Chapter 1 **Diagnostic Testing in Emergency Care**

COPYRIGHTED MATERIAL

As emergency department (ED) physicians, we spend a good deal of our time ordering, interpreting, and waiting for the results of diagnostic tests. When it comes to determining who needs a test to rule out potentially life-threatening conditions, ED physicians are the experts. There are several reasons for this expertise. First and foremost, we see a lot of patients. Especially for those working in busy hospitals, the expectation is to see everyone in a timely way, provide quality care, and ensure patients have a good experience. If we order time-consuming tests on everyone, ED crowding and efficiency will worsen, costs of care will go up, and patients will experience even longer waits than they already do. In addition, the way ED physicians in the United States are paid may be changing over the coming years through mechanisms such as accountable care organizations and payment bundling. There may be more pressure to carefully choose who needs and who does not need tests in an evidence-based manner.

Differentiating which patients will benefit from further testing in the ED is a complex process. Over the past 30 to 40 years, science and research in diagnostic testing and clinical decision rules in emergency care have advanced considerably. Now, there is a greater understanding of test performance regarding the reliability, sensitivity, specificity, and overall accuracy of tests. Validated clinical decision rules exist to provide objective criteria to help distinguish who does and does not need a test. Serious, potentially life-threatening conditions such as intracranial bleeding and cervical spine (C-spine) fractures can be ruled out based on clinical grounds alone. There are also good risk stratification tools to determine a probability of disease for conditions like pulmonary embolism before any tests are even ordered.

How do we decide who to test and who not to test? There are some patients who obviously need tests, such as the head-injured patient who has altered mental status and who may have a head bleed where the outcome

Evidence-Based Emergency Care: Diagnostic Testing and Clinical Decision Rules, Second Edition.
Jesse M. Pines, Christopher R. Carpenter, Ali S. Raja and Jeremiah D. Schuur.
© 2013 John Wiley & Sons, Ltd. Published 2013 by John Wiley & Sons, Ltd.

may be dependent upon how quickly the bleeding can be detected with a computed tomography (CT) scan. There are also patients who obviously do not need tests, such as patients with a simple toothache or a mild upper respiratory tract infection. Finally, there is a large group of patients in the middle for whom testing decisions can sometimes be challenging. This group of patients may leave you feeling "on the fence" about testing. In this large middle category, it may not be clear whether to order a test or even how to interpret a test once you have the results. And when we receive unexpected test results, it may not be clear how best to use those results to guide the care of an individual patient.

Let's give some examples of how diagnostic testing can be a challenge in the ED. You are starting your shift and are signed out a patient for whom your colleague has ordered a D-dimer assay (a test for pulmonary embolism). She is 83 years old and developed acute shortness of breath, chest pain, and hypoxia (room air oxygen saturation = 89%). She has a history of a prior pulmonary embolism and her physical examination is unremarkable, except for mild left anterior chest wall tenderness and notably clear lung sounds. The D-dimer comes back negative. Has pulmonary embolism been satisfactorily ruled out? Should you perform a pulmonary angiogram or a CT scan of the chest, or maybe even consider a ventilation–perfusion (V/Q) scan? Was D-dimer the right test for her to begin with?

Let's consider a different scenario. Consider a positive D-dimer assay in a 22-year-old male with atypical chest pain, no risk factors, and normal physical examination including a heart rate of 70 beats per minute and an oxygen saturation of 100% on room air. What do you do then? Should he be anticoagulated and admitted? Does he have a pulmonary embolism? Should you move forward with further confirmatory testing before initiating treatment? Or is he so low risk that he's probably fine anyway? Of course, you might wonder why, if he was so low risk, was the D-dimer ordered in the first place?

As a third example, you are evaluating a 77-year-old female who has fallen down, has acute hip pain, and is unable to ambulate. The hip radiograph is negative. Should you pursue it? Possibly get a CT or magnetic resonance imaging (MRI)? But even though the hip radiograph is negative, will she be able to go home?

These are examples of when test results do not confirm your clinical suspicion. What do you do in those cases? Should you believe the test result or believe your clinical judgment before ordering the test? Were these the optimal tests in the first place? Remember back to conversations with your teachers in emergency medicine on diagnostic testing. Didn't they always ask, "How will a test result change your management?" and "What will you do if it's positive, negative, or indeterminate?"

The purpose of diagnostic testing is to reach a state where we are adequately convinced of the presence or absence of a condition. Test results are interpreted in the context of the prevalence of the suspected disease state: your clinical suspicion of the presence or absence of disease in the individual patient. For example, coronary artery disease is common. However, if we look for coronary disease in 25 year olds, we are not likely to find it because it is very uncommon in that population. There are also times when your clinical suspicion is so high that you do not need objective testing. In certain patients, you can proceed with treatment. For example, some emergency physicians may choose to treat a dislocated shoulder based on the clinical examination rather than first obtaining a radiograph. However, testing is often needed to confirm a diagnosis or to rule out more severe, life-threatening diseases.

The choice over whether to test or not test in the ED also depends upon the resources of the hospital and of the patient. Some hospitals allow easy access to radiographic testing and laboratory testing. In other hospitals, obtaining a diagnostic test may not be as easy. Some hospitals don't have CT scanners. Others do not have the staff available for certain types of tests at night or on weekends (like MRIs and ultrasounds). Sometimes patients may not need a test if you believe that they are reliable to return if symptoms worsen. For others, you may believe that a patient's emergency presentation may be the only time he or she will have access to diagnostic testing. For example, saying to a patient, "Follow up with your doctor this week for a stress test" may be impractical if the patient does not have a primary doctor or does not have good access to medical care. Many providers practice in environments where they cannot order a lot of tests (like developing countries). You also may practice in an office environment that simply does not have easy access to testing. However, regardless of the reason why we order tests in the ED or other acute settings, what is certain is that the use of diagnostic testing in many cases can change how you manage a patient's care.

Sometimes, you may question your choice of whether to test, to not test, or to involve a specialist early. Should you get a CT scan first or just call a surgeon in for a young male with right lower quadrant pain, fever, nausea, and possible appendicitis? How many cases have you seen where the CT scan has changed your management? What if the patient is a young, nonpregnant female? Does that change your plan?

How about using clinical decision rules in practice? By determining if patients meet specific clinical criteria, we can choose not to test some patients if they are low risk. Do all patients with ankle sprains need X-rays? Can you use the Ottawa ankle rules in children? What are the limits of clinical decision rules? Is it possible to apply the Canadian C-spine rules to a 70-year-old female? What is sufficiently "low risk"? These questions come

up daily in the practice of emergency medicine. In fact, a major source of variability among physicians is whether or not they order tests. Remember back to your training when you were getting ready to present a patient to the attending physician. Weren't you trying to think to yourself, "What would she do in this case? What tests would she order?"

Access to test results helps us decide whether to treat a disease, initiate even more testing, or no longer worry about a condition. The cognitive psychology of clinical decision making has evolved rapidly over the last several decades. As ED physicians, we gain confidence in this process with experience. Much of the empirical science and mathematics behind testing that are described in this book become instinctive and intuitive the longer you practice emergency medicine. Sometimes we may think a patient does not need to be tested because the last hundred patients who had similar presentations all had negative tests. Maybe you or your colleagues were "burned" once when a subtle clinical presentation of a life-threatening condition was missed (like a subarachnoid hemorrhage). The next patient who presents with those symptoms is probably more likely to get a head CT followed by a lumbar puncture. Is this evidence based? Recognizing our individual diagnostic biases is one way to decrease the likelihood of erroneous decision making while increasing efficiency and effectiveness.

Step back for a moment and think about what we do when ordering a test. After evaluating a patient, we come away with a differential diagnosis of both the most common and the most life-threatening possibilities. The following approach to medical decision making was derived by Pauker and Kassirer in 1980.[1] Imagine diagnostic testing as two separate thresholds, each denoted as "I" (for *indeterminate*). The scale at the bottom of Figure 1.1 denotes pretest probability, which is the probability of the disease in question before any testing is employed. In practice, it is often a challenge to come up with a pretest probability, and frequently opinions on pretest probability differ considerably between experienced physicians. However, for the moment, assume that pretest probability is a known quantity.

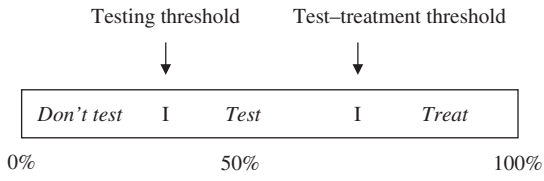

Figure 1.1 Pretest probability of disease. (Source: Data from Pauker and Kassirer (1980)).

In Figure 1.1, the threshold between "don't test" and "test" is known as the *testing threshold*. The threshold between "test" and "treat" is known as the *test–treatment threshold*. In this schema, treatment should be withheld if the pretest probability of disease is smaller than the testing threshold, and no testing should be performed. Treatment should be given without testing if the pretest probability of disease is above the test–treatment threshold. And, when our pretest probability lies between the testing and test–treatment thresholds, the test should be performed and the patients treated according to the test results. That is the theory. But now let's make this more clinically relevant.

Sometimes disease is clinically apparent and we do not need confirmatory testing before proceeding with treatment. If you are evaluating a patient with an obvious cellulitis, you may choose to give antibiotics before initiating any testing. How about a 50-year-old male with acute chest pain who on his electrocardiogram (ECG) has large inferior "tombstone" ST-segment elevations consistent with acute myocardial infarction (AMI)? Cardiac markers will not be very helpful in the acute management of this patient. This is an example of a situation in which it is important to treat the patient first: give the patient aspirin, anticoagulation, beta blockers, and oxygen, and send him off to the cardiac catheterization lab if your hospital has one or provide intravenous thrombolysis if cardiac catheterization is not readily available. Well, now imagine that the patient has a history of Marfan's syndrome and you think he is having an AMI, but you want to get a chest X-ray or even a CT scan to make sure he doesn't have an aortic dissection before you anticoagulate him. That might put you on the "test" side of the line.

Now imagine the scenario of the potential use for tissue plasminogen activator (tPA) in stroke, a situation frequently encountered in the ED. When a patient comes to the ED within the first few hours of the onset of her stroke symptoms, you rush to get her to the CT scanner. Why? The primary reason is to differentiate between ischemic and hemorrhagic stroke, which will make a major difference in whether or not the patient is even eligible to receive tPA.

Now imagine cases that fall below the testing threshold. You have a 32-year-old male with what sounds like musculoskeletal chest pain. Many would argue that the patient doesn't need any emergency tests at all if he is otherwise healthy and the physical examination is normal. Others might get a chest X-ray and an ECG to rule out occult things like pneumothorax and heart disease, while some others may even get a D-dimer to rule out pulmonary embolism. What is the right way to manage the patient? Is there any evidence behind that decision, or is it just the physician's preference? In some patients, at the end of the ED evaluation, you may not have a

definitive answer. Imagine a 45-year-old female with atypical chest pain, a normal ECG, and normal cardiac markers, who you are evaluating at a hospital that does not perform stress testing from the ED. Does she need a hospital admission to rule out acute myocardial infarction and a stress test?

The way that Pauker and Kassirer designed the test–treatment thresholds more than 30 years ago did not account for the proliferation of ''confirmatory'' diagnostic testing in hospitals. While the lower bound testing threshold is certainly lower than it has ever been, the upper bound threshold has also increased to the point where we are sometimes loath to treat before testing, even when the diagnosis seems apparent. The reason for this is that Occam's razor often does not hold true in emergency medicine. What is Occam's razor? Fourteenth-century philosopher William of Occam stated, ''Plurality must not be posited without necessity,'' which has been interpreted to mean, ''Among competing hypotheses, favor the simplest one.''[2] When applied to test–treatment thresholds, what we find is that a patient with objective findings for what might seem like pneumonia (e.g., hypoxia, infiltrates, and a history of cough) likely does have pneumonia, and should be treated empirically, but may also have a pulmonary embolism. While finding that parsimony of diagnosis is important, often the principle of test–treatment thresholds means that if you're above the test–treatment threshold, then you should certainly treat the patient but also consider testing more, particularly in patients with objective signs of additional disease.

Think about how trauma surgeons practice. In the multi-injured trauma patient, isn't their approach to test, test, test? In a seriously injured patient trauma surgeons often default to scanning everything (aka the *pan-scan*). Some surgeons order CT scans of areas in which the patient has no complaints. They argue that this approach is not illogical. When a patient has been in a major car accident and has a broken left femur, a broken left radius, and mild abdominal tenderness, do they need more CT scans to rule out intra-abdominal injuries and intracranial injuries? Where Occam's razor dulls is that while the most parsimonious diagnosis (just radial and femoral fractures) is possible, patients with multiple traumatic injuries tend to have not only the obvious ones, but also occult injuries. This may necessitate a diagnostic search for the occult intra-abdominal, intrathoracic, and intracranial injuries in a patient with an obviously broken arm and leg.

Risk tolerance refers to the posttest probability at which we are comfortable with excluding or confirming a disease. That is, risk tolerance is where we are comfortable setting our own testing and test–treatment thresholds; it guides where we draw these thresholds and how much we do or do not search for the occult. When deciding on care plans, we develop our own risk tolerance based on our training, clinical expertise, and experiences, as well as local

standard practice and the attitudes of the patient, family, or other physicians caring for the patient.

For example, consider possible acute coronary syndrome. After your ED evaluation with cardiac markers, an ECG, and a chest X-ray, you estimate that your patient has a 2% risk of having an unexpected cardiac event within 30 days if he is sent home without additional testing. Is it OK to send him home with this level of risk? Isn't 2% the published rate for missed AMI? What if the risk is 1%, or 0.5%, or 0.1%?

How do you make the decision about when to order a test or just treat? How do you assign a pretest probability? How do you apply test results to an individual patient? This is where research and the practice of evidence-based medicine (EBM) can influence practice by taking the best evidence in the literature about diagnostic testing or clinical decision rules and using that information to make an informed decision about how to care for patients. Chapters 2 and 3 provide an updated overview of the process of EBM as well as examples of the application of EBM to individual patients in the ED, levels of evidence, and how to evaluate a body of literature on diagnostic testing. Chapter 4 is a revised discussion of how we derive, validate, and study the impact of clinical decision rules in practice. Chapter 5, a new chapter in the second edition of this book, reviews recent trends in health policy that may force us to reduce test ordering and use clinical decision rules. Chapter 6 describes various forms of bias that can skew estimates of diagnostic accuracy in research settings.

Understanding the evidence behind diagnostic testing and using clinical decision rules to decide when not to test is at the core of emergency medicine practice. Think back to your last shift in the ED: how many tests did you order?

The purpose of this book is to demystify the evidence behind diagnostic testing and clinical decision rules in emergency care by carefully evaluating the evidence behind our everyday decision making in the ED. This book is written to provide objective information on the evidence behind these questions and our opinion on how we manage our patients with specific clinical problems given the best available evidence. It should be noted that we are writing this from the perspective of academic emergency physicians. We all work in academic EDs with abundant (although not always quick) access to consultants, state-of-the-art laboratories, and high-resolution imaging tests. As you read this, realize that not all emergency medicine practice is the same and you should interpret the literature yourself in the context of your own clinical practice environment.

We have designed each chapter around clinical questions that come up in everyday emergency medicine practice. In the second edition of the book, we have added more chapters and updated all of the old chapters to include new

and relevant studies or insights that have emerged in the literature since the first edition was published in 2008. For each question, we present the objective data from published studies and then provide our "expert" comments on how we use these tests in our practice. While we try to provide insight into how we interpret the literature for each testing approach, again, our comments should not be interpreted as the standard of care in emergency medicine. Standard of care is based on practice guidelines and local practice patterns. Instead, these chapters should serve as a forum or basis for discussion. If you are a researcher, you can also think of this book as a roadmap to what is really "known" or "not known" with regard to diagnostic testing in emergency medicine and what needs further study. Finally, rigorous and sound research often takes months to years to accomplish, and sometimes longer to publish. Therefore the discussions we present are likely to change as newer, larger, more comprehensive studies are published, as new prediction or decision rules are validated and replicated, and as newer diagnostic technology is introduced.

References

1. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. New England Journal of Medicine. 1980; 302: 1109–17.
2. Drachman DA. Occam's razor, geriatric syndromes, and the dizzy patient. Annals of Internal Medicine. 2000; 132: 403–4.

Additional Reading

1. Bossuyt PMM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L *et al*. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. Annals of Internal Medicine. 2003; 138: W1–W12.
2. Brownlee S. Overtreated: Why too much medicine is making us sicker and poorer. New York: Bloomsbury; 2007.
3. Empey M, Carpenter C, Jain P, Atzema C. What constitutes the standard of care? Annals of Emergency Medicine. 2004; 44: 527–31.
4. Kovacs G, Croskerry P. Clinical decision making: An emergency perspective. Academic Emergency Medicine. 1999; 6: 947–52.
5. Schünemann AHJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE *et al*. GRADE: Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. British Medical Journal. 2008; 336(7653): 0.3. Available from: http://www.bmj.com/content/336/7654/0.3
6. Whiting P, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB *et al*. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Annals ofInternal Medicine. 2011; 155: 529–36.