Annals of Internal Medicine

SPECIAL ARTICLE

Annals Clinical Decision Making: Avoiding Cognitive Errors in Clinical Decision Making

Daniel Restrepo, MD; Katrina A. Armstrong, MD, MS; and Joshua P. Metlay, MD, PhD

Making a diagnosis is perhaps the most frequently performed and important procedure in an internist's wheelhouse. Indeed, a patient's therapeutic options and prognosis often center around the right diagnosis being made in a timely fashion. Unfortunately, diagnostic errors are common, with frequency estimates ranging from 10% to 40% in some studies (1, 2). Furthermore, misdiagnosis is more likely to be considered negligent and have proportionately higher morbidity and morbidity than other errors (3, 4).

Cognitive psychology research over the past 40 years has led to the development of dual process theory, wherein 2 central cognitive systems, system 1 and system 2, help us process and respond to incoming data (5). System 1 uses intuitive and effortless recognition of previously seen patterns, and system 2 engages effortful, slower analytical thinking. Out of necessity, our minds gravitate toward pattern recognition. The accuracy of this pattern recognition system grows through the acquisition of experience. This is true in clinical care; as experience grows, clinicians develop cognitive "shortcuts." We refer to these shortcuts as "heuristics" (Glossary). A cognitive bias (Glossary) arises when these heuristics fail.

To date, over 100 cognitive biases and heuristics have been identified (6). When made, errors in diagnosis can be summarized into 2 themes: errors of prevalence estimation and errors of ego. The former can be traced to a failure of weighing the a priori probability of a diagnosis. The latter centers around the inherently human tendency to want to be right and fall in love with our initial impressions.

Given that cognitive biases are overrepresented in cases of diagnostic failure (7, 8), in this article, we focus on how to both recognize these heuristics and avoid the resulting errors by framing our discussion around 2 cases.

Glossary

Heuristics: Cognitive shortcuts taken on the basis of prior experience and recognition of previously seen patterns.
Cognitive bias: In clinical reasoning, cognitive bias occurs when a mental shortcut, or heuristic, results in misdiagnosis or inappropriate treatment.
Base rate: The underlying frequency or prevalence of a diagnosis, and a correlate to pretest probability.
Representativeness heuristic: The tendency to search for patterns and select a diagnosis because a constellation of findings matches the respective pattern, regardless of the actual probability of the underlying diagnosis.
Confirmation bias: Disproportionately believing facets of a case that confirm or support initial theories.
Availability heuristic: Diagnoses or recent outcomes that easily come to mind and as such may be overrepresented in a clinician's reasoning.
Anchoring: The tendency to focus on a singular facet of a case, creating the potential to negate other disconfirming aspects of the case.
Premature closure: Closure of the diagnostic process in a premature fashion before all relevant information is obtained.

Case 1

A 40-year-old woman with a history of obesity is admitted to the hospital because of malaise, right upper-quadrant pain, and anorexia. She is afebrile and has otherwise unremarkable vital signs. Her examination is notable for a negative Murphy sign, no scleral icterus, and mild tenderness on palpation of the right upper quadrant. Laboratory analysis shows marked elevations in her aminotransferase levels. Right upper-quadrant ultrasonography shows a mildly dilated common bile duct, but otherwise no gallstones and patent hepatic and portal veins. Given her young age and biochemical evidence of liver inflammation, an evaluation for autoimmune hepatitis (AIH) is undertaken; positive results are found, with an antinuclear antibody titer of 1:320 and an anti–smooth-muscle antibody titer of 1:160. While hospitalized, her liver enzyme levels remain elevated, and she continues to have pain. Because of her serologic evaluation, she undergoes liver biopsy, which shows ductular dilation and neutrophilic reaction, a pattern most consistent with macroscopic biliary ductal obstruction and commonly seen with passage of a gallstone. There is no evidence of AIH on biopsy.

Why Do We Make Mistakes in Estimating the Probability of Diagnoses?

In this real-life case, the physician was drawn to the patient's young age and the degree of her hepatitis, which matched their prototype for autoimmunity, prompting a serologic evaluation. The mildly positive anti–smooth-muscle antibody titer then triggered an invasive test for AIH that ultimately revealed an unrelated and much more prevalent diagnosis. In this case, the diagnostic process went awry because the physician compared the findings in the case with the prototype of a condition in the differential diagnosis, a shortcut often called a "representativeness heuristic" (Glossary). In doing so, she neglected the underlying prevalence of this disorder and went with a good fit in the face of diagnostic uncertainty. This culminated in a diagnostic test that has the potential for procedural complications.

See also:

This article is 1 of 6 within the *Annals* Clinical Decision Making series. The series is available at Annals.org.

This article was published at Annals.org on 21 April 2020.

An alternate approach could have been that of watchful waiting to see whether the patient's pain would resolve or her liver enzyme levels would come down, both of which would argue for a stone as the cause of her symptoms. Although the representativeness heuristic allows physicians to identify common diagnoses quickly and be right most of the time, it comes at the risk of neglecting the underlying prevalence, or base rate (Glossary), of diagnoses in the general population with different clinical presentations.

Furthermore, our minds often gravitate toward information that aligns well with our leading theory, and it becomes tempting to pursue or believe the facts that support our point while casting aside those that don't. This is known as "confirmation bias" (Glossary), or disproportionately believing data points that support our initial theories.

How Can I Avoid Errors of Base Rate Neglect?

Physicians begin crafting a differential diagnosis for a problem from the moment a chief complaint is known. Upon obtaining enough clinical information to formulate a differential diagnosis, the physician should perform an initial reshuffling of the deck on the basis of the relative prevalence of each of the items on the list. This was not done in the case presented above, where the disease prevalence was neglected and the representativeness heuristic fell short.

The prevalence of cholelithiasis in adults living in the United States is somewhere between 175 and 685 cases per 100 000 people (9). Although we do not have recent data in the United States, the prevalence of AIH in European countries is somewhere between 10 to 15 cases per 100 000 people (10)—making cholelithiasis an order of magnitude more common than AIH. To prevent cognitive pitfalls, it behooves us to return to the old adage that "common things are common" when putting our nickel down on a diagnosis. Indeed, it is still more likely to diagnose an atypical manifestation of a common condition—in our case, biliary obstruction due to gallstones, resulting in high aminotransferase levels—over a somewhat typical manifestation of an uncommon disease, such as AIH. Thus, even if the representativeness heuristic brought AIH to mind first, reviewing the actual frequency of different diagnoses among similar patients with this presentation should lead cholelithiasis to be placed above it on the list of potential diagnoses. Of course, it is important to consider and even test for serious "can't miss" diagnoses that may be less common. However, the information obtained from diagnostic testing, including positive results, must be considered within the confines of the test's characteristics and the pretest probability of the diagnosis.

In our case, the physician was initially seduced by the constellation of clinical and laboratory findings as representative of AIH. Indeed, a younger woman with high aminotransferase levels could fit the narrative for the diagnosis well. The test results of positive ANA and anti–smooth-muscle antibody titers further fueled the initial diagnostic impression. However, the diagnostic weight of these tests depends on the prevalence of the disease. Unfortunately, few, if any, tests are perfect, and the interpretation of autoimmune serologies is potentially fraught with false-positives. Understanding how diagnostic test characteristics influence clinical reasoning is discussed in further detail in another article in this series (11).

How Does the Recency or Vividness of a Clinical Encounter or Diagnosis Affect My Estimation of the Pretest Probability of a Condition? How Can I Avoid Cognitive Errors Related to This?

Recent experiences, events, or outcomes can have a powerful effect on our assessment of probability. This is known as the "availability heuristic" (Glossary) and predisposes us to favor diagnoses or outcomes that come to mind easily. Sometimes this can be helpful, because things that come to mind easily are often because we see them frequently (12). However, easily available information doesn't always equate with a correct diagnosis and can lead us astray. The clinical significance, proximity, or vividness of an adverse outcome; a missed diagnosis; or simply encountering a rare disease can be powerful modifiers of subsequent event rate estimation. Indeed, house officers have been shown to commit more diagnostic errors when relying on intuitive reasoning related to the availability heuristic. In a study of internal medicine residents, Mamede and colleagues (12) exposed the participants to cases and had residents ascertain the diagnosis. Shortly thereafter, the same residents were then reexposed to cases with similar presenting syndromes but different underlying diagnoses to promote nonanalytical reasoning and the availability bias, resulting in a high percentage of errors. When the same sequence was repeated but the participants were asked to pause and reflect on aspects of the case that didn't fit with their first impression, the frequency of errors significantly dropped and the correct diagnosis was achieved, supporting the theory that the "effortful pause" can prove instrumental in the diagnostic process (12, 13).

To combat the availability heuristic, clinicians must consciously reflect and seek to recall an instance or outcome that contrasts with the first example that comes to mind. This may be indispensable when a recent experience, negative or otherwise, is influencing whether to administer or withhold therapy. The physician whose patient recently bled while receiving warfarin for atrial fibrillation thromboprophylaxis is well served by recalling the preceding 5 people who did well with therapeutic anticoagulation, or to perhaps consider the patient with an unfortunate stroke while not receiving anticoagulation. The clinician who is seeing many patients with a certain diagnosis in a day should recall that the preceding patient's diagnosis has no effect on the probability of the following patient's diagnosis. As you will see throughout this article, the acts of conscious reflection and deliberate, effortful analytical thought are central in combating cognitive biases.

How Should My Own Accumulated Experience Influence My Base Rate Estimation?

Knowledge of prevalence and years of clinical experience allow clinicians to treat, and subsequently recognize, the immense degree of variability in the presentation of common diagnoses, such as heart failure or bacteremia. As clinicians progress throughout their career, their ability to appreciate the widening variability of presentations within a diagnosis increases, and this allows them to reach a conclusion by identifying key and salient data points. Because of this, their reliance on holding out for "textbook presentations" diminishes (Figure 1). However, this is not to say that experienced clinicians are immune to the same pitfalls as their junior colleagues, and we would all benefit from the same effortful pauses and deliberate reflection. Of note, the process of adjusting epidemiologic data to conform with one's own clinical experience must be based on cumulated experience in a certain practice setting, not just the past 1 or 2 cases that a physician encounters. Otherwise, one runs the risk of ignoring probability and falling prey to the representativeness heuristic and the availability bias.

CASE 2

A 65-year-old man with a history of essential hypertension and hyperlipidemia and a 100–pack-year smoking history presents to his primary care clinic with subacute exertional dyspnea. He has no chest pain and reports a minimally productive cough. His assessment includes an oxygen saturation of 90% on exertion, and his physical examination is notable for a mildly prolonged expiratory phase. Chest radiography shows hyperinflation. He undergoes pulmonary function testing that does not demonstrate an obstructive deficit; however, given the preponderance of findings, he is started on bronchodilators for empirical treatment of chronic obstructive pulmonary disease (COPD). After his symptoms do not abate, treatment with inhaled and later systemic glucocorticoids is started, again without improvement. Eventually, when the patient's condition fails to improve, he undergoes pharmacologic stress testing, which reveals evidence of reversible ischemia.

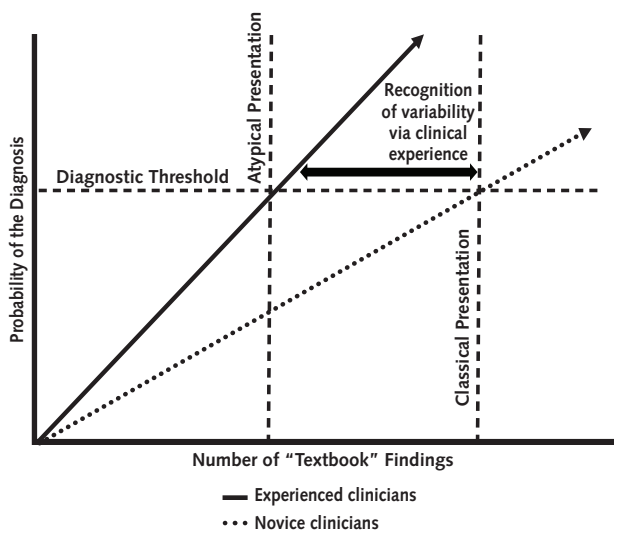
Clearly, elements of representativeness heuristic and availability bias are again demonstrated in this hypothetical case. However, the case also illustrates examples of what we previously termed "errors of ego": specifically, 2 other cognitive biases, known as "anchoring" and "premature closure" (Glossary). In diagnostic reasoning, anchoring is the tendency to adhere to an initial hypothesis, leading us to disregard or discount evidence that disproves our favored theory. A close correlate to anchoring, premature closure occurs when a clinician accepts a diagnosis as final and concludes the diagnostic process before the true diagnosis is identified. This may stem from failing to acquire a critical piece of information or not revising the diagnosis on the basis of this information. In our case, the physician decided the patient had COPD by anchoring on the high pack-year history of smoking and the chest radiograph. The physician then prematurely closed the diagnostic evaluation without exploring the possibility of other causes of exertional dyspnea in a heavy smoker, especially when the patient's condition failed to improve with appropriate management of obstructive lung disease or when the pulmonary function tests did not suggest significant obstruction.

The medical record can be a powerful asset in terms of accessing a wealth of clinical data; however, caution should be exercised with documentation of diagnoses that are untrue or unproven, because they will serve as a substrate for clinicians who may care for the patient in the future. Consider the effect of seeing a diagnosis of COPD in the patient's medical history, perhaps derived from a prior billing code, when the patient presents with dyspnea to an acute care setting. This could act as an anchoring point in the future and could result in physicians narrowing the differential diagnosis for dyspnea and closing prematurely when deciding to treat for an acute exacerbation of COPD.

How Can the Use of Diagnostic Schema or Algorithms Help Prevent Anchoring and Premature Closure?

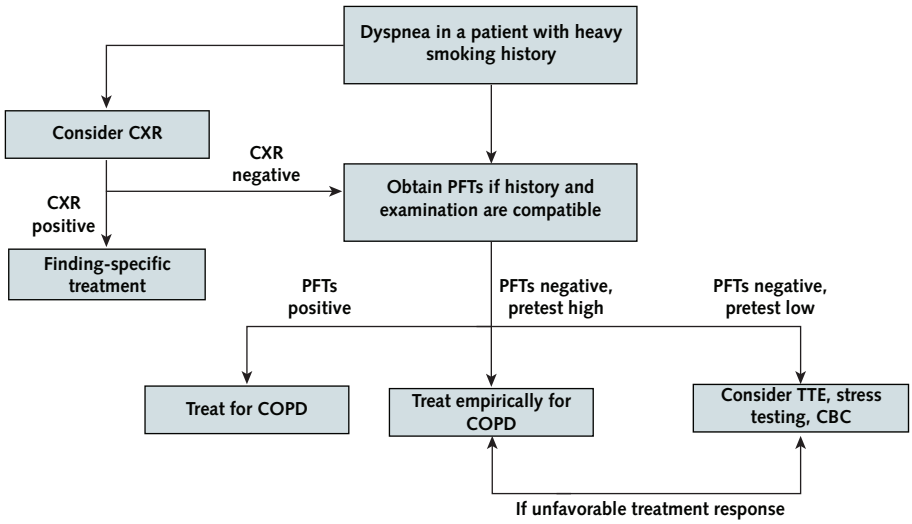
Accessing diagnostic schema or algorithms for common clinical problems can be time-efficient and invaluable in mitigating anchoring and premature clo-

Figure 1. Relationship between experience and threshold at which a diagnosis can be made.



The dotted line represents more novice clinicians, who rely on classic or prototypical presentations of diagnoses, with many symptoms, signs, or findings needed to recognize the typical manifestation of a common diagnosis. Through clinical experience and deliberate practice, seasoned clinicians (*black line*) will recognize the inherent variability in these conditions and begin to appreciate the atypical manifestations of common conditions. If one considers heart failure as an example, a medical student may require orthopnea, lower extremity edema, an S3 gallop, and elevated natriuretic peptide levels to make the diagnosis, whereas a more experienced attending could make the diagnosis from a history of anorexia, nausea, and distended neck veins without the other signs or findings of decompensated heart failure.

Figure 2. Sample diagnostic schema for a diagnostic and therapeutic approach to breathlessness in a patient with a heavy smoking history.



In this schema, physicians would consider obtaining CXR and, in those with a compatible history and physical examination, PFTs. If positive, this is confirmatory of COPD, and the patient should start treatment. If negative, the clinician should recall the patient's pretest probability for the diagnosis of COPD. If this probability was low, then other items on the differential diagnosis, such as heart failure or coronary disease, should be explored. If pretest probability was high, a reasonable strategy would be to treat empirically and monitor for response to treatment, with a plan to proceed with additional diagnostic testing should the patient not improve. CBC = complete blood cell count; COPD = chronic obstructive pulmonary disease; CXR = chest radiography; PFT = pulmonary function test; TTE = transthoracic echocardiography.

sure by acting as a diagnostic "checklist" before accepting a final theory as true. These tools can be cognitive frameworks created by individual physicians, a product of clinical practice guidelines, or found online, and they are a useful means to expand a differential diagnosis or provide additional counterhypotheses to the initial impression. Drawing from their success in other sectors of health care, checklists have been proposed to mitigate or minimize diagnostic error (14).

For example, let's explore how a diagnostic algorithm (Figure 2) for the outpatient with dyspnea and a heavy smoking history could have helped with the preceding case. If confronted with a smoker with dyspnea, a physician should consider whether testing for COPD with formal pulmonary function tests is appropriate. If positive, the patient should be treated with smoking cessation and bronchodilator therapy. If negative, the algorithm would encourage physicians to simultaneously broaden their differential diagnosis, perhaps with consideration of coronary artery disease and other COPD/asthma mimics, such as vocal cord dysfunction and heart failure, alongside simultaneous treatment with empirical bronchodilators if suspicion for COPD remains high. This would have helped curtail anchoring and premature closure by allowing the physician to revisit atypical parts of the case once disconfirming evidence (such as normal pulmonary function findings) became available.

Figure 3. Steps in the diagnostic process to avoid cognitive errors in diagnosis.

1. Recognize the tendency to over- and underweight diagnoses on the basis of the similarity of the presentation to the "textbook" presentation.
2. Rearrange diagnoses on the differential diagnosis on the basis of information about the prevalence of different conditions by using local epidemiology and other sources.
3. Access diagnostic schema or diagnostic checklists if uncertainty exists regarding validated approaches to a symptom complex, finding, or syndrome.
4. Before settling on a final diagnosis, perform a conscious pause to seek disproving evidence and ask, "What doesn't fit?"
5. Confer with clinical teammates or, if working alone, peers or superiors regarding the case in order to reduce the risk for biases.
6. Once a diagnosis and treatment are instituted, continuously reassess response to therapy as further evidence of diagnostic accuracy.

How Can the Medical Note Help Me to Avoid Cognitive Biases?

The effort devoted to generating a good admission or progress note can be a useful built-in cross-check because it provides the physician a purposeful moment in time to delineate the differential diagnosis, highlight salient features of the case, and provide reasoning as to why one diagnosis is favored over the others. It is a useful time to put words to the underlying thinking as the physician's mind runs through prevalence rates, diagnostic schema, and therapeutic algorithms—all of which can help stave off heuristics and potentially cognitive biases. This strategy is, of course, time-consuming but can be an invaluable moment of reflection and encapsulation, wherein our previous strategies can be applied. Moreover, because it is visible to other members of the health

care team, it provides an opportunity for additional diagnoses and strategies to be added to the list. However, with increasing frequency of copying and pasting to generate notes, this strategy can only work if notes are thoughtfully written and not autopopulated.

How Does Working in Teams Help With the Diagnostic Reasoning Process?

The presence of large clinical teams, as is often the case in academic hospitals with residency training programs, can itself be a powerful boon to diagnostic reasoning. Each case may be viewed from many different vantage points: for example, by medical students well versed in physiology all the way up to senior attending physicians with years of clinical experience. Presentation of new cases to a medical team results in multiple simultaneous first impressions with many different lines of diagnostic reasoning. By encouraging input across the team, these different lines of reasoning further reduce the risk for diagnostic biases. In its 2015 report focusing on diagnosis in health care, the National Academy of Medicine called special attention to the role that teamwork plays in diagnosis and highlighted the vital role that nursing and nonphysician members have in ensuring accuracy of potential diagnostic intervention as well as in preventing potentially harmful interventions from being administered (15).

SUMMARY

Cognitive biases and subsequent errors in diagnostic medicine are pervasive and costly. Furthermore, they are made even by seasoned physicians, and although experience can aid in efficiency of practice, it does not alone grant expertise. Ultimately, most strategies to prevent these errors center around knowledge of epidemiology; common cognitive pitfalls; and the process of integrating these factors into deliberate, thoughtful practice. When our minds reach a diagnosis quickly on the basis of pattern recognition or singular facets of the case, it is essential to always look for the potential fallacies of our reasoning, specifically the tendency to cast aside probability for a good fit. Indeed, our differentials should and must be reshuffled to account for known prevalence of different diagnoses even when a case is perfectly representative of a rare diagnosis. The act of pausing and asking "What doesn't make sense?" or "Am I missing something?" and the practice of seeking disconfirming evidence before adopting a final diagnosis can combat our tendency to love our first hypothesis, the perennial error of ego (illustrated in an accompanying interactive graphic [16]). To think about one's thinking, known as "metacognition," is integral in the internist's diagnostic process. It is what ultimately lets us decide when we must pause and engage in effortful analysis or when we safely can use intuitive processes to get us to the right answer efficiently and keep ourselves from making potentially tragic errors (Figure 3).

From Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts (D.R., K.A.A., J.P.M.).

Disclosures: Authors have disclosed no conflicts of interest. Forms can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M19-3692.

Corresponding Author: Daniel Restrepo, MD, Core Educator Faculty and Hospital Medicine Unit, Department of Medicine, Massachusetts General Hospital, 50 Staniford Street, Suite 550, Boston, MA 02114; e-mail, drestrepo1@mgh.harvard.edu.

Current author addresses and author contributions are available at Annals.org.

Ann Intern Med. 2020;172:747-751. doi:10.7326/M19-3692

References

1. **Berner ES, Graber ML.** Overconfidence as a cause of diagnostic error in medicine. Am J Med. 2008;121:S2-23. [PMID: 18440350] doi:10.1016/j.amjmed.2008.01.001
2. **Shojania KG, Burton EC, McDonald KM, et al.** Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. JAMA. 2003;289:2849-56. [PMID: 12783916]
3. **Leape LL, Brennan TA, Laird N, et al.** The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. N Engl J Med. 1991;324:377-84. [PMID: 1824793]
4. **Brennan TA, Leape LL, Laird NM, et al.** Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med. 1991;324:370-6. [PMID: 1987460]
5. **Evans J.** Heuristic and analytic processes in reasoning. Br J Psychol. 1984;75:451-68. doi:10.1111/j.2044-8295.1984.tb01915.x
6. **Croskerry P.** The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med. 2003;78:775-80. [PMID: 12915363]
7. **Graber ML, Franklin N, Gordon R.** Diagnostic error in internal medicine. Arch Intern Med. 2005;165:1493-9. [PMID: 16009864]
8. **Norman GR, Monteiro SD, Sherbino J, et al.** The causes of errors in clinical reasoning: cognitive biases, knowledge deficits, and dual process thinking. Acad Med. 2017;92:23-30. [PMID: 27782919] doi:10.1097/ACM.0000000000001421
9. **Everhart JE.** Gallstones. In: Everhart JE, ed. Digestive Diseases in the United States: Epidemiology and Impact. NIH publication no. 94-1447. Washington, D.C.: U.S. Government Printing Office, 1994: 647-69.
10. **Werner M, Prytz H, Ohlsson B, et al.** Epidemiology and the initial presentation of autoimmune hepatitis in Sweden: a nationwide study. Scand J Gastroenterol. 2008;43:1232-40. [PMID: 18609163] doi:10.1080/00365520802130183
11. **Armstrong KA, Metlay JP.** *Annals* clinical decision making: using a diagnostic test. Ann Intern Med. 2020;172:604-9. doi:10.7326/M19-1940
12. **Mamede S, van Gog T, van den Berge K, et al.** Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. JAMA. 2010;304:1198-203. [PMID: 20841533] doi:10.1001/jama.2010.1276
13. **Graber ML, Kissam S, Payne VL, et al.** Cognitive interventions to reduce diagnostic error: a narrative review. BMJ Qual Saf. 2012;21:535-57. [PMID: 22543420] doi:10.1136/bmjqs-2011-000149
14. **Ely JW, Graber ML, Croskerry P.** Checklists to reduce diagnostic errors. Acad Med. 2011;86:307-13. [PMID: 21248608] doi:10.1097/ACM.0b013e31820824cd
15. **National Academies of Sciences, Engineering, and Medicine.** Improving Diagnosis in Health Care. Washington, DC: National Academies Pr; 2015.
16. **Green MJ, Croskerry P, Rieck R.** Annals Graphic Medicine - Bed blocker. Ann Intern Med. 2020;172:W142-W148. doi:10.7326/G20-0001

Current Author Addresses: Dr. Restrepo: Core Educator Faculty and Hospital Medicine Unit, Department of Medicine, Massachusetts General Hospital, 50 Staniford Street, Suite 550, Boston, MA 02114.
Dr. Armstrong: Massachusetts General Hospital, Gray-Bigelow 7, 55 Fruit Street, Boston, MA 02114.
Dr. Metlay: Division of General Internal Medicine, Massachusetts General Hospital, 100 Cambridge Street, 16th Floor, Boston, MA 02114.

Author Contributions: Conception and design: D. Restrepo, K.A. Armstrong, J.P. Metlay.
Analysis and interpretation of the data: J.P. Metlay.
Drafting of the article: D. Restrepo, K.A. Armstrong, J.P. Metlay.
Critical revision of the article for important intellectual content: K.A. Armstrong, J.P. Metlay.
Final approval of the article: D. Restrepo, K.A. Armstrong, J.P. Metlay.
Administrative, technical, or logistic support: K.A. Armstrong.