Evidence-Based Medicine, Part 5. An Introduction to Critical Appraisal of Articles on Prognosis

Roberto Cardarelli, DO, MPH
Joseph R. Oberdorfer, OMS IV

This article provides an introductory step-by-step process to appraise an article on prognosis. The authors introduce these principles using a systematic approach and case-based format. The process of assessing the validity of an article on prognosis, determining its importance, and applying it to an individual patient is reviewed. The concepts of study population homogeneity, sufficient follow-up periods, and completeness are discussed to help physicians determine an article’s validity. Instruction on how to evaluate confidence intervals is provided. Finally, information that is learned from the previously mentioned steps is applied to patient care. Study generalizability and the role of patient values, expectations, and concerns are also addressed. The skills learned from appraising an article on prognosis in the manner outlined provides a solid basis for life-long learning and improved patient care.

[Editor’s note: This article is part 5 of a six-article series intended to introduce the principles of evidence-based medicine (EBM) to busy clinicians, physician residents, and medical students. Because the application of EBM is a career-long process, further training is needed beyond the information provided within this article and series. A foundation of knowledge about research methods is critical in understanding EBM; however, such details, though introduced, are beyond the scope of this series.]

J Am Osteopath Assoc. 2007;107:315-319

After patients receive a diagnosis, a common question they ask their physicians is, “What is my prognosis?” Clear prognoses are important for patient care and counseling, but they also guide the physician and patient in selecting the most effective treatment. This complex process requires physicians to connect certain clinical predictors of disease and associated comorbidities to the temporal outcome of that disease.

However, not every physician can confidently provide appropriate prognoses in terms of the evidence available in the most current medical literature. For most physicians, prognostic skills develop with practice and experience and alongside clinical research. A framework exists in which any physician can master the art of making a prognosis and helping patients make informed decisions on the basis of available evidence provided by the current literature, the use of one’s clinical expertise, and the unique circumstances relevant to a patient. This approach to medicine, called evidence-based medicine (EBM), is a way for physicians to accelerate their effectiveness in practicing various clinical skills, such as prognosis.

Applying EBM to prognoses requires that physicians thoroughly and critically review the current pertinent literature to verify the accuracy and validity of the evidence and its applicability to a specific clinical case.

In this article, we introduce a strategy for busy physicians, physician residents, and medical students to critically assess the medical literature on prognosis. In-depth details of research methods are beyond the scope of this introductory series on EBM. Readers are encouraged to seek further training on these topics with supplemental learning opportunities and continuing medical education. Finally, the clinical scenario described has been simplified to provide readers with an illustrative example for the general concepts introduced.

Searching the Evidence

Choosing a high quality article to evaluate is a critical first step in practicing EBM. The preferred articles to support prognoses are systematic evidence-based reviews of studies of the specific disease or disorder in peer-reviewed journals. Unfortunately, systematic reviews on prognosis are sometimes hard to find. In such instances, individual research articles can be used—with caution.

Often the best study designs for prognostic articles are cohort studies (Figure 1). In such studies, investigators identify a group of individuals with a previously specified characteristic of interest and follow up with them over a predetermined period of time. The results of the group with the disease of interest are then compared with normal subjects who do not have the disease. Other study designs, such as randomized controlled trials (RCTs) and case-control studies,
Validity of Articles on Prognosis

The process of critically appraising an article on prognosis begins with evaluating the quality of information provided in the selected article. To ascertain the validity of an article, physicians need to determine not only if the methods used to arrive at the conclusions were free of error and bias but also if the study’s results and conclusions were accurately deduced. Physicians may use the following questions to help them determine the validity of an article on prognosis:

- Are the study subjects truly representative of the population of interest?
  A study of an entire population with a specific disease or disorder would be impractical. Instead, studies analyze a representative sample of a patient population and generalize the results to a larger population. The representative sample used in a study can have a direct impact on the usefulness of the results. For example, if the study’s population has substantially different characteristics (eg, age, race or ethnicity, comorbidity) compared with a clinician’s patient, the applicability of the results in this particular instance becomes questionable. Therefore, a study population may not be completely representative of the general population or a particular clinician’s patients.

- Were the study subjects chosen at a common point in the disease course?
  It is important to determine the stage of the disease that researchers began to evaluate the study’s participants because the stage of the disease can impact whether the results are valid and applicable to a specific patient. For example, if the subjects in a study are all in an advanced stage of the disease (eg, New York Heart Association [NYHA] class III or IV for congestive heart failure [CHF]), a physician whose patient is in an earlier stage of the disease (eg, NYHA class I or II for CHF) might incorrectly deduce that his or her patient has a small amount of time before the patient will experience an expected endpoint. To assess the applicability of the article regarding the disease stage of recruited study participants, physicians should look at the article’s abstract and its “Methods” section (Figure 2).

- Was the follow-up period sufficiently long and complete?
  Follow-up in a valid study should be sufficiently long and complete. Short follow-up periods may allow too little time for an outcome of interest to occur. The appropriate length of the study is dependent on the study question, the intervention used, the outcomes of interest, and any special circumstances (eg, funding).

Clinical Scenario

A 65-year-old woman visits your primary care clinic for her annual physical. She tells you that she has noticed shortness of breath after walking up the stairs to work, after walking from her car to her home, and even while sitting at her desk. She presents to you with dyspnea and restlessness. The patient’s medical history is negative for major illnesses or surgeries. The family history reveals that her father died of a disorder related to his heart and her mother died while she was very young. She reports that she drinks a glass of wine every night and denies smoking. On physical examination, her blood pressure is 183/95 mm Hg and her heart rate is 102 beats per minute. There is bilateral pitting edema of both ankles. During the pulmonary examination, you hear crackles in the right and left lower lobes of her lungs, and her cardiac examination is significant with an S3 gallop on auscultation. You suspect class III heart failure based on New York Heart Association classification and begin a full workup.

After the blood work and echocardiograph you mention to the patient that she is suffering from congestive heart failure (CHF) and that you used a test that measured B-type natriuretic peptides (BNP) to support your diagnosis.

Your patient is visibly worried about having something called “heart failure” and wishes to know what will happen to her and what to expect from this disease. You decide to use EBM techniques to evaluate the usefulness of BNP in determining the likelihood that she will have a cardiac event in the future.

You are interested in studies that address the benefit of BNP to predict future cardiac events in patients with CHF. A PubMed search of the terms “future cardiac events,” “dyspnea,” and “B-type natriuretic peptide” yields a cohort study titled “B-type natriuretic peptide predicts future cardiac events in patients presenting to the emergency department with dyspnea.” You decide to review this study.

(continued)
Clinical Scenario (continued)

In the “Methods” section of the study, the authors report that patients presenting to an emergency department were eligible to participate in the study if they presented with shortness of breath at rest, on exertion, or lying down as their chief complaint. The demographic characteristics include a history of chronic heart failure (CHF), chronic obstructive pulmonary disease, and coronary artery disease, and a mean age of 65 years.

This study has limitations in its applicability to your patient because it is not specific to CHF. In addition, the outcome of this study looked at the common chief complaint of shortness of breath, not diagnosis of CHF. Also, patients with different presentations of dyspnea were included. You decide to continue reviewing the article but will make sure to inform the patient of these limitations and how they affect any conclusions that can be made.

(continued)

Figure 2. Clinical scenario (continued).

A clinical investigation is considered complete when all study participants are accounted for at the study’s completion. Subjects who dropped out of the study early because of adverse effects or death should be included and analyzed with their original group assignment. This is called an intention-to-treat analysis.

Failure to perform an intention-to-treat analysis may direct researchers to provide readers with misleading results. The “5-and-20 rule” can be used by a critical reader to evaluate a study’s completeness. If less than 5% of the study population is lost to follow-up, one can be assured that the loss minimally impacted the results. If, however, more than 20% of the study population is lost to follow-up, caution is advised when making clinical decisions based on study findings. An attrition rate of 5% to 20%—and its impact on the researchers’ results—must be determined by the reader based on other specifics of the study (Figure 3). For example, a study design that is complex with high participant burden or a long follow-up period can contribute to a large number of patients lost to follow-up.

Were the study subjects and investigators blinded to the measures of interest? Assessment should also be blinded to prevent observer bias, which occurs when the researcher unconsciously (or consciously) looks harder for outcomes in the study group than in the control group. Blinding means that the researcher is unaware of the participants’ group assignment. Double-blinding means that study participants are also masked to group allocation. In prognosis articles, which are often presented as cohort studies, clinicians are usually single-blinded. In other words, clinicians are blinded to the method by which the outcome is measured, but not to the characteristics of the patients involved in the study (Figure 4).

Were adjustments made for important prognostic factors? Studies on prognosis often stratify the study group into cohorts based on comorbid conditions that influence prognostic outcomes. These factors can include the stage of disease or other diseases, such as coronary artery disease, that may affect the prognosis. In addition, factors that may confound the prognostic outcomes, such as age, sex, and environmental risk factors (eg, living conditions, pollution) must be accounted for. It is important to make sure that the authors of any study have identified factors applicable to their study’s participants and determined if any of these factors influenced the predictive value of their conclusions (Figure 5).

Study Results
After the quality of a study on prognosis has been ascertained, physicians must determine what the results are. Evaluating the endpoints and projected outcomes can help physicians analyze the results and determine their importance.

How likely are the projected outcomes over time?

The study shows that none of the 325 patients who entered the study were lost to follow-up. Although this result is initially reassuring, you soon realize that no patients were lost—most likely because the follow-up period was only 6 months. This relatively short follow-up period limits the applicability of the study’s results to your patient to only those months where patients were observed.

In addition, the authors define endpoint as repeat emergency department visits for chronic heart failure, hospital admission (cardiac), or death (cardiac and noncardiac). The high percentage of patients who reached an endpoint during a short follow-up leads you to read further and test the design against more standards. (continued)
This question deals with the common question most physicians receive from their patients: “What are my chances of having another event?” Valid and accurate studies allow physicians to respond to this question quickly and within a reasonable amount of confidence. Often these findings can be located in the “Results” section of an article. One caveat that must be considered is the study endpoint. As presented in Figure 6, the endpoint can be defined as (1) an emergency department visit, (2) hospitalization as a result of CHF, or (3) death.

Many times, studies provide a more stratified endpoint, such as only death caused by CHF, which may make the prognosis results more clinically applicable by allowing the clinician to provide information that is specific to the patient’s concern. Moreover, composite endpoints vary in the severity of the outcome, making the results less clinically applicable. For example, an emergency department visit is substantially different from death.

### What is the precision of the relationship between the endpoint and the outcome?
Odds ratios and relative risks are estimates of some unknown “true” value. If one were to repeat the study with different subjects, it would yield similar, but not identical, results. The results of repeat studies are dispersed above and below the “true” value. This sampling variation is referred to as precision. The principle measure of precision is the 95% confidence interval (CI). The 95% CI quantifies the uncertainty of a measurement by reporting a range of values within which we can be 95% certain that the true value lies for the entire population. Results are said to be significant if the 95% CI does not include 1.0. If the 95% CI includes 1.0, there is a 95% chance that there is no difference in the outcome between the comparison groups (Figure 7).

### Practical Use
Having verified the validity and importance of a study on prognosis, a physician’s ultimate consideration of the research would then become how the study’s results and the authors’ conclusions can be applied in practice. Making an evidence-based decision requires skills that integrate the patient’s unique values and circumstances with current and valid evidence. One of the goals of EBM is to assist physicians in developing a system for backing up clinical decisions with sound evidence. Those who criticize EBM presumably focus on the first goal of treating patients according to the evidence. However, the goal of EBM is twofold. The second, and most important, goal of EBM is to shift physicians’ focus from the disease to the patient, restoring the patient to the center of the clinical decision-making process—and better aligning the healthcare professions with longstanding osteopathic principles and practice.
Clinical Scenario (continued)

Looking at the study, which lists the relative risk of reaching a chronic heart failure (CHF) endpoint, you see that participants with B-type natriuretic peptides (BNP) levels over 230 pg/mL have a 15.5% greater risk of a CHF event than subjects with a BNP below 230 pg/mL. The confidence interval (CI) for this value is 6.2-43.7. This wide range of the CI brings into question how sure you can be about the results of this study.

(continued)

Figure 7. Clinical scenario (continued).

The result of your patient’s B-type natriuretic peptides is 360 pg/mL, which places her at increased risk for a future chronic heart failure event. After evaluating the study, you question the applicability of the results to your patient. The subjects of the study were predominantly men, and many of them were on some form of treatment or had a history of chronic obstructive pulmonary disease or coronary artery disease. These characteristics do not apply to your patient. You explain to the patient that the study’s duration was only 6 months and there is no evidence beyond that point. You also explain that you searched the literature and that this study was fairly limited in its impact on treatment.

Figure 8. Clinical scenario (continued).

- Are the study subjects substantially different from your specific patient?
  The findings of a valid, important study are considered applicable if the patient to be treated is similar to those described in the study. As mentioned earlier, such information can be accessed by reviewing the study population characteristics.

- How important is the evidence for your specific patient?
  If the evidence from the study shows that the prognosis for the patient is disastrous without treatment, then this information should be presented to and discussed with the patient (Figure 8). The risks and benefits must be weighed while taking into account the patient’s special circumstances (eg, financial concerns, health insurance coverage, and personal wishes). Any evidence becomes moot if the patient is not informed or willing to undergo further diagnostic tests or treatment.

Conclusion

Although most clinicians are already incorporating EBM principles in their practices, often instinctively, some physicians may require a more organized approach to integrating this relatively new model of self-education. Improved comfort levels and true expertise in the practice of EBM are the result of additional education, repetition, and self-assessment. The principles of EBM allow physicians to stay informed while also improving the quality of the information communicated to patients during patient encounters. The systematic approach that is used to appraise an article on prognosis is but one step in practicing EBM. Remember, the goal is always to provide the best care possible to patients—using one’s clinical expertise to address patient values and expectations for treatment.

References