

# Users' Guides to the Medical Literature

## XV. How to Use an Article About Disease Probability for Differential Diagnosis

W. Scott Richardson, MD

Mark C. Wilson, MD, MPH

Gordon H. Guyatt, MD, MSc

Deborah J. Cook, MD, MSc

James Nishikawa, MD

for the Evidence-Based Medicine Working Group

### CLINICAL SCENARIO

You are an experienced clinician working at a hospital emergency department. One morning, a 33-year-old man presents with palpitations. He describes the new onset of episodes of fast, regular chest pounding, which come on gradually, last 1 to 2 minutes, and occur several times a day. He reports no relation of symptoms to activities and no change in exercise tolerance. He is very anxious and tells you he fears heart disease. He has no other symptoms, no personal or family history of heart disease, and takes no medications. His heart rate is 90 bpm and regular, and physical examination of his eyes, thyroid gland, lungs, and heart is normal. His 12-lead electrocardiogram is normal, without arrhythmia or signs of pre-excitation.

You suspect his anxiety explains his palpitations, that they are mediated by hyperventilation, and are possibly part of a panic attack. While he has no findings of cardiac arrhythmia or hyperthyroidism, you wonder if these disorders are common enough in the emergency department setting to consider seriously. You reject pheochromocytoma as too unlikely. Thus, you

can list causes of palpitations, but want more information about the frequency of these causes to choose a diagnostic work-up. You ask the question: "In patients presenting with palpitations, what is the frequency of underlying disorders?"

### THE SEARCH

Your emergency department computer networks with the medical library, where MEDLINE is on CD-ROM. In the MEDLINE file for current years, you enter 3 text words: *palpitations* (89 citations), *differential diagnosis* (7039 citations), and *cause or causes* (71 848 citations). You combine these sets, yielding 17 citations, including an article by Weber and Kapoor<sup>1</sup> that promises to have what you want.

Sick persons seldom present with the diagnosis already made; instead, they present with 1 or more symptoms. These symptoms prompt the clinician to gather information through history and physical examination, identifying clinical findings that suggest explanations for the symptom(s). For example, in an older woman presenting with generalized pruritis, the clinician could identify recent anorexia and weight loss, along with jaundice and the absence of a rash. For most symptoms, the clinician must consider multiple causes for the patient's findings.

*Differential diagnosis* is the method by which the clinician considers the possible causes of a patient's clinical

findings before making a final diagnosis.<sup>2,3</sup> Experienced clinicians often group the findings into meaningful clusters, summarized in brief phrases about the symptom, body location, or organ system involved, such as "generalized pruritis," "painless jaundice," and "constitutional symptoms" for the older woman mentioned earlier. We call these clusters *clinical problems*<sup>3,4</sup> and include problems of biological, psychological, or sociological origin.<sup>5</sup> It is for these clinical problems, rather than for the final diagnosis, that the clinician selects a patient's differential diagnosis.

When considering a patient's differential diagnosis, how is the clinician to decide which disorders to pursue? If the clinician were to consider all known causes equally likely and test for them all simultaneously (the *possibilistic* ap-

**Author Affiliations:** Departments of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario (Drs Guyatt, Cook, and Nishikawa); the Department of Ambulatory Care, Audie L. Murphy Memorial Veterans Hospital and the Department of Medicine, University of Texas Health Sciences Center at San Antonio (Dr Richardson); and the Department of Medicine, Bowman Gray School of Medicine and Wake Forest University Baptist Medical Center, Winston-Salem, NC (Dr Wilson).

The original list of members (with affiliations) appears in the first article of this series (*JAMA*. 1993;270:2093-2095). A list of new members appears in the 10th article of the series (*JAMA*. 1996;275:1435-1439). The following members contributed to this article: Les Irwig, MBBCh, PhD; Virginia Moyer, MD, MPH; Thomas B. Newman, MD, MPH; David L. Sackett, MD, MSc; Jack Sinclair, MD; and John W. Williams, Jr, MD, MHS.

**Corresponding Author and Reprints:** Gordon H. Guyatt, MD, MSc, McMaster University Health Sciences Centre, 1200 Main St W, Room 2C12, Hamilton, Ontario, Canada, L8N 3Z5.

**Users' Guides to the Medical Literature Section Editor:** Drummond Rennie, MD, Deputy Editor (West), *JAMA*.

proach), then the patient would undergo unnecessary testing. Instead, the experienced clinician is selective, considering first those disorders that are more likely (a *probabilistic* approach), more serious if left undiagnosed and untreated (a *prognostic* approach), or more responsive to treatment if offered (a *pragmatic* approach). Prior articles in this series showed how to use evidence about prognosis<sup>6</sup> and therapy,<sup>7-9</sup> so this article will focus on using evidence about disease probability.

Wisely selecting a patient's differential diagnosis involves all 3 considerations (probabilistic, prognostic, and pragmatic), as depicted in TABLE 1. The clinician's single best explanation for the patient's clinical problem(s) can be termed the *leading hypothesis* or *working diagnosis* (group 1 in Table 1). A few (usually 1-5) other diagnoses, termed *active alternatives* (group 2 in Table 1), may be worth considering further at the time of the initial work-up because they are likely, or serious, or treatable. Additional causes of the clinical problem(s), termed *other hypotheses* (group 3 in Table 1), may be too unlikely to consider at the time of the initial diagnostic work-up, but remain possible and could be considered further if the working diagnosis and active alternatives are later disproved.

Using this framework for the patient with palpitations in the scenario, you consider anxiety the working diagnosis, and you wonder whether cardiac arrhythmias, hyperthyroidism, or pheochromocytoma are active alternatives (group 2 in Table 1) or other hypotheses (group 3 in Table 1).

Selecting a patient-specific differential diagnosis should strongly influence diagnostic testing. For the leading hypothesis, the clinician may choose to confirm the diagnosis, using a highly specific test with a high likelihood ratio for a positive result.<sup>10,11</sup> For the active alternatives, the clinician would choose to exclude these diagnoses, using highly sensitive tests with low likelihood ratios for negative results. Usually, the clinician would not order tests initially for the other hypotheses.

How can information about disease probability help clinicians select patients' differential diagnoses? We will illustrate with some brief cases. First, consider a patient who presents with a painful eruption of grouped vesicles in the distribution of a single dermatome. An experienced clinician would make a diagnosis of herpes zoster and turn to thinking about whether to offer the patient therapy. Using Table 1, the working diagnosis is zoster and there are no active alternatives. In other words, the probability of zoster is so high (near 1.0 or 100%) that it is above a threshold where no further testing is required.

Next, consider a previously healthy athlete who presents with lateral rib cage pain after being accidentally struck by an errant baseball pitch. Again, the experienced clinician might rapidly recognize the clinical problem (posttraumatic lateral chest pain), quickly list a leading hypothesis (rib contusion) and an active alternative (rib fracture), and plan a test (radiograph) to exclude the latter. If asked, the clinician could also list disorders that are too unlikely to

consider further (such as myocardial infarction). In other words, while not as likely as rib contusion, the probability of a rib fracture is above the threshold for testing, while the probability of myocardial infarction is below the threshold for testing.

These cases illustrate how clinicians can estimate the probability of disease from the patient's clinical findings, risk factors, exposures, and the like, and then compare disease probabilities to 2 thresholds. The probability above which the diagnosis is sufficiently likely to warrant therapy defines the upper threshold. This threshold is termed the *test-treatment* or simply the *treatment threshold*.<sup>12</sup> In the case of shingles above, the clinician judged the diagnosis of zoster to be above this treatment threshold. The probability below which the clinician believes a diagnosis warrants no further consideration defines the lower threshold. This threshold is termed the *no test-test* or simply the *test threshold*. In the case of post-traumatic torso pain above, the diagnosis of rib fracture fell above, and the diagnosis of myocardial infarction fell below, the test threshold.

Clinicians begin with pretest estimates of disease probability and then adjust the probability as new diagnostic information arrives. Test results are useful when they move our pretest probabilities across 1 of these 2 thresholds. For a disorder with a pretest probability above the treatment threshold, a confirmatory test that raises the probability further would not aid diagnostically. On the other end of the scale, for a disorder with

**Table 1.** Selecting a Patient-Specific Differential Diagnosis

Diagnostic Hypotheses	Description of Hypotheses	Implications for Choosing Diagnostic Tests*	Implications for Choosing Initial Therapy
(1) Leading hypothesis or "working diagnosis"	Single best overall explanation of this patient's problem(s)	Choose test(s) to confirm this disorder, emphasizing high specificity and LR+ much larger than 1	Start initial therapy for this disorder, <i>unless</i> special circumstances exist
(2) Active alternatives or "rule outs"	Not as good as No. 1, but likely, serious, or treatable enough to be actively sought in this patient	Choose test(s) to exclude these disorders, emphasizing high sensitivity and LR- much smaller than 1	Consider starting initial therapy for 1 or more of these, <i>if</i> special circumstances exist
(3) Other hypotheses	Not likely, serious, or treatable enough to be pursued at this point, but not yet excluded	Hold off on tests for these disorders at this point	Hold off on initial therapy of these disorders at this point
(4) Excluded hypotheses	Causes of the problem that have been disproved	No further tests necessary	No treatment necessary

\*LR+ indicates likelihood ratio for a positive result; LR-, likelihood ratio for a negative result.

**Table 2.** Users' Guides for Articles About Disease Probability for Differential Diagnosis

<b>Are the results valid?</b>
Primary guides
Did the study patients represent the full spectrum of those who present with this clinical problem?
Were the criteria for each final diagnosis explicit and credible?
Secondary guides
Was the diagnostic work-up comprehensive and consistently applied?
For initially undiagnosed patients, was follow-up sufficiently long and complete?
<b>What were the results?</b>
What were the diagnoses and their probabilities?
How precise are these estimates of disease probability?
<b>Will the results help in caring for my patients?</b>
Are the study patients similar to my own?
Is it unlikely that the disease possibilities or probabilities have changed since this evidence was gathered?

a pretest probability below the test threshold, an exclusionary test that lowers the probability further would not aid diagnostically. When the clinician believes the pretest probability is high enough to test for and not high enough to begin treatment (ie, between the 2 thresholds), a test could be diagnostically useful if it moves the probability across either threshold.

Where can clinicians get pretest estimates of disease probability? We can use our memories of prior cases with the same clinical problem(s), so that disorders we have diagnosed frequently would have higher probability in the current patient's illness than diagnoses we have made less frequently. Remembered cases are always available and are calibrated to our local practices. Yet our memories are imperfect, and the probabilities we estimate based on them are subject to various biases and errors.<sup>13-15</sup>

Original research constitutes another source of information about disease probability. For example, in a study of diagnostic tests for anemia in the aged, investigators compared blood tests with bone marrow results in 259 elderly persons, finding iron deficiency in 94 patients (36%).<sup>16</sup> Thus, while this study focused on evaluating tests for iron deficiency, it also provides information about disease frequency. Some research studies focus more directly on the

frequency of diseases that cause symptoms,<sup>17</sup> such as the article by Weber and Kapoor<sup>1</sup> on palpitations. This Users' Guide will help you understand direct studies of disease probability, judge their validity, and decide whether to use them for estimating pretest probability for your own patients (TABLE 2).

## THE GUIDES

### Are the Results Valid?

**Did the Study Patients Represent the Full Spectrum of Those Who Present With This Clinical Problem?** This question asks about 2 related issues. First, how do the investigators define the clinical problem? The definition of the clinical problem for study determines the population from which the study patients should be selected. Thus, investigators studying "hematuria" might include patients with microscopic and gross hematuria, with or without symptoms. On the other hand, investigators studying "asymptomatic, microscopic hematuria" would exclude those with symptoms or with gross hematuria.

Such differing definitions of the clinical problem can yield different frequencies of underlying diseases. Including patients with gross hematuria or urinary symptoms might raise the frequency of acute infection as the underlying cause relative to those without symptoms. So assessing the validity of an article about differential diagnosis begins with a search for a clear definition of the clinical problem.

Having defined the target population by clinical problem, investigators next assemble a patient sample. Ideally, the study sample mirrors the target population, so that the frequency of underlying diseases in the sample approximates that of the target population. Such a patient sample is termed *representative*, and the more representative the sample is, the more accurate the resulting disease probabilities. Investigators seldom are able to use the strongest method of ensuring representativeness, obtaining a random sample of the entire population. The next strongest methods are either to include all patients with the clinical prob-

lem from a defined geographic area, or to include a consecutive series of all patients with the clinical problem who receive care at the investigators' institution(s). To the extent that a nonconsecutive case series opens the study to the differential inclusion of patients with different underlying disorders, it compromises study validity.

You can judge the representativeness of the sample by examining the setting from which patients come. Patients with ostensibly the same clinical problem can present to different clinical settings, resulting in different services seeing different types of patients. Typically, patients in secondary or tertiary care settings have higher proportions of more serious diseases or more uncommon diseases than those patients seen in primary care.<sup>18</sup> For instance, in a study of coronary artery disease in 1074 patients with chest pain, a higher proportion of referral practice patients had coronary artery disease than the primary care practice patients, even in patients with similar clinical histories.<sup>19</sup>

To evaluate representativeness, you can also note the methods by which patients were recruited. By considering how investigators identified their patients, how they avoided missing patients, and who was included and who was excluded, you can judge whether important subgroups appear to be missing. The wider the spectrum of patients in the sample, the more representative the sample should be of the whole population, and therefore the more valid the results. For example, in a study of *Clostridium difficile* colitis in 609 patients with diarrhea, the patient sample consisted of adult inpatients whose diarrheal stools were tested for cytotoxin, thereby excluding any patients whose clinicians chose not to test.<sup>20</sup> The inclusion of only those tested is likely to raise the probability of *C difficile* in relation to the entire population of patients with diarrhea.

Weber and Kapoor<sup>1</sup> defined palpitations broadly, as any one of several patient complaints (eg, fast heartbeats, skipped heartbeats, and the like) and included patients with new and recurrent palpitations. They obtained patients from

3 clinical settings (emergency department, inpatient floors, and a medical clinic) in 1 university medical center in a middle-sized North American city. Of the 229 adult patients presenting consecutively for care of palpitations at their center during the study period, 39 refused participation, and the investigators included the remaining 190 patients, including 62 from the emergency department setting. No important subgroups appear to have been excluded, so these 190 patients probably represent the full spectrum of patients presenting with palpitations.

**Were the Criteria for Each Final Diagnosis Explicit and Credible?** Clinicians often disagree about a patient's diagnosis. Such disagreement about final diagnosis could threaten the validity of a study's conclusions about disease frequency. To minimize this threat, investigators can use a set of explicit criteria when assigning each of the final diagnoses. Ideally, these criteria should include not only the findings needed to confirm each diagnosis, but also those findings useful for rejecting each diagnosis. For example, published diagnostic criteria for infective endocarditis include criteria for both verifying the infection and for rejecting it.<sup>21,22</sup> Investigators can then classify patients into diagnostic groups that are mutually exclusive, with the exception of patients whose symptoms are from more than 1 cause. This allows clinicians to understand which diagnoses remain possible for any undiagnosed patients after the investigators have ruled out alternatives.

Ideally, studies of disease probability would also measure the agreement beyond chance for the clinicians assigning diagnoses, as was done in a study of the causes of dizziness.<sup>23</sup> The greater the agreement, the more reproducible and credible are the diagnostic assignments.

While reviewing the diagnostic criteria, keep in mind that "lesion finding" is not necessarily the same thing as "illness explaining." In other words, by using explicit and credible criteria, a patient may be found to have 2 or more disorders that might explain the clinical prob-

lem, causing some doubt as to which disorder is the culprit. Better studies of disease probability will include some assurance that the disorders found actually do explain the patients' illnesses. For example, in a sequence of studies of syncope, investigators required that the symptoms occur simultaneously with an arrhythmia before that arrhythmia was judged to be the cause.<sup>24</sup> Alternatively, in a study of chronic cough, investigators gave cause-specific therapy and required positive treatment responses to confirm the final diagnoses.<sup>25</sup>

Weber and Kapoor<sup>1</sup> developed a priori explicit and credible criteria for confirming each possible disorder causing palpitations and listed their criteria in an appendix along with supporting citations. They evaluated study patients prospectively and assigned final diagnoses using the explicit criteria. Wherever relevant, such as for cardiac arrhythmias, they required that the palpitations occur at the same time as the arrhythmias for that cause to be diagnosed. They do not report on agreement for these assignments.

**Was the Diagnostic Work-up Comprehensive and Consistently Applied?** This criterion addresses 2 closely related issues. First, have the investigators evaluated their patients thoroughly enough to detect any of the important causes of this clinical problem? Within reason and ethics, the more comprehensive the work-up, the lower the chance that invalid conclusions about disease frequency will be reached. For example, in a retrospective study of stroke in 127 patients with mental status changes, the diagnostic work-up was reported to include neurological examination and neuroimaging; a comprehensive search for other causes of delirium was not described, and 118 cases remained unexplained.<sup>26</sup> As the investigators do not describe a complete and systematic search for the causes of delirium, the disease probabilities appear less credible.

The second issue is how consistently the diagnostic work-up was applied. This does not mean that every patient must undergo every test. Instead, for many clinical problems, the clini-

cian performs a detailed, yet focused, history; a problem-oriented examination of the involved organ systems; and a few initial tests. Then, depending on the clues discovered, further inquiry proceeds down one of multiple-branching pathways. Ideally, investigators would evaluate all patients with the same initial work-up, and then "follow the clues" using prespecified testing sequences. Once a definitive test result confirms a final diagnosis, then further confirmatory testing is unnecessary and unethical.

You may find it easy deciding whether the patients' illnesses have been well investigated if they were evaluated prospectively using a predetermined diagnostic approach. It becomes harder to judge when patients are studied only after their investigation is complete or when investigation is not standardized. For example, in a study of precipitating factors in 101 patients with decompensated heart failure, while all patients underwent history and physical examination, the lack of standardization of subsequent testing makes it difficult to judge the accuracy of the disease probabilities.<sup>27</sup>

Weber and Kapoor<sup>1</sup> evaluated their patients' palpitations prospectively using 2 principal means, a structured interview completed by one of the investigators, and the combined diagnostic evaluation (ie, history, examination, and testing) chosen by the physician seeing the patient at the index visit. In addition, all patients completed self-administered questionnaires designed to assist in detecting various psychiatric disorders. A majority of patients (166/190) had electrocardiograms, and large numbers had other testing for cardiac disease as well. Thus, the diagnostic work-up was reasonably comprehensive for common disease categories, although not exhaustive. Since the subsequent testing ordered by the physicians was not fully standardized, some inconsistency may have been introduced, although it does not appear likely to have distorted the probabilities of common disease categories such as psychiatric or cardiac causes.

**For Initially Undiagnosed Patients, Was Follow-up Sufficiently Long and Complete?** Even when investigators use explicit diagnostic criteria after a comprehensive evaluation that is consistently applied, some patients' clinical problems may remain unexplained. The higher the number of undiagnosed patients, the greater the chance of error in the estimates of disease probability. For example, in a retrospective study of various causes of dizziness in 1194 patients at an otolaryngology clinic, about 27% remained undiagnosed.<sup>28</sup> With more than a quarter of patients' illnesses unexplained, the disease probabilities for the overall sample might be inaccurate.

If the study evaluation leaves any patients undiagnosed, investigators can follow these patients over time, searching for additional clues leading to eventual diagnoses and observing the prognosis. The longer and more complete the follow-up, the greater our confidence in the benign nature of the condition in patients who remain undiagnosed yet unharmed at the study's end. How long is long enough? No answer would correctly fit all clinical problems, but we would suggest 1 to 6 months for symptoms that are acute and self-limited and 1 to 5 years for chronically recurring or progressive symptoms.

Weber and Kapoor<sup>1</sup> identified a diagnosable cause of palpitations in all but 31 (16.3%) of their 190 patients. The investigators followed nearly all of the study patients (96%) for at least a year, during which 1 additional diagnosis was made in those initially undiagnosed (symptomatic correlation with ventricular premature beats). None of the 31 undiagnosed patients had a stroke or died.

### What Were the Results?

**What Were the Diagnoses and Their Probabilities?** The authors of many studies of disease probability display the main results in a table listing the diagnoses made, and the numbers and percentages of patients with those diagnoses. For some symptoms, patients may have more than 1 disease coexisting and contributing to the clinical problem. In these situations, authors often identify

the major diagnosis for such patients and separately tabulate contributing causes. Alternatively, authors sometimes identify a separate, "multiple-cause" group.

Weber and Kapoor<sup>1</sup> present a table that tells us that 58 patients (31%) were diagnosed as having psychiatric causes, 82 patients (43%) had cardiac disorders, while 5 patients (2.6%) were found to have thyrotoxicosis and none had pheochromocytoma. This distribution differed across clinical settings: for instance, cardiac disorders were more than twice as likely in patients presenting to the emergency department compared with patients presenting to the outpatient clinic.

**How Precise Are These Estimates of Disease Probability?** Even when valid, these disease probabilities are only estimates of the true frequencies. You can examine the precision of these estimates using their confidence intervals (CIs). If the authors do not provide the CIs for you, you can calculate them yourself using the following formula (for 95% CIs):

$$95\%CI = p \pm 1.96\sqrt{(p[1-p])/n}$$

where  $p$  is the proportion of patients with the cause of interest, and  $n$  is the number of patients in the sample.<sup>29</sup> This formula becomes inaccurate when the number of cases is 5 or fewer, and approximations are available for this situation.<sup>30,31</sup>

For instance, consider the category of psychiatric causes of palpitations in the study by Weber and Kapoor.<sup>1</sup> Using the above formula, we would start with  $p = 0.31$ ,  $(1 - p) = 0.69$ , and  $n = 190$ . Working through the arithmetic, we find the CI to be  $0.31 \pm 0.066$ . Thus, while the most likely true proportion is 31%, it may range between 24.4% and 37.6%.

Whether you will consider the CIs sufficiently precise depends on where the estimated proportion and CI fall in relation to your test or treatment thresholds. If both the estimate and the entire 95% CI are on the same side of your threshold, then the result is precise enough to allow firm conclusions about disease probability for use in planning tests or treatments. Conversely, if the confidence limit around the estimate crosses your threshold, the result may

not be precise enough for definitive conclusions about disease probability. You might still use a valid but imprecise probability result, while keeping in mind the uncertainty and what it might mean for testing or treatment.

Weber and Kapoor<sup>1</sup> do not provide the 95% CIs for the probabilities they found. However, as we just illustrated, if you were concerned about how near the probabilities were to your thresholds, you could calculate the 95% CIs yourself.

### Will the Results Help in Caring for My Patients?

**Are the Study Patients Similar to My Own?** This question concerns whether the clinical setting and patient characteristics are similar enough to yours to allow you to extrapolate the results to your practice. The closer the match, the more confident you can be in applying the results. We suggest you ask yourself whether the setting or patients are so different from yours that you should discard the results.<sup>32</sup> For instance, consider whether your patients come from areas where 1 or more of the underlying disorders are endemic, which could make these disorders much more likely in your patients than was found in the study. Also, consider whether your patients have different cultural patterns of illness behavior or health practices that might cause important differences in the disease probabilities when compared with the patients in the study.

Weber and Kapoor<sup>1</sup> recruited the 190 palpitation patients from those presenting to the outpatient clinics, the inpatient medical and surgical services, and the emergency department (62 of the 190 patients) in 1 university medical center in a middle-sized North American city. Thus, these patients are likely to be similar to the patients seen in your hospital emergency department, and you can use the study results to help inform the pretest probabilities for the patient in the scenario.

**Is It Unlikely That the Disease Possibilities or Probabilities Have Changed Since This Evidence Was Gathered?** As time passes, evidence about disease frequency can become obsolete. Old dis-

eases can be controlled or eliminated. New diseases can arise. Such events can so alter the spectrum of possible diseases or their likelihood that previously valid studies may no longer be applicable to current practice. For example, consider how much the arrival of human immunodeficiency virus disease has transformed the list of possibilities for such clinical problems as generalized lymphadenopathy, chronic diarrhea, and unexplained weight loss.

Similar changes can occur as the result of progress in medical science or public health. For instance, in studies of fever of unknown origin, new diagnostic technologies have substantially altered the proportions of patients with malignancy or unexplained fevers.<sup>33-35</sup> Treatment advances that improve survival, such as chemotherapy for childhood leukemia, can bring about shifts in disease likelihood because the treatment might cause complications, such as secondary malignancy years after cure of leukemia. Public health measures that control some diseases, such as cholera, can alter the likelihood of the remaining causes of the clinical problems that the prevented disease would have caused, in this example, acute diarrhea.

The palpitations study by Weber and Kapoor<sup>1</sup> was published in 1996, and the study period was in 1991. You know of no new developments likely to cause a change in the spectrum or probabilities of disease in patients with palpitations.

## RESOLUTION OF THE SCENARIO

Using the structure outlined in Table 1, your "leading hypothesis" is that acute anxiety is causing your patient's palpitations. You offer the patient a more in-depth discussion of his psychosocial situation as the next test to explore this diagnosis (ie, the pretest probability is above your test threshold). At the same time, you do not feel that anxiety is so certain that you can stop considering other disorders (ie, the pretest probability is below your threshold for treatment without testing). After reviewing the palpitations study by Weber and

Kapoor,<sup>1</sup> you decide to include in your "active alternatives" some cardiac arrhythmias (as common, serious, and treatable) and hyperthyroidism (less common but serious and treatable), so you arrange testing to exclude these disorders (ie, these are above your test threshold). Finally, given that none of the 190 study patients had pheochromocytoma, and since your patient has none of the other clinical features of this disorder, you place it into your "other hypotheses" category (ie, below your test threshold) and decide to hold off on testing for this condition.

We recommend applying these Users' Guides to identify good evidence on which to base initial estimates of disease probability for use in differential diagnosis. As you apply this evidence, keep in mind that selecting a patient's differential diagnosis wisely includes not only considering how likely various disorders are, but also how serious are the various diseases if left undiagnosed and untreated, and how much other clinical actions, like treatment or public health measures to reduce disease spread, could help the patient or the community.

## REFERENCES

- Weber BE, Kapoor WN. Evaluation and outcomes of patients with palpitations. *Am J Med.* 1996;100:138-148.
- Sox HC, Blatt MA, Higgins MC, Marton KI. *Medical Decision Making.* Boston, Mass: Butterworth; 1988.
- Barondess JA, Carpenter CCJ, eds. *Differential Diagnosis.* Philadelphia, Pa: Lea & Febiger; 1994.
- Glass RD. *Diagnosis: A Brief Introduction.* Melbourne, Australia: Oxford University Press; 1996.
- Engel GL. The need for a new medical model: a challenge for biomedicine. *Science.* 1977;196:129-136.
- Laupacis A, Wells G, Richardson WS, Tugwell P, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, V: how to use an article about prognosis. *JAMA.* 1994;272:234-237.
- Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, A: are the results of the study valid? *JAMA.* 1993;270:2598-2601.
- Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, B. *JAMA.* 1994;271:59-63.
- Dans AL, Dans LF, Guyatt GH, Richardson WS, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, XIV: how to decide on the applicability of clinical trial results to your patient. *JAMA.* 1998;279:545-549.
- Jaeschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, III: how to use an article about a diagnostic test, A. *JAMA.* 1994;271:389-391.
- Jaeschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, III: how to use an article about a diagnostic test, B. *JAMA.* 1994;271:703-707.
- Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med.* 1980;302:1109-1117.
- Tversky A, Kahneman D. Judgment under uncertainty. *Science.* 1974;185:1124-1131.
- Dawson NV, Arkes HR. Systematic errors in medical decision making: judgment limitations. *J Gen Intern Med.* 1987;2:183-187.
- Kassirer JP, Kopelman RI. Cognitive errors in diagnosis. *Am J Med.* 1989;86:433-441.
- Guyatt GH, Patterson C, Ali M, et al. Diagnosis of iron deficiency anemia in the elderly. *Am J Med.* 1990;88:205-209.
- Kroenke K. Symptoms and science: the frontiers of primary care research [editorial]. *J Gen Intern Med.* 1997;12:509-510.
- Fletcher RH, Fletcher SW, Wagner EH. *Clinical Epidemiology: The Essentials.* 3rd ed. Baltimore, Md: Williams & Wilkins; 1996:61-64.
- Sox HC, Hickam DH, Marton KI, et al. Using the patient's history to estimate the probability of coronary artery disease. *Am J Med.* 1990;89:7-14.
- Katz DA, Bates DW, Rittenberg E, et al. Predicting *Clostridium difficile* stool cytotoxin results in hospitalized patients with diarrhea. *J Gen Intern Med.* 1997;12:57-62.
- von Reyn CF, Levy BS, Arbeit RD, et al. Infective endocarditis: an analysis based on strict case definitions. *Ann Intern Med.* 1981;94:505-517.
- Durack DT, Lukes AS, Bright DK, and the Duke Endocarditis Service. New criteria for diagnosis of infective endocarditis. *Am J Med.* 1994;96:200-209.
- Kroenke K, Lucas CA, Rosenberg ML, et al. Causes of persistent dizziness: a prospective study of 100 patients in ambulatory care. *Ann Intern Med.* 1992;117:898-904.
- Kapoor WN. Evaluation and outcome of patients with syncope. *Medicine.* 1990;69:160-175.
- Pratter MR, Bartter T, Akers S, et al. An algorithmic approach to chronic cough. *Ann Intern Med.* 1993;119:977-983.
- Benbadis SR, Sila CA, Cristea RL. Mental status changes and stroke. *J Gen Intern Med.* 1994;9:485-487.
- Ghali JK, Kadakia S, Cooper R, Ferlinz J. Precipitating factors leading to decompensation of heart failure. *Arch Intern Med.* 1988;148:2013-2016.
- Katsarkas A. Dizziness in aging—a retrospective study of 1194 cases. *Otolaryngol Head Neck Surg.* 1994;110:296-301.
- Altman DG. Confidence intervals [Appendix]. In: Sackett DL, Richardson WS, Rosenberg WMC, Haynes RB, eds. *Evidence-Based Medicine: How To Practice and Teach EBM.* New York, NY: Churchill Livingstone; 1997.
- Hanley JA, Lippman-Hand A. If nothing goes wrong, is everything all right? interpreting zero numerators. *JAMA.* 1983;249:1743-1745.
- Newman TB. If almost nothing goes wrong, is almost everything all right? interpreting small numerators. *JAMA.* 1995;274:1013.
- Glaziov P, Guyatt GH, Dans AL, Dans LF, Straus SE, Sackett DL. Applying the results of trials and systematic reviews to individual patients [editorial]. *ACP Journal Club.* November-December 1998;129:A15-A16.
- Petersdorf RG, Beeson PB. Fever of unexplained origin: report on 100 cases. *Medicine.* 1961;40:1-30.
- Larson EB, Featherstone HJ, Petersdorf RG. Fever of undetermined origin: diagnosis and follow-up of 105 cases, 1970-1980. *Medicine.* 1982;61:269-292.
- Knockaert DC, Vanneste LJ, Vanneste SB, Bobbaers HJ. Fever of unknown origin in the 1980s. *Arch Intern Med.* 1992;152:51-55.