Diagnostic Errors and the Bedside Clinical Examination

Bennett W. Clark, MDa,*, Arsalan Derakhshan, MDa, Sanjay V. Desai, MDb

INTRODUCTION

In 2014, a 48-year-old woman with a history of stroke and uncontrolled diabetes presented to her local hospital for evaluation of a lesion on the left side of her face (Fig. 1). Previous swabs of the lesion had grown methicillin-resistant Staphylococcus aureus, so her doctors diagnosed her with cellulitis and sent her home with a peripherally inserted central catheter (PICC) line and a 10-day course of intravenous (IV) vancomycin. Unfortunately, the lesion did not improve, and she returned to the same hospital.

KEYWORDS
• Diagnostic error • Clinical reasoning • Clinical decision-making
• Heuristics and biases • Dual-processing theory • Medical education

KEY POINTS
• Diagnostic errors are common in clinical practice and result in adverse patient outcomes.
• Diagnostic errors are frequently unrecognized and under-reported because of individual and systematic factors.
• Deficiencies or omissions in the bedside clinical examination and in disease-specific content knowledge are among the most common causes of diagnostic errors.
• Unconscious heuristics and biases contribute to diagnostic errors.
• Research in clinical settings suggests that education in clinical content knowledge and bedside history and physical examination skills can reduce diagnostic errors.

INTRODUCTION

In 2014, a 48-year-old woman with a history of stroke and uncontrolled diabetes presented to her local hospital for evaluation of a lesion on the left side of her face (Fig. 1). Previous swabs of the lesion had grown methicillin-resistant Staphylococcus aureus, so her doctors diagnosed her with cellulitis and sent her home with a peripherally inserted central catheter (PICC) line and a 10-day course of intravenous (IV) vancomycin. Unfortunately, the lesion did not improve, and she returned to the same hospital.
twice over the next year. Both times, her doctors sent her home with a PICC line for more IV vancomycin. Convinced that the woman had refractory cellulitis, her outpatient doctors gave her additional courses of oral antibiotics. Despite these treatments, the lesion on her face never improved.

More than a year later, she was admitted to the general medicine service of a teaching hospital. Her neurologic examination revealed decreased sensation on the right side of her body and a left-sided Horner syndrome consistent with a prior lateral medullary stroke, a diagnosis confirmed by review of a prior MRI scan. Additionally, a punch biopsy of the facial lesion showed no evidence of cancer, infection, or autoimmune pathology. This, combined with evidence of injury to the left spinal trigeminal nucleus led to the diagnosis of trigeminal trophic syndrome—a rare, noninfectious condition caused by neuropathic itch, decreased facial sensation, and chronic skin abrasion from scratching in the distribution of the trigeminal nerve.¹

In the end, it took more than a year to give the woman an accurate diagnosis. Why did it take so long, and what explains the tenacity of the cellulitis diagnosis despite abundant evidence against it? Finally, and most importantly, how can it be done better?

Diagnostic error is a central concern in medicine and has had increased focus from stakeholders across the professional community and the public over the last 20 years. This article aims to orient readers to this complex field, with particular attention to

1. The impact of diagnostic errors on patient outcomes
2. Controversies in defining and studying diagnostic errors
3. Diagnostic errors common in clinical practice
4. Conditions, both environmental and cognitive, that predispose doctors to making diagnostic errors
5. Methods for improving diagnostic accuracy
THE IMPACT OF DIAGNOSTIC ERRORS ON PATIENT OUTCOMES

Diagnosis is at the heart of a doctor’s craft. It is the precondition of effective treatment and the foundation of trust between doctor and patient. It is also a point of professional pride. When doctors realize they have missed a diagnosis, they feel guilt and remorse.

But diagnostic errors take a far greater toll on patients’ lives than on doctors’ psyches. “To Err is Human,” a landmark study published by the Institute of Medicine (IOM) in 1999, estimated that diagnostic errors were responsible for 17% of preventable adverse hospital events. A review of more than 30,000 New York hospital records found that 14% of hospital errors were diagnostic in nature, and that most diagnostic errors were not only preventable but negligent. The problem is no less serious in the outpatient setting. Observational studies suggest that primary care doctors miss about 12 million diagnoses each year, and that about half of these misses cause patients significant harm.

Advanced medical technology appears to make only a marginal impact on diagnostic accuracy. Studies comparing the frequency of missed diagnoses before and after the advent of modern diagnostic imaging found little improvement in diagnostic accuracy. A more recent analysis argues that this lack of improvement is likely an artifact of clinical selection bias. Autopsies are far less common than they were prior to the use of cross-sectional imaging, and cases that do proceed to autopsy tend to be complex. Controlling for this selection bias, the rate of major diagnostic error is likely around 8%, in line with recent reviews of intensive care unit (ICU) autopsy cases. Even at this modestly improved error rate, as many as 35,000 patients die in US hospitals each year because of a missed diagnosis.

CONTROVERSIES IN DEFINING AND STUDYING DIAGNOSTIC ERRORS

Diagnosis can refer to the explanation for a patient’s condition, or the process of arriving at this explanation. This ambiguity has contributed to a lack of systematicity in research on diagnostic error. Newman-Toker helped resolve these semantic problems by distinguishing between failures in the diagnostic process and failures in diagnostic labeling (Fig. 2). Most clinicians can easily recall cases in which these 2 types of error were linked, when flawed thinking led to an incorrect or delayed diagnosis. However, it is also possible to get the process wrong but the label right, such as when a radiologist misses a malignant tumor on chest radiograph, but the cancer is

![Fig. 2. Schema for the classification of diagnostic errors. (Data from Newman-Toker DE. A unified conceptual model for diagnostic errors: underdiagnosis, overdiagnosis, and misdiagnosis. Diagnosis (Berl) 2014;1(1):43–8.)](image-url)
identified by another member of the health care team before the malignancy progresses in stage. In this case, the patient receives the correct label despite a flaw in the process. In Newman-Toker’s updated taxonomy of diagnostic errors, these instances of flawed diagnostic reasoning leading to an accurate diagnostic label are called near misses.

The reverse can also happen. Kassirer and Kopelman described a 53-year-old woman who returned from an overseas trip during which she had eaten at unsanitary restaurants and developed diarrhea. Microscopic examination of her stool revealed multiple parasites, and she was diagnosed with intestinal parasitosis. However, her diarrhea worsened after treatment for parasites, and she was ultimately diagnosed with a vasoactive intestinal peptide (VIP)-secreting tumor. Newman-Toker calls these cases, along with conditions that cannot be diagnosed using current medical technology, as unavoidable diagnostic errors. Although this is an important conceptual distinction, the practicing clinician may wonder, justifiably, whether something unavoidable should be considered an error at all. In keeping with the preponderance of current research on diagnostic error, this article focuses on avoidable errors.

DIAGNOSTIC ERRORS COMMON IN CLINICAL PRACTICE

Diagnosis Label Failures

Doctors have limited insight into their diagnostic skills. They have similar confidence with common, standardized clinical cases, which they diagnose correctly more than half of the time, as with unusual cases, which they solve correctly only 5% of the time. A retrospective review of autopsy cases from a medical intensive care unit found that doctors who were completely certain of their diagnosis were wrong 40% of the time. Overconfidence is not unique to the medical profession, and examples of this better-than-average effect are widely reported in social psychology literature. Put simply, without external feedback, doctors rarely predict the accuracy of their diagnoses. This phenomenon is reflected in reviews of error-reporting systems, in which computerized error identification turns up 10 times as many errors as physician self-report.

Individualized data on diagnostic error are lacking, so most information on missed diagnoses often comes from pooled data sets. For example, missed cases of cancer account for more than half of malpractice claims against outpatient internal medicine physicians. Singh and colleagues performed a retrospective review of 209 missed diagnoses in the ambulatory setting, in which the most common missed diagnosis was pneumonia, at 7% of the total. Missed primary cancer accounted for 6% of the total missed diagnoses in this study. Voluntary surveys of doctors, which are susceptible to recall biases, report primary cancers as the most common category of missed diagnosis, highlighting the challenge of measuring the rates of diagnostic errors accurately.

Diagnostic Process Failures

Failures in diagnostic processing and clinical reasoning are more difficult to identify than failures in diagnostic labeling. Advances in cognitive psychology over the last 50 years have uncovered some of the reasons why, beginning with the fact that, as doctors gain experience, they rely heavily on rapid, unconscious processes to make diagnoses. Thus, the specific processes a doctor uses to arrive at a diagnosis are hidden not only from researchers but also from the doctor. Even when doctors take the artificial step of thinking aloud about their diagnostic process, their descriptions are unreliable. Moreover, once doctors or other observers know the outcome of a
case, they consistently overestimate what could have been known prior to the diagnosis being firmly established, a phenomenon known as hindsight bias.24

Useful frameworks to examine the features and failures in diagnostic processing have been developed (Fig. 3). Kassirer and Kopelman divided the process into 4 steps:

1. Hypothesis generation, which they called triggering
2. Framing the patient’s problem
3. Gathering and processing information, such as findings on the clinical examination and laboratory tests
4. Verifying the diagnosis by making sure that competing hypotheses can be reasonably excluded

They found that errors in gathering and processing information were the most common, followed by errors in triggering.16

SYSTEMS, COGNITIVE AND PRACTICE-RELATED CONTRIBUTIONS TO DIAGNOSTIC ERROR

System-Related Factors

Organizational and environmental factors play an important role in diagnostic errors. Among these factors are

- Reimbursement structures that discourage consultation
- Incomplete medical records
- Cultural and logistical barriers to communication between doctors
- High physician workloads
- Patient failure to follow up
- Community hospital settings (compared with teaching hospitals)25

Cognitive Factors

Heuristics are methods used to solve problems quickly. Doctors use heuristics to make diagnoses all the time. However, heuristics can also lead to errors, because they can introduce unconscious biases. In the early 1970s, the cognitive psychologists Amos Tversky and Daniel Kahneman demonstrated that heuristics lead to predictable errors in judgment. In 1 experiment, they played a tape-recorded list of names, then asked participants to estimate whether the list included more women or men. When
the list included famous women (eg, Elizabeth Taylor) and less famous men (eg, William Fulbright), 80% of participants erroneously believed the list had more women.26 This phenomenon of overestimating the frequency of things that come to mind readily is now called availability bias.

Heuristics and their associated biases have been subsumed into an overarching model of cognitive reasoning called dual-processing theory. This theory describes 2 systems for making judgments and decisions. System 1 is rapid, instinctive, automatic, and driven by networks of associations. Take a moment to look at the face in Fig. 4. After just a few seconds, you will have made numerous inferences about the person’s age, mood, and background. These rapid judgments represent the outcome of System 1 processing. System 2, on the other hand, is deliberate, sequential, logical, and demands cognitive energy. Try to solve the following problem without pen or paper: 673 x 779. To have any chance of success, one must block out distractions and come up with a plan for keeping track of multiplied values. This is a System 2 task.27 Not surprisingly, physicians rely more heavily on System 1 as they become more experienced.28

Croskerry has described how patterns of bias can corrupt diagnostic reasoning. His survey29 of 32 common errors in diagnostic reasoning includes habits of thought that most doctors will find familiar, such as the sunk costs phenomenon—“the more clinicians invest in a particular diagnosis, the less likely they may be to release it and consider alternatives”—and anchoring—“the tendency to perceptually lock onto salient features in the patient’s initial presentation too early in the diagnostic process, failing to adjust this initial impression in light of later information.” An abbreviated version of Croskerry’s list is presented in Table 1.

---

Although many writers and researchers have described these flawed decision processes, few have undertaken rigorous studies of how they influence patient care. Researchers in the Netherlands have demonstrated the power of availability bias in skewing clinical cases presented to trainees in a booklet.\textsuperscript{30,31} A prospective study of trainees managing simulated emergencies found that premature closure (31\%) and confirmation bias (30\%) were the most common cognitive contributors to erroneous diagnosis.\textsuperscript{32} By contrast, a prospective analysis by Voytovich of student and physician efforts to solve 3 written clinical cases found that more than 90\% of participants missed a diagnosis because of premature closure.\textsuperscript{33} To the authors’ knowledge, no prospective studies have examined cognitive bias in patient care environments.

The discrepancies in these studies’ findings reflect a general lack of standardization among researchers investigating the cognitive psychology of diagnosis. Did Voytovich’s study differ from the study of simulated emergencies because it was conducted in a different setting, because of different methods for obtaining data about participants clinical reasoning, or because of different definitions of what counts as premature closure? The current literature on the cognitive psychology of clinical diagnosis is not mature enough to answer these questions.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Common biases that lead to diagnostic errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate bias</td>
<td>The tendency to believe that aggregated data, such as those used to develop clinical practice guidelines, do not apply to individual patients (especially their own)</td>
</tr>
<tr>
<td>Anchoring</td>
<td>The tendency to perceptually lock onto salient features in the patient’s initial presentation too early in the diagnostic process failing to adjust this initial impression in the light of later information; this CDR may be severely compounded by the confirmation bias</td>
</tr>
<tr>
<td>Availability bias</td>
<td>The disposition to judge things as being more likely, or frequently occurring, if they readily come to mind; recent experience with a disease may inflate the likelihood of its being diagnosed, and conversely, if a disease has not been seen for a long time (is less available), it may be underdiagnosed</td>
</tr>
<tr>
<td>Confirmation bias</td>
<td>The tendency to look for confirming evidence to support a diagnosis rather than look for disconfirming evidence to refute it, despite the latter often being more persuasive and definitive</td>
</tr>
<tr>
<td>Outcome bias</td>
<td>The tendency to opt for diagnostic decisions that will lead to good outcomes, rather than those associated with bad outcomes, thereby avoiding chagrin associated with the latter</td>
</tr>
<tr>
<td>Overconfidence bias</td>
<td>A universal tendency to believe one knows more than he or she does; overconfidence reflects a tendency to act on incomplete information, intuitions, or hunches</td>
</tr>
<tr>
<td>Premature closure</td>
<td>The tendency to apply premature closure to the decision-making process, accepting a diagnosis before it has been fully verified</td>
</tr>
<tr>
<td>Psych-out error: Representative restraint</td>
<td>The tendency to attribute presenting symptoms to psychiatric etiologies, especially in patients carrying a psychiatric diagnosis; serious medical conditions can be misdiagnosed as psychiatric conditions</td>
</tr>
</tbody>
</table>

Adapted from Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med 2003;78(8):777–8; with permission.
Some psychologists, postulating that heuristics are hardwired into the decision-making process by evolution, argue that the biases described are difficult if not impossible to unlearn. Others have attempted to show that reminding physicians about their biases and fallibility is a promising way to reduce diagnostic error. These efforts are described in the next section under the heading Reflective Practice.

The Missing Bedside Examination

When doctors recall cases in which they missed a diagnosis, they frequently report performing an incomplete bedside examination.\(^{34}\) Association between faulty bedside assessment and diagnostic error is corroborated in a systematic review of medical malpractice cases, in which failure to perform an adequate history and physical examination contributed to 42% of missed diagnoses. The same review showed almost 70% of missed cancer cases were because of an inadequate history and physical examination.\(^{34}\) In reviews of computer-identified diagnostic errors in outpatient internal medicine clinics, more than half of cases involved a shortcoming in the history or physical examination.\(^{26}\)

Disease-specific studies examining trends in missed diagnoses among grave conditions with benign mimics suggest that deficiencies in bedside clinical evaluation often contribute to diagnostic delay. A retrospective review of cases of ruptured aortic abdominal aneurysm (AAA) found that doctors missed this diagnosis 61% of the time. Unfamiliarity with the cardinal signs of a contained AAA rupture—urinary retention, flank pain, abdominal distension, leukocytosis, and an absence of shock or anemia—was common among doctors missing the diagnosis, as was a failure to palpate large AAAs, even in patients without abdominal distension.\(^{35}\) Kowalski and colleagues\(^{36}\) reviewed 56 cases of missed subarachnoid hemorrhage and found that doctors’ unfamiliarity with the phenomenon of sentinel headaches correlated with missed diagnosis.

Clinical mimicry, combined with deficiencies in the clinical examination, also contributes to overdiagnosis. A retrospective review of patients referred to a Lyme disease specialty clinic found that 77% did not have active Lyme disease, and that many patients would have avoided misdiagnosis if their referring physicians had been able to distinguish degenerative arthritis from Lyme arthritis.\(^{37}\)

**HOW TO IMPROVE DIAGNOSTIC ACCURACY**

*System-Based Interventions*

The literature examining real-world methods for avoiding diagnostic error is limited, especially in light of the abundant theoretic writing on the topic in the past decade. Although data are sparse, experts are optimistic about electronic medical records’ (EMRs) ability to reduce errors, not only by catching them before they become clinically consequential, but also by actively steering clinicians toward accurate diagnoses.\(^{38}\) Nonrandomized prospective studies have demonstrated effectiveness of electronic interventions using several approaches:

- A diagnostic decision support system (DDSS) that generated diagnostic suggestions based on preliminary clinical data improved the accuracy of pediatrics residents.\(^{39}\)
- Computer-guided patient histories can provide important clinical data that physicians fail to elicit.\(^{40}\)
- An automated electrocardiogram (ECG) interpretation program improved interns’ accuracy in the diagnosis of acute myocardial ischemia.\(^{41}\)
The most common system-based intervention aimed at enhancing patient outcomes by reducing diagnostic error involves building redundancy into the interpretation of diagnostic tests. Redundancy is especially relevant for clinical pathology and radiology, where data suggest that interpretation of a study by more than 1 physician can improve test sensitivity. For example, review of computed tomography (CT) colonoscopies by 2 radiologists instead of one increased sensitivity for underlying colon cancer, although it also decreased specificity. Although such strict methods of redundancy may not seem relevant for internal medicine, Graber and colleagues note that the impulse to seek help from colleagues, whether through curbside consultation or a formal second opinion, can also improve diagnostic accuracy, a premise that has been confirmed in simulated clinical problem solving.

Reflective Practice

Cognitive psychologists spent much of the twentieth century cataloging System 1’s habitual inaccuracy, so it should come as no surprise that much of the current clinical reasoning research aims at getting diagnosticians out of System 1 and into System 2. The idea is that one can reduce diagnostic error by thinking slow instead of fast.

For example, a group of psychologists in the Netherlands presented internal medicine residents with a booklet of clinical vignettes that had been selected to activate availability bias, a well-described bias of System 1. Residents who solved the cases in an unstructured manner were less accurate than those who went through a prescribed process of listing data for and against their initial diagnosis before deciding on a final diagnosis. A prospective study evaluating undergraduate students’ ability to learn ECG interpretation found that they were more accurate when receiving this prompt: “Don’t jump the gun; consider the feature list before providing a final diagnosis.”

Common sense suggests that a slow, deliberative approach to diagnosis is not needed in all cases. For an experienced clinician, common syndromes with clear signs—decompensated heart failure, cirrhosis, psoriasis—do not require a second thought. A study comparing diagnostic strategies in simple and complex cases supports this premise; for simple cases, deliberation does not boost diagnostic accuracy. For complex cases, it does.

Unfortunately, no pro-System 2 study has been conducted in a clinical setting. The result is research that confirms something most diagnosticians already understand: when posed with cases that are complex, challenging, or misleading, clinicians do a better job when they have the time, space, and resources to slow down and organize their thoughts. The more difficult question is how clinicians can better recognize cases that require a System 2 approach, and how medical technology, workflows, and infrastructure can better support clinicians when such cases arise. Consider the example of the woman at the beginning of this article. For more than a year, her physicians approached her case with a decidedly System 1 approach, diagnosing her over and over with cellulitis. The fact that she had completed multiple courses of IV antibiotics without getting better should have prompted reconsideration of the diagnosis, but it did not. What training could her doctors have received to help them toggle from System 1 to System 2? How could their working environment have been improved to help them do so? These are active areas of research for proponents of cognitive debiasing and reflective practice.

The Clinical Examination and Clinical Content Knowledge

Imagine that the patient described previously had been treated by doctors capable of recognizing the signs of a lateral medullary stroke. If they had known what they were
looking at in the first place, an effortful System 2 analysis might not have been needed to diagnose trigeminal trophic syndrome. It is also possible that they were able to recognize such signs, but never stopped to look. Reviews of malpractice claims and surveys of physicians suggest that failure to perform a complete bedside examination underlies many diagnostic errors. To correct this failing, 1 expert has recommended that clinicians use a diagnostic checklist, the first 2 steps of which are to obtain a complete history and complete a purposeful physical examination.

Education in bedside medicine has the potential to reduce diagnostic errors. The evidence for such interventions is more robust than the evidence for cognitive debiasing and reflective practice, having been proven in real-world clinical settings. Several prospective studies have demonstrated that provider-specific feedback and disease-specific education improve diagnostic accuracy. Providing emergency room physicians with intensive, real-time feedback on the outcomes of discharged patients decreased physicians’ rate of adverse events.48 Similar interventions with attending psychiatrists and trainees in clinical psychology resulted in improvements in diagnostic accuracy.49 A regional program to educate primary care physicians on the clinical presentation of subarachnoid hemorrhage resulted in more timely diagnosis for patients with this life-threatening syndrome.50 A renewed emphasis on the bedside examination and deliberate feedback loops to physicians may be the most effective way to reduce diagnostic errors.

SUMMARY

Diagnostic errors cause patients serious harm. Such errors arise from a complex set of factors at both the system and the clinician level. Although most interventions to reduce diagnostic error have focused on system-level improvements, recent advances in cognitive psychology have prompted debate on how best to improve clinicians’ diagnostic reasoning. A limited body of evidence suggests that adjusting habits of thought can lead to more accurate diagnosis. However, the preponderance of evidence in this field points to the importance of improving bedside skills and receiving detailed clinical feedback.

REFERENCES


