

# Chapter 1 **Diagnostic Testing in Emergency Care**

Jesse M. Pines<sup>1,2</sup> and Christopher R. Carpenter<sup>3</sup>

<sup>1</sup>US Acute Care Solutions, Canton, OH, USA

<sup>2</sup>Department of Emergency Medicine, Drexel University, Philadelphia, PA, USA

<sup>3</sup>Department of Emergency Medicine, Washington University School of Medicine, St. Louis, MO, USA

## Highlights

- Emergency physicians are experts in diagnostic testing
- The choice of ED-based testing depends on the resources of the hospital
- Validated clinical decision rules can help guide ED testing decisions
- Pauker and Kassirer test-treatment thresholds are a helpful tool in determining the use and value of diagnostic tests

As emergency department (ED) physicians, we spend a good deal of our time ordering, interpreting, and waiting for the results of diagnostic tests. ED physicians are the experts when it comes to determining who needs a test to rule out a potentially life-threatening condition. There are several reasons for this expertise. First and foremost, we see a lot of patients with undifferentiated symptoms in a decision-dense and time-constrained environment. 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. Some patients and consultants value lab or imaging tests more than the history and physical exam tests that formulate clinical intuition, a phenomenon called “technological tenesmus.”<sup>1</sup> However, if we order time-consuming tests on everyone, ED crowding and inefficiency will worsen, costs of care will go

---

*Evidence-Based Emergency Care: Diagnostic Testing and Clinical Decision Rules*, Third Edition.

Edited by Jesse M. Pines, Fernanda Bellolio, Christopher R. Carpenter, and Ali S. Raja.

© 2023 John Wiley & Sons Ltd. Published 2023 by John Wiley & Sons Ltd.

## 4 Chapter 1: Diagnostic Testing in Emergency Care

up, and patients will experience even longer waits than they already do. In addition, there is increased pressure to carefully choose who needs and who does not need tests in an evidence-based manner, particularly as costs of care have risen so dramatically in recent years particularly in the United States.<sup>2,3</sup>

Differentiating which patients will benefit from ED testing is a complex process. Over the past 40 years, science and research in ED diagnostic testing and clinical decision rules have advanced considerably. Today, there is a greater understanding of test performance, specifically 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 safely ruled out based on clinical grounds alone, with acceptable accuracy and precision. There are also accurate risk stratification tools to estimate the probability for conditions like pulmonary embolism (PE) before any tests are even ordered. Since the second edition of this textbook, *Academic Emergency Medicine* created the “Evidence-Based Diagnostics” series to synthesize the ever-expanding volume of emergency medicine-specific research around history, physical exam, labs, and imaging for common diagnoses like subarachnoid hemorrhage, congestive heart failure, urinary tract infection, and mesenteric ischemia.<sup>4,5</sup> Similarly, the Society for Academic Emergency Medicine launched Guidelines for Reasonable and Appropriate Care to provide emergency medicine’s first Grading of Recommendations Assessment Development and Evaluation (GRADE)-based diagnostic recommendations that contemplate issues like costs and health inequities.<sup>6–8</sup> This third edition will summarize the key recommendations from these two new resources.

How do we decide who to test and who not to test? There are some patients who clearly need tests, such as the head-injured patient who has altered mental status and who may have a head bleed. In such a case, the outcome 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 at an individual point in time, such as patients with a simple toothache or a mild headache without concerning features. 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 unexpected test results come back, it may not be clear how best to use those results to guide patient care.

Let us 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 PE). 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 PE 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 PE been satisfactorily ruled out? Should you order 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? Would he benefit from a CT scan of the chest to further evaluate the possibility of PE? What are the potential harms of liberally obtaining CT on every patient in whom the physician or the patient is concerned about PE just to be absolutely certain? Or is he so low risk that he's probably fine anyway? Of course, you might wonder, if he was so low risk, why 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 CT or magnetic resonance imaging (MRI) for a radiographically occult hip fracture? While you contemplate time-consuming advanced imaging, you also consider that regardless of whether or not CT or MRI demonstrates no fracture will she be able to go home?

These are examples of when test results may not confirm you're *a priori* 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 professors 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 must be 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 a young healthy population, we are unlikely to find it because it is not common in young people. 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, particularly in

## 6 Chapter 1: Diagnostic Testing in Emergency Care

patients with a history of prior dislocations. 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. Most hospitals allow easy access to radiographic testing and laboratory testing. In other hospitals, obtaining a diagnostic test may not be as easy. Some hospitals may 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 clinicians 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? What is the differential diagnosis for these symptoms in your patient and how likely are each possible diagnosis? How knowledgeable are your consultants about the additive value (or lack thereof) for the different test options and multidisciplinary consensus recommendations?<sup>9</sup>

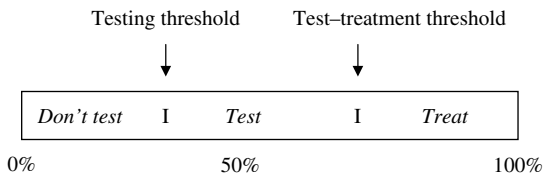
How about using clinical decision rules in practice? By determining if patients meet specific clinical criteria, we can choose not to test some patients if they are low risk. Do all patients with ankle sprains need X-rays? Can you use the Ottawa ankle rules in children? What are the limits of clinical 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 also 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. This is discussed later in Chapter 60 on Cognitive Bias in more detail, which is new to the third edition of the book.

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.<sup>10</sup> 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.

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



**Figure 1.1** Pretest probability of disease. (Data from Pauker and Kassirer [10].)

## 8 Chapter 1: Diagnostic Testing in Emergency Care

thresholds, the test should be performed and the patients treated according to the test results.

Calculation of the testing threshold and the test–treatment threshold include seven variables:

1.  $P_{\text{pos/d}}$  – The probability of a positive test in patients with disease (i.e., sensitivity)
2.  $P_{\text{neg/d}}$  – The probability of a negative test in patients with disease (1-sensitivity)
3.  $P_{\text{pos/nd}}$  – Probability of a positive test result in patients without disease (1-specificity)
4.  $P_{\text{neg/nd}}$  – Probability of a negative test result in patients without disease (i.e., specificity)
5.  $B_{\text{rx}}$  – The benefit of treatment in patients with disease
6.  $R_{\text{rx}}$  – The risk of treatment in patients with disease
7.  $R_t$  – The risk of the diagnostic test.

Using these variables together, Figure 1.2 demonstrates the formula for the testing and test–treatment thresholds.

The *Academic Emergency Medicine* “Evidence-Based Diagnostics” series described above provides estimates of pretest probability for common diagnoses based on a synthesis of ED research and then presents an interactive test–treatment threshold calculator based on this Pauker–Kassirer theory.

But now let us make this more clinically relevant. Sometimes the 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 and/or other antiplatelet agents, anticoagulation, and other resuscitative treatments and send him off to the cardiac catheterization lab if your hospital has one, arrange for transfer, 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

$$T_t \text{ (testing threshold)} = (P_{\text{pos/nd}}) \times (R_{\text{rx}}) + R_t / (P_{\text{pos/nd}}) \times (R_{\text{rx}}) + (P_{\text{pos/d}}) \times (B_{\text{rx}})$$

$$T_{\text{trx}} \text{ (test-treatment threshold)} = (P_{\text{neg/nd}}) \times (R_{\text{rx}}) - R_t / (P_{\text{neg/nd}}) \times (R_{\text{rx}}) + (P_{\text{neg/d}}) \times (B_{\text{rx}})$$

**Figure 1.2** The formulas for the testing and test–treatment threshold. (Adapted from Pauker and Kassirer [10].)

want to get a chest X-ray or even a CT scan to make sure he does not 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 does not 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 PE. 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 or observation to rule out AMI and a stress test?

The way that Pauker and Kassirer designed the test–treatment thresholds more than 40 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.”<sup>11</sup> 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 PE. 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, is not their approach to test, test, test? In a seriously injured patient trauma surgeons often default to scanning everything (aka the *pan-scan*) despite evidence demonstrating no patient-centric benefit.<sup>12</sup> 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, but the balance between careful trauma imaging and over-testing harms continues to be defined.<sup>13</sup>

*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.<sup>14,15</sup> 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.<sup>16</sup> 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?<sup>17</sup> What if the risk is 1%, or 0.5%, or 0.1%? If you send someone home with a HEART score less than 3, what is their actual risk of a major adverse cardiovascular event within 6 weeks?<sup>18</sup> Does that differ if their HEART score is zero or if it's 3?

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 and communicate our clinical impression, level of certainty, and subsequent management options with those patients?<sup>19</sup> 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. Chapter 59 is a new chapter that explores the science and potential of shared decision-making in ED 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 through the lens of seasoned ED clinicians who are also diagnostic testing researchers, while providing 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 mostly academic EDs or well-resourced community 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 third 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 second edition was published in 2013. 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

## 12 Chapter 1: Diagnostic Testing in Emergency Care

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. Fatovich DM. The inverted U curve and emergency medicine: Overdiagnosis and the law of unintended consequences. *Emergency Medicine Australasia*. 2016; 28(4): 480–2.
2. Carpenter CR, Raja AS, Brown MD. Overtesting and the downstream consequences of overtreatment: Implications of “preventing overdiagnosis” for emergency medicine. *Academic Emergency Medicine*. 2015; 22(12): 1484–92.
3. Venkatesh AK, Scofi JE, Rothenberg C *et al*. Choosing wisely in emergency medicine: Early results and insights from the ACEP emergency quality network (E-QUAL). *The American Journal of Emergency Medicine*. 2021; 39: 102–8.
4. Carpenter CR, Hussain AM, Ward MJ *et al*. Spontaneous subarachnoid hemorrhage: A systematic review and meta-analysis describing the diagnostic accuracy of history, physical examination, imaging, and lumbar puncture with an exploration of test thresholds. *Academic Emergency Medicine*. 2016; 23(9): 963–1003.
5. Martindale JL, Wakai A, Collins SP *et al*. Diagnosing acute heart failure in the emergency department: A systematic review and meta-analysis. *Academic Emergency Medicine*. 2016; 23(3): 223–42.
6. Musey PJ Jr, Bellolio F, Upadhye S *et al*. Guidelines for reasonable and appropriate care in the emergency department (GRACE): Recurrent, low-risk chest pain in the emergency department. *Academic Emergency Medicine*. 2021; 28(7): 718–44.
7. Schünemann HJ, Oxman AD, Brozek J *et al*. GRADE Working Group. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008; 336(7653): 1106–10.
8. Carpenter CR, Bellolio MF, Upadhye S, Kline JA. Navigating uncertainty with GRACE: Society for Academic Emergency Medicine’s guidelines for reasonable and appropriate care in the emergency department. *Academic Emergency Medicine*. 2021; 28(7): 821–5.
9. Moore CL, Carpenter CR, Heilbrun ME *et al*. Imaging in suspected renal colic: Systematic review of the literature and multispecialty consensus. *Annals of Emergency Medicine*. 2019; 74(3): 391–9.
10. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *New England Journal of Medicine*. 1980; 302: 1109–17.
11. Drachman DA. Occam’s razor, geriatric syndromes, and the dizzy patient. *Annals of Internal Medicine*. 2000; 132(5): 403–4.
12. Sierink JC, Treskes K, Edwards MJ *et al*. REACT-2 study group. Immediate total-body CT scanning versus conventional imaging and selective CT scanning in patients with severe trauma (REACT-2): A randomised controlled trial. *Lancet*. 2016; 388(10045): 673–83.

13. Hoffman JR, Carpenter CR. Guarding against overtesting, overdiagnosis, and overtreatment of older adults: Thinking beyond imaging and injuries to weigh harms and benefits. *Journal of the American Geriatrics Society*. 2017; 65(5): 903–5.
14. Pines JM, Szyld D. Risk tolerance for the exclusion of potentially life-threatening diseases in the ED. *The American Journal of Emergency Medicine*. 2007; 25(5): 540–4.
15. Pines JM, Hollander JE, Isserman JA *et al*. The association between physician risk tolerance and imaging use in abdominal pain. *The American Journal of Emergency Medicine*. 2009; 27(5): 552–7.
16. Pines JM, Isserman JA, Szyld D, Dean AJ, McCusker CM, Hollander JE. The effect of physician risk tolerance and the presence of an observation unit on decision making for ED patients with chest pain. *The American Journal of Emergency Medicine*. 2010; 28(7): 771–9.
17. Schull MJ, Vermeulen MJ, Stukel TA. The risk of missed diagnosis of acute myocardial infarction associated with emergency department volume. *Annals of Emergency Medicine*. 2006; 48(6): 647–55.
18. Fernando SM, Tran A, Cheng W *et al*. Prognostic accuracy of the HEART score for prediction of major adverse cardiac events in patients presenting with chest pain: A systematic review and meta-analysis. *Academic Emergency Medicine*. 2019; 26(2): 140–51.
19. Hess EP, Grudzen CR, Thomson R, Raja AS, Carpenter CR. Shared decision-making in the emergency department: Respecting patient autonomy when seconds count. *Academic Emergency Medicine*. 2015; 22(7): 856–64.