1C

THE PROCESS OF DIAGNOSIS

W. Scott Richardson, Mark Wilson, and Gordon Guyatt

The following EBM Working Group members also made substantive contributions to this section: Peter Wyer, Jonathan Craig, Roman Jaeschke, Jeroen Lijmer, Luz Maria Letelier, Virginia Moyer, C. David Naylor, and Deborah Cook

IN THIS SECTION

CLINICAL SCENARIO

Generating a Differential Diagnosis

It is another busy day in the emergency department and one of your nurse colleagues tells you that a 60-year-old man has presented with a severe cough of 1 day's duration. Immediately, you think, "upper respiratory tract infection; perhaps pneumonia." When you enter the room, you find the patient appears short of breath and is in more distress than you were expecting. Other possible diagnoses spring to mind: could the patient be suffering from acute airflow obstruction, myocardial infarction with pulmonary edema, a pneumothorax, or a pulmonary embolus? You sit down beside the patient and begin taking a history. You ask the nurse to place him on cardiac and pulse oximetry monitors, to start an intravenous line, and to obtain a 12-lead electrocardiogram and a portable chest radiograph.

The patient appears moderately tachypneic but is able to speak in complete sentences. Vital signs show a regular heart rate of 96 bpm, a blood pressure of 140/90 mm Hg, and a respiratory rate of 24/min. In view of his tachypnea, you request a rectal temperature. Oximetry shows a saturation of 93% and you ask that he receive 4 L/minute of oxygen by nasal cannula.

The patient reports that he was previously in excellent health, but began to suffer from a cough about 24 hours previously. There was no preceding or accompanying fever, runny nose, sore throat, headache, or muscular discomfort. However, he did experience several hours of central chest discomfort at the time of the onset of the cough, a discomfort that subsequently resolved. The cough has been productive of only small amounts of clear sputum which, during the past 2 hours, has been flecked with small amounts of bright red blood. During the past 12 hours, the patient has felt increasingly short of breath on minimal activity and now feels short of breath at rest.

Cardiac auscultation reveals no extra heart sounds or murmurs. Abnormal findings on physical examination are limited to decreased breath sounds and crackles at the left base on chest auscultation. The electrocardiogram confirms mild sinus tachycardia, but is otherwise normal, and the chest radiograph shows only a small left pleural effusion with minimal associated opacification. The nurse reports the patient's temperature to be 38.1°C. You draw arterial blood gases and arrange for an urgent ventilation-perfusion scan. The room air blood gas results show a normal PCO_2 and a PO_2 of 70 mm Hg with a saturation of 93%.

While waiting for the results of the ventilation-perfusion scan, you consider how likely the diagnosis of pulmonary embolism is, given the available information. On the one hand, the patient lacks risk factors, cough is a very prominent symptom, and highly suggestive findings for clinical examination (such as pleuritic chest pain) or further investigation (such as a typical electrocardiographic pattern) are absent. On the other hand, you believe you have

ruled out a number of competing diagnoses, including asthma, pulmonary edema, and pneumothorax, and the clinical picture is not typical of pneumonia. You ultimately decide the probability is intermediate, and you mentally commit yourself to a 30% likelihood of pulmonary embolus. When the ventilation-perfusion scan reveals an unmatched segmental defect that you know is associated with a likelihood ratio of 18,[1] you use your likelihood ratio card (see Part 1C2, "Diagnostic Tests") to generate a posttest probability of approximately 90%, and you begin anticoagulation.

THE DIAGNOSTIC PROCESS

Making a diagnosis is a complex cognitive task that involves both logical reasoning and pattern recognition.[2,3] Although the process happens largely at an unconscious level, we can identify two essential steps.

Step 1. In the first step, you enumerate the diagnostic possibilities and estimate their relative likelihood.[4] 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 pruritus," "painless jaundice," and "constitutional symptoms." These clusters, or clinical problems, may be of biologic, psychologic, or sociologic origin, and they are the object of the differential diagnosis. In the opening scenario, we considered a previously healthy 60-year-old man with a clinical problem encompassing a day-long history of cough and dyspnea. The differential diagnosis included a respiratory infection, acute airflow obstruction, myocardial infarction with pulmonary edema, a pneumothorax, and pulmonary embolus.

Step 2. In the second step in the diagnostic process, you incorporate new information to change the relative probabilities, rule out some of the possibilities, and, ultimately, choose the most likely diagnosis. For each diagnostic possibility, the additional information increases or decreases the likelihood. In our scenario, the absence of manifestations that usually accompany an infectious process reduces the likelihood of an upper respiratory tract infection or pneumonia. The central chest discomfort increased the possibility that we could be observing an atypical presentation of a myocardial infarction and prompted the timely electrocardiogram. Physical examination made heart failure a much less likely possibility; pneumonia and pulmonary embolus remained as the competing diagnoses. The chest radiograph failed to provide definitive evidence of pneumonia, necessitating an additional test, the ventilation-perfusion scan.

Thus, with each new finding, we moved, albeit intuitively and implicitly, from one probability, the *pretest probability,* to another probability—the *posttest probability*. Some findings, such as the absence of any sign of pneumothorax on

the chest radiograph, eliminated one of the possibilities (a posttest probability of 0). Prior to the last test, our approach became explicitly quantitative: we committed to a pretest likelihood of 30% and subsequently used information from the literature to arrive at a final, 90% posttest likelihood of pulmonary embolus.

If we know the properties of each of piece of information (and, in the case of pulmonary embolism, if we have strong data for many elements of the diagnostic workup; see Part 2C, "Diagnosis, Examples of Likelihood Ratios"), we can be highly quantitative in our sequential move from pre- to posttest probability. Later in this section, we will show you how.

Because the properties of the individual items of history and physical examination often are not available, you must rely on clinical experience and intuition to predict the extent to which many pieces of information modify your differential diagnosis. For some clinical problems, including the diagnosis of pulmonary embolism, clinicians' intuition has proved remarkably accurate.[1]

CHOICES IN THE DIAGNOSTIC PROCESS

When considering a patient's differential diagnosis, how can you decide which disorders to pursue? If you were to consider all known causes to be equally likely and test for them all simultaneously (the possibilistic approach), 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).

Wisely selecting a patient's differential diagnosis involves all three considerations (probabilistic, prognostic, and pragmatic). Your single best explanation for the patient's clinical problem(s) can be termed the leading hypothesis or working diagnosis. In the opening scenario, a respiratory infection was the leading diagnosis until the final test result became available. A few (usually one to five) other diagnoses, termed active alternatives, may be worth considering at the time of initial workup because of their likelihood, seriousness if undiagnosed and untreated, or responsiveness to treatment. In the scenario, pulmonary embolus entered the differential diagnosis early because of its seriousness and responsiveness to treatment.

Additional causes of the clinical problem(s), termed other hypotheses, may be too unlikely to consider at the time of initial diagnostic workup, but remain possible and could be considered further if the working diagnosis and active alternatives are later disproved. In our scenario, remote possibilities such a pulmonary hemorrhage or collagen vascular disease never entered the active differential diagnosis, but might eventually have done so if we had not confirmed one of the active alternatives.

DIAGNOSTIC AND THERAPEUTIC THRESHOLDS

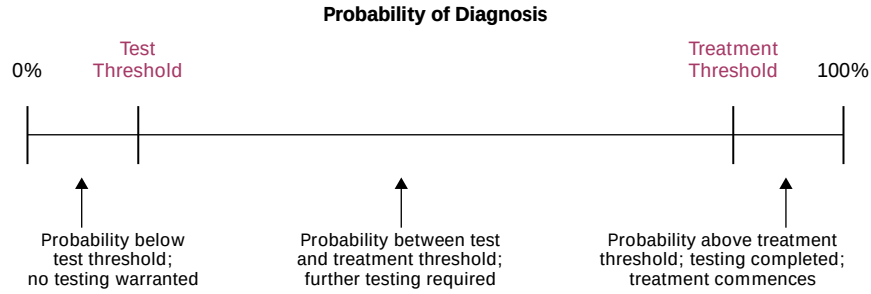
Consider a patient who presents with a painful eruption of grouped vesicles in the distribution of a single dermatome. In an instant, an experienced clinician would make a diagnosis of herpes zoster and would consider whether to offer the patient therapy. In other words, the probability of herpes zoster 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, an experienced clinician would recognize the clinical problem (posttraumatic lateral chest pain), identify 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 a threshold for testing, while the probability of myocardial infarction is below the threshold for testing.

These cases illustrate how you can estimate the probability of disease and then compare disease probabilities to two thresholds (Figure 1C-1). The probability above which the diagnosis is sufficiently likely to warrant therapy defines the upper threshold. That is, if a clinician believes that the diagnosis is sufficiently likely that she is ready to recommend treatment, she has crossed the upper threshold. This threshold is termed the *treatment threshold*.[5] In the case of shingles described above, the clinician judged the diagnosis of herpes zoster to be above this treatment threshold of probability. In our scenario, with the results of the ventilation-perfusion scan we crossed the treatment threshold only after we arrived at a probability of 90% for one of the competing causes, pulmonary embolus.

FIGURE 1C–1

Test and Treatment Thresholds in the Diagnostic Process



The probability below which the clinician decides a diagnosis warrants no further consideration defines the lower threshold. This threshold is termed the no test-test threshold or, simply, the *test threshold*. In the case of posttraumatic torso pain described above, the diagnosis of rib fracture fell above the test threshold and the diagnosis of myocardial infarction fell below it. In our opening scenario, heart failure dropped below the diagnostic threshold when we received the results of the chest radiograph; we did not, for instance, order an echocardiogram. Immune-mediated pulmonary hemorrhage remained below the test threshold throughout the entire investigation.

For a disorder with a pretest probability above the treatment threshold, a confirming test that raises the probability further would not assist diagnostically. On the other end of the scale, for a disorder with a pretest probability below the test threshold, an exclusionary test that lowers the probability further would not help diagnostically. When the clinician believes the pretest probability is high enough to test for and not high enough to warrant beginning treatment (ie, when probability is between the two thresholds), testing will be diagnostically useful, and it will be most valuable if it moves the probability across either threshold.

What determines our treatment threshold? The greater the adverse effects of treating, the more we will be inclined to choose a high treatment threshold. For instance, because a diagnosis of pulmonary embolus involves long-term anticoagulation with appreciable risks of hemorrhage, we are very concerned about falsely labeling patients. The invasiveness of the next test we are considering will also impact our threshold. If results from the next test (such as a ventilation-perfusion scan) are benign, we will be ready to choose a high treatment threshold. We will be more reluctant to institute an invasive test associated with risks to the patient, such as pulmonary angiogram, and this will drive our treatment threshold downward. That is, we will be more inclined to accept a risk of a false-positive diagnosis because a higher treatment threshold implies putting some patients through the test unnecessarily.

Similar considerations bear on the test threshold. The more serious a missed diagnosis, the lower we will set our test threshold. Since a missed diagnosis of a pulmonary embolus could be fatal, we would be inclined to set our diagnostic threshold low. However, this is again counterbalanced by the risks associated with the next test we are considering. If the risks are low, we will be comfortable with our low diagnostic threshold. The higher the risks, the more it will push our threshold upward.

USING SYSTEMATIC RESEARCH TO AID IN THE DIAGNOSTIC PROCESS

How do clinicians generate differential diagnoses and arrive at pretest estimates of disease probability? They remember prior cases with the same clinical problem(s), so that disorders diagnosed frequently have higher probability than diagnoses

made less frequently. Remembered cases are easily and quickly available, and they are calibrated to our local practices. Yet our memories are imperfect, and the probabilities that result are subject to bias and error.[6-8]

Two sorts of systematic investigations can inform the process of generating a differential diagnosis. One type of study addresses the manifestations with which a disease or condition presents (see Part 2C, "Diagnosis, Clinical Manifestations of Disease"). The second—and more important—type of study directly addresses the underlying causes of a presenting symptom, sign, or constellation of symptoms and signs (see Part 1C1, "Differential Diagnosis"). In our opening scenario, the question would be: When patients present with acute cough and shortness of breath, what are the ultimate diagnoses and the relative frequency of these diagnoses?

Having generated an initial differential diagnosis with associated pretest probabilities, how can you incorporate additional information to arrive at an ultimate diagnosis? For each finding, you must implicitly ask: How frequently will this result be seen in patients with one particular diagnostic possibility (or target condition) in relation to the frequency with which it is seen in the competing diagnostic conditions? Once again, you may intuitively refer to your own past experience. Alternatively, you may use data from research studies focusing on test properties. For instance, in our scenario, the Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED) study of ventilation-perfusion scanning in the diagnosis of pulmonary embolism[1] provided the likelihood ratio that allowed calculation of the posttest probability of 90% (see Part 1C2, "Diagnostic Tests").

Some articles provide evidence about differential diagnosis as well as diagnostic test properties. For example, in a study of diagnostic tests for anemia in aged persons, investigators compared blood tests with bone marrow results in 259 elderly persons, finding iron deficiency in 94 (36%).[9] The investigators also reported a diagnosis of anemia in the remaining 165 patients. Thus, although this study focused on evaluating tests for iron deficiency, it also provides information about disease frequency.

In the following sections of the book, we provide guidelines for you to assess the validity of both types of formal investigations related to diagnosis: studies that focus on a constellation of presenting symptoms or signs and determine patients' ultimate diagnoses, and studies that explore the properties of a diagnostic test. In each case, we suggest that validity will depend on the answers to questions regarding two key design features: Did the investigators enroll the right group of patients; and did they undertake the appropriate investigations to determine the true diagnosis? As we deal in sequence with each of the three types of study, we will explain how you can use the results to improve the accuracy of diagnosis in your clinical practice. As for therapy, prognosis, and harm, the systematic reviews of all diagnostic test articles addressing a particular issue will provide the strongest inferences (see Part 1E, "Summarizing the Evidence"). To understand and interpret such reviews, we must use the principles of assessing primary diagnostic studies.

References

1. The PIOPED Investigators. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED). *JAMA*. 1990;263:2753-2759.
2. Sox HC, Blatt MA, Higgins MC, Marton KI. *Medical Decision Making*. Boston: Butterworths; 1988.
3. Glass RD. *Diagnosis: A Brief Introduction*. Melbourne: Oxford University Press; 1996.
4. Barondess JA, Carpenter CCJ, eds. *Differential Diagnosis*. Philadelphia: Lea & Febiger; 1994.
5. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med*. 1980;302:1109-1117.
6. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. *Acad Med*. 1990;65:611-621.
7. Bordage G. Elaborated knowledge: a key to successful diagnostic thinking. *Acad Med*. 1994;69:883-885.
8. Regehr G, Norman GR. Issues in cognitive psychology: implications for professional education. *Acad Med*. 1996;71:988-1001.
9. Guyatt GH, Patterson C, Ali M, et al. Diagnosis of iron-deficiency anemia in the elderly. *Am J Med*. 1990;88:205-209.