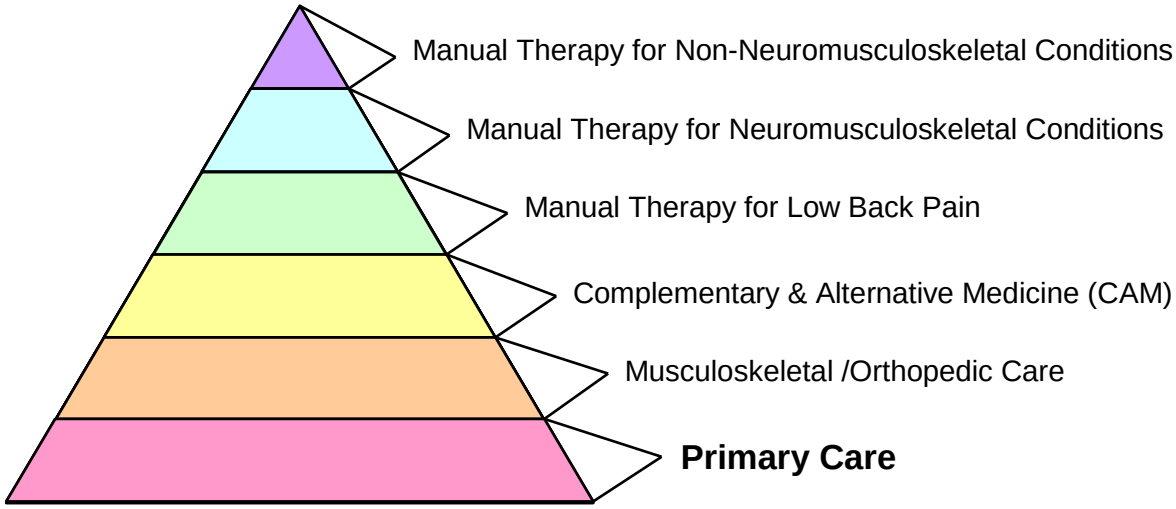
References
Hogarth R. Judgment and choice. 2nd ed. New York: John Wiley & Sons; 1987.

Ellis J, Mulligan I, Rowe J, Sackett DL. Impatient general medicine is evidence-based. Lancet 1995;346:407-10.

Gross CP, Anderson GF, Powe NR. The relation between funding by the National Institutes of Health and the burden of disease. N Engl J Med 1999;340:1881-7.

Furthermore, the amount of data available varies greatly depending on the filed or domain of knowledge. This document divided knowledge necessary for a chiropractic physician into 4 domains: primary care, musculoskeletal care, complementary and alternative medicine (CAM), and Manual therapy.

The Information Pyramid



Curricular Suggestions: It is easier to find high quality research in many areas of primary care (e.g., the role of hypertension and heart disease) than it is in the arena of manipulation for a visceral complaint like asthma). Because of the emphasis and interest in chiropractic in questions regarding manual therapy, students often will formulate questions to which there is no or little patient-oriented outcome research. Unless early assignments are divided into domains where the student will meet some success in accessing the literature, they may quickly become disenchanted with the entire process. Early assignments in the curriculum must both be relevant and assure success so that foundation skills are built and students see the value of the new skills we are asking them to acquire. [RL]

c. Understands that the generalizability of research evidence may be limited in practice settings. (1 dp, 1 rl, 1 mh, 1 jt, 1 rg, 1 cn) (1.0)
d. Understands that there may be conflicting studies or systematic reviews. (1 dp, 1 rl, 1 mh, 1 jt, 1 rg, 1 cn) (1.0)
e. Understands the pitfall of responding to the limitations of evidence-based care with a general nihilistic view. (1 dp, 2 rl, 1 mh, 1 rg, 1 jt, 1 cn) (1.2)
f. Understands that there is limited research demonstrating whether EBP itself actually improves patient outcomes. (2 rl, 2 jt, 2 dp, 1 mh, 2 rg, 1 cn) (1.7)

Commentary: The potential for seeing not just the process of EBP as irrelevant, but chiropractic care to be insufficiently substantiated is a particularly big problem. Unlike most medical programs where students early on are exposed to the phenomenology of patient care, replete with successes and the failures, students at WSCC having very

limited clinical exposure until late in the program. It is easy to become dejected because of the lack of evidence or the flaws in existing evidence without counter-balancing the positive aspects of the phenomenon of patient care.

Skepticism can also creep in because flaws or limitations are easy to find in most studies. Here we see a clash in professional cultures. Researches by nature are trained to be skeptical, see why something doesn't work, and to challenge the research assumption. Clinicians by nature are optimistic, trained to want to see how things might work, especially if those things already coincide with their personal assumptions. The evidence-based practitioner needs to harmonize both inherent tendencies.

Students should be trained to under the following axiom: All studies are flawed, not all flaws are fatal, flawed studies can still be useful studies. As Dawes (2005) writes, although no study is perfect, "this does not mean that you should automatically throw the study away. Rather, the results need to be interpreted in the light of the bias(es) that might have been introduced. In general, weaknesses in study design tend to lead to overestimates of test accuracy. Lijmer et al (1999) found that the two design weaknesses that were associated with the greatest inflation of estimates of test accuracy were when case-control designs were used, and when differential verification bias was present (different reference standards used for positive and negative test results)."

Teaching Tips: It is critical that this concept be communicated to students or it is too easy to dismiss the entire endeavor of accessing and assessing research as useless. However, the nuanced judgment of how useful is a flawed study is a difficult one. Although there is no absolute rubric to aid us, instructor discussion and modeling may be critical here. How is this done? One specific example comes from Mant ("Is this test effective?" in Daws 2005):

"At the beginning of appraisal many people new to it are surprised at the number of flaws in papers, even from established journals. It is therefore quite easy to 'rubbish' a paper. This will give you confidence to begin with. The skill of appraisal is not only to answer these quality questions, but later to evaluate how these flaws might influence the results. Would 78% follow-up significantly alter the results in this paper? By examining critically you seek to assess the inference of bias produced during the research, on the eventual results. It is possible to value and use results that contain bias. That is the real skill of appraisal."

ASK

STANDARD 2

The EBP competent practitioner can translate an issue of clinical uncertainty into an answerable question.

2. The EBP competent practitioner can translate an issue of clinical uncertainty into an answerable question.

2.1. The practitioner understands the issues relating to clinical ambiguity and uncertainty. (1 dp, 1 rl, 1 mh, 1 jt, 1 rg, 1 cn) (1.0)

Curricular Suggestions: This is a very important concept and one which students may be very uncomfortable with. It may be important to deal with this issue formally in a lecture format and then encourage instructors throughout the curriculum to remain sensitive to this issue from a student's perspective.

1. Understands the role of probability (and, therefore, uncertainty) in establishing provisional or differential diagnoses, predicting prognoses, and assessing risks. (1 dp, 1 rl, 1 mh, 1 rg, 1 cn, 1 jt) (1.0)

Commentary: Whereas students may be comfortable with the issue of probability in terms of risk factor assessment, they may be surprised of the role it plays in making a diagnosis. The terms differential diagnoses, provisional diagnoses and working diagnoses should be introduced to them. More importantly is that diagnoses rarely come with certainty. Especially in the realm of diagnosis in the chiropractic setting we rarely have gold standard tests. Diagnostic uncertainty is a daily reality. Uncertainty is introduced in Introduction to EBP video in Phil 1 (Q1) and is taught more substantively as it applies to diagnosis in EBP 1 (Q4).

2. Understands the challenges in linking cause and effect regarding therapy or harm in clinical settings or in research studies. (1 dp, 1 rl, 1 jt, 1 mh, 1 rg, 1 cn) (1.0)
   a. Can describe the rules of evidence regarding causality in clinical research (e.g., using Hill's/Koch's postulates). (1 dp, 2 cn, 1 rg, 1 rl, 1 mh, 1 jt) (1.2)

Commentary: Students should be familiar with the basic tents necessary to establish a cause and effect relationship. This understanding can be applied to issues of etiology as well as the classic EBP categories of diagnosis, treatment and harm.

Evaluation of Cause & Effect: Koch's (1882)/ Bradford-Hill's (1965) Postulates for Evaluating Causation:

{1} Postulate – of a *Temporal Order effect:* cause precedes effect.
{2} Postulate – of a *Biological Gradient (or Dose/Response) effect:* larger exposure to cause will lead to greater effects.
{3} Postulate – of a *Consistency / Repeatability effect* (scientific replication): repeatedly observed by different people, in different circumstances, and times.
{4} Postulate – of an *Interventional ('dechallenge / rechallenge') effect:* the association between cause and effect is reversible.
{5} Postulate – of *Biological Plausibility:* makes sense, according to biologic knowledge of the time.

2.2. The practitioner can translate uncertainty or a knowledge gap into a question the answer to which is best found in reliable sources. (1 dp, 1 rl, 1 jt, 1 mh, 1 rg, 1 cn) (1.0)
1. Can differentiate a *foreground* from a *background question*. (1 dp, 1 rl, 1 jt, 1 mh, 1 cn , 1 rg) (1.0)
   a. Can identify the best types of resouces to answer background questions, such as textbooks and narrative reviews. (2 dp, 1 rg, 1 cn, 1 rl, 1 mh, 1 jt) (1.2)
2. Can determine the type of clinical question that is being posed. (1 dp, 1 rl, 1 jt, 1 mh, 1 rg, 1 cn) (1.0)
   a. Recognizes t a therapy-related question. (1 cn, 1 dp, 1 rg, 1 rl, 1 mh, 1 jt) (1.0)
   b. Recognizes a harm-related question in terms of risk factors and prevention as well as side effects. (1 cn, 1 dp, 1 rg, 1 rl, 1 mh, 1 jt) (1.0)
   c. Recognizes a diagnosis-related question, both in terms of differential diagnosis and test accuracy. (1 cn, 1 dp, 1 rl, 1 mh, 1 jt, 1 rg) (1.0)
   d. Recognizes a prognosis-related question. (1 cn, 1 dp, 1 rg, 1 rl, 1 mh, 1 jt) (1.0)

Commentary: The four major categories of questions are taken from Sachet (2004). However, these categories can be further divided. Below is a listing quoted in Dawes (2005) from Richardson et al (1995) in their concise summary of clinical question formulation, checked the main question types to help us formulate questions. We can adapt their categories, adding others if we need to, to help us 'locate' our questions, ready for the next stage. (Dawes 2005)

- Intervention: Is this intervention (treatment, test, exposure, etc.) more effective (in terms of stated outcome /s) than another / others/ doing nothing, etc.?
- Prevention: How do we reduce the risk of this disease?
- Harm / risk: What are the side-effects, risks, etc. of this intervention? Does it do more harm than good?
- Cause / etiology: What are the causes of this condition or state of affairs?
- Differential diagnosis: How do we distinguish condition *a* from condition *b*?
- Diagnostic testing: How accurate (sensitive / specific) is this diagnostic test (compared with another)?
- Prognosis: What is the likely outcome, course, progression, or survival time of this condition?
- Cost effectiveness: Is intervention *x* more cost-effective than intervention *y*?
- Quality of life: What will be the quality of life for the patient(s) following (or without) this intervention, with this condition, etc.?

Curricular Suggestions: Formulating questions is a skill that should be introduced early in the curriculum (the first year). Basic Science courses may be able to play a role here. It will also be a skill that will need to be re-visited throughout the 2nd through 4th years. [RL]

3. Can frame a foreground question into its critical "PICO" components (i.e., the relevant population of patients or the problem of interest (P); the type of intervention/exposure/prognostic indicator (I), or comparison of one intervention to a standard intervention (C), and the specific outcome of interest (O)). (1 cn, 1 dp, 1 rl, 1 mh, 1 jt, 1 rg) (1.0)
4. Can construct an effective search string based on the components of a PICO questioni (1 jt, 1 cn, 1 mh, 1 dp, 1 rl , 1 rg) (1.0).
   a. Can choose appropriate Boolean operators and search punctuation (e.g., parentheses, asterisk) (1 sb, 1jt, 1rl, 1 tw, dp 1, lh 1)
   b. Can choose appropriate synonyms as search terms (1 sb, 1jt, 1rl, 1 tw, dp 1, lh 1)
5. Can demonstrate a strategy to capture patient-related questions while working in a clinic setting. (1 cn, 1 dp, 1 rl, 1 mh, 1 jt, 1 rg) (1.0)

Teaching Tips: In the clinic milieu, specific behavior patterns should be re-enforced. Below are a series of tip[s from Dawes (2005).

Tip 1 Ask questions

Try asking one question per patient:

- sticky label or name –
- the problem – COPD
- the question – is spirometry an effective predictor of clinical outcome (mortality – length hospital stay)?

Put them in your pocket and look at them at the end of the week.
Select one question because there is likely to be an answer
The question has arisen:

- More than once or
- Is important

Tip 2 Searching

Search one question every 2 weeks or every month or every quarter!
Search logically – 1st Clinical Evidence, 2nd Journal Evidence-Based Medicine, 3rd Cochrane, 4th MEDLINE
Often you will find:

- Too few articles
- They will not be in your library or they may take a long time to get (this is not so true anymore)
- There are too many and a systematic review is needed.

Unless you have time and the question is desperately important, move onto the next question and let someone else answer this one!
Appraise the articles that answer your question, offer the highest level of evidence and are readily available.

Tip 3 Appraisal

Look for letters about the article in subsequent issues of the journal.

- Appraise with others until confident
- Appraise using worksheets
- Or use software – CATmaker or www.gpfaqs.com

Mark (highlight) on the printed article where you found the important data.
Get someone else to check it for you.
Practice writing declarative headings – use the word 'may' a lot.

Tip 4 Share your knowledge

Try sharing uncertainty with your colleagues:

- Discuss your questions with colleagues (maybe they have answered it!)
- Find fault with the article(s) – never your colleagues
- Seek improvement in your own care
- Strive to do no harm.

ACCESS

STANDARD 3

The EBP competent practitioner can effectively and efficiently access, retrieve and manage useful and up-to-date healthcare information and evidence.

3. STANDARD 3: The EBP competent practitioner can effectively and efficiently access, retrieve and manage useful, up-to-date health care information and evidence.

3.1. Can choose appropriate sources to access information/research evidence based on need, time restrictions, and the nature and depth of the information/evidence being sought. (1 dp, 1 rl, 1 jt, 1 mh, 1 rg, 1 cn) (1.0)

1. Can define and discuss the suitability of research-based professional journals, open-source journals, peer-reviewed journals, trade journals and lay publications depending on the nature of the information required. (1 dp, 1 rl, 1 rg, 1 jt, 1 mh, 1 cn) (1.0)
2. Can select an appropriate *data base or other electronic resource.* (1 dp, 1 rl, 1 rg, 1 jt, 1 mh, 1 cn) (1.0) (See 2.5.1)
   a. Recognizes that each database/resource has unique characteristics and selects appropriately. (1 dp, 1 rl, 1 rg, 1 jt, 1 mh, 1 cn) (1.0)
   b. **Can demonstrate familiarity with and the ability to use the following identified best electronic tools and resources:**
      1. DynaMed
      2. EBSCOHost platform databases : MEDLINE COMPLETE, CINAHL, Cochrane Library, SportDiscus, Rehabilitation and Sports Medicine, DARE, AMED
      3. Other proprietary databases: Natural Standard, Natural Medicines Comprehensive Database
      4. Publically available web-based databases and websites: PubMed, PubMed Clinical Queries , TRIP, BestBets, Index to Chiropractic Literature, PEDro, Medline Plus
      5. Guidelines: U.S. Preventive Services Task Force, Canadian Task Force on Preventive Health Care, Guidelines.gov

Commentary: "Conducting a literature search using a software package such as *Grateful ed* to answer specific clinical questions is another approach to obtaining relevant information.

You may also access information from Web sites such as *HYPERLINK* (<http://www.gacguidelines.ca>), which provide guidelines on many clinical topics that have been assessed as the most evidence-based and least biased guideline on the subject currently available. P. 78

3. Can select sources based on limited time considerations. (1 dp, 1 rg, 1 rl, 1 jt, 1 mh, 1 cn) (1.0)
   a. Understands the need to access quality information in a busy practice setting. (1 cn, 1 dp, 1 rg, 1 mh, 1 jt) (1.0)
   b. Can demonstrate a strategy of how to use pre-filtered/pre-appraised sources (e.g., guidelines, synopses, point of service resources) to aid in rapid acquisition and assessment. (1 rl, 1 cn, 1 dp, 1 mh, 1 rg) (1.0)

Teaching Tip: There was strong agreement within the committee that students should be taught to start with pre-filtered EBP reviews to answer foreground questions (e.g., OVID EBM including ACP Journal, DARE, Cochrane Library). They should start with a well referenced, recent and frequently updated textbook or narrative review to answer background questions.

4. Can select sources when greater depth and comprehensiveness is desired or a search of pre-filtered resources is unproductive.
   a. Understands where to search for primary studies and systematic reiews (e.g., MEDLINE, PUBMED, Cochrane Library, CINAHL) (1 dp, 1 rl, 1 jt, 1 mh, 1 rg, 1 cn) (1.0)
5. Can select appropriate sources based when the goal is browsing (e.g., "*foraging*" for useful information from journals or "push" services) rather than problem-solving (*"hunting*" for an answer to a specific clinical question). (1 dp, 1 rl, 1 jt, 1 mh, 1 cn, 1 rg) (1.0)
   a. Can define and identify push services. [to be revised and voted on]
   b. Knows criteria useful in selecting appropriate push services.
   c. Knows how to set up alerts to have targeted material pushed (e.g., alerts or RSS feeds).

Commentary: The flowing list for what makes for a high quality hunting foraging tool was taken from Slawson DC, Shaughnessy AF, Teaching evidence-based medicine: should we be teaching information management instead? Academic Medicine 2005 Jul;80(7): 685-9.*

A high-quality foraging tool employs a transparent process that

- filters out disease-oriented research and presents only patient-oriented research outcomes,
- demonstrates that a validity assessment has been performed using appropriate criteria,
- assigns levels of evidence, based on appropriate validity criteria, to individual studies,
- provides specific recommendations, when feasible, on how to apply the information, placing it into clinical context,
- comprehensively reviews the literature for a specific specialty or discipline, and
- coordinates with a high-quality hunting tool.

A high-quality hunting tool employs a transparent process that

- uses a specific, explicit method for comprehensively searching the literature to find relevant *and* valid information,
- provides key recommendations supported by patient-oriented outcomes when possible and, when not, specified as preliminary when supported only by disease-oriented outcomes,
- assigns levels of evidence† or strength of recommendation‡ to key recommendations using appropriate criteria, and coordinates with a high-quality foraging tool.

6. Can differentiate primary research literature from pre-appraised/pre-filtered and other secondary sources.(1 rl, 1 jt, 1 dp, 1 rg, 1 cn, 1 mh) (1.0)
   a. Understands the why primary research literature is the best first choice in answering foreground questions. (1 mh, 1 rl, 1 jt, 1 dp, 1 rg, 1 cn) (1.0)
      i. Can recognize primary research literature. (1 dp, 1 rl, 1 jt, 1 rg, 1 cn, 1 mh) (1.0)
      ii. Can cite the advantages and disadvantages of using primary research literature . (1 dp, 1 rl, 1 jt, 1 rg, 1 cn, 1 mh) (1.0)
   b. Understands the role of pre-appraised/pre-filtered sources in answering foreground questions. (1 mh, 1 rl, 1 jt, 1 dp, 1 rg, 1 cn) (1.0) SB wants wording also "the best first choice"?
      i. Can recognize pre-appraised/pre-filtered sources and cite examples (e.g., guidelines, synopses, systematic reviews, point of service resources) . (1 dp, 1 rl, 1 jt, 1 rg, 1 cn, 1 mh, 1 jt) (1.0)
      ii. Can cite the advantages and disadvantages of using pre-appraised/pre-filtered literature. (1 dp, 1 rg, 1 cn, 1 rl) (1.0)
   c. Understands the role of textbooks, narrative reviews and similar resources in answering background questions. (1 mh, 1 rl, 1 jt, 1 dp, 1 rg, 1 cn) (1.0)
      i. Can define what a narrative review is and distinguish it from a systematic review (1 cn, 1 dp, 1 rg, 1 rl) (1.0)
      ii. Can cite the advantages and disadvantages of using textbooks and narrative reviews. (1 cn, 1 rg, 1 dp, 1 rl) (1.0)

Commentary: The strengths of a primary source are the strengths inherent in scientific, controlled experimentation. The weakness relate to distilling clinically useful generalities from inherently complex and variable investigative results. The strengths of a secondary source include its ability to offer a synthesis of data and information based on rigorous rules of evidence, and consensus practice recommendations based on evidence synthesis. The weaknesses relate to problems of accuracy of information portrayal because of content expert's "filter bias." (Submitted by Rich Gillette, PhD.)

---

* These are currently available tools that enable clinicians to remain up to date with new valid information that is relevant to patient care and is accessible while taking care of patients.

† Oxford Center for Evidence-Based Medicine. Levels of evidence and grades of recommendation <http://www.cebm.net/levels_of_evidence.asp>. Accessed 13 December 2004.

‡ Ebell MH, Siwek J, Weiss BD, Woolf SH, Susman J, Ewigman B, Bowman M. Strength of recommendation taxonomy (SORT) : a patient-centered approach to grading evidence in the medical literature. J Am Board Fam Pract 2004;17:59-67.

3.1. Is familiar with recommended "best" resources for finding evidence in a variety of circumstances. (1 jt, 1 mh, 1 dp, 1 rg, 1 cn, 1 rl) (1.0)

1. Recognizes a hierarchy of *information sources* and *services* (e.g., *Hanes pyramid* / *i.e. secondary vs. primary sources*). (1 mh, 1 jt, 1 dp, 1 rl, 1 rg, 1 cn) (1.0)
   a. Can define and describe the utility of *decision support system*s (i.e., computerized decision-making programs). (1 dp, 1 rg, 1 cn, 1 rl) (1.0)
   b. Can define and describe the utility of recommended information synopses (e.g., summaries of individual studies or systematic reviews). (1 dp, 1 rg, 1 cn, 1 rl) (1.0)
   c. Can define and describe the utility of recommended information syntheses (e.g., *clinical review articles*, *systematic reviews, meta-analysis*). (1 dp, 1 rg, 1 cn, 1 rl) (1.0)
   d. Understands the hierarchy of primary studies with respect to cause and effect (e.g., RCT vs. *cohort*). (1 dp, 2 rl, 1 rg, 1 cn) (1.3)
2. Can access appropriate sources and services based on the type of question posed (e.g., diagnosis, therapy, harm or prognosis). (1 mh, 1 jt, 1 dp, 1 rg, 1 cn, 1 rl) (1.0)
3. Can access appropriate sources and services based on the health care discipline being mined (i.e. primary care/general medicine, neuromusculoskeletal health care, complementary and alternative medicine (CAM) and manual therapy). (1 mh, 1 jt, 1 dp, 1 rg, 1 cn, 1 rl) (1.0)
4. Can describe the characteristics and content focus of a variety of evidence-based databases. (1 mh, 1 jt, 1 dp, 1 rl, 1 rg, 1 cn) (1.0)
   a. Free databases (e.g., PubMed, ICL, clinicaltrials.gov). (1 dp, 1 rg, 1 cn, 1 rl) (1.0)
   b. Proprietary databases (e.g., Medline, CINAHL, ICL). (1 dp, 1 rg, 1 cn, 1 rl) (1.0)
5. Can access "best" pre-filtered resources which have the greatest likelihood of being clinically useful.(1 mh, 1 jt, 1 dp, ? rg, 1 cn, 1 rl) (1.0)
6. Is familiar with and can access important evidence-based electronic sources of information. (1 mh, 1 jt, 1 dp, 1 rg, 1 cn, 1 rl) (1.0)
   a. For questions in the domain of primary care/general medicine (e.g., American Family Physician, (*AFP*) http://www.aafp.org/afp/, Cochrane Collaboration www.cochrane.com). (2 dp, 1 cn, 1 rg, 1 rl, 2 jt, 1 mh) (1.3)
   b. For questions in the domain of NMS health care. (1 dp, 2 rg, 1 cn, 1 rl, 2 jt) (1.4)
   c. For questions in the domain of CAM (e.g., National Center for Complementary and Alternative Medicine). (1 dp, 1 rg, 1 cn, 1 rl, 2 jt) (1.2)
   d. For questions in the domain of manual therapy. (2 dp, 2 rg, 1 cn, 1 rl, 2 jt) (1.6)
7. Is familiar with and can access important sources for evidence-based clinical guidelines. (1 dp, 1 rg, 1 cn, 1 rl) (1.0)
   a. For questions in the domain of primary care/general medicine (e.g., Canadian Task Force on the Periodic Health Care www.ctfphc.org, US Preventive Services Task Force www.uspstf.org,). (2 dp, 2 rg, 1 cn, 1 rl, 2 jt) (1.6)
   b. For questions in the domain of NMS health care (e.g., CCGPP). (1 dp, 1 rg, 1 cn, 1 rl, 2 jt) (1.2)
   c. For questions in the domain of CAM. (2 dp, 2 rg, 1 cn, 1 rl, 2 jt, 1 mh) (1.5)
   d. For questions in the domain of manual therapy (e.g., CCGPP, Canadian Practice Guidelines). (1 dp, 1 cn, 1 rl, 1 rg) (1.0)
8. Is familiar with and can access the important general sources for systematic reviews. (1 mh, 1 jt, 2 dp, 1 rg, 1 cn, 1 rl, 1 jt) (1.1)
   a. For questions in the domain of primary care/general medicine (e.g., Canadian Task Force on the Periodic Health Care, US Preventive Services Task Force, Cochrane Library). (1 dp, 1 rg, 1 cn, 1 rl) (1.0)
   b. For questions in the domain of NMS health care (e.g., Cochrane Library, ACP Journal club). (1 dp, 1 rl, 1 rg, 1 cn, 1 jt) (1.0)
   c. For questions in the domain of CAM (e.g., Cochrane Library). (1 dp, 1 rg, 1 cn, 1 rl, 1 jt) (1.0)
   d. For questions in the domain of manual therapy (e.g., Cochrane Library, Spine, JMPT). (2 dp, 2 rg, 1 cn, 1 rl, 1 mh, 1 jt) (1.3)
9. Is familiar with important journal sources for primary research articles and evidence-based review articles. (1 mh, 1 jt, 2 dp, 1 rg, 1 cn, 1 rl) (1.2)
   a. For questions in the domain of primary care/general medicine (e.g., Annals of Family Practice, Annals of Internal Medicine). (1 jt, 1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
   b. For questions in the domain of NMS health care (e.g., Spine, JMPT). (1 jt, 1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)

c. For questions in the domain of CAM (e.g., Alternative and Complementary Medicine). (1 jt, 1 dp, 1 rg, 1 cn, 1 r, 1 mh) (1.0)
d. For questions in the domain of manual therapy (e.g., Manual Therapy, JMPT). (1 jt, 1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)

10. Can determine and access best evidence-based textbooks. (1 mh, 1 jt, 1 dp, 1 rg, 1 cn, 1 rl) (1.0)
    a. Can utilize the criteria listed below to identify which type of textbook would be most relevant for the question being asked. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
        3.1.10.a.i.1. Based on foreground (e.g., CMDT, Mosby's 5 Minute Consult series) or background questions (e.g., Harrison's Principles of Internal Medicine). (1 rg, 1 dp, 1 cn, 1 rl, 1 mh, 1 jt) (1.0)
        3.1.10.a.i.2. Based on type of knowledge: signs and symptoms (e.g., Souza's Differential Diagnosis for the Chiropractor), specific conditions diagnosis in primary care (e.g., Harrison's Principles of Internal Medicine), orthopedic tests (e.g., Magee's Orthopedic Physical Assessment), physical examination (e.g., McGee's Evidence-Based Physical Diagnosis), and specialty issues (e.g., Liebenson's Rehabilitation of the Spine). (1 dp, 1 cn, 1 rg, 1 rl, 1 mh, 1 jt) (1.0)
        3.1.10.a.i.3. Based on domain of knowledge: CAM (e.g ?), manual therapy (e.g., Peterson's Chiropractic Technique Principles and Procedures, 2nd edition), NMS/orthopedics, or general medicine/health care. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh, 1 jt) (1.0)
    b. Can utilize specific criteria to assess a textbook relative to its quality and usefulness for evidenced-based information. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh, 1 jt) (1.0)
        3.1.10.b.i.1. How recent and how often the text is updated. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh, 1 jt) (1.0)
        3.1.10.b.i.2. Discussion of diagnostic strategies and processes. (2 dp, 2 rg, 2 rl, 1 cn, 1 mh) (1.6)
        3.1.10.b.i.3. Information on accuracy and reliability. (2 dp, 2 rg, 1 rl, 1 cn, 1 mh) (1.4)
        3.1.10.b.i.4. Accuracy of specific signs and symptoms provided. (2 dp, 2 rg, 1 rl, 1 cn, 1 mh) (1.4)
        3.1.10.b.i.5. References provided. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh, 1 jt) (1.0)
        3.1.10.b.i.6. Frequency of disease or clinical finding. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh, 1 jt) (1.0)
        3.1.10.b.i.7. Above categories are rated based on whether the concept is consistently explained and applied through the text along with specific examples. (2 dp, 2 rg, 2 rl, 2 cn, 1 mh, 1 jt) (1.7)

Commentary: The list of the attributes of a high quality ext book are from EMB notebook 10:October 2005.

3.2. Can design an effective search.
1. Can effectively use limiters in a variety of data bases (e.g., Clinical Queries)

3.3. Can conduct an effective and efficient search
1. Can modify searches to respond to search "feasts" and "famines"
2. Can quickly scan search results for currency, relevancy, and quality
3. Can scan abstracts for clues of relevancy and quality
4. Can navigate to full text using a variety of methods (e.g, using the A-Z list, linking directly from a data base, using inter-library loan)

3.4. Has the knowledge and skills necessary to coalesce, organize, store and previously searched health care information. (1 dp, 1 rg, 1 rl, 1 mh, 1 jt, 1 cn) (1.0)
1. Can generate and manage data bases of health care references and articles [?] (can demonstrate familiarity with a commercial product such as Refworks, Reference Manager, EndNote, or ProCite. (2 dp, 2 rg, 3 rl, 3 jt, 2 mh, 2 cn) (2.3)
2. Can generate a *critically appraised topic* (CAT) or other type of summary from a single source for later retrieval. (1 dp, 1 rl, 1 jt, 2 mh, 1 cn, 1 rg) (1.2)
3. Can synthesize evidence from a variety of resources into a coherent and balanced summary.

APPRAISE

STANDARD 4

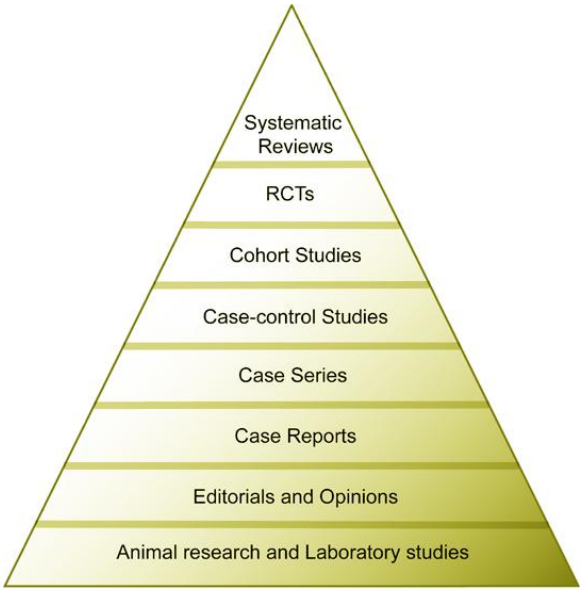
The EBP competent practitioner can critically appraise the validity and clinical significance of relevant evidence.

4. The EBP competent practitioner can critically appraise the validity and clinical significance of relevant evidence.

4.1. Understands the inherent strengths and weaknesses of different levels of evidence and can rate their quality. (1 dp, 1 mh, 1 rl, 1 rg, 1 cn) (1.0)

1. Can outline and define levels of evidence. (1 dp, 1 rg, 1 cn, 1 mh, 1 rl) (1.0)
2. Can identify and contrast the differences between narrative reviews and systematic reviews. (1 dp, 1 cn, 1 mh, 1 rl, 1 rg) (1.0)

Commentary:



Levels of Evidence

From the Centre for Evidence-Based Medicine, Oxford

For the most up-to-date levels of evidence, see http://www.cebm.net/levels_of_evidence.asp

Therapy/Prevention/Etiology/Harm:

1a: Systematic reviews (with homogeneity) of randomized controlled trials
1a-: Systematic review of randomized trials displaying worrisome heterogeneity
1b: Individual randomized controlled trials (with narrow confidence interval)
1b-: Individual randomized controlled trials (with a wide confidence interval)
1c: All or none randomized controlled trials
2a: Systematic reviews (with homogeneity) of cohort studies
2a-: Systematic reviews of cohort studies displaying worrisome heterogeneity
2b: Individual cohort study or low quality randomized controlled trials (<80% follow-up)
2b-: Individual cohort study or low quality randomized controlled trials (<80% follow-up / wide confidence interval)
2c: 'Outcomes' Research; ecological studies
3a: Systematic review (with homogeneity) of case-control studies
3a-: Systematic review of case-control studies with worrisome heterogeneity
3b: Individual case-control study
4: Case-series (and poor quality cohort and case-control studies)
5: Expert opinion without explicit critical appraisal, or based on physiology, bench research or 'first principles'

Diagnosis:

1a:	Systematic review (with homogeneity) of Level 1 diagnostic studies; or a clinical rule validated on a test set.
1a-:	Systematic review of Level 1 diagnostic studies displaying worrisome heterogeneity
1b:	Independent blind comparison of an appropriate spectrum of consecutive patients, all of whom have undergone both the diagnostic test and the reference standard; or a clinical decision rule not validated on a second set of patients
1c:	Absolute SpPins And SnNouts (An Absolute SpPin is a diagnostic finding whose Specificity is so high that a Positive result rules-in the diagnosis. An Absolute SnNout is a diagnostic finding whose Sensitivity is so high that a Negative result rules-out the diagnosis).
2a:	Systematic review (with homogeneity) of Level >2 diagnostic studies
2a-:	Systematic review of Level >2 diagnostic studies displaying worrisome heterogeneity
2b:	Any of: 1)independent blind or objective comparison; 2)study performed in a set of non-consecutive patients, or confined to a narrow spectrum of study individuals (or both) all of whom have undergone both the diagnostic test and the reference standard; 3) a diagnostic clinical rule not validated in a test set.
3a:	Systematic review (with homogeneity) of case-control studies
3a-:	Systematic review of case-control studies displaying worrisome heterogeneity
4:	Any of: 1) reference standard was unobjective, unblinded or not independent; 2) positive and negative tests were verified using separate reference standards; 3) study was performed in an inappropriate spectrum of patients.
5:	Expert opinion without explicit critical appraisal, or based on physiology, bench research or 'first principles'

Prognosis:

1a:	Systematic review (with homogeneity) of inception cohort studies; or a clinical rule validated on a test set.
1a-:	Systematic review of inception cohort studies displaying worrisome heterogeneity
1b:	Individual inception cohort study with > 80% follow-up; or a clinical rule not validated on a second set of patients
1c:	All or none case-series
2a:	Systematic review (with homogeneity) of either retrospective cohort studies or untreated control groups in RCTs.
2a-:	Systematic review of either retrospective cohort studies or untreated control groups in RCTs displaying worrisome heterogeneity
2b:	Retrospective cohort study or follow-up of untreated control patients in an RCT; or clinical rule not validated in a test set.
2c:	'Outcomes' research
4:	Case-series (and poor quality prognostic cohort studies)
5:	Expert opinion without explicit critical appraisal, or based on physiology, bench research or 'first principles'

Key to interpretation of practice guidelines

Agency for Healthcare Research and Quality:

A: There is good research-based evidence to support the recommendation.
B: There is fair research-based evidence to support the recommendation.
C: The recommendation is based on expert opinion and panel consensus.
X: There is evidence of harm from this intervention.

USPSTF Guide to Clinical Preventive Services:

A: There is good evidence to support the recommendation that the condition be specifically considered in a periodic health examination.
B: There is fair evidence to support the recommendation that the condition be specifically considered in a periodic health examination.
C: There is insufficient evidence to recommend for or against the inclusion of the condition in a periodic health examination, but recommendations may be made on other grounds.
D: There is fair evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination.
E: There is good evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination.

University of Michigan Practice Guideline:

A: Randomized controlled trials.
B: Controlled trials, no randomization.

C: Observational trials.

D: Opinion of the expert panel.

Other guidelines:

A: There is good research-based evidence to support the recommendation.

B: There is fair research-based evidence to support the recommendation.

C: The recommendation is based on expert opinion and panel consensus.

X: There is evidence that the intervention is harmful.

a. Can differentiate types of systematic reviews. (1 dp, 1 mh, 1 rl , 1 cn, 1 rg, 1 mh) (1.0)
   i. Knows the key characteristics of a meta-analysis (e.g., pooling of data from similar studies, use of formal quantitative analysis). (1 dp, 1 rg, 1 rl , 1 cn, 1 mh) (1.0)
   ii. Knows the key characteristics of a qualitative systematic review/ best-evidence synthesis (e.g., qualitative nature, often composed of studies too heterogeneous to pool for statistical meta-analysis). (1 dp, 1 rl , 1 cn, 1 rg, 1 mh) (1.0)

Commentary: Clinical/narrative review articles are probably best to get background information on a topic. In that regard they are like textbooks. Often one can find a review article that is more up to date than a text. These reviews are excellent resources for students. However, they are not rigorous enough to be the first choice of physicians when trying to get an analysis of the best evidence to help one make an important clinical decision in practice. Qualitative systematic reviews and meta-analyses carry more clout. On the other hand, because they tend to be very focused, they are not useful to provide an overview of a condition (i.e., etiology, diagnosis, treatment, etc).

One definition describes a systematic review as: "A review of a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyze data from studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyze and summarize the results of the included studies.

Another definition: A systematic review is: "An attempt to minimize the element of arbitrariness...by making explicit the review process, so that, in principle, another reviewer with access to the same resources could undertake the review and reach broadly the same conclusions. Dixon et al 1997:157" in Dawes, M Evidence Based Practice, 2005

"Three key features of such a review area strenuous effort to locate all original reports on the topic of interest critical evaluation of the reports

The following is useful resource material. "The meta-analysis of a number of small trials should be more generalizable to primary care practice populations than results from a single large trial P. 33."

Differences between Clinical/Narrative Review Articles, Qualitative Systematic Reviews, and Meta-analyses*

Feature	***Clinical Review***	***Qualitative Systematic Review***	***Meta-analysis***
Question	Often broad in scope	Often a focused clinical question	Usually a focused clinical question
Sources and search	Not usually specified, potentially biased	Comprehensive sources and explicit search strategy	Comprehensive sources and explicit search strategy
Selection	Not usually specified, potentially biased	Criterion-based selection, uniformly applied	Criterion-based selection, uniformly applied
Appraisal	Variable	Rigorous critical appraisal	Rigorous critical appraisal
Synthesis	Often a qualitative summary	Qualitative summary	Statistical analysis of pooled data
Potential method strength	Least strong	Strong	Strongest

Adapted from Table 1.1 in Cook D, Mulrow C, Haynes B. Synthesis of best evidence for clinical decisions. In: Mulrow C, Cook D, editors. Systematic reviews: synthesis of best evidence for health care decisions. Philadelphia: American College of Physicians; 1998. p. 5-12.

*All reviews are subject to systematic and random error, and the quality of a review depends on the extent to which scientific methods have been used to minimize error and bias. This table describes articles in which rigorous methods are applied to meet standards appropriate for the type of review

Curricular Suggestions: The strategy is to teach students to be able to perform rapid assessments as well as more detailed assessments. To accomplish this, checklists or instruments should be agreed upon to aide students in these two different approaches. [RL 10/2/06]

"The Oxman, et. al (1994) checklist is also presented here, following reports from some students that they find it more accessible when starting to critically appraise systematic reviews for the first time. Oxman et al (1994) break down their approach into three sections (Answers are typically, yes, no or cannot tell in Dawes, M Evidence Based Practice, 2005):

- are the results valid?
- if they are, what are the results?
- will the results help in my patient care?

Another instrument that is used as a guide for assessing a meta-analysis is QUOROM. It or a similar instrument might be used for in-depth assessments. [RL 9/28/06]

QUOROM Guidelines for Meta-Analyses and Systematic Reviews of RCTs*

TITLE	***Identify the study as a meta-analysis (or systematic review) of RCTs***
ABSTRACT	***Use the journal's structured format***
INTRODUCTION	***Present***
	• The clinical problem • The biological rationale for the intervention • The rationale for the review • An explicit statement of objectives which includes the study population, the condition of interest, the exposure or intervention, and the outcome(s) considered
SOURCES	***Describe***
	• The information sources in detail (e.g., databases, registers, personal files, experts, agencies, hand-searching) • Any restriction (years considered, publication status, language of publication)
STUDY SELECTION	***Describe***
	• Inclusion and exclusion criteria (defining population, intervention, main outcomes, and study design) • How clinical heterogeneity was assessed • Methods used for validity assessment • The criteria and process used for validity assessment (e.g., masked conditions, quality assessment) • The data abstraction process (e.g., completed independently, in duplicate) • Study characteristics and how clinical heterogeneity was assessed • The principal measures of effect (e.g., relative risk) • Method of combining results (statistical testing and confidence intervals) • Handling of missing data • How statistical heterogeneity was assessed • Rationale for any a-priori sensitivity and subgroup analyses
RESULTS	***Present***
	• A meta-analysis profile summarizing trial flow • Descriptive data for each trial (study design, participant characteristics, sample size, details of intervention, outcome definitions, length of follow-up) • Agreement on the selection and validity assessment • Simple summary results (for each treatment group in each trial, for each primary outcome) • Data needed to calculate effect sizes and confidence intervals in intention-to-treat analyses
DISCUSSION	***Discuss***
	• Key findings • Clinical inferences based on internal and external validity • The results in light of the totality of available evidence • Strengths and weaknesses • Potential biases in the review process (e.g., publication bias) • Future research agenda

*Modified from Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet 1999;354:1896–900.

3. Can evaluate the quality of narrative clinical review articles. (1 dp, 1 mh, 1 rg, 1 rl , 1 cn) (1.0)
   a. Can identify if an article has the characteristics of a higher quality clinical review. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
      i. Can determine if the review cites original research, not just other reviews. (1 dp, 1 rl , 1 rg, 1 cn, 1 mh) (1.0)
      ii. Can determine if it is primarily (but not necessarily exclusively) composed of the highest levels of evidence. (1 dp, 1 rl , 1 rg, 1 cn, 1 mh) (1.0)
      iii. Can determine if it cites latest studies and also landmark studies. (1 dp, 1 rg, 1 rl , 1 cn, 2 mh) (1.2)
      iv. Can determine if it cites peer-reviewed journals. (1 dp, 1 rl , 1 rg, 1 c, 1 mh) (1.0)
   b. Can identify and discuss potential weaknesses of a narrative clinical review. (1 dp, 1 rg, 1 mh, 1 rl , 1 cn) (1.0)
      i. Understands the potential for selection bias inherent in narrative reviews. (1 dp, 1 rg, 1 rl , 1 cn, 1 mh) (1.0)
      ii. Understands the limitations of an unsystematic search strategy. (1 dp, 2 rl , 1 rg, 1 cn, 1 mh) (1.2)
      iii. Understands the potential for a review to be influenced by funding, author or journal bias. (1 dp, 2 rl, 1 rg, 1 cn, 1 mh) (1.2)

Commentary: The following quotes can be used as resource material.

"The most common form of review article finds the author describing his/her approach to diagnosis and management using a few selected references. From an evidence-based perspective, this style of review article has some value in providing an expert's approach to a common problem, but the methodology is not rigorous enough to ensure that the conclusions represent an objective and systematic review of the current literature.

"Clinical review articles, or updates, selectively review the medical literature and incorporate the most important, most relevant research findings about disease management. (Siwek 2002) Like other types of evidence summaries, clinical review articles discuss a topic broadly. However, clinical reviews evaluate the literature less comprehensively and use less structured search methodologies than systematic reviews. Whereas meta-analyses statistically analyze pooled research data, clinical reviews organize research findings without using statistical analysis to draw conclusions. (Table 6-1).

"… The bottom line here is that clinical review articles are done on a much smaller scale than systematic reviews and meta-analyses, and for a number of reasons, they are more subject to bias. You have to be a smart shopper to be able to find the best review articles out there, and you have to know when they are appropriate and when they are not. p. 32

"Look for a reasonable list of references (about 20 to 30 for a limited topic, 50 to 60 for a broad topic) that are up to date and from reputable journals. Rather than rely on other review articles, a clinical review should cite high-quality original research studies. The highest levels of evidence available on a topic should be well represented; in particular, if many relevant randomized controlled trials (RTCs) and meta-analyses exist, the article should highlight the most important recent studies and the landmark studies with which readers should be familiar to understand the current literature. However, the selected evidence should not be restricted to RTCs and meta-analyses. In some cases, RTCs are not essential or are not yet available. In other cases, RTCs are not the best type of study to answer a clinical question; for example, the accuracy of a diagnostic test should be assessed in cross-sectional studies, and questions about prognosis are best addressed in follow-up studies of patients observed from an early stage of disease. (Sackett 1996) Where potentially harmful exposures are under study, RTCs are neither practical nor ethical, and observational study design is warranted. (McKibbon 2002) p. 34

"Some evidence-based clinical review articles describe their search methodology in a "data sources" section. At minimum, this section should list (1) databases searched, (2) dates of articles searched, (3) medical subject heading search terms used, and (4) inclusions and exclusions of studies (eg, RCTs only, case reports excluded). p. 35

"Event the best clinical reviews have inherent limitations. Keep the following points in mind as you read this type of article:

"Selection Bias

The average literature search on a common disease condition yields thousands of articles of widely varying clinical relevance. But there is more. "Hidden" sources of evidence include unpublished studies, abstracts, articles in foreign languages, articles that are inaccessible through common search methods, and personal communication. Excluding any portion of this data from a search will impose bias. For example, it is well known that positive research results are more likely than negative ones to enhance the career of the investigator and to be submitted for publication (publication bias). Hence, excluding unpublished studies in a clinical review will result in a higher representation of

articles with positive results. (Kelch 2002) Because traditional clinical review articles are usually limited to published articles that are easily accessible through common search venues, the pool of data that authors search will always be subject to a degree of bias.

"Unsystematic Search Strategy

To prepare a systematic review that answers a clinical question, a large team of experienced reviewers (e.g., US Preventative Services Task Force, Cochrane Collaboration) performs a comprehensive search of the relevant literature for many sources. Reviewers then sort through a large portion of the body of evidence using strict selection criteria and appraisal methods before formulating clinical recommendations.

"In contrast, traditional clinical review articles are written by one author or a small team of authors. Authors usually search a number of reliable sources of evidence and choose 20 to 40 articles based on informal, subjective criteria. The quality of the review article depends on the authors' skill in choosing the highest-quality evidence, their thoroughness, and their ability to accurately interpret and translate studies into practical recommendations for readers. Even a well-done traditional review is based on a limited collection of data, without a guarantee of being the highest-quality data available on a topic."

Publication Restrictions

"Authors of traditional review articles are subject to many editorial constraints on structure and, to some degree, emphasis of their article. For example, a primary care journal may encourage citation of research studies involving a primary care population rather than patients referred to a specialty clinic. Length is usually the most restrictive factor; it is impossible to thoroughly summarize the literature on some important clinical topics within the average 2,000-word limit for a clinical review article."

References

Kelch RP. Maintaining the public trust in clinical research. N Engl J Med 2002;346:285-7.

McKibbon A, Hunt D, Richardson WS, et al. Finding the evidence. In: Guyatt G, Rennie D, editors. User's guides to the medical literature. Chicago: AMA Press; 2002. p. 13-47.

Sackett DL, Rosenberg WMC, Gray JAM, et al. Evidence-based medicine: what it is and what it isn't. BMJ 1996; 312:71-2.

Siwek J, Gourlay M, Slawson DC, Shaughnessy AF. How to write an evidence-based clinical review article. Am Fam Physician 2002;65:251-8.

4. Can evaluate the quality of systematic reviews. (1 dp, 1 rg, 1 rl , 1 cn, 1 mh) (1.0)
   a. Can identify if an article has the characteristics of a higher quality systematic review. (1 dp, 2 mh, 1 rl , 1 rg, 1 cn) (1.2)
      i. Can determine if the methodology has adequate transparency (e.g., citation of search techniques, data synthesis, conflicts of interest). (1 dp, 1 rg, 1 rl , 1 cn, 1 mh) (1.0)
      ii. Can determine if the types of studies selected were appropriately matched to the type of question asked in the realms of diagnosis, harm, therapy or prognosis (e.g., RCTs are preferred for questions of therapy). (1 dp, 1 rg, 1 rl , 1 cn, 1 mh) (1.0)
      iii. Can determine if there was a comprehensive and detailed search for relevant studies (e.g., appropriate key words and data bases, a wide range of sources including personal communications with researchers, discussions at scientific meetings, or other less formal resources). (1 dp, 1 rl , 1 cn, 1 mh, 1 rg) (1.0)
      iv. Can determine if all of the individual studies included were assessed for methodological quality. (1 dp, 1 rl , 1 rg, 1 cn, 2 mh) (1.2)
      v. Can determine if the author addresses whether the individual studies were sufficiently similar for meaningful synthesis. (1 dp, 1 rg, 1 rl , 1 cn, 1 mh) (1.0)
      vi. Can determine if there is any significant funding, author and journal bias. (2 dp, 2 rg, 2 rl , 1 cn, 2 mh) (1.8)

Commentary: Background material comes from Straus.

Is the evidence from the systematic review valid? (Straus 2005, Table 5.9)

1. Is this a systematic review of randomized trials?
2. Does it describe a comprehensive and detailed search for relevant trials?
3. Were the individual studies assessed for validity?

A less frequent point:

4. Were individual patient data (or aggregate data) used in the analysis?

Is the valid evidence from this systematic review important? (Straus 2005, Table 5.10)

1. Are the results consistent across studies?
2. What is the magnitude of the treatment effect?
3. How precise is the treatment effect?

What are the results?

"What this does illustrate is that, like understanding a systematic review, critically appraising these reviews is not an exact science, but there are many subjective decisions along the way. Just like understanding a systematic review, making explicit your decisions in the critical appraisal is therefore very important."
in Dawes, M Evidence Based Practice, 2005

"Crombie and McQuay (1998) point out that it is important not to overstate potential limitations, as systematic reviews are a major advance on both the traditional review and the use of selected evidence that went before. However, there are some limitations. Crombie and McQuay (1998) raise the possibility that sometimes a review may mislead. They argue this can occur when the quality of the review is poor or when publication bias occurs. Thus a review may overestimate the effectiveness of the treatment/intervention. If the results of a review (which may combine several small trials) are compared with those of a large RCT, the results may not be the same. For example, Cappelleri et al (1996) found over 80% of 79 reviews agreed with large trials. LeLorier et al. (1997) found a 65% agreement, although this article has been criticized by Naylor and Davey Smith (1998)" in Dawes, M Evidence Based Practice, 2005 (1 dp, 1 rl, 1 rg, 1 cn) (1.0)

b. Can evaluate the usefulness of a systematic review. (1 dp, 1 rg, 1 mh, 1 rl , 1 cn) (1.0)
   i. Can determine whether the evidence is of sufficient quality. (1 dp, 1 rg, 1 rl , 1 cn, 1 mh) (1.0)
   ii. Can determine if there is a consistency of results across studies. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
   iii. Can determine if the evidence was of sufficient magnitude and precision to impact practice. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)

Commentary: Background material: Is this valid and important evidence from a systematic review applicable to our patient? (Straus 2005 Table 5.14)

1. Is our patient so different from those in the study that its results cannot apply?
2. Is the treatment feasible in our setting?
3. What are our patient's potential benefits and harms from the therapy?
4. What are our patient's values and expectations for both the outcome we are trying to prevent and the adverse effects we may cause?

c. Can summarize the inherent weaknesses and controversy pertaining to systematic reviews. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
   i. Understands that systematic reviews of the same pool of evidence can reach different conclusions (based on quantitative vs. qualitative methods, the degree of consensus among the reviewers, different quality scales, rules for inclusion, or rules of evidence). (1 dp, 2 rg, 2 mh, 2 rl, 2 cn) (1.8)
   ii. Understands the importance of using appropriate quality scales based on type of research (e.g., ratings of physical medicine studies may be affected by using quality scales more appropriate for medicine).(2 dp, 2 rg, 2 rl , 2 cn, 1 mh) (1.8)
   iii. Understands that patients, comparison groups, outcomes, and follow-up time points from various studies may not be similar enough to be pooled for the quantitative methodology used in meta-analysis. (1 cn, 2 mh, 1 dp, 1 rl, 1 rg) (1.2)

5. Can evaluate the quality of clinical practice guidelines. (1 dp, 1 rg, 1 mh, 1 cn, 2 mh, 1 rl) (1.2)
   a. Can determine whether a guideline includes a comprehensive, reproducible literature review current within a reasonable timeframe (recommendations range from 1-3 years). (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
   b. Can determine whether individual recommendations are both tagged by the level of evidence (based on type, quality, and quantity) and linked to specific citations. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
   c. Can assess the relative quality based on the transparency of the methodology, the make up and qualifications of the authors or consensus group, the consensus process, and the opinions offered in any appended minority report. (1 dp, 1 rl, 1 cn, 1 rg, 3 mh) (1.4)

Curricular Suggestions: The strategy is to teach students to be able to perform rapid assessments as well as more detailed assessments. To accomplish this, checklists or instruments should be agreed upon to aide students in these two different approaches. [RL 10/2/06]

Below are a couple of ideas for brief assessments.

- Who produced the guideline?
- Is the guideline relevant to family and general practice?
- What was the approach to obtaining evidence to support the guidelines? p. 81

Guides for deciding whether a guideline is valid (Straus 2005, Table 5.23)

1. Did its developers carry out a comprehensive, reproducible literature review within the past 12 months?
2. Is each of its recommendations both tagged by the level of evidence upon which it is based and linked to a specific citation?

Another instrument for more in-depth assessments that is used as a guide for assessing a clinical guideline is AGREE [RL 9/28/06]

6. Can identify and evaluate the quality of clinical decision making tools. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
   a. Can identify decision-making instruments and formats such as clinical decision-making rules (e.g., Ottawa rules for acute ankle radiographs), algorithms/decision-making trees, and quantitative clinical decision analyses. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)

Commentary: At the time of the writing of this guide, some of the clinical decision making rules that should be discussed in the curriculum include the Well's criteria for DVP, the Ottawa rules of ankle and knee radiographs, the rules regarding applying manipulation to low back cases. [RL]

   b. Can explain in general the strengths and weaknesses of diagnostic and treatment decision-making trees/algorithms. (2 dp, 2 rg, 1 rl, 1 cn, 3 mh) (1.8)
   c. Can explain in general the strengths and weaknesses of clinical decision-making rules. (2 dp, 2 rg, 1 rl, 1 cn, 2 mh) (1.6)
   d. Can define and discuss in general quantitative clinical decision analysis. (3 dp, 3 mh, 3 rg, 3 rl, 3 cn) (3.0)

Commentary: 6b, c, and d all have inserted the words "in general." The committee felt that in most cases the ability of a student finishing the program to assess an individual decision-making tree, rule, or CDA would be limited at best, students should be able to at least grasp the generic strengths and weaknesses of each of these types of tools. [RL 7/30/07] For those interested in more, below is a checklist for evaluating a CDA.

Is this valid evidence from a CDA important? (Straus 2005 Table 5.16)

1. Did one course of action lead to clinically important gains?
2. Was the same course of action preferred despite clinically sensible changes in probabilities and utilities?

   e. Can assess the quality of decision-making tools in general. (1 dp, 1 rg, 2 rl, 3 mh, 1 cn) (1.6)
      i. Considers the level of content expertise of the authors. (2 dp, 2 rg, 2 cn, 2 rl, 3 mh) (2.2)
      ii. Considers the rigor of the methodology. (2 dp, 2 rg, 2 cn, 1 rl, 2 mh) (1.8)
      iii. Considers the levels of evidence utilized. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
      iv. Considers if it has verified clinical efficacy/validity in actual clinical trials. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
      v. Considers the ease of use. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
      vi. Considers the Intended end user (e.g., chiropractor specifically, manual therapist, medical specialist). (1 dp, 2 rg, 2 cn, 1 rl, 1 mh) (1.4)
      vii. Considers whether it includes significant diagnostic and therapeutic alternatives. (2 dp, 2 rg, 2 cn, 2 rl, 2 mh) (2.0)
      viii. Considers whether each branch of a quantitative-based decision-making tree contains valid and credible outcome probabilities (leading to a particular result). (3 dp, 2 cn, 3 rg, 3 rl, 3 mh) (2.8)

ix. Considers whether each branch of a quantitative-based decision-making tree contains valid and credibly assigned weightings of clinical utility (based on an estimation of the risk-benefit impact on the patient). (3 dp, 3 rg, 3 rl, 3 mh, 2 cn) (2.8)
x. Considers whether the gains associated with one course of action opposed to another are clinically important enough to justify its application. (1 dp, 1 rg, 1 cn, 2 rl, 1 mh) (1.2)

 **Commentary:** There was considerable disagreement in the committee about the value of considering the level of expertise of an author when assessing the quality of expert opinion. MH and RG felt that it was essentially impossible to determine an author's real expertise so identifying indicators was useless (degrees, publication history, and standing in the community may all be misleading). Others on the committee felt that although it should always be born in mind that these indicators can be misleading, nonetheless degrees/ training, background or affiliation may be useful in initially choosing among expert opinions to consider and that this was important because many times the highest level of evidence available was only expert opinion. [RL 7/30/07]

Is this valid and important evidence from a CDA applicable to our patient? (Straus 2005 Table 5.17)
1. Do the probabilities in this CDA apply to our patient?
2. Can our patient state his/her utilities in a stable, usable form?

7. Can evaluate the clinical applicability of expert opinion. (1 dp, 1 rg, 1 mh, 1 cn, 1 rl) (1.0)
   a. Can assess the expert's content expertise (based on credentials, publications, frequency of being cited, etc.) and EBP competence (e.g., there is reason to believe the personal clinical opinion is offered within the context of current evidence). (1 cn, 1 dp, 1 rg, 1 rl, 3 mh) (1.4)
   b. Considers whether the expert opinion might be generalizable to other patient populations and clinical environments outside of the expert's own clinical populations. (1 cn, 1 rg, 1 dp, 1 rl, 3 mh) (1.4)
   c. Appreciates that opinions may be highly variable even among equally qualified experts. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)

 **Commentary:** "This could be a specialist who is well versed and experienced in the area of interest. This experience can be used to interpret confusing diagnostic findings in the care of a difficult patient with an atypical or unusual disease.

"A content expert may also be able to pass on clinical pearls that can fill in gaps that are not covered by current outcomes-based research. This is where years of contact with patients with a particular condition can generate anecdotes that can help guide decisions in diagnosis.

"Do not assume that an expert is skilled in evaluating medical research just because of the "expert status." Many may not be any better than you at determining the validity of research findings. Techniques on how to critically evaluate the medical literature are just beginning to be taught in clinical training programs. Few currently practicing clinicians have had the luxury of benefiting from this "information age" expertise. "

Patient Selection
"An expert's experientially based knowledge is often developed through contact with a highly selected patient population, and this may not apply as well to the general population or to the population that a primary care physician sees in the office. …"

Conflict of Interest
"There is also the potential for conflict of interest. Treatment recommendations are frequently biased by a physician's training and source of income. When evaluating recommended treatment for patients with upper gastrointestinal bleeding, Chalmers found that surgeons are more likely to recommend surgical approaches whereas internists are more likely to recommend more conservative management. (Chalmers 1982) Like the old phrase says, "Never ask a barber if you need a haircut!" p. 48"

Variability
"Another concern in evaluating information from a content expert is inherent in human nature: variability. Several studies have documented high rates of both interobserver (not agreeing with others) and intraobserver (not agreeing with oneself when presented with the same information at a different time) variability. For example, one study found that radiologists, when given the same radiographs, disagreed with each other 29% of the time and disagreed with their own earlier interpretations in about 20% of cases. (Garland 1959) A study evaluating eight pathologists who are experts in the diagnosis of melanoma found a lack of agreement in 62% of cases. Using the most extreme comparison, two of these well-respected experts disagreed on whether the specimens were benign or malignant in one-third of the cases! No one mention this to a lawyer! ("Pathology as art appreciation…" Bandolier 2002) p. 49"

Improper Generalization

"There is often a tendency for clinicians to develop general rules out of a patient-specific recommendation made by a specialist. For example, you might think, "The last time in this situation, the cardiologist recommended amlodipine, so I'll use it here again." Most likely, the expert did not think that **[he or]** she was giving a wide-ranging endorsement of a particular therapy when answering your specific question."

"As a result, experts may value their personal experience or beliefs on a topic over more recent evidence that comes from outcomes-based studies, a situation know as "reverse gullibility."

References

Chalmers TC. Informed consent, clinical research and the practice of medicine. Trans Am Clin Climatol Assoc 1982; 94:204-12.

Garland LH. Studies on the accuracy of diagnostic procedures. AJR Am J Roentgenol 1959;82:25-38.

Pathology as art appreciation: melanoma diagnosis. Bandolier [Serial online] 1997;37-2. http://www.jr2.ox.ac.uk/bandolier/band37/b37-2/html (accessed Apr 9, 2002).

8. Can evaluate the clinical applicability of consensus statements based on practitioner surveys. (2 dp, 2 rg, 2 cn, 2 mh, 2 rl) (2.0)
   a. Can describe a Delphi process. (2 dp, 2 rg, 2 rl, 2 cn, 2 mh) (2.0)
   b. Can articulate the limitations of such methodologies in terms of validity and usefulness. (2 dp, 2 rl, 2 rg, 2 cn, 1 mh) (1.8)
   c. Can articulate the role of surveys in documenting common practice behaviors. (2 dp, 2 rl, 2 cn, 2 rg, 1 mh) (1.8)

4.2. Demonstrate a basic conceptual understanding of biostatistics as they apply to EBP. (1.0)

1. Demonstrate a basic understanding of the role and importance of statistical analysis in the generation, interpreting, and reporting of research results. (1.0)

   Teaching tip: The primary emphasis here (and for these learning objectives) should be on being able to read and interpret what the researchers are trying to convey in a particular article when communicating in the language of statistics (for example, how good is a treatment, how accurate a diagnostic test, how potent a risk factor, etc) as well as how accurate is the data itself (e.g., expressed in terms of precision, confidence intervals, P value, etc) of secondary interest is the ability to read and have a basic understanding of raw data when displayed in charts, graphs, forest plots; of tertiary interest is to be able to make a reasonable match of the statistical test to the type of data presented in the paper.
2. Recognize the terms biostatistics and epidemiology. (1.2)

   Teaching tip: Biostatistics is the application of statistics to biological situations. Epidemiology is the study of epidemics. The study of epidemics has motivated much of the funding for the development of biostatistics methodology in medical applications, thus many medical researchers are trained in epidemiology so they may learn how to conduct research properly.
3. Distinguish population parameters from descriptive statistics (1.4) and descriptive statistics from inferential statistics (i.e., population estimates). (1.0)

   Teaching tip: A population parameter is the actual measurement of some aspect of an entire population. For example, if you wanted to know the mortality rate in a given year in the state of Oregon, you could simply count the deaths and you would not need a statistical estimate (this is the actual population parameter). If, on the other hand, you wanted to count the deaths of those people who had never seen a chiropractor and compare to those who had – this is not data that the state regularly collects – you would have to collect information on a subsample and make an *estimate* that you could then generalize to the larger population. A statistic gives an estimate of the real population parameter. If you have the population parameter, you don't need the statistic. Depending on the selection of the subpopulation, the type of statistical analysis selected, and the assumptions that the results are based on, statistics can either closely mirror the actual parameter in the population or give a false picture of the real situation. This is why having a basic understanding of the role of statistics in clinical research is so critical.

4. Demonstrate a basic knowledge of how data can be distributed or shaped (when graphically displayed).
   a. Define variability (1.0) and related terms (i.e., dispersion (1.0) standard deviation (1.0) interquartile range (2.2) and variance (3.0)).

      
      Teaching tip: Students should have a general understanding of what interquartile range means and be given a couple of examples of how it is used in the literature.

   b. Recognize normal distribution (1.0) (AKA Gaussian distribution/ bell shaped curve).
   c. Recognize skewed distribution. (1.0)
5. Define descriptors of central tendency: mean, (1.0) median (1.0) and mode. (2.0)

   
   Teaching tip: The mean is the average, the median is the middle result, and the mode is the most common result in a set of values. Students can get practice understanding these descriptive statistics through grade distributions or clinical examples.
6. Recognize the difference between categorical (e.g. nominal, ordinal) and continuous (e.g., interval, ratio) data. (1.0)

   
   Teaching tip: Students should be able to look at examples of data, correctly characterize the type, and select a reasonable choice of statistical test from a table of options. The goal here is primarily to acquaint them with significance that certain types of statistical formula must be used with certain types of data.
7. Explain the difference between sampling and randomization as it applies to a study design. (1.0)
   a. Define *sampling* (1.0)
   b. Define *sample mean* (central tendency) (1.0)[§]
   c. Define *random error* (1.0)
   d. Define variability (e.g., standard error). (1.2)

      
      Teaching tip: Students should understand the difference between the sampling step of a study and the randomization step; they should understand that they should look at how the initial pool of potential subjects is gathered for a study, that since the reader is going to infer that the sample that is chosen represents a larger group of people. That s/he needs to judge whether there is anything importantly different about the group (examples should be offered: patients in an exercise study may be more motivated than usual patients if they were allowed to self select from a newspaper advertisement for the study**).** Learners should understand that since multiple subjects and results are pooled, the statistical analysis often centers around what the central tendency of the group was, and that there are multiple ways of capturing that information
8. Use the concepts of precision and point estimate in interpreting research results. (1.0)
   a. Recognize a point estimate. (1.0)
   b. Define the precision of a point estimate using a standard error or confidence interval. (1.0)

      
      Teaching tip: Students should understand the concept of a point estimate and that there are different ways to measure and express how precise that estimate actually is. Students should understand how a reported confidence interval or standard error sheds additional light on where the true value of a reported treatment effect, test validity, prognostic indicator, relative risk, etc. actually lies.
   c. Use confidence intervals or standard error in interpreting the precision of research results. (1.0)
9. Recognize common ways used to display data in charts and graphs (1.0)
   a. Read a scatter plot bar graph (1.0) a line graph (1.0), a forest plot (1.0), a box plot (1.0), a histogram (1.5), an ROC curve (2.0) and a survival plot (2.0).

[§] Indicate most common tests to see in the literature.

b. Recognize the difference between a standard error bar (precision) versus a standard deviation bar (estimate of variability in the population) when presented in a plot (2.0)

10. Use the concept of statistical significance to better understand the results of a study. (1.0)
    a. Define the concept of *P values* (i.e., the tolerable amount of chance intrusion) (1.0)

Teaching tip: Students should be able to read a P value in a study and understand what it is saying about the intrusion of chance on the results. They should be able to read two P values and recognize which one is statistically more significant. They should be able to distinguish the meaning of P values from confidence intervals.

    b. Recognize that an "acceptable" level of probability/chance error is set before the study begins and that it is usually set ≤ .05 in clinical trials. (1.0)

Teaching tip: Students should have a general sense of how and why the .05 number has been set. A more advanced understanding would be to relate p value to balancing the risk of an false positive or false negative result from the study (i.e., an alpha or beta error.

    c. Demonstrate simple ways to estimate whether or not sample size was adequate in a particular study (based on the concepts of p values, confidence intervals, and power).
    d. Recognize some of the basic concepts associated with the *power* of a study (1.0)
        i. Define power as the probability of a study to detect a statistically significant difference between groups when there really is a difference in the study population. (1.2)
        ii. Recognize that a study is too small if the power to detect a clinically meaningful benefit is less than 80% (in studies with negative results). (1.0)

Commentary: Studies are designed to have an 80% or 90% probability of being able to detect a clinically important difference between groups. [MH 8/10/07]

11. Demonstrate familiarity with a variety of descriptive and inferential statistics.
    a. Define common descriptive statistics including mean (1.0), median (1.2), mode (1.6), standard deviation (1.0), standard error*(1.0), odds ratio*(1.0), relative risk* (1.0), and hazard ratio (1.4).
    b. Recognize a variety of methods to compare groups statistically (inferential statistics). (1.0)

Teaching tip: Students need to understand that different statistical tests are used depending on characteristics of data. They should be able recognize the broad category of data that is being presented and, consulting a table, be able to recognize tests that are commonly used within a particular research article.

    c. Recognize Chi-square.* (1.4)
    d. Recognize T-test.* (1.5)
    e. Recognize non-parametric tests: Wilcoxon, Mann-Whitney, Kruskal-Wallis, Friedman's, median, and sign tests. (2.4)
    f. Recognize post hoc tests. (1.8)

Teaching tip: Students should understand that post hoc calculations are useful for finding potentially meaningful information from which to generate new hypothesis and new experiments. It should be explained why ironclad conclusions cannot be made from post hoc calculations. As a general principle, they are not as trustworthy as those planned for and used in the planning of the research project.

    g. Recognize analysis of variance (ANOVA)*. (1.4)
    h. Recognize analysis of covariance (ANCOVA)*. (1.4)
    i. Recognize other tests, that like ANCOVA, correct for baseline differences between groups: regression, logistic regression, general linear models, generalized linear models, mixed effects models, and generalized estimating equations; proportional hazards models and Cox regression (time to event analysis). (2.4)
    j. Recognize common measures of correlation. (1.6)

i. Recognize Pearson's correlation coefficient (Pearson's r)*. (1.6)
ii. Recognize Spearman's rho. (2.0)

k. Define and demonstrate a basic understanding of regression analysis used for the purpose of prediction. (1.4)

Teaching tip: Students should have a recognition level understanding of this term and are not require to be exposed to the statistical formula or how it is derived. They are expected to know that when they a see this referenced, it refers to a statistical approach in which multiple variables are assessed to see if any of them have an association with an outcome. Students should be provided with concrete examples to illustrate how and when they will encounter this term in the literature: such as studies which identify independent risk factors for a condition (such as heart disease or low back pain), predictors of treatment outcomes in Clinical Prediction Rule studies (such as factors are most likely to affect the outcome of a particular treatment).

i. Recognize linear regression*. (1.6) Recognize multiple regression. (1.8)
ii. Recognize logistic regression. (2.0)

l. Recognize if treatment and control groups are similar at baseline in terms of important prognostic predictor variables or, if not, the predictor variables are adjusted for in the analysis. (1.0)

m. Recognize if analysis of covariance (ANCOVA) or equivalent (including general linear models or regression) was conducted. (2.0)

Teaching tip: The basic concept here is that when patient cohorts are compared, they may not be a perfect match in terms of their characteristics. In such cases, the consumer of the study should see if any baseline differences that were deemed to be clinically important were adjusted for in the statistical analysis.

4.3. Understands the design and hierarchy of different types of primary studies along with their inherent strengths and weaknesses. (1 mh, 1 jt, 1 cn, 1 dp, 1 rg, 1 rl) (1.0)

1. Can demonstrate a basic understanding of hypothesis testing. (2 mh, 2 dp, 3 rg, 2 cn, 2 rl) (2.2)
   a. Can explain the terms research hypothesis (alternative hypothesis, H1) (2 mh, 2 dp, 3 rg, 2 cn, 3 rl) (2.4) and the Null hypothesis (Ho). (2 mh, 2 dp, 3 rg, 2 cn, 2 rl) (2.2)
   b. Understands the basic difference between a Type I/alpha error (the probability of incorrectly rejecting the null hypothesis) and a Type II/beta error (the probability of incorrectly accepting the null hypothesis). (2 mh, 2 dp, 3 rg, 2 cn, 3 rl) (2.4)
2. Can explain the differences in design and methodology of various types of primary studies. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
   a. Can define and differentiate prospective vs. retrospective, observational vs. experimental, randomized vs. non-randomized comparisons (quasi-experimental), between subjects (nomothetic) vs. within subject (idiographic), and qualitative vs. quantitative studies. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
   b. Can define and explain basic terminology used in research studies. (1 rl, 1 rg, 1 cn, 1 mh, 1 dp) (1.0)
      i. Can define basic terms and concepts used in RCTs including intervention/treatment group vs. control group, sham treatment, nonspecific treatment effect and placebo effect. (1 rg, 1 rl, 1 cn, 1 mh, 1 dp) (1.0)
      ii. Can define basic terms and concepts regarding participants in a research study to include population, target population, sample (including random and nonrandom), and cohort. (1 rg, 1 rl, 1 cn, 1 mh, 1 dp) (1.0)
      iii. Can recognize if appropriate randomization occurred in a study, based on method (e.g., sealed envelopes, computer generated, and coin flip) and type (e.g., simple, block, stratified, and design adaptive). (2 rl, 2 cn, 2 rg, 2 mh, 2 dp) (2.0)
      iv. Can explain the need for concealing the study group prior to allocation (i.e., to prevent selection bias). (1 mh, 1 rl, 1 cn, 1 rg, 1 dp) (1.0)
   c. Can define and describe a randomized controlled trial (RCT). (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
      i. Can differentiate pragmatic from explanatory (fastidious) trials. (1 dp, 2 cn, 2 rg, 1 rl, 1 mh) (1.4)
      ii. Can differentiate placebo-controlled vs. comparison trials. (1 dp, 1 cn, 1 rl, 2 rg, 1 mh) (1.2)
      iii. Can define and compare crossover, single-blind, double-blind, triple blind, and assessor-blind randomized controlled trials. (2 cn, 1 mh, 1 dp, 1 rl, 2 rg) (1.4)

d. Can cite and discuss the strengths and weaknesses inherent in the design of RCTs. (1 cn, 1 mh, 1 rl, 1 rg, 1 dp) (1.0)
    i. Can discuss the following advantages of RCTs relative to other study designs: able to establish causality, able to diminish the effects of random chance, and the potential offers more trustworthy data. (1 rl, 1 cn, 1 mh, 1 dp, 1 rg) (1.0)
    ii. Can discuss the following limitations of RCTs: too difficult or unethical to design for some questions, possible problems with generalizability to practice (particularly for explanatory trials). (1 rl, 1 cn, 1 mh, 1 dp, 1 rg) (1.0)
e. Can cite and discuss the strengths and weaknesses inherent in nonrandomized comparison studies. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
    i. Can explain a variety of design strengths including usefulness for large practice-based studies, useful for generating hypotheses, better at accumulating large amounts of data than an RCT, results are more generalizable than those of RCTs. (1 cn, 1 rg, 1 rl, 1 mh, 1 dp) (1.0)
    ii. Appreciates the limitations of this design including the data are less reliable (trustworthy) than that of an RCT and are limited by lack of blinding, lack of randomization, and inherent susceptibility to selection bias. (1 cn, 1 rg, 1 rl, 1 mh, 1 dp) (1.0)
f. Can describe a variety of observational studies and discuss their inherent strengths and weaknesses. (1 cn, 1 rl, 1 mh, 1 rg, 1 dp) (1.0)
    i. Can cite the differences between a cohort design, a case-control design, and a cross-sectional design. (1 dp, 1 cn, 1 rg, 1 mh, 1 rl) (1.0)
    ii. Understands that they are considered to be the strongest design after RCTs and can cite examples when they would be more appropriate than an RCT (e.g., when an RCT is not possible or advisable due to ethical considerations). [Guyatt 1] (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    iii. Understands that observational studies have a tendency to overestimate intervention effects compared to an RCT. [Guyatt 1] (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
g. Can cite and discuss the inherent strengths and weaknesses of a cohort design (e.g., confounding variables may not be controlled). (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
    i. Can discuss the following advantages of a cohort design relative to other study designs: ability to identify large group trends, better reflects actual practice environment, may be a more ethical design than an RCT for some questions of harm, has the potential to identify cause and effect relationships suitable for further research. (1 rl, 1 cn, 1 mh 1 dp, 1 rg) (1.0)
    ii. Can discuss the following limitations of the cohort design: cannot establish casuality, lack of randomization increases the possibility of results being influenced by a variety of confounders. (1 rl, 1 cn, 1 mh, 1 dp, 1 rg) (1.0)
h. Can cite and discuss the inherent strengths and weaknesses of a case-control design. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
    i. Can explain their usefulness in identifying potential causes of rare diseases. (1 dp, 1 rl, 1 rg, 1 mh, 1 cn) (1.0)
    ii. Can cite design difficulties such as finding appropriately matched controls, establishing temporal linkages from the past (e.g., recall bias) and the inability to control for other confounding biases and causal factors. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
i. Can cite and discuss the inherent strengths and weaknesses of cross-sectional studies. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
    i. Can explain the problems of exposure. (1 dp, 2 rg, 2 rl, 2 cn, 2 mh) (1.8)
    ii. Can explain the potential effect of "recall bias." (1 dp, 1 rg, 2 rl, 1 cn, 1 mh) (1.2)
    iii. Can explain the difference between the association/correlation identified in cross-sectional studies compared to questions of direct causation. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
    iv. Understands that uncontrolled confounders may be present. (1 dp, 1 rg, 2 rl, 1 mh, 1 cn) (1.2)
j. Can define the role and inherent weaknesses of a case series design. (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0)
    i. Understands that they are principally useful for hypothesis generation to prompt further research. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
    ii. Understands their lack of control groups introduces many potential confounders. (rg) (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
k. Can define the role and inherent weaknesses of case studies/case reports. (1 dp, 1 rl, 1 cn, 1 rg) (1.0)

i. Can describe their usefulness for hypothesis generation and to share unique observations with the profession. (1 dp, 1 rg, 1 rl, 1 mh, 1 cn) (1.0)
ii. Understands their lack of control groups introduces many potential confounders. (rg) (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
iii. Understands findings are isolated to a single patient and are not generalizable. (1 dp, 1 rg, 1 rl, 1 mh, 1 cn) (1.0)

Commentary: The following thoughts may be helpful from Linda L. Isaacs, MD, Evaluating Anecdotes and Case Reports, Alternative Therapies, Mar/Apr 2007, Vol 13, No 2

"As one author puts it, 'The term anecdotal evidence' connotes secondhand or poorly documented fact and should not be confused with case studies of individual patients that involve careful observation and recording of detail.'" (Doyle)

"Anecdotes and case reports cannot be used to definitively prove a therapy is effective. But case reports cannot be dismissed entirely. As a recent article stated, 'Case reports and series have a high sensitivity for detecting novelty and therefore remain one of the cornerstones of medical progress; they provide many new ideas in medicine.'" (Vandenbrouke)

"A well written case report should provide clear evidence of the patient's problem or condition and its treatment. In addition, it should provide a clear explanation of why the reader should be surprised by the outcome of the case, with appropriate references. "(Vandenbrouke)

"In a case report, then, it should be clear exactly how the diagnosis was made. It should also be clear what treatment a patient might have received before embarking on the treatment that is being credited with an unusual outcome."

Why the Outcome is Unusual

"For a case report to be worth reporting, the outcome of the patient in question must be remarkable or unusual in some way. In the case of cancer, unusual results can be prolonged survival or stabilization, shrinkage, or disappearance of the tumor mass. Cancer by its nature grows and spreads; stabilization over a prolonged period, shrinkage, and disappearance are all unusual for a biopsy-proven cancer."

"A well-written case history should describe the typical outcome and the reference(s) from which this information was obtained."

Limitations of Case Studies

"Case studies are good for picking up novelty, but they have limitations. Generally speaking, a case report cannot prove that the treatment described is actually what created or caused the desired result. And a case report cannot indicate if the experience described is typical; only statistical analysis of a larger treatment group, compared to a clearly defined control group, can do that."

"The outcome described in a case report may not be the typical experience for patients pursuing a particular treatment. As an example, the drug Iressa (gefitinib) created great excitement when it was first introduced for lung cancer because some patients in initial case reports had amazing resolution of their disease. (Fujiwara) (Villano) The US Food and Drug Administration approved it for use outside of research studies in May 2003 under its accelerated approval regulations. But when the drug was more extensively tested in controlled clinical trials, it was found that very few patients actually had any response. (US Food & Drug Administration) Overall, there was no improvement in survival. (Thatcher)"

References

- Doyle RP. The Medical Wars. New York. William Morrow & Co. Inc. 1983.
- Fujiwara K, Kiura K, Ueoka H, Tabata M, Hamasaki S, Tanimoto M. Dramatic effect of ZD1839 ('Iressa') in a patient with advanced non-small-cell lung cancer and poor performance status. Lung Cancer 2004;40:73-6.
- Thatcher N, Chang A, Parikh P, et al. Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: Results from a randomized, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). Lancet 2005;366:1527-37.
- US Food and Drug Administration Center for Drug Evaluation and Research. Questions and answers on Iressa (gefitinib). Available at: http://www.fda.gov/cder/drug/infopage/iressa/iressaQ&A2005.htm. Accessed February 7, 2007.
- Vandenbrouke JP. In defense of case reports. Ann Intern Med 2001;134:330-4.
- Vandenbrouke JP. Case reports in an evidence-based world. JR Soc Med 1999;92:159-63.
- Villano JL, Mauer AM, Vokes EE. A case study documenting the anticancer activity of ZD1839 ('Iressa') in the brain. Ann Oncol 2003;14:656-8.

l. Can describe the design and the strengths and weaknesses of an N-of-1 randomized trial. (1 rg, 2 dp, 1 cn, 2 rl, 1 mh) (1.4)
   i. Can explain how an N-of-1 study is conducted. (2 rl, 2 rg, 2 mh, 1 cn, 2 dp) (1.8)
   ii. Can explain the potential usefulness for an individual patient in a real patient setting. (2 dp, 1 rg, 1 mh, 2 rl, 1 cn) (1.4)
   iii. Can explain why the results provide no evidence of generalizability beyond the case under study. (2 dp, 1 mh,1 rg, 2 rl, 1 cn) (1.4)
   iv. Understands the controversy surrounding the value of the research design (e.g., criticized by some epidemiologists as being quasi-experimental). (2 dp, 2 rl, 1 rg, 3 mh, 2 cn) (2.0)

Commentary: Guides for n-of-1 randomized trials (Straus 2005 Table 5.25)

1. Is an n-of-1 trial indicated for our patient?
   - Is the effectiveness of the treatment really in doubt for our patient?
   - Will the treatment, if effective, be continued long-term?
   - Is our patient willing an eager to collaborate in designing and carrying out an n-of-1 trial?
2. Is an n-of-1 trail feasible in our patient?
   - Does the treatment have a rapid onset?
   - Does the treatment cease to act soon after it is discontinued?
   - Is the optimal treatment duration feasible?
   - Can outcomes that are relevant and important to our patient be measured?
   - Can we establish sensible criteria for stopping the trial?
   - Can an unblended run-in period be conducted?
3. Is an n-of-1 trial feasible in our practice setting?
   - Is there a pharmacist available to help?
   - Are strategies for interpreting the trial data in place?
4. Is the n-of-1 study ethical?
   - Is approval by our medical research ethics committee necessary?

3. Can identify a hierarchy of research designs based on the type of clinical question posed. (1 dp, 2 cn, 1 rg, 1 rl, 2 mh) (1.4)
   a. Can identify the best research designs for questions of differential diagnosis. (2 dp, 1 rg, 1 cn, 2 rl, 2 mh) (1.6)
   b. Can identify the best research design for questions involving diagnosis. (1 dp, 2 cn 1 rg, 1 rl, 1 mh) (1.2)
      i. Can identify the best research designs regarding reliability and validity (i.e., cross-sectional with randomization and blinding). (1 rl, 1 dp, 1 rg, 2 cn, 1 mh) (1.2)
      ii. Can identify the best research designs regarding utility and efficacy of specific diagnostic tests (i.e., RCT, non-radmized comparison study). (2 rl, 2 dp, 1-2 rg, 1 mh, 2 cn) (1.7)
      iii. Can identify the best research designs regarding test responsiveness (i.e., prospective observational study). (3 rl, 2 dp, 1 rg, 2 cn, 1 mh) (1.8)
   c. Can discuss the recommended hierarchy (along with its variations) of research designs for questions of therapy, to include N-of-1 randomized controlled trials, systematic reviews of randomized trials, individual randomized controlled trials, systematic reviews of observational studies (e.g., cohort, case-control), individual observational studies, physiologic/ biomechanical research (e.g., studies of blood pressure, stress strain curve analysis of joint loading), and unsystematic clinical observations. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)

Comments: Not all sources site the exact same hierarchy, but most sources are in general agreement. Below is a commonly acceptable hierarchy.

Levels of evidence for therapy studies
(Straus 2005 Table 5.26)

1a. Systematic review with homogeneity of RCTs[a]
1b. Individual RCT with narrow confidence interval[b]
1c. All or none[c]

2a. Systematic review (with homogeneity) of cohort studies
2b. Individual cohort study (including low-quality RCT; e.g. <80% follow-up)

3a. Systematic review (with homogeneity) of case-control study
3b. Individual case-control study

4. Case series (and poor quality cohort and case-control studies)[d]

5. Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

---

[a] By homogeneity we mean that a systematic review is free of worrisome variations (heterogeneity) in the directions and degrees of results between individual studies. Not all systematic reviews with statistically significant heterogeneity are worrisome, and not all worrisome heterogeneity need be statistically significant.
[b] For example, if the confidence interval excludes a clinically important benefit or harm.
[c] Met when all patients died before the treatment became available, but some now survive on it, or when some patients died before the treatment became available, but now none die on it.
[d] By poor-quality cohort study, we mean one that failed to clearly define companion groups and/or failed to measure exposures and outcomes in the same (preferably blinded) objective way in both exposed and non-exposed individuals, and/or failed to identify or appropriately control known confounders, and /or failed to carry out a sufficiently long and complete follow-up of patients. By poor-quality case-control study, we mean one that failed to clearly define comparison groups, and/or failed to measure exposures and outcomes in the same blinded, objective way in both cases and controls, and/or failed to identify or appropriately control known confounders.

Teaching Tips: The committee suggests that instructors may wish to explain that the evidence pyramid concept is a somewhat oversimplification. For example, a well designed cohort study (which is lower on the pyramid) can render more accurate and useful information than a poorly designed RCT. Systematic reviews are rated higher than individual RCTs, but the methods selected to construct the review can significantly alter the conclusions. That is, different experts analyzing a group of RCTs may arrive at different conclusions regarding what the RCTs in toto suggest, depending on how they rate and synthesize the data.

d. Can identify the best research designs for questions of treatment side effects (i.e., RCTs and observational studies such as cohort or case control). [Guyatt 1] (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
e. Can identify the best research designs for harm questions regarding health risk factors (i.e., observational studies such as cohort or case control). [Guyatt 1] (1 dp, 1 rg, 2 cn, 1 rl, 1 mh) (1.2)

Commentary: The following quote is useful. **"**Case-control and other cohort studies really come into their own when the question involves harm. For example, does air pollution cause or worsen asthma in children? Does eating meat increase the risk of cancer? It is usually not feasible or ethical to conduct a randomized controlled trial to answer this sort of question, so alternative designs must be used. Cohort studies are also of particular value in addressing questions of prognosis and natural history. For example, what is the chance that someone who is HIV positive will develop AIDS in a given period of time?" in Dawes, M Evidence Based Practice, 2005

f. Can identify the best research designs for questions regarding prognosis (i.e., observational studies such as cohort and case control). (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)

4.4. Can describe the basic characteristics that determine the quality of research studies. (1 dp, 1 cn, 1 mh-4.3, 1 rl, 1 rg) (1.0)
1. Can define the broad concepts of *external validity* (i.e., generalizability of evidence from a research study population to an actual practice population), internal validity (i.e., the degree to which a study is measuring what it set out to) and experimental bias. (1 dp, 1 cn, 2 rg, 1 rl, 1 mh) (1.2)
2. Can define and discuss the key determinants of external validity. (1 dp, 2 rg, 1 rl, 1 mh, 2 cn) (1.4)
   a. Can discuss the importance of the patient population in the study. (1 rl, 1 rg, 1 mh, 1 cn, 1 dp) (1.0)
      i. Can describe the following methods of sampling: random sampling, stratified random sampling, cluster sampling, and convenience sampling. (1 dp, 1 rl, 1 mh, 1 rg, 1 cn) (1.0)

ii. Can describe the selection process and the impact of inclusion/exclusion criteria. (1 rl, 1 rg, 1 dp, 1 cn, 1 mh) (1.0)
iii. Can discuss the potential effects of subpopulations. (1 dp, 1 cn, 2 rl, 1 rg, 2 mh) (1.4)

b. Can discuss the role of provider and assessor characteristics (including the degree to which they are blinded and the potential for a variety of biases). (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0)
c. Can discuss the impact of the research setting (including the differences between hospital vs. private practice settings, primary care vs. specialist practice settings, and chiropractic vs. allopathic practice settings). (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0)

3. Can define and discuss the key determinants of internal validity. (1 dp, 1 rl, 1 rg, 1 mh, 1 cn) (1.0)
   a. Can discuss the potential impact of unplanned events that affect the history of the study as it unfolds (e.g., care sought outside the study, additional treatment that is not part of the study design, data from resentful respondents receiving less desirable treatment). (1 dp, 1 rl, 1 rg, 2 cn, 2 mh) (1.4)
   b. Understands the importance of factoring in the role of natural history ("maturation") of the subject's condition. (1 dp, 1 rl, 1 mh, 1 cn, 1 rg) (1.0)
   c. Understands the effect of the attrition rate (i.e., number of dropouts and noncompliant subjects). (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0)
   d. Understands that the very act of measuring a phenomenon may change it, influencing the outcomes and conclusions. (1 dp, 2 rl, 1 rg, 1 cn, 2 mh) (1.4)
   e. Understands that the quality of the data is influenced by the quality and characteristics of the outcome measures used in the study (i.e., issues of test reliability, validity and responsiveness). (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
   f. Understands the phenomenon of regression to the mean. (1 dp, 1 rl, 1 rg, 1 mh, 1 cn) (1.0)
      i. Understands the natural tendency of signs, symptoms and physiological systems to return to a natural mean value even without intervention. (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0)
      ii. Understands the concepts of central tendency in measurements with random error. (1 dp, 2 rg, 2 rl, 2 cn, 1 mh) (1.6)
   g. Undestands the importance of appropriate *allocation* (i.e., assuring that the characteristics of participants are the same across comparison groups). (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   h. Understands the inherent ambiguity of differentiating cause from effect and the potential for drawing erroneous conclusions. (1 rg, 1 dp, 1 rl, 1 cn, 1 mh) (1.0)
   i. Understands the confounding role of patient expectations and actions (e.g., the *Hawthorne effect*, *placebo effect*, *non-specific treatment effect*, *recall bias*). (1 dp, 1 rl, 1 mh, 1 cn, 1 rg) (1.0)
   j. Understands the potential effects of experimenter's/provider's expectations and actions, i.e., trying harder or greater enthusiasm because of participation in the study (*attention* and *expectation bias*). (1 dp, 1 cn, 1 mh, 1 rl, 1 rg) (1.0)

Commentary: A number of other types of bias can be discussed in this context. They include, but are not limited to, channeling effect or channeling bias, surveillance bias, verification bias, and detection bias. See glossary for definitions. The committee did not wish to indicate exactly which of these biases instructors should choose to elaborate on, nor whether the knowledge of the concepts alone was sufficient or whether familiarity with the actual terms was also necessary. [RL 8/30/07]

   k. Can explain the problem of diffusion of information or imitation of treatments (e.g., one group gets information that only the other group should have). (3 dp, 3 rg, 3 rl, 3 mh, 3 cn) (3.0)
   l. Can explain the importance of accounting for the ceiling and floor effects. (2.2)

4.5. Demonstrate an understanding of the basic characteristics of DIAGNOSTIC tests. (1.0)
1. Can explain the differences between normal and abnormal vs clinically significant or clinically insignificant in the context of diagnostic testing. (1.0)
   a. Explain a clinical, evidence-based definition. (1.4)
   b. Explain a statistical norm-based definition. (1.4))
   c. Explain an opinion-based definition. (2.0)


Teaching tip: Learners need to understand that the difference between what is considered normal and abnormal is based, in part, the purpose of the test or measurement. One method is to look at a population and identify the outliers as abnormal (e.g., cholesterol levels very low vs very high relative to a sample representing the general pubic. Another is to establish "normal " and "abnormal" based on optimal vs suboptimal (e.g., the average lipid level in the general population might be 200, but less than 150 might be decided to be more compatible with better health). An orthopedic test result might be considered normal in one population but not another based on clinical experience and opinion.

d. Discuss various methods that help determine cut points and test thresholds used to divide normal from abnormal. (1.8)
   i. Recognize that ROC curves can be used to establish optimal statistical performance of a diagnostic test. (2.6)
   ii. Explain what a cut point is relative to designating clinically important test results. (2.2)
   iii. Explain the choice of cut points based on whether the test is used for screening, case finding, or confirming a diagnosis. (2.2)
   iv. Explain the choice of cut points based on population normative values (e.g., within 2 standard deviations). (2.6)
   v. Understands the choice of cut points based on definitions of normal and abnormal as they apply to the state of optimum health. (2.2)

2. Demonstrate a basic understanding of common measures of reliability. (1.0)
   a. Explain the concept of reliability measures.


Teaching tip: Reliability, although a simple concept, is nonetheless very often confused by students. Explaining that the way the word is used in layman's language is sometimes a bit different than its strict use in science may help. Reliability simply means the repeatability of a procedure or test. But in everyday language a reliable car or a reliable friend or a reliable witness seems to imply that this is a car, friend or witness that you can trust and that you can expect it/him/her to do the job it is suppose to do and to do it accurately and appropriately. These additional inferences begin to sound more like test validity. The student's implicit understanding of the word from regular usage is constantly in conflict with its narrower meaning in diagnosis. Occasionally, course instructors, clinicians and even the literature also misuse the word further fueling the inherent confusion.

   b. Define the following types of reliability: *inter-examiner* (1.0) *intra-examiner* (1.0) and *test-retest* . (2.0)
   c. Recognize and interpret the results of Kappa (1.0) and intraclass correlation coefficient (ICC) (1.8) in terms of excellent, good, fair and poor.


Teaching tip: An important discussion should occur around the fact that the cut offs for each qualitative descriptor are somewhat arbitrary. For context, it may be useful to inform students that a kappa of 0.40 is often set as the minimal acceptable reliability in physical medicine, although this threshold has its critics.

3. Demonstrate a basic understanding of common measures of validity. (1.0)
   a. Demonstrate an understanding of test sensitivity. (1.0)


Teaching tip: Even before launching into discussions of false positives and negatives and how sensitivity is calculated, it is important to impart a good basic understanding of what sensitivity is. It should be carefully explained that the sensitivity of a test is only one characteristic of a test. In clinical terms, its greatest importance is indicating the probability of missing something you are looking for. Analogies such as metal detectors at the airport or metal detectors used at the beach to find coins can be useful in illustrating that the more sensitive the setting the less likely it is to miss what you are looking for. These analogies are also useful when trying to illustrate the difference between the sensitivity and specificity of a test. Because although in some ways a simple concept, it is rapidly lost when the discussions become more complex with false positives and negatives, specificity, predictive values, etc. It is critical to continuously remind students of the basic concept.

i. Explain how test sensitivity is determined. (1.2)
ii. *Define* sensitivity in terms of percent of true positives. (1.6)
iii. *Use* test sensitivity to rule out conditions based on its rate of false negatives. (1.2)

Teaching tip: This an area of constant confusion. A highly sensitive test rarely misses the condition tested for (few false negatives) but is *defined* in terms of the number of cases it identifies (true positives). The problem is that the sensitivity was established in a controlled cohort that was comprised only of true positives. There was no such thing as a false positive. But in real clinical settings not all of the positive results will be "true." What ***doesn't*** change in the clinical setting compared to the research setting is the number of false negatives (i.e., the number of misses). And so while we define sensitivity based on true positives, it is more common clinically to talk about and focus on its low rate of (false) negatives. The low rate of false negatives may be more useful in helping to rule out a condition rather than its high rate of positives (many of which may be false and may or may not be useful at ruling the condition in).

iv. Calculate sensitivity using a 2X2 table. (1.4)

b. Demonstrate an understanding of test specificity. (1.0)

Teaching tip: The concept of specificity is much more challenging than sensitivity for students to comprehend and then to hang on to as they study more about validity. Even before launching into discussions of false positives and negatives and how specificity is calculated, it is important to impart a good basic understanding of what specificity is. It should be carefully explained that the specificity of a test is only one characteristic of a test. In clinical terms, its greatest importance is indicating the probability that a test will cross react with a healthy patient or a condition other than the one being targeted. From this perspective, it should be explained that a test with 95 specificity actually means that it will falsely cross react (become positive) about 5 % of the time that the test is used. If using the analogy of metal detectors, the correlation is with a metal detector falsely identifying that something has a significant amount of metal in it (i.e., it may not be enough metal for either the security personnel or the coin collector to care about). It is often helpful to remind students that specificity is calculated based on a group of subjects who do NOT have the condition. One of the most common thinking errors is that a test with 95% specificity predicts that a patient has a 95% chance of having a particular condition (positive predictive value and specificity are notoriously confused by students). If is critical to continuously remind students of the basic concept of sensitivity, it is doubly important to continuously clarify what specificity implies and what it does NOT imply.

i. Explain how test specificity is determined. (1.2)
ii. *Define* test specificity in terms of true negatives.
iii. *Use* test specificity to rule in a condition based on its rate of true positives. p. (1.0)
iv. Calculate specificity using a 2X2 table. (1.2)

c. Demonstrate an understanding of the relationship between sensitivity and specificity. (1.4)
i. Explain the inverse relationship between sensitivity and specificity when establishing cut points. (1.0)
ii. Recognize that if the specificity and specificity of a test adds up to 100% the diagnostic test is of no clinical value.

4. Explain the meaning of test results when expressed as *test accuracy* as well as the limitations of this mode of expression. (2.0)
5. Demonstrate an understanding of pre-test probability and its application to diagnostic testing. (1.0)
a. Define incidence and prevalence. (1.0)
b. Explain how prevalence affects the results of screening an asymptomatic population. (1.0)

Commentary: "Despite extreme differences in prevalence of disease compared with primary care, almost all medical education and most research on diagnostic tests take place in teaching hospitals. Specialist researchers often extrapolate their experience of dealing with highly selected referral populations to family and general practice. As a result, the ability of diagnostic tests to correctly identify a disease outside a tertiary care center is markedly overestimated."

Prevalence vs. Predictive Value (based on test "x" with 91% sensitivity and specificity)

Prevalence of Diabetes	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Normal population 2%	91%	91%	17%	99%
Obese elderly population 10%	91%	91%	53%	99%
Cree 32%	91%	91%	83%	96%

"Men with moderately aggressive disease, who represent approximately 8 to 12% of patients with prostate cancer, appear to respond to some therapies, including radical prostatectomy. This procedure, however, carries with it several risks, which include mortality less than 1%, complete incontinence (7%), any incontinence (27%), and impotence (32%). (Sant'Ana AM)"

Sant' Ana AM, Rosser W, Talbot T. Five years of health care in Sao Jose. Fam Pract 2002;19:410-5.

"Once men are true-positives for cancer of the prostate have been confirmed, 80 to 85% will undergo radical treatment. Three percent of men with confirmed prostate cancer will die from the disease or treatment, and one-third will have a diminished quality of life in the absence of any benefit. Eight to 12% of men may benefit from early detection and treatment, but for those with more aggressive disease, no intervention appears to alter their rapid disease progression"

*Excerpts from Rosser, Slawson, Shaughnessy. Information Mastery: Evidence-Based Family Medicine, 2nd Ed. 2004

   c. Explain how pre-test probability affects testing a symptomatic patient. (1.0)
6. Demonstrate an understanding of positive and negative predictive values. (1.2)
   a. Define positive and negative predictive values in relationship to false positives and false negatives. (2.2)
7. Demonstrate an understanding of likelihood ratios. (1.0)
   a. Define positive and negative likelihood ratios. (1.0)
   b. Explain their relationship to sensitivity and specificity. (1.2)
   c. Explain their relationship to establishing predictive values for a particular condition being tested (post-test vs. pretest odds ratios). (1.4)
   d. Calculate positive and negative likelihood ratios from sensitivity and specificity numbers. (1.2)
   e. Use a nomogram to calculate post test probabilities. (1.0)
8. Define and discuss the clinical significance of test responsiveness. (1.0)
   a. Explain the concept of test responsiveness (i.e., evaluation of clinical change). (1.0)
   b. Explain the significance of determining the minimally clinically important change for an instrument. (1.0)

4.6. Can appraise the validity and usefulness of a primary study of DIAGNOSTIC tests. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
1. Can assess the common characteristics of a valid study of a diagnostic test. (1 dp, 1 mh, 1 rl, 1 rg, 1 cn) (1.0)
   a. Can ascertain if an appropriate case mix is used (e.g., a representative patient spectrum or a subgroup). [Dawes] (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   b. Can ascertain if subjects are blinded to all test findings. [Dawes] (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   c. Can determine if assessors are blinded to confounding information (e.g., other exam findings that might influence the interpretation). (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   d. Can ascertain if a complete test or complete test battery (cluster) was evaluated (i.e., partial test characteristics alone do not evaluate the reliability or validity of a complete test). (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   e. Can ascertain whether the procedure is described clearly enough to be reproduced. [Dawes] (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   f. Can identify the key elements of a valid study on test reliability (clinical agreement). (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   g. Can assess if proper methodology was used (i.e., a randomized order of assessors and blinding of assessors to each others' findings). (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
   h. Can assess if proper statistical tools were used. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)

i. Can define the following statistical tools and scales of measurement: NOI/R (nominal ordinal interval/ratio), KAPPA, weighted KAPPA, and ICC. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
ii. Can match the appropriate statistic with type of data: nominal, ordinal, interval, or ratio. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
iii. Can interpret the magnitude of a kappa or intraclass correlation coefficient. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)

2. Can identify the key elements of a valid study on test validity. (1 dp, 1 cn, 1 mh, 1 rl, 1 rg) (1.0)
   a. Can define the following terms: *gold standard* (AKA, *reference/criterion standard*), *face validity*, *content validity*, *construct validity*, and *discriminative validity*. (2 mh, 2 rg, 2 cn, 2 dp, 2 rg, 1 rl) (1.8)
   b. Can determine if an appropriate gold standard was used. [Dawes] (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
   c. Can determine if all the patients were compared to the reference (gold) standard. [Dawes] (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)

Curricular Suggestions: The strategy is to teach students to be able to perform rapid assessments as well as more detailed assessments. To accomplish this, checklists or instruments should be agreed upon to aide students in these two different approaches. [RL 10/2/06]

A brief assessment about a the validity of a diagnostic test valid (Straus 2005 Table 3.2)

1. **Measurement:** was the reference ("gold") standard measured independently, i.e. blind to our target test?
2. **Representative:** was the diagnostic test evaluated in an appropriate spectrum of patients (like those in whom we would use it in practice)?
3. **Ascertainment:** was the reference standard ascertained regardless of the diagnostic test result?

(Fourth question to be considered for clusters of tests of clinical prediction rules: was the cluster of tests validated in a second, independent group of patients?)

An instrument that is used as guides for researchers who write up diagnostic papers for publication can also be useful for readers assessing the research. A commonly used instrument is STARD [RL 9/28/06]

STARD checklist of items to improve the reporting of studies on diagnostic accuracy. Test version, November 2001. *For evaluation purposes only*

Section and topic	Item	Describe
TITLE/ABSTRACT/ KEYWORDS	1	The article as a study on diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity')
INTRODUCTION	2	The research question(s), such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups
METHODS		
Participants	3	The study population: the inclusion and exclusion criteria, setting(s) and location(s) where the data were collected
	4	Participant recruitment: was this based on presenting symptoms, results from previous tests, or the fact that the participants had received the index test(s) or the reference standard?
	5	Participant sampling: was this a consecutive series of patients defined by selection criteria in (3) and (4)? If not specify how patients were further selected.
	6	Data collection: were the participants identified and data collected before the index test(s) and reference standards were performed (prospective study) or after (retrospective study)?
Reference standard	7	The reference standard and its rationale
Test methods	8	Technical specification of material and methods involved including how and when measurements were taken, and/or cite references for index test(s) and reference standard
	9	Definition and rationale for the units, cutoffs and/or categories of the results of the index test(s) and the reference standard
	10	The number, training and expertise of the persons (a) executing and (b) reading the index test(s) and the reference standard
	11	Whether or not the reader(s) of the index test(s) and reference standard were blind (masked) to the results of the other test(s) and describe any information available to them
Statistical methods	12	Methods for calculating measures of diagnostic accuracy or making comparisons, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals)
	13	Methods for calculating test reproducibility, if done
RESULTS		
Participants	14	When study was done, including beginning and ending dates of recruitment
	15	Clinical and demographic characteristics (e.g. age, sex, spectrum of presenting symptom(s), co morbidity, current treatment(s), recruitment center)
	16	How many participants satisfying the criteria for inclusion did or did not undergo the index test and/or the reference standard; describe why participants failed to receive either test (a flow diagram is strongly recommended)
Reference standard	17	Time interval and any treatment administered between index and reference standard
	18	Distribution of severity of disease (define criteria) in those with the target condition; describe other diagnoses in participants without the target condition
Test results	19	A cross tabulation of the results of the index test(s) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard
	20	Indeterminate results, missing responses and outliers of index test(s) stratified by reference standard result and how they were handled
	21	Adverse events of index test(s) and reference standard
Estimation	22	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals)
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done
	24	Measures of test reproducibility, if done
DISCUSSION	25	The clinical applicability of the study findings

3. Can identify the key elements of a valid study on test utility and efficacy (i.e., the same criteria used for studies on treatment). (1 mh, 1 cn, 2 rl, 1 cn, 1 rg, 1 dp) (1.2)
4. Knows the criteria for a useful study of a diagnostic test. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    a. Can determine exactly how the test was performed (operationally defined). (1 rl, 1 rg, 1 cn, 1 dp, 1 mh) (1.0)
    b. Can determine if the test was evaluated in a clinically meaningful manner. (1 dp, 1 cn, 1 mh, 1 rg, 1 rl, 1 mh) (1.0)
    c. Can determine if a relevant patient population was used. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    d. Can determine if a relevant assessor population was used. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    e. Can determine if the reliability and validity of a test or procedure are relevant to the condition or clinical question being posed. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    f. Can distinguish test scale accuracy from test diagnostic validity. (1 dp, 1 cn, 1 rg, 1 mh, 2 rl) (1.2)
    g. Can determine if the evidence supports whether the test can accurately distinguish patients who do and do not have a specific disorder. (1 dp, 1 mh, 1 cn, 1 rg, 1 rl) (1.0)

Commentary: "To be able to apply the results of the study to your own clinical practice, you need to be confident that the test is performed in the same way in your setting as it was in the study. In the study by Wells et al [on evaluating DVT] (1995) 'clinical assessment' was not left as an implicit judgment, but clearly defined criteria were written down as to what constituted high, moderate and low probability of DVT. Therefore, it should be feasible to use the same clinical model in your own practice and achieve similar results." (Mant J. Is this test effective? in Dawes, M Evidence Based Practice, 2005)

"To be able to apply the results of the study to your own clinical practice, you need to be confident that the test is performed in the same way in your setting as it was in the study. In the study by Wells et al (1995) 'clinical assessment' was not left as an implicit judgment, but clearly defined criteria were written down as to what constituted high, moderate and low probability of DVT. Therefore, it should be feasible to use the same clinical model in your own practice and achieve similar results.

"At the beginning of appraisal many people new to it are surprised at the number of flaws in papers, even from established journals. It is therefore quite easy to 'rubbish' a paper. This will give you confidence to begin with. The skill of appraisal is not only to answer these quality questions, but later to evaluate how these flaws might influence the results. Would 78% follow-up significantly alter the results in this paper? By examining critically you seek to assess the inference of bias produced during the research, on the eventual results. It is possible to value and use results that contain bias. That is the real skill of appraisal." RL 9/28/06] From Jonathan Mant, Evidence Based Practice "Is it clear how the test was carried out?

4.7. Can appraise the validity and usefulness of research on the process of DIFFERENTIAL DIAGNOSIS. (1 dp, 1 cn, 1 mh, 1 rg, 1 rl) (1.0)
1. Can demonstrate an understanding of the diagnostic process. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    a. Understands the role of pattern recognition. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    b. Understands the role of individual tests and test clusters in narrowing down the diagnostic possibilities. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    c. Understands the process of identifying a working/provisional diagnosis out of a set of differential diagnoses. [Guyatt 1 p. 104] (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
    d. Understands how criteria are derived for some diagnostic entities (e.g., IHS criteria for cervicogenic headache, American College of Rheumatism's criteria for SLE). (2 dp, 2 cn, 2 rg, 2 rl, 1 mh) (1.8)
2. Can differentiate an article on diagnostic procedures from an article on differential diagnosis. (2 rg, 2 rl, 1 mh, 1 dp, 2 cn) (1.6)
3. Can determine if the patients enrolled in a differential diagnosis study are representative of typical patients with the clinical problem. (2 rl, 2 rg, 2 mh, 1 cn, 2 dp) (1.8)
    a. Can ascertain if the clinical problem assessed was clearly defined. (2 rl, 2 rg, 1 cn, 1 dp, 2 mh) (1.6)
    b. Can ascertain if the study's patient population is representative of those with the clinical problem. (2 rl, 1 rg, 1 cn, 1 dp, 2 mh) (1.4)
        i. Can determine if subjects were from a consecutive series design or from a specific geographical location. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)
        ii. Can identify the inclusion and exclusion criteria for the study. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)

iii. Can determine if all subjects were assessed in a similar setting (e.g., a specialty clinic vs. a primary care clinic) or represent a broader cross section of settings. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)
iv. Can determine if the authors identified and addressed any subjects who dropped out of the study or who had incomplete follow-up. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)

4. Can ascertain if the definitive diagnostic standard used in the study was appropriate and whether the differential diagnostic process was credible. (2 rl, 2 rg, 2 cn, 1 dp, 2 mh) (1.8)
   a. Can determine if explicit diagnostic criteria were used, described, and referenced. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)
   b. Can determine if findings were described and used to both confirm and exclude a diagnosis. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)
   c. Can determine if the diagnostic criteria were based on a comprehensive search to identify all causes of the clinical problem. (2 rl, 2 rg, 1 cn, 2 dp, 2 mh) (1.8)
   d. Can determine if the interexaminer reliability of the assessment procedures used in the study was cited and adequate. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)
   e. Can determine if the process was clear, sufficiently described, and standardized to replicate their design. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)
   f. Can determine if the diagnostic criteria were applied consistently among examiners. (2 rl, 2 rg, 2 mh, 2 cn, 2 dp) (2.0)
5. Can determine if the follow-up period was of sufficient time and completeness for initially undiagnosed patients. (2 rl, 2 rg, 2 cn, 2 dp, 2 mh) (2.0)
   a. Understands that a higher number of undiagnosed patients increases the chance of error in estimating disease probability. (2 rl, 2 rg, 2 mh, 2 cn, 2 dp) (2.0)
   b. Understands that longer follow-up periods have a better chance of determining if a patient has a diagnosable disorder which was initially missed. (2 rl, 2 rg, 2 mh, 2 dp, 2 cn) (2.0)
6. Can determine if the study reported all diagnoses identified and their probabilities. (2 rl, 2 rg, 2 mh, 2 dp, 2 cn) (2.0)
   a. Can determine the percentages of the established diagnoses. (2 rl, 2 rg, 2 cn, 2 mh, 2 dp) (2.0)
   b. Can determine how precise the estimates of the probability of each disease were by evaluating the reported confidence intervals. (2 rl, 2 rg, 2 mh, 2 cn, 2 dp) (2.0)

4.8. Can appraise the validity and usefulness of a primary study on THERAPY (e.g., an RCT). (1 mh, 1 dp, 1 cn, 1 rl, 1 rg) (1.0)
1. Knows the criteria for a valid study on a therapeutic intervention. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   a. Can determine if patients were properly identified and appropriate sampling was done to help ensure external validity. [Guyatt 1] (1 dp, 1 mh, 1 cn, 1 rl, 1 rg) (1.0)
   b. Can determine if proper subject randomization was conducted to ensure control of internal validity (control for allocation bias). [Guyatt 1] (1 dp, 1 mh, 1 cn, 1 rl, 1 rg) (1.0)
   c. Can determine if proper blinding of experimenters, patients and therapists was conducted to ensure internal validity. [Guyatt 1] (1 dp, 1 mh, 1 cn, 1 rl, 1 rg) (1.0)
      i. Can determine if there was potential for selection bias. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
      ii. Can define and describe study designs that are single-blind, double-blind, triple-blind assessor-blind, blinding to the degree possible, and the use of naiveté in lieu of blinding. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
      iii. Can determine if there was a concealment of group assignment prior to acceptance into study. (1 dp, 1 cn, 2 rl, 1-2 rg, 1 mh) (1.1)
   d. Can determine if treatment and control groups are similar at baseline in terms of important prognostic predictor variables or, if not, the predictor variables are adjusted for in the analysis. [Guyatt 1] (1 dp, 1 mh, 1 cn, 1 rl, 1 rg) (1.0)
      i. Can determine if analysis of covariance (ANCOVA) or equivalent (including general linear models or regression) was conducted. (2 dp, 2 rl, 2 rg, 2 cn, 2 mh) (2.0)
      ii. Can determine if the baseline values of outcome measures were treated as a covariate in the analysis. (2 dp, 2 rl, 2 rg, 2 cn, 1 mh) (1.8)
   e. Can determine if appropriate outcome measures were used. (1 dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.0)
      i. Can determine if patient-centered outcomes were included as primary outcomes. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
      ii. Can determine if there were biased outcomes and/or treatment effects. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)

iii. Can determine if outcomes were measured at appropriate follow-up time points. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)

f. Understands the importance of experimental and control groups being treated equally aside from the main intervention (expectation bias). (1 dp, 1 mh, 1 cn, 1 rg, 1 rl) (1.0)
   i. Can determine if outside care is evaluated and balanced across groups. (1 dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.0)

g. Can determine if there are missing data or dropouts and whether these concerns are addressed (attrition bias). [Guyatt 1] (1 dp, 1 rg, 1 mh, 2 cn, 1 rl) (1.2)
   i. Can determine if percentages of missing data were small and balanced in each group. (1 cn, 1 rl, 1 rg, 1 mh, 1 dp) (1.0)
   ii. Can determine if the reasons for missing data are reported by each group. (1 dp, 2 cn, 1 rg, 2 rl, 2 mh) (1.6)
   iii. Can determine if missing data are addressed in the statistical analysis. (2 dp, 2 rg, 2 cn, 2 rl, 2 mh) (2.0)
   iv. Knows how to use the "5 and 20" rule (i.e., fewer than 5% loss is a low threat to validity, more than 20% significant is a threat) along with the limitations to this rule. (1 rl, 1 dp, 1 rg, 1 cn, 1 mh) (1.0)

h. Can assess if appropriate analysis was performed. (1 dp, 1 rl, 1 rg, 1 mh, 1 cn) (1.0)
   i. Understands the need for *intention-to-treat* analysis. [Guyatt 1] (1 dp, 1 rl, 1 rg, 1 mh, 1 cn) (1.0)
   ii. Understands the need for adjusting p-values for multiple comparisons, multiple outcome measures, and multiple looks at the data. (2 dp, 1 mh, 2 cn, 2 rg, 2 rl) (1.8)
   iii. Understands the difference between primary and secondary outcomes as well as the role of each (i.e., the potential for drawing major conclusions vs. simply generating hypotheses). (2 dp, 2 mh, 2 cn, 2 rg, 2 rl) (2.0)
   iv. Can determine if authors' conclusions are justified based on the study design, how it was conducted, method of analysis, and how robust are the actual results (author filter bias). (1 dp, 1 rl, 1 cn, 1 mh, 1 rg) (1.0)

Curricular Suggestions: The strategy is to teach students to be able to perform rapid assessments as well as more detailed assessments. To accomplish this, checklists or instruments should be agreed upon to aide students in these two different approaches. [RL 10/2/06]

Guidelines for appraising a therapeutic article (Dawes 2005, Box 5.1)

1. Did the authors answer the question?
2. What were the characteristics of the patients?
3. Were the groups similar at the start of the trial
4. Aside from the experimental treatment, were the groups treated equally?
5. What was the treatment?
6. What was the comparison (placebo?)
7. Were all patients who entered the trial accounted for at its conclusion? Were they analyzed in the groups to which they were randomized?
8. Was the assignment of patients to treatments randomized?

8b. Was the randomized list concealed?

9. Were patients and clinicians kept "blind" to which treatment was being received?
10. Was the length of the study appropriate?
11. Is the context of the study similar to your own?
12. Did the treatment work?

Another instrument that is used as a guide for researchers who write up RCTs for publication can also be useful for readers assessing the research. A commonly used instrument ss CONSORT [RL 9/28/06]

CONSORT Checklist of items to include when reporting a randomized trial

PAPER SECTION and TOPIC	Item	DESCRIPTION
TITLE & ABSTRACT	1	How participants were allocated to interventions (e.g., "random allocation", "randomized", or "randomly assigned").
INTRODUCTION Background	2	Scientific background and explanation of rationale.
METHODS Participants	3	Eligibility criteria for participants and the settings and locations where the data were collected.
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered.
Objectives	5	Specific objectives and hypotheses.
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors).
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.
Randomization -- Sequence generation	8	Method used to generate the random allocation sequence, including details of any restrictions (e.g., blocking, stratification)
Randomization -- Allocation concealment	9	Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.
Randomization -- Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. When relevant, how the success of blinding was evaluated.
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); Methods for additional analyses, such as subgroup analyses and adjusted analyses.
RESULTS Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.
Recruitment	14	Dates defining the periods of recruitment and follow-up.
Baseline data	15	Baseline demographic and clinical characteristics of each group.
Numbers analyzed	16	Number of participants (denominator) in each group included in each analysis and whether the analysis was by "intention-to-treat". State the results in absolute numbers when feasible (e.g., 10/20, not 50%).
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (e.g., 95% confidence interval).
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory.
Adverse events	19	All important adverse events or side effects in each intervention group.
DISCUSSION Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes.
Generalizability	21	Generalizability (external validity) of the trial findings.
Overall evidence	22	General interpretation of the results in the context of current evidence.

2. Apply criteria to determine if a study on THERAPY may be clinically useful. (1.0)
   a. Explain the concept of treatment effect magnitude. (1.0)
      i. Define and interpret appropriate expressions of treatment efficacy to include treatment effect (difference between groups) (1.0), relative risk (1.0), relative risk reduction (1.25), absolute risk (1.0), absolute risk reduction (1.2), ORs (1.0), NNT (1.0), and effect size (standardized difference between groups) (2.2).

Possible Teaching tip: Relative risk reduction = absolute risk reduction / control risk. It is abstract and not used so much… There are two effect sizes. The first is the "treatment effect" which is simply the difference in outcomes. The student should know that 10 points (of 100) is a clinically significant difference between groups. The other is more abstract but seen in systematic reviews. This is the standardized treatment effect size which is the absolute effect size / SD. There are standard rules from Cohen on small, moderate, and large. Last, beware of "reduction" because often we talk about risk of improvement as in the headache literature.

      ii. Calculate NNT if the absolute risk is provided. (1.2)
      iii. Distinguish within person, within-group and between-group effect magnitudes in identifying a clinically important effect. (1.6)
   b. Explain the concept of clinical importance/significance (1.0)
      i. Explain the distinction between a statistically significant and a minimal clinically important difference (MCID).
      ii. Recognize that interpreting the magnitude of the treatment effect depends, in part, on what the intervention is being compared to (e.g., placebo, no-treatment, a validated treatment, a non- validated treatment). (1.4)
      iii. Recognize the factors involved in deciding if an NNT is judged to be clinically important (such as patient profile, phase of the condition, definition of treatment success, what it is compared to). (1.8)

Teaching tip: Learners should realize that when they read a reported NNT (e.g., the use of splints for carpal tunnel syndrome has an NNT of 5), this number refers to how many patients could be potentially helped, it does not indicate how much they will be helped. To more fully appreciate the usefulness of an NNT, a number of important facts need to be known. What counted as treatment success (full resolution? pain reduction?)? What was the intervention compared to (no treatment? placebo?)? How severe must the disorder be (does the NNT apply only to mild cases? to severe cases? a broad cross section of cases?)? What phase of the disorder does the NNT apply to (acute LBP vs chronic LBP?)? How long will the benefit last (short term palliation? Curative?)? Therefore, a more complete context for appreciating a published NNT might be "Combined cervico-thoracic manipulation and exercise therapy for reducing headache frequency in patients with persistent headache had an NNT of 2 when compared to self-care instruction." This critical contextual information is often loss when NNT are bandied about.

Teaching Tip: The concept "number needed to treat" (NNT) can potentially be confusing. It is a measure of the effectiveness of a particular therapy. An easy misinterpretation is that an NNT of 10 means that one would need to treat 10 patients to get only one better. Actually, it usually denotes how many patients would need to be treated to get *one additional patient* better compared to placebo or no treatment at all. The number of patients out of 10 who actually get better could be much higher. It would be a combination of those that would get better due to natural history/placebo plus one more due to the therapy. Furthermore, the context of the study in which the NNT is reported also makes a difference. It can be used to compare therapies. In one early study (Fochet 2002) comparing duct tape for the treatment of warts versus cyrotherapy, duct tape actually removed more warts. The NNT was reported as 4. In this context that meant that for every four patients treated with duct tape one more wart would have been removed than if cryotherapy had been used instead.

Another way to explain it to students is by way of absolute risk reduction. When NNT is calculated, it is the inverse of the absolute risk reduction comparing two different treatments (again, one is usually placebo). In the wart example, 85% were cured with the duct tape, 60% cured with cyrotherapy, resulting in an absolute risk reduction of 25% for the duct tape approach. In other words, for every 100 people treated, 25 more people will be cured with duct tape than with cryotherapy. If you need to treat 100 people to cure an extra 25, then you need to treat 4 people to cure an extra 1.

Bottom line: An NNT seen without a context usually means how many patients would need to be treated to get one person better *compared to doing nothing or using a placebo*. Occasionally, it is used to compare the effectiveness of two different treatments. But it does not necessarily answer the broader question that a patient may ask, "What are my chances of getting better?" That number could be relatively close to the NNT or considerably higher. [RL

and JM 3/20/07] (Fochet DR III, Spicer C, Fairchoke MP. The efficacy of duct tape vs. cryotherapy in the treatment of verruca vulgaris (the common wart). Arch Pediatr Adolesc Med. 2002;156:971-974.)

Another difficult concept revolving around NNT to teach revolves around what is a good number. It depends partly on what the intervention is vs. the side effect vs. the cost vs. the outcome of not treating. Below presents interesting numbers for calcium supplementation and statins. [RL 8/24/08]

Calcium Supplements Linked to Lower Fracture Risk in Older Adults

Calcium supplementation lowers fracture risk among older adults, according to a meta-analysis published in *Lancet.* Data were extracted from 29 placebo-controlled trials of calcium supplementation (with or without vitamin D) that enrolled people aged 50 or older. Seventeen trials reporting fracture as an endpoint found a 12% reduction in risk with calcium or calcium plus vitamin D. The treatment effect was largest among adults older than 70, as well as for calcium doses of 1200 mg or more, or vitamin D doses of 800 IU or more. In 24 trials reporting on bone mineral density, supplementation was associated with a significant reduction in bone loss at the hip and spine.The authors say that to prevent one fracture, 63 patients would need to receive calcium supplements for 3.5 years — making calcium "comparable to other preventive treatments such as statins." To prevent one fracture among "elderly" adults, they note, the NNT dropped to 30 or fewer. (From Physician's First Watch for August 24, 2007)

More Commentary: Attempts to mount RCTs were considered unethical because they would deprive the control group of a widely accepted and perceived beneficial intervention. Fortunately, an RCT is being conducted, and the preliminary results question the benefit of inserting tympanostomy tubes in all but a few children with chronic otitis media. A 3-year follow-up study has demonstrated poorer hearing in children with tubes compared to those who when untreated. (Maw R, Bawden R. Spontaneous resolution of severe chronic glue ear in children and the effects of adenoidectomy, tonsillectomy, and insertion of ventilation tubes. BMJ 1993;306:750-60.)

It is strongly recommended that you visit the Web site http://www.cebm.net/scratching_post.asp and become comfortable with demanding the NNT or calculating them yourself to improve your own and your patients' understanding of the benefit of therapies.

iv. Recognize the challenges around establishing what degree of improvement is necessary to be meaningful to researchers vs. clinicians vs. patients. (2.0)

c. Recognize whether the outcome has a patient-centered, clinically meaningful effect (e.g., decreased pain, improved activities of daily living or quality of life) or is based on a surrogate measure (e.g., improved muscle test, range of motion, cholesterol level). (1.0)

4.9. Can appraise the validity and usefulness of a study on PROGNOSIS. (1 dp, 1 rg, 1 mh, 1 rl, 1 cn) (1.0)

1. Knows the criteria for a valid study on prognosis. (1 dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.0)
   a. Can determine if defined, representative patient samples were recruited (to avoid *referral filter bias*). (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
   b. Can determine if subjects were assembled at a common point in the disease process. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   c. Can assess if there was appropriate length and completeness of follow-up. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      i. Can determine if the percentages of missing data are small and balanced in each group. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      ii. Can determine if the number and reasons for missing data are reported and whether these omissions are likely to have a significant impact on the conclusions. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      iii. Can determine if missing data are included in the statistical analysis. (2 dp, 1 cn, 2 rg, 1 rl, 1 mh) (1.4)
      iv. Can determine if the follow-up was too brief to provide useful information. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      v. Can determine if appropriate periodic sampling was conducted and whether there might be a significant problem due to recall bias. (1 dp, 1 cn, 1 rg, 1 rl) (1.0)
   d. Can determine if the study contains objective outcome criteria applied in a blind fashion. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
   e. Can determine if subgroups were adjusted for important prognostic indicators. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
   f. Can determine if subgroups were validated by an independent group of "test-set" patients. (2 dp, 2 cn, 2 rl, 2 rg, 1 mh) (1.8)

 **Commentary:** The following is background material from Dawes, M Evidence Based Practice, 2005, adapted from Laupacis et al (1994).

Criteria for assessing validity of a cohort study giving information on prognosis

Key issues

1. Was there a representative sample of patients?
2. Were the patients at a similar point in the course of their illness?
3. Was follow-up complete?

Secondary issues

1. Was the follow-up over a sufficient period of time?
2. Were the outcomes used objective and unbiased?
3. Was adjustment made for important prognostic factors?

2. Knows the criteria for a useful study on PROGNOSIS. (2 dp, 2 mh, 2 rg, 2 cn, 2 rl) (2.0)
   a. Can determine the likelihood of predicted outcomes and the likelihood that these outcomes can be sustained over time. (2 dp, 2 cn, 2 rg, 2 mh, 2 rl) (2.0)
   b. Can explain the use of regression coefficients for predictors of outcomes and the standard error of estimate (precision of predicted outcomes). (2 dp, 2 cn, 2 rg, 2 mh, 2 rl) (2.0)

4.10. Can appraise the validity and usefulness of a study on HARM. (1 dp, 1 cn, 1 rg, 1 mh, 1 rl) (1.0)

1. Can describe two types of harm studies: risk factors related to prevention and side effects from treatments. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
2. Knows the criteria for a valid study on harm. (1 dp,1 rg, 1 mh, 1 cn, 1 rl) (1.0)
   a. Can determine if comparison groups were clearly defined and were similar in all important ways other than exposure to the treatment or risk factor. (1 dp, 1 rg, 1 rl, 1 cn, mh 1) (1.0)
      i. Can rule out selection and information bias. [Guyatt 1] (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      ii. Can determine if any remaining differences between the groups were adequately accounted for. [Guyatt 1] (1 dp, 1 cn, 2 rg, 1 rl, 1 mh) (1.2)
   b. Can determine if treatments/exposures and clinical outcomes were measured in the same way in both groups. [Guyatt 1] (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
   c. Can determine if assessment of outcomes was either objective or blinded to the exposure variables. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
   d. Can determine if patient follow-up of the study was sufficiently long enough for the measured outcome to occur. [Guyatt 1] (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
   e. Understands that the results of a harm study are influenced by the choice of research design (e.g., larger changes in risk are needed to be significant in observational studies than in RCTs). [Guyatt 1] (1 rl, 1 rg, 1 dp, 1 mh, 1 cn) (1.0)
   f. Can determine if the study demonstrates a cause and effect relationship. (1 dp, 1 rl, 1 rg) (1.0)
      i. Can determine if exposure precedes the onset of the outcome. (1 cn, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
      ii. Can determine if regression analysis establishes a link between a particular factor and harm. [Guyatt 2] (2 dp, 1 cn, 2 rg, 1 rl, 1 mh) (1.4)
      iii. Can determine if there is a dose-response gradient (e.g., increased exposure links to increase magnitude of effect). [Guyatt 1] (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      iv. Can determine if there is positive evidence from a “dechallenge-rechallenge” study (in the case of risk factors). [Dawes] (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      v. Can determine if there is consistent association from study to study (i.e., a repeatable effect). (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      vi. Can determine whether there is a biologically plausible association. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
      vii. Can determine if alternate explanations have been adequately addressed. [Guyatt 1] (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)

Commentary: The following is background material from Dawes, M Evidence Based Practice, 2005 regarding whether the association between exposure and disease is causal.

1. How strong is the association?	How large is the odds ratio (case control study) or relative risk (cohort study)? Have different types of study in different places and at different times shown the same association between exposure and disease?
2. How consistent is the evidence?	
3. Is the temporal relationship correct?	Does exposure precede onset of the disease?
4. Is causation biologically plausible?	Does a casual link fit with what we know already from our understanding of the basic sciences such as pathology and physiology in relation to the disease process?
5. Is there a dose response relationship?	Are people who have had greater exposure at greater risk of the disease?
6. Is there evidence of reversibility?	If the risk factor is removed, does the incidence of the disease fall?
7. Might confounding still explain the association?	Is it plausible that the association is due to confounding factors that have been inadequately dealt with in the studies?

Another issue is whether the evidence on harm is valid (Straus 2005 Table 6.1)

1. Were there clearly defined groups of patients, similar in all important ways other than exposure to the treatment or other cause?

2. Were treatments/exposures and clinical outcomes measured in the same way in both groups? (Was the assessment of outcomes either objective or blinded to exposure?)

3. Was the follow-up of the study patients sufficiently long (for the outcome to occur) and complete?

4. Do the results of the harm study fulfill some of the diagnostic tests for causation?
- Is it clear that the exposure preceded the onset of the outcome?
- Is there a dose-response gradient?
- Is there any positive evidence from a "dechallenge-rechallenge" study?
- Is the association consistent from study to study?
- Does the association make biological sense?

3. Knows the criteria for a useful study on HARM. (1.0)
   a. Can determine the magnitude of the association between the exposure and outcome. (1.0)
   b. Can demonstrate an understanding of the various terms used to communicate the degree of risk in Harm studies. (1.0)
      i. Can define absolute risk (AR) (1.0) relative risk (RR) (1.0), relative risk reduction (RRR) (1.6) odds ratios OR (1.0) and numbers needed to harm NNH (1.0).
      ii. Can interpret the clinical significance of reported RRR (2), RR (1), OR (1), AR (1), and NNH values. (1) (1.0)
      iii. Recognizes the relationship between case control studies and odds ratio (OR) and cohort studies and relative risk (RR). (2.2)

Commentary: The following is background info. "What do these odds and relative risks mean in plain English? Let's say a study looked at the odds of getting inadequate pain relief with relaxation compared with the odds of getting inadequate pain relief without relaxation. If the odds ratio was 0.70, this would mean you had 30% reduction in the odds of having inadequate pain relief with relaxation, compared with without relaxation. If the relative risk was 0.82, this would mean patients' risk of having inadequate pain relief is 18% less if they had relaxation. in Dawes, M Evidence Based Practice, 2005

Are the valid results of this harm study important? (Straus 2005, Table 6.3)

1. What is the magnitude of the association between the exposure and outcome?
2. What is the precision of the estimate of the association between the exposure and the outcome?

4.11. Can appraise the validity and usefulness of a study on COST EFFECTIVENESS. (2 rl, 1 rg, 2 cn, 1 mh, 1 dp) (1.4)

1. Knows the criteria for a valid study on cost effectiveness. (1 dp, 1 mh, 1 cn, 1 rg, 1 rl) (1.0)
   a. Understands the need for comparable patients and/or correcting for differences between study groups. (1 dp, 1 cn, 2 rl, 1 rg, 1 mh) (1.2)
   b. Understands the need for a fair comparison between interventions or tests (e.g., inclusion of comparable costs across comparison groups). (1 dp, 3 rl, 1 cn, 1 rg, 1 mh) (1.4)
   c. Understands the difference between cost-effectiveness, cost-benefit, and cost-utility. (1 dp, 2 cn, 2 mh, 2 rl, 2 rg) (1.8)
   d. Understands the concept of quality-adjusted life years. (2 dp, 2 cn, 2 rl, 2 rg, 2 mh) (2.0)
   e. Can define direct and indirect health care cost and understands the need to assess their relevance. (3 dp, 2 cn, 2 rl, 3 rg, 3 mh) (2.6)

Commentary: In cost effectiveness studies it is important to look at the complete cost of interventions. For example, in some studies comparing MD to DC, the additional cost of the MD sending the patient to a PT is not included in the comparison.

2. Knows the criteria for a useful study on cost-effectiveness. (1 dp, 1 cn, 1 rg, 2 rl, 1 mh) (1.2)
   a. Can determine if the procedures under study are relevant to practice. (1 dp, 1 cn, 1 rg, 2 rl, 1 mh) (1.2)
   b. Can determine if the study setting (e.g., HMO, PPO, out-of-pocket) is relevant to practice. (1 dp, 1 cn, 2 rl, 1 rg, 1 mh) (1.2)
   c. Can understand the significance of marginal cost-effectiveness ratios. (2 dp, 2 rg, 2 cn, 2 rl, 3 mh) (2.2)

APPLY

STANDARD 5

The EBP competent practitioner applies the relevant evidence to practice.

5. The EBP competent practitioner applies the relevant evidence to practice.

5.1. Assesses the relevance of the appraised evidence to the clinical problem at hand (*clinical applicability*). (1 dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.0)

1. Can distinguish research papers and reviews intended to change clinical decision-making from those papers proposing theoretical models or studies intended only as a basis for further research (e.g., pilot studies, animal studies, studies with insufficient power, studies with trends identified only in secondary outcomes). (1 dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.0)
2. Can determine if the study subjects were sufficiently similar to the practitioner's patient. (1 dp, 1 cn, 1 mh, 1 rg, 1 rl) (1.0)
    a. Can determine if the study setting is similar to their practice setting. (1 rl, 1 rg, 1 mh, 1 dp, 1 cn) (1.0)
    b. Can determine if the disease frequency (pre-test probability) for the conditions evaluated in the study are similar to their practice. (1 rl, 1 rg, 1 mh, 1 dp, 1 cn) (1.0)
3. Understands the importance of weighing the strength of the evidence. [Guyatt 1] (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
4. Can determine whether the action taken based on a study will have a significant impact on the patient based on degree of efficacy (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0), cost (1 dp, 1 rl, 1 rg, 2 cn, 1 mh) (1.2), cost-effectiveness (2 rl, 1 rg, 2 cn, 1 mh, 1 dp) (1.4), safety (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0), or patient preference. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)

5.2. Can select and interpret diagnostic tests appropriate to a particular patient's problem. (? dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.0)

Commentary: Many of the criteria we selected came from Straus 2005, Table 3.5

Questions to answer in applying a valid diagnostic test to an individual patient (Straus 2005)

1. In the diagnostic test available, affordable, accurate, and precise in our setting?

2. Can we generate a clinically sensible estimate of our patient's pre-test probability?
    • From personal experience, prevalence statistics, practice databases, or primary studies?
    • Are the study patients similar to our own?
    • Is it unlikely that the disease possibilities or probabilities have changed since this evidence was gathered?

3. Will the resulting post-test probabilities affect our management and help our patient?
    • Could it move us across a test-treatment threshold?
    • Would our patient be a willing partner in carrying it out?
    • Would the consequences of the test help our patient reach his or her goals in all this?

Also used is Dawes, M Evidence Based Practice, 2005

"An evidence-based approach to deciding whether a test is effective for your patient involves the following steps:

1. Frame the clinical question (see Chapter 2)
2. Search for evidence concerning the accuracy of the test (see Chapter 3)
3. Assess the methods used to determine the accuracy of the test (see Chapter 6)
4. Find out the likelihood ratios for the test
5. Estimate the pre-test probability of disease in your patient
6. Apply the likelihood ratios to this pre-test probability using the nomogram to determine what the post-test probability would be for different possible test results.
7. Decide whether or not to perform the test on the basis of your assessment of whether it will influence the care of the patient, and the patient's attitude to different possible outcomes."

1. Understands prevalence and pre-test probability as it applies to diagnostic testing of a particular patient. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
    a. Understands the multiple factors involved in estimating a patient's pre-test probability for a given problem. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
        i. Knows how to access prevalence based on authoritative sources (national, state, primary studies, etc.). (2 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.2)

ii. Understands that the pre-test probability may be different in his/her specific practice setting (primary care vs. secondary/tertiary care vs. chiropractic settings). (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
iii. Understands that the pre-test probability may be different from published prevalence estimates based on the patient's constellation of signs and symptoms. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
iv. Understands that the pre-test probability continues to change based on the results of prior testing. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)

2. Takes into consideration test reliability when choosing a diagnostic procedure and interpreting the results for a particular patient. (1 dp, 1 rl, 1 rg, 1 mh, 1 cn) (1.0)
   a. Can distinguish experimental reliability from clinically acceptable reliability. (2 or 3 dp, 2 rl, 3 rg, 1 cn, 1 mh) (1.9)
3. Demonstrate how to apply likelihood ratios to diagnosis. (1.0)
   a. Recognize likelihood ratios which are potentially useful vs. those of little to no value. (1.0)
   b. Explain how to use likelihood ratios to compare examination procedures to each other when selecting the best test. (1.0)
   c. Recognize that there are circumstances when likelihood ratios cannot be multiplied in sequence to predict post test probability.

Commentary: information was taken from Dawes M, Evidence Based Practice, 2005

"As a rule of thumb, diagnostic tests with positive likelihood ratios greater than 10 and/or negative likelihood ratios less than 0.1 can be thought of as fairly powerful tests. A likelihood ratio of 10 means, literally, that the odds of disease are 10 times greater than they were before the test was performed. A likelihood ratio of 0.1 means that the odds of disease are one-tenth what they were before the test was performed."

Some rules about likelihood ratios can help guide their application in practice

A relatively high likelihood ratio (5 to 10) will significantly increase the probability of a disease, given a positive test. A relatively low likelihood ratio (0.1 to 0.5) will significantly decrease the probability of a disease, given a negative test. Likelihood ratios of 2, 5, and 10 are associated with an increase in the probability of disease in the presence of a positive test, as follows:

LR+ = 2 increases the probability of the disease by ~15 percent
LR+ = 5 increases the probability of the disease by ~30 percent
LR+ = 10 increases the probability of the disease by ~45 percent

Likelihood ratios of 0.5, 0.2, and 0.1 are associated with a decrease in the probability of a disease in the presence of a negative test, as follows:

LR- = 0.5 decreases the probability of the disease by ~15 percent
LR- = 0.2 decreases the probability of the disease by ~30 percent
LR- = 0.1 decreases the probability of the disease by ~45 percent

Saint S, Drazen J, Solomon C. The New England Journal of Medicine Clinical Problem-Solving.McGraw Hill, 2006.

   d. Understands how to use likelihood ratios in comparing examination procedures to each other when selecting the best test. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
   e. Understands the sequential application of multi-level likelihood ratios in predicting a particular diagnosis in a particular patient. (2 dp, 1 rl, 2 rg, 2 cn, 2 mh) (1.8)
4. Understands how to choose tests to <u>rule in</u> a condition. (1 dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.0)
   a. Knows how to select a test based on its specificity (e.g., the mnemonic +SPin, "if positive, high specificity helps to rule in"). (1 dp, 1 rl, 1 rg, 1 mh, 1 cn) (1.0)
   b. Knows how to select a test on its positive predicative value. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
   c. Knows how to select a test based on its positive likelihood ratio. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
5. Understands how to choose tests to <u>rule out</u> a condition. (1 dp, 1 rl, 1 mh, 1 rg, 1 cn) (1.0)
   a. Knows how to select a test based on its sensitivity (e.g., the mnemonic -SNout, "if negative, high sensitivity helps to rule out). (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
   b. Knows how to select a test based on its negative predictive value. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
   c. Knows how to select a test based on its negative likelihood ratio. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)

6. Understands role of serial testing vs. parallel testing strategies. (3 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.4)
7. Identifies and understands the concepts of utility and test efficacy and their application to diagnostic testing. (1 dp, 2 rl, 1 rg, 1 mh, 2 cn) (1.4)
   a. Can define clinical utility and test efficacy. (1 dp, 1 rg, 2 rl, 2 cn) (1.5)
   b. Understands that when applying a test to a patient a determination must be made whether the test makes an important contribution to treatment selection or clinical outcome. (2 dp, 2 rg, 1 rl, 2 cn, 1 mh) (1.6)
   c. Understands the importance of balancing risks and benefits within the context of the individual patient when selecting a test to diagnose a condition. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
   d. Can balance the potential harm of being labeled with a disorder or risk compared to the likelihood of compliance with a management plan. (2 dp, 2 rl, 2 rg, 2 cn, 1 mh) (1.8)
8. Understands how to use evidence to make clinical decisions regarding screening and case finding. (2 mh, 1 dp, 1 rg, 1 rl, 1 cn) (1.2)
   a. Can explain the difference between screening and case finding. (1 dp, 2 rl, 1 cn, 1 rg, 2 mh) (1.4)
   b. Understands the importance of balancing risks and benefits when choosing a screening strategy for asymptomatic populations with various levels of risk. (1 dp, 1 cn, 1 rl, 1 rg, 2 mh) (1.2)
   c. Can make an informed judgment if the frequency and severity of the target disorder warrants the time and resources necessary to screen in a particular practice setting. (2 dp, 2 rg, 2 rl, 1 cn, 2 mh) (1.8)
   d. Can establish a system to incorporate screening and case finding into his/her own practice. (1 dp, 1 rl, 1 cn, 1 rg, 2 mh) (1.2)

5.3. Understands how to decide if a potential therapy is likely to be appropriate and effective for a particular patient. (2 dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.2)
1. Understands treatment effect and effect size of a particular therapy. (1 dp, 2 mh, 2 cn, 1 rg, 1 rl) (1.4)
2. Understands the use of surrogate endpoints and class effect (p. 416) when comparing therapies. [Guyatt 2] (1 mh, 2 dp, 2 rl, 1 rg, 2 cn) (1.6)

Commentary: Guyatt considers this a more advanced topic. Page 416 of his texctbook offers a good discussion.

3. Understands how to implement an N-of-1 trial study. (2 dp, 2 rg, 2 rl, 2 mh, 2 cn) (2.0)
   a. Understands the possible indications for conducting an N-of-1 trial (e.g., the likely lack of effectiveness of conventional treatment, the likelihood that the alternative treatment, if effective, will be continued long-term, and the willingness of the patient to collaborate in designing and carrying out the trial). (2 dp, 2 rl, 1 rg, 2 mh, 2 cn) (1.8)
   b. Can determine the feasibility of conducting a formal N-of-1 trial on a patient in his/her own practice (e.g., based on if the treatment has a rapid enough effect, the treatment ceases to act soon after it is discontinued, the optimal treatment duration is feasible, the relevant outcomes can be measured, sensible criteria for stopping the trial are established, an unblended run-in period can be conducted, the patient is willing and capable of participating, and strategies for interpreting the trial data are in place). (2 dp, 2 rl, 1 rg, 2 cn, 2 mh) (1.8)
   c. Can determine if there are ethical obstacles to conducting an N-of-1 trial. (2 dp, 2 rl, 1 rg, 2 cn, 2 mh) (1.8)
   d. Can determine if the mode of therapy is so experimental that it is necessary to seek approval by a medical research ethics committee or local Chiropractic Board is necessary. (2 dp, 2 rg, 2 rl, 1 cn) (1.8)
4. Understands how to choose and apply clinical decision rules, (1.2) clinical guidelines (1.2) and quantitative clinical decision analysis (CDA) tool to management decisions. (2 dp, 1 rl, 2 rg, 1 cn, 2 mh) (1.6)

5.4. Can apply pertinent evidence to a particular patient situation when estimating potential harm from health care decisions (diagnostic test, treatments, lifestyle choices, etc.). (1 dp, 1 rg, 1 mh, 1 cn, 1 rl) (1.0)
1. Can use appropriate evidence to estimate the patient's risk vs. benefit for a particular procedure. [Guyatt 1] (1 dp, 1 rl, 2 mh, 1 cn, 1 rg) (1.2)

a. Understands numbers needed to harm as it applies to the individual patient. (1 dp, 1 rg, 1 cn, 1 rl, 2 mh) (1.2)
b. Understands the importance of weighing the magnitude of harm. [Guyatt 1] (1 dp, 1 cn, 1 rg, 1 rl, 2 mh) (1.2)
c. Understands the importance of weighing the option of alternative treatment. [Guyatt 1] (1 dp, 1 rg, 1 cn, 1 rl, 2 mh) (1.2)
d. Understands the importance of weighing any corresponding loss of benefit. [Guyatt 1] (1 dp, 1 cn, 1 rg, 1 rl, 2 mh) (1.2)

2. Considers the patient's preferences, concerns and expectations regarding potential harm when choosing a diagnostic or treatment procedure. (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)

Commentary: Material was taken from Straus 2005 (Table 6.6).

Guides for deciding whether valid important evidence about harm can be applied to our patient.

1. Is our patient so different from those included in the study that its results cannot apply?
2. What is our patient's risk of benefit and harm from the agent?
3. What are our patient's preferences, concerns, and expectations from this treatment?
4. What alternative treatments are available?

5.5. Understands and applies prognostic indicators to help predict a patient's outcome. (1 cn, 1 rg, 1 rl, 1 mh, 1 dp) (1.0)
1. Understands the role of natural history on prognosis. (1 dp, 1 mh, 1 rg, 1 cn, 1 rl) (1.0)
2. Can identify risk factors for poorer outcome (e.g., "yellow flags," "red flags" for diseasae, pain severity). (1 dp, 1 mh, 1 cn, 1 rl, 1 rg) (1.0)

5.6. Understands how to select appropriate outcome measures. (1 dp, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)
1. Knows how to choose an outcome measure based on validly, reliability, and responsiveness. (1 dp, 1 rl, 1 rg, 1 c, 1 mh) (1.0)
2. Knows how to match an outcome measure to the health parameter to be monitored. (1 dp, 1 cn, 1 rl, 1 rg, 1 mh) (1.0)
3. Knows how to select an outcome measure based on patient compliance. (1 dp, 1 cn, 1 rg, 2 rl, 1 mh) (1.2)
4. Knows how to select an outcome measure based on ease of administration. (1 dp, 1 cn, 1 rg, 2 rl, 1 mh) (1.2)
5. Knows how to administer and score a variety of commonly used outcome questionnaires (e.g., PSFS, NDI, Oswestry, Roland Morris). (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0)

Commentary: The following material was taken from Liebenson 2007. Outcome Assessment. Steven Yeomans, Craig Liebenson, Jennifer Bolton, and Howard Vernon

Validity: The extent to which a measure is a true estimate of the underlying property.

Longitudinal validity: The capacity of a measure to detect true change over time.

Minimal clinically important difference: The change score that maximizes the accurate classification of those patients who change (improved) an important amount from those who do not.

How is the Minimal Clinically Important Change in an Outcome Determined? A key dimension of responsiveness is the minimal clinically important change in an outcome in a specific patient population. This is the smallest change in the OA score that the patient perceives as beneficial. A patient's own global impression of change (PGIC) (improvement/deterioration) is the most commonly used external criterion to compare the outcome against.

PGIC scores are calculated on the basis of the patients own perception of change with care. A PGIC may ask if the patient is very much improved, much improved, slightly improved, unchanged, or worse with care. The PGIC for improvement has been defined by subtracting the mean OA score of "unchanged" from "much improved" or "very much improved" the PGIC for deterioration has been defined by the subtracting the mean OA score of "unchanged" from "worse".

Another common way responsiveness is determined is by the effect size. This is the size of an effect from a treatment intervention. It is determined from a comparison of different instruments measuring the same thing. The larger the effect size,

the greater the treatment effect (signal) as related to the variability (noise) in the sample. An effect size of 0.2 is small, 0.5 is moderate, and 0.8 or more is large.

Different methods are used to calculate effect size. They each use a ratio with the same numerator of the mean pretreatment score minus post-treatment score across the study population. The denominator is usually the range of scores or standard deviation of the entire group.

In individuals who classify themselves as having improved greatly, a responsive instrument should have a large effect size. Whereas in individuals who classify themselves as not improving, the effect size should be small. Thus, it would be expected that in chronic patients (who are less likely to show improvement) an instruments effect size would be much smaller than in acute patients (who are more likely to show improvement).

Another way of determining when meaningful change in an outcome instrument has occurred is from the minimal detectable change (MDC). This is the amount of error associated with a multiple measures on stable patients (expressed in the same units as a measure). For a change to be significant, it must be equal to or greater than the MDC.

Ceiling and Floor Effects

A ceiling effect occurs when a respondent begins at a high level of function and therefore if they improve, the instrument cannot accurately detect this improvement. An example would be an athlete. A floor effect occurs when a respondent begins at a low level of function and further deterioration in function cannot be detected by the measure. An example is a frail or postoperative person. Ceiling or floor effects are caused by the inability of the instrument to discriminate at the higher or lower end of the dimension being measured. The impact of ceiling and floor effects is that clinically important change will not be measured or detected.

Practicality

An outcome tool should be simple to administer and understand, time-efficient, and easy to score and interpret. Disability questionnaires should have wording that is simple and unambiguous so that patients will easily be able to complete the entire form. Scoring should be possible with a simple computer program that shows a percent improvement over time. “Yes” and “no” responses are ideal for research questionnaires because they are easier to administer with telephonic follow-up. However, HCPs may prefer forms with 0-to 10 visual analog scales that give patients more options for their answers. A practical tool is time- and cost-effective as well as valid, reliable, and responsive.

5.7. Can develop and employ a plan to apply new evidence to the patient’s situation. (1 dp, 1 mh, 1 cn, 1 rg, 1 rl) (1.0)

1. Understands the necessity of blending research evidence with clinical experience and patient’s values and goals (cultural/personal). (1 dp, 1 rl, 1 mh, 1 rg, 1 cn) (1.0)

Commentary: Although perhaps beyond the scope of our basic curriculum, instructors may wish to be aware of another movement that complements EBP by trying to create models which quantify and include patient values. See the following.

From Guadagnino, C, Moving from evidence-based to value-based medicine, Published July 2006

From an interview with Brown, M, author of Evidence-Based to Value-Based Medicine

“Value-based medicine is the practice of medicine based on the value conferred by a systematic intervention. Value is the ability to measure improvement in both length of life and quality of life.

“We ask what length of time one might expect to live and how much of that time one would trade to get a particular outcome, such as perfect vision, or perfect ambulation, or perfect gastrointestinal function. When you ask these questions of many patients the confidence intervals become very small, these numbers become very solid, we can compare them across specialties and across different fields and also use them in economic analyses.

“For example, you start out with a clinical trial about cataract extraction where you take someone from a 20/100 vision to a 20/30 vision. Then you convert those numbers – 20/100 or 20/30 – to value. What quality-of-life standard does the patient have with a vision of 20/100 or 20/30? If you ask this over many patients, you’ll get what the utility is of having vision at 20/30 or 20/100.

“Right now, evidence-based medicine looks at the positive effects of the treatment from a standpoint of a particular function, but when we look at a valued assessment, we look at the value of the adverse effects as well.” [RL 10/9/06]

2. Can appropriately educate, motivate and negotiate patient participation in an evidence-based management plan. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
    a. Understands the basic elements of motivational psychology (e.g., understanding that explaining the facts may not be the most important aspect in changing behavior, coercion typically fails, being sympathetic and supportive of the patient's ideas and attitudes is important, and realizing that the patient, not the doctor, has ultimate control). (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)
    b. Can employ a step by step systematic process to engage the patient in the management plan.
        i. Knows how to introduce the idea of change openly, educating the patient about the evidence in language readily understandable by the patient. (1 dp, 1 rl, 1 cn, 1 r, 1 mh) (1.0)
        ii. Can assess the patient's readiness to change. (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0)
        iii. Demonstrates the ability to notice and take seriously any resistance and obstacles to change. (1 dp, 1 rl, 1 cn, 1 rg, 1 mh) (1.0)
        iv. Demonstrates the ability to negotiate with the patient. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
        v. Can create a plan to circumvent the obstacles to the assessment and management recommendations. (1 dp, 1 rl, 1 rg, 1 cn, 1 mh) (1.0)

 **Commentary:** A number of sources give explicit suggestions on how to approach getting patient compliance.

"When advising patients to make meaningful lifestyle changes, remember these 4 "Ps": Participatory, Personalized, Practical, and Persistent. First, engage the patients in a conversation about their lifestyle habits and partner with them to develop specific, personalized strategies to make improvements. For example, target significant sources of sodium in the specific foods they eat and find practical opportunities for physical activity in the context of their own schedule and circumstances.

"Most importantly, persist in your advice by revisiting lifestyle recommendations and the patients' progress at each visit, and modify as needed. Often, once medications are prescribed, patients disregard the lifestyle changes, and may need repeated encouragement to adopt regular, healthful habits." (Linda N. Meurer, MD, MPH From the J Fam Pract 2006 Nov;55(11):991-3.Clinical inquiries (www.jfponline.com))

Ground rules of Motivational Interviewing

Facts and "truth" are not the most important things in helping people change their behavior.
Coercive pressure toward a particular outcome does not work.
The patient has his or her own ideas, often quite strong ideas, about what the doctor is suggesting, requesting, or prescribing.
The patient's ideas and attitudes toward the doctor's suggestion are extremely important and need to be understood.
The most important single issue is where the patient is relative to our ideas, not how strongly we believe in them.
The patient has the ultimate control because it is the patient who has to enact a particular behavior.

3. Understands the role of the PARQ conference (i.e., a discussion of the procedures, alternatives, risks and an opportunity for questions) and applies it in practice. (1 dp, 1 rl, 1 mh, 1 cn, 1 rg) (1.0)

SELF ASSESS

STANDARD 6

The EBP competent practitioner engages in self evaluation of his/her process for accessing, appraising, and incorporating new evidence into practice.

6. The EBP competent practitioner engages in self evaluation of his/her process for accessing, appraising and incorporating new evidence into practice.

6.1. Demonstrates the behavior necessary to maintain and improve EBP skills. (1 dp, 1 rl, 1 mh, 1 rg, 1 jt, 1 cn) (1.0)

1. Understands the necessity of devoting sufficient time to keep current with expanding health care information and EBP skills. (1 dp, 1 rg, 1 rl, 1 cn, 1 mh) (1.0)
2. Understands to stay current with EBP skills an ongoing financial investment for training and technology is required. (2 dp, 2 rl, 2 rg, 1 cn, 1 mh) (1.6)
3. Understands the need for EBP skills to be efficient and pragmatic. (2 dp, 1 rg, 2 rl, 1 cn, 1 mh) (1.4)
4. Can establish a plan to address the time constraints imposed by a busy clinical practice. (1 dp, 1 rg, 2 rl, 1 cn, 1 mh) (1.2)
5. Understands the need for adequate physical space and hardware to support information searching. (1 dp, 2 rl, 1 rg, 1 cn) (1.3)
6. Understands how to acquire and maintain adequate access to health care information resources and data bases. (1 dp, 2 rl, 1 rg, 1 c, 1 mh) (1.2)

6.2. Reflects on how well these activities are performed and continues to improve them. (1 jt, 1 cn, 1 rg, 1 rl, 1 mh) (1.0)

1. Generates a plan for maintaining and improving EBP competency through regular attendance at EBP workshops. (1 jt, 1 dp, 1 cn, 2 rg, 1 rl, 1 mh) (1.2)
2. Improves information resources as necessary. (1 rl, 1 rg, 1 dp, 1 cn, 1 mh) (1.0)
   a. Considers acquiring *"*push*"* services. (1 rl, 1 jt, 1 dp, 1 rg, 1 cn, 1 mh) (1.0)
   b. Understands how create a system of support utilizing free and propriety data bases and local resources (local chiropractic colleges, medical libraries, etc.). (1 dp, 1 rg, 1 cn, 1 rl, 1 mh) (1.0)
3. Keeps reflective journals to record impression of application of EBP methods. (3 dp, 3 rg, 2 rl, 1 mh, 2 cn) (2.2)

Commentary: The following long excerpt from the Center for EBM is useful [rl 5/4/07]: Practicing EBM – Evaluation. The fifth step in practicing EBM is self-evaluation and we've suggested some approaches for doing this in the tables that follow.

Self-evaluation in asking answerable questions

1. Am I asking any clinical questions at all?
2. Am I asking well-formulated (3-part) questions?
3. Am I using a "map" to locate my knowledge gaps and articulate questions?
4. Can I get myself unstuck when asking questions?
5. Do I have a working method to save my questions for later answering?
6. Is my success rate of asking answerable questions rising?
7. Am I modeling the asking of answerable questions for my learners?
8. Am I writing any educational prescriptions in my teaching?
9. Are we incorporating question asking and answering into everyday activities?
10. How well am I guiding my learners in their question asking?
11. Are my learners writing educational prescriptions for me?

Self-evaluation in finding the best external evidence

1. Am I searching at all?
2. Do I know the best sources of current evidence for my clinical discipline?
3. Have I achieved immediate access to searching hardware, software and the best evidence for my clinical discipline?
4. Am I finding useful external evidence from a widening array of sources?
5. Am I becoming more efficient in my searching?
6. Am I using MeSH headings, thesaurus, limiters, and intelligent, free text when searching MEDLINE?
7. How do my searches compare with those of research librarians or other respected colleagues who have a passion for providing best current patient care?

Self-evaluation in critically appraising the evidence for its validity and potential usefulness

1. Am I critically appraising external evidence at all?
2. Are the critical appraisal guides becoming easier for me to apply?
3. Am I becoming more accurate and efficient in applying some of the critical appraisal measures? (such as likelihood ratios, and NNTs)
4. Am I creating and CATs?

Self-evaluation in integrating the critical appraisal with clinical expertise and applying the result in clinical practice

1. Am I integrating my critical appraisals into my practice at all?
2. Am I becoming more accurate and efficient in adjusting some of the critical appraisal measures to fit my individual patients? (such as pretest probabilities, NNTs etc.)
3. Can I explain (and resolve) disagreements about management decisions in terms of the integration?
4. Have I conducted any clinical decision analyses?
5. Have I carried out any audits of my diagnostic, therapeutic or other EBM performance?

Self-evaluation in teaching EBM

1. When did I last issue an educational prescription?
2. Am I helping my trainees learn how to ask answerable questions?
3. AM I teaching and modeling searching skills?
4. Am I teaching and modeling critical appraisal skills?
5. Am I teaching and modeling the generation of CATs?
6. Am I teaching and modeling the integration of best evidence with my clinical expertise and my patients' preferences?
7. Am I developing new ways of evaluating the effectiveness of my teaching?
8. Am I developing new EBM educational material?

If so, please share them with others and add them to the bank of resources available on this site

Self-evaluation of continuing professional development

1. Am I a member of an EBM-style journal club?
2. Have I participated in or tutored at one of the workshops on how to practice or teach EBM?
3. Have I joined the evidence-based health e-mail discussion group?
4. Have I established links with other practitioners or teachers of EBM?

REFERENCES

Bakken S. An informatics infrastructure is essential for evidence-based practice. J Am Med Inform Assoc 2001 May–Jun; 8(3):199–201.

Cohen AM, Starvi PZ, Hersh WR. A categorization and analysis of the criticisms of evidence-based medicine, Medical Informatics 2004;73;35-43.

Colle 2003

Dawes, M, Summerskill W, Glasziou P et al. Sicily statement on evidence-based practice. BMC Medical Education 2005;5:1-7

Dawes M, Davies P, Gray A, Mant J, Seers K, Snowball R. Evidence-based Practice: A Primer for Health Care Professionals, 2nd ed. Edinburgh: Churchill Linvingstone;2005.

GMC: Tomorrow's Doctors**.** London, General Medical Council' 2002.

Guyatt G, Rennie D (Eds.) User's Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice. The Evidence-Based Medicine Working Group. Chicago, IL: AMA Press;2002.

Hatala R, Keitz SA, Wilson MC, Guyatt G. Beyond journal clubs: moving toward an integrated evidence-based medicine curriculum. Journal of General Internal Medicine, 2006; 21(5):538-541(4)

Rosser WW, Slawson DC, AF Shaughnessy. Information Mastery: Evidence-Based Family Medicine. 2nd Edition. Hamilton, Ontario: BC Decker Inc; 2004.

Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS: Evidence based medicine: what it is and what it isn't**.** *BMJ* 1996,312:71-72.

Straus SE, Richardson WS, Glasziou P, Haynes RB. Evidence-Based Medicine: How to Practice and Teach EBM., 3rd ed. Edinburgh: Churchill Livingstone;2005.

Villanueva-Russel Y. Evidence-based medicine and its implications for the profession of chiropractic. Social Sci & Med 2005;60:545-61.

Evidence-Based Websites

American Family Physician (*AFP*)
http://www.aafp.org/afp/

Canadian Task Force on the Periodic Health Care
www.ctfphc.org

Cochrane Collaboration
www.cochrane.com

Evidence-Based Medicine Librarian
http://emblibrarian.wetpaint.com/

Guideline Advisory Committee
www.gacguidelines.ca

Journal of Family Practice
http://www.jfponline.com

Journal of the American Medical Association (*JAMA*)
http://jama.ama-assn.org/

Med Consult
http://www.mdconsult.com

National Library of Medicine Web site
http://www.nlm.nih.gov/hinfo.html

Netting the Evidence
A ScHARR Introduction to Evidence Based Practice on the Internet
http://www.med.unr.edu/medlib/netting.html

New England Journal of Medicine
http://nejm.org/

ARTICLE
Pathology as art appreciation: melanoma diagnosis. Bandolier [Serial online] 1997;37-2.
http://www.jr2.ox.ac.uk/bandolier/band37/b37-2/html (accessed Apr 9, 2002).

Websites on Evidence-Based Practice

Centre for Evidence-Based Medicine (CEBM) in Oxford www.cebm.net

http://www.infopoems.com/concept/ebm_loe.cfm

EBM websites

http://www.cebm.net/

http://www.cebm.utoronto.ca/syllabi/