Perspectives on Diagnostic Failure and Patient Safety

Pat Croskerry

The Institute of Medicine (IOM) report *To Err Is Human* (Kohn et al. 1999), published at the turn of the century, signalled the beginning of the modern era of patient safety. Although a variety of good work followed, in hindsight what appears to have been missing at the beginning was sufficient pause to ask which of the many possible directions should be taken to achieve the most value for effort; only recently has this question been framed (Pronovost and Faden 2009). At the outset, it was understandable that the focus should be put on the tangible, obvious and most easily fixed problems. Medication error, for example, was quickly identified, and considerable work followed in this area. Other less tangible areas attracted less attention. Diagnostic failure appears to have been a major oversight and remains a major threat to patient safety. Wachter (2010) notes that the term medication error is mentioned more than 70 times in the IOM report (Leape et al. 1991), whereas diagnostic error appears only twice – even though physician errors that lead to adverse events are significantly more likely to be diagnostic (14%) than medication related (9%) (Leape et al. 1991). In both Canada (Canadian Medical Protective Association, personal communication, 2011) and the United States (Chandra 2005), diagnostic error is the leading cause of malpractice litigation. At the end of a long career in medical decision-making, Elstein (2009) estimated that diagnostic failure across the board in medicine was in the order of 15%.

The Invisibility of Diagnostic Failure
A major part of the problem is the inherent obscurity, complexity and irreducible uncertainty associated with human illness. In turn, the diagnostic process that attempts to identify it is similarly obscure. These issues have been well articulated by Montgomery (2006). As she notes, if illness were to follow scientific rules, our difficulties in this area would have been solved long ago. Instead, physicians often have to deal with a process that has “mushy but unavoidable ineffabilities” (Montgomery 2006). This is the process that takes place at the dynamic interface between the patient’s and the physician’s thoughts, feelings, reasoning and decisions.

Part of the problem, too, has been the tight control that physicians have exerted over diagnosis as well as an attendant culture of silence and secrecy – both have led to further obfuscations of the process. Other healthcare workers have been largely excluded, sharing little or no responsibility or accountability, with hospitals usually more than willing to distance themselves from the decisions of their physicians. From a medical-legal standpoint, diagnosis is largely a physician responsibility. Another important reason has been the invisibility of diagnostic error. At the outset of the modern patient safety era, its impact was obscured through the principal methodology used in the benchmark studies (Brennan et al. 1991; Thomas et al. 2000; Wilson et al. 1995), retrospective chart review, which did not readily permit the detection of diagnostic failure. Whereas
medication errors or failed surgical procedures are often clearly visible using this methodology, diagnostic errors are not. They would not be discussed in progress notes or discharge summaries, and there are few triggers that would help reviewers detect them. Even today, the Global Trigger Tool developed by the Institute for Health Improvement (Classen et al. 2011), acclaimed as a more sensitive and efficacious method of detecting adverse events, does not capture diagnostic error. A variety of reasons can be posited for the failure to recognize the important role of diagnostic failure in patient safety (Table 1).

The Psychology of Decision-Making
A major factor in diagnostic error has been that understanding the complex psychological processes that underlie clinical decision-making has not been seen as relevant to medical training. There remains a tacit belief by medical educators that becoming a well-calibrated decision-maker is a process that occurs through exposure to the decision-making of mentors or other exemplars and through experience, but there is usually no specific training. While sound formal training does exist for analytical, quantitative, rational decision-making of the type described by Rao (2007), there is little available on intuitive decision-making, where clinicians spend most of their time (Figure 1). Parenthetically, rational decision-making must also include, although it is rarely acknowledged, a consideration of the elements of argumentation, logic and critical thinking that are implicit in sound reasoning (Jenicek and Hitchcock 2005).

A further problem has been that rational clinical decision-making is far more amenable to experimental testing than is intuitive reasoning. The properties of rational thought are easily described, conforming as they do to fundamental laws of science, objectivity and probability. They have scientific rigour and are reproducible. In contrast, intuitive decision-making is often much more dependent on the individual, non-scientific, unpredictable and less reproducible. It is difficult to re-create in a laboratory the rich, complex context under which many clinical decisions are made. It is not surprising, therefore, that rational decision-making has emerged as the main focus of medical decision-making researchers.

Neither have there been adequate models of decision-making to explain the process of diagnostic reasoning, until quite recently. The history of clinical decision-making in medicine has followed a fairly conservative and sterile path, largely confining itself to a respectable, orthodox, rational approach. Unquestionably, this is an important part of decision-making; but given that we spend most of our thinking time in the intuitive mode (Lakoff and Johnson 1999), there is now a strong imperative to better understand its nature and consequences and to incorporate it into a more comprehensive approach toward medical decision-making.

Two important developments have facilitated our understanding of medical reasoning and provided insight into diagnostic failure. They came about through the “cognitive revolution” in psychology in the latter part of the past century, where “mental processes” were rediscovered and seen fit for study. The first was an important new way of looking at everyday thinking, termed heuristics and biases (Tversky and Kahneman 1971, 1974). A substantial literature has revealed that much of our thinking is vulnerable to a variety of affective and cognitive biases that exert their influence, particularly under conditions of uncertainty, when we are using heuristics (thinking shortcuts). This is an inherent feature of human reasoning. An example is illustrated in the sidebar.

However, initial efforts to apply the findings of psychologists to medical decision-making did not attract sufficient interest in the broader clinical community; they were viewed as lacking in ecological validity and therefore clinically irrelevant (i.e., experimental studies in the laboratory were too far removed from clinical reality). A turning point came with the publication of a book on clinical reasoning in which two prominent internists showed how psychological findings could be directly applied in the clinical context (Kassirer and Kopelman 1991).

The second development was the emergence (Dawson 1993; Elstein 1976) and progressive refinement (Evans and Frankish 2009) of dual process models of thinking; these held that the multiple varieties of thinking could be organized into two main modes: intuitive decision-making, where we spend most of our time, and analytical thinking, which is used more explicitly for deliberate problem solving. Over 35 years ago, the important distinction between rational and intuitive approaches was well articulated as a statistical-clinical polarization by Elstein (1976) in an article prophetic of modern dual process theory.

Most of our thinking occurs in the intuitive mode (Lakoff and Johnson 1999), where heuristics predominate (Gilovich et al. 2002). Dual process theory proper was introduced into medicine in the early 1990s (Dawson 1993) and subsequently provided a useful platform for a universal model for diagnostic reasoning (Croskerry 2009). Experienced clinicians make much of their diagnostic decision-making quickly and effectively in the intuitive mode using pattern recognition, but need to access the analytical mode for novel, ambiguous and more challenging cases. The particular operating characteristics of the dual process approach in diagnostic reasoning have been described (Croskerry 2009). In the past decade, there has been an increased willingness from medical quarters to accept the relevance and importance of the research findings of cognitive psychologists (Redelmeier 2005).

Knowledge Gaps or Thinking Deficits?
A strong consensus has emerged that thinking failures occur predominantly in the intuitive mode, having their origins in failed heuristics and the maladaptive influence of cognitive
TABLE 1.
Reasons for lack of focus on diagnostic failure

<table>
<thead>
<tr>
<th>Reason</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amorphous nature of many illnesses</td>
<td>Illness and disease do not follow precise rules. While some illnesses are straightforward, many are complex and have covert features that are difficult to detect and quantify. A &quot;scientific&quot; approach is not always possible.</td>
</tr>
<tr>
<td>Invisibility</td>
<td>Critical parts of diagnostic failure are the physician’s clinical reasoning and decision-making processes, which are largely covert and poorly understood.</td>
</tr>
<tr>
<td>Accountability</td>
<td>Diagnosis, for the most part, is the responsibility of physicians. There is little incentive for others to concern themselves about the process if they are not immediately involved or have no direct accountability.</td>
</tr>
<tr>
<td>Attitude of silence and culture of silence</td>
<td>Historically, it has been relatively easy for one group (physicians) to control and suppress information about a specific topic, especially when not to do so might reveal weakness or incompetence. Generally, physicians are not forthcoming about their failures and are unwilling to discuss those of their colleagues. The result is that diagnostic failure has not reached the light of day. This is now improving, with morbidity and mortality rounds taking a more open approach toward error.</td>
</tr>
<tr>
<td>Denial: discounting and distancing</td>
<td>When diagnostic failure occurs, physicians are often unwilling to confront their own failings; they use various forms of denial: discounting the failure by blaming the illness or circumstances they could not control, or blaming the patients (for failing to give a full history, non-compliance etc.), other healthcare workers or the system.</td>
</tr>
<tr>
<td>Non-disclosure</td>
<td>A significant proportion of diseases resolves spontaneously or is “cured” through placebo effects. Disclosing failures to diagnose might reduce the confidence of patients in their doctors and weaken the power of the placebo. Thus, diagnostic failure is generally not sufficiently broadcast.</td>
</tr>
<tr>
<td>Investigation</td>
<td>The main methodological genre for investigating adverse events has been retrospective chart review. However, diagnostic failure is often not apparent from what is recorded in a hospital chart. Trigger tool methodology is similarly flawed due to the invisibility of diagnostic failure. Thus, diagnostic failure has been less visible and has attracted less attention.</td>
</tr>
<tr>
<td>Natural course of illness</td>
<td>Of the 12,000 or so diagnostic entities that currently exist, probably more than 80% resolve without the help of physicians. Even if a physician has the wrong diagnosis and prescribes an inappropriate treatment, the patient’s illness might still improve. Thus, diagnostic performance often looks better than it is, and people have more faith in it than is actually warranted. Alternative medicine owes much of its popularity to the natural course of most illnesses to resolve.</td>
</tr>
<tr>
<td>Lack of awareness</td>
<td>Outcome feedback has traditionally been poor in medicine, and physicians may be unaware that they have failed in their diagnosis, especially if their patients go to another institution or provider or, worse still, die. Patients themselves may “protect” well-liked physicians to spare their feelings.</td>
</tr>
<tr>
<td>Lack of understanding of the decision-making process</td>
<td>The psychological aspects of decision-making are not formally addressed in medical undergraduate curricula. Physicians receive little or no training on decision-making or on the influence of cognitive and affective biases. They do not generally reflect upon or view introspectively their own decision-making behaviours. There has been little incentive to change something that is poorly understood.</td>
</tr>
<tr>
<td>Unavailability of tangible solutions</td>
<td>Whereas other sources of error have been amenable to tangible solutions (checklists in surgery, computerized order entry for medications, safety bundles for specific error-prone practices), there are few options currently available for improving diagnostic performance, although some current clinical decision-support systems and other strategies (e.g., cognitive debiasing) show promise.</td>
</tr>
<tr>
<td>Tolerance</td>
<td>True costs of diagnostic errors are difficult to assess, and it has been easy for healthcare systems and payers to quietly tolerate them providing that physicians are not producing wild outliers.</td>
</tr>
<tr>
<td>Poor business case</td>
<td>Historically, healthcare systems and insurers have seen the investigation and remediation of intangible and obscure diagnostic adverse events as a poor business case, preferring to develop interventions in other areas of medical error. This may change with new data emerging that show that diagnostic adverse events may increase hospital stays by more than 80%.</td>
</tr>
</tbody>
</table>
and affective biases. As Stanovich notes (2011), there is now a voluminous psychology literature on the topic, to which can be added a recently burgeoning medical literature (Berner and Graber 2008; Croskerry 2005; Lawson et al. 2011; Neale et al. 2011; Pennsylvania Patient Safety Authority 2010; Redelmeier 2005; Scott 2009; Wachter 2010; Zwaan et al. 2010). An important study by Graber et al. (2005) implicated cognitive factors in about 75% of diagnostic errors; and in a subsequent and very much larger population-based Dutch study, this was estimated at 96% (Zwaan et al. 2010). Efforts to challenge this consensus are few, being largely confined to proponents of naturalistic decision-making (NDM; Klein et al. 1993) and others (Norman and Eva 2010). Researchers in NDM have relied on cognitive field research methods to study skilled performance and have largely dismissed the heuristics and biases approach as a laboratory artifact (Klein et al. 1993). They have placed a strong emphasis on intuitive decision-making, despite repeated demonstrations of its vulnerability to failure (Kahneman 2011). In view of protean demonstrations of heuristics and biases in all walks of everyday life, it is extremely unlikely that similar phenomena do not prevail in medical decision-making (Ruscio 2006; Whaley and Geller 2007), and the once popular criticism that they are only found in the laboratory has become a “cartoon” (Kahneman and Klein 2009).

The view has also been advanced that diagnostic failure may be explained by physicians’ knowledge deficits (Norman and Eva 2010), although this too is unsupported by clinical findings. In a study of 6,000 patients in primary care who were cared for by mostly internal medicine residents, diagnostic errors were found in “a wide spectrum of diseases commonly seen in an outpatient general medical setting” (Singh et al. 2007). And

**FIGURE 1. The two major categories of medical decision-making**

<table>
<thead>
<tr>
<th>Medical decision making</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rational</td>
</tr>
<tr>
<td>Informal</td>
</tr>
<tr>
<td>Formal</td>
</tr>
<tr>
<td>Subjective</td>
</tr>
<tr>
<td>Objective</td>
</tr>
<tr>
<td>Context dependent</td>
</tr>
<tr>
<td>Quantitative</td>
</tr>
<tr>
<td>Qualitative</td>
</tr>
<tr>
<td>Verifiable</td>
</tr>
<tr>
<td>Dynamic</td>
</tr>
<tr>
<td>Verifiable</td>
</tr>
<tr>
<td>Flexible</td>
</tr>
<tr>
<td>Rigorous</td>
</tr>
</tbody>
</table>

**Diagnostic Failure Due to Intuitive Biases**

A 28-year-old female patient is sent to an emergency department from a nearby addictions treatment facility. Her chief complaints are anxiety and chest pain that have been going on for about a week. She is concerned that she may have a heart problem. An electrocardiogram is routinely done at triage. The emergency physician who signs up to see the patient is well known for his views on “addicts” and others with “self-inflicted” problems who tie up busy emergency departments. When he goes to see the patient, he is informed by the nurse that she has gone for a cigarette. He appears angry, and verbally expresses his irritation to the nurse.

He reviews the patient’s electrocardiogram, which is normal. When the patient returns, he admonishes her for wasting his time and, after a cursory examination, informs her she has nothing wrong with her heart and discharges her with the advice that she should quit smoking. His discharge diagnosis is “anxiety state.” The patient is returned to the addictions centre, where she continues to complain of chest pain but is reassured that she has a normal cardiogram and has been “medically cleared” by the emergency department. Later in the evening, she suffers a cardiac arrest from which she cannot be resuscitated. At autopsy, multiple small emboli are evident in both lungs, with bilateral massive pulmonary saddle emboli.

**Comment:** This case illustrates several intuitive biases that lead the physician to an inadequate assessment of the patient and, ultimately, an incorrect diagnosis. First, he stereotypes and labels the patient, failing to provide the assessment that he might otherwise do for a non-addicted patient. He focuses on the disposition of the patient rather than the circumstances that have led her into addiction (fundamental attribution error); she has had a long history of physical and sexual abuse. These biases lead to premature closure of his diagnostic reasoning process. Further, his anger at the patient produces visceral arousal and distracts him from his usual practice of reviewing medications and noticing on her triage chart that she is on a birth control pill and is a smoker – both risk factors for pulmonary embolus.

The biases described in this case reside in the intuitive mode of reasoning and lead to diagnostic failure. An accurate diagnosis is more likely in the analytical mode; but in order to get to it, the physician must decouple from the intuitive mode through the process known as cognitive debiasing.
in a population-based Netherlands study involving close to 8,000 patients, the majority of diagnostic errors were associated with common conditions such as pulmonary embolism, sepsis, myocardial infarction and appendicitis (Zwaan et al. 2010). Further, in a study by Schiff et al. (2009) that surveyed experienced clinicians, it was found that the top 20 missed diagnoses were common, well-known illnesses. At the top of the list was pulmonary embolus (PE), the pathophysiology of which, and its associated risk factors, are very well known to clinicians. Yet, PE appears particularly difficult to reliably diagnose; in a study of fatal cases of PE, the diagnosis was not suspected in 55% of cases (Pineda et al. 2001). In my experience of three decades of attendance at morbidity and mortality rounds in emergency medicine, at which the majority of cases concern diagnostic failure, a knowledge deficit is very rarely identified as a reason for missing a diagnosis. Thus, it seems clear that how physicians think and not what physicians know is primarily responsible for diagnostic failure. This is an important distinction because it directs us to where remedial action should be taken to improve the calibration of diagnostic reasoning and patient safety.

**Extent of Diagnostic Failure**

Given that diagnostic failure is usually associated with physicians’ reasoning and thinking, it is present in every domain of medicine where decisions are required. It would be expected that diagnostic failure would be highest where uncertainty and ambiguity are high, and this appears to be the case. In the benchmark studies of medical error, diagnostic error was found to be high in emergency medicine, family practice and internal medicine (Brennan et al. 1991; Thomas et al. 2000; Wilson et al. 1995). In contrast, it would be expected to be lower in those areas where there is less uncertainty and ambiguity and clinical presentations are more pathognomonic, such as dermatology, orthopedics and plastic surgery. In an extensive review, Berner and Graber (2008) concluded that the overall rate of diagnostic failure in medicine is probably around 15% but that in the perceptual specialties that rely on visual interpretation (pathology, dermatology, radiology), it is considerably lower, at about 2–5%. Interestingly, the overall failure rate accords well with an ecologically valid study using standardized patients in a general internal medicine outpatient clinic in which the diagnostic failure rate was 13% (Peabody et al. 2004).

**Where Are We Now?**

The first articles specifically focusing on clinical diagnostic error and cataloguing it began to appear in the 2000s (Croskerry 2002, 2005; Elstein and Schwarz 2002; Graber et al. 2005; Redelmeier et al. 2001) and by the end of that decade were more robust (Berner and Graber 2008; Croskerry and Nimmo 2011; Schiff et al. 2009; Singh et al. 2007; Zwaan et al. 2010). In 2007, the Agency for Healthcare Research and Quality identified diagnostic error as an area of special emphasis, and much of the activity on its web-based morbidity and mortality site (http://webmm.ahrq.gov/) deals with cases containing diagnostic errors. In 2008, a working group of the World Health Organization Alliance for Patient Safety identified diagnostic error as a research priority, and in the same year the American Journal of Medicine produced a special supplement on the topic (Berner and Graber 2008). The first international conference on Diagnostic Error in Medicine was held in Phoenix, Arizona, in 2008 and the conference has been held annually since. The first conference on diagnostic error in the United Kingdom was held in 2011 in Edinburgh, Scotland. Society to Improve Diagnosis in Medicine (SIDM), has recently been established.

**Not all cognitive biases are created equal, and different cognitive pills might be needed for our different ills.**

**Where Do We Go from Here?**

For the multiple reasons outlined here, it is not surprising that diagnostic error has attracted insufficient attention in patient safety. Its covert nature and complex features have made it a tough problem to solve. One of the major impediments lies in medical education, where the theory and practice of decision-making have not been adequately addressed. However, the problem is now unmasked, and a variety of initiatives are under way. The next steps involve finding appropriate solutions to avoid failure, which remains a tall order. Diagnostic reasoning is mostly an invisible process – we infer failure usually after the fact. We can, however, offer plausible explanations, as well as combine what we know from research in cognitive psychology. A primary goal should be the development of a variety of educational initiatives aimed at teaching about decision-making, specifically, cognitive debiasing, advancing critical thinking and educating intuition (Croskerry and Nimmo 2011; Hogarth 2001).

Given that most diagnostic errors are due to cognitive failure, strong efforts should be made to develop a variety of debiasing strategies to mitigate the adverse consequences of cognitive and affective biases. We should acknowledge that not all cognitive biases are created equal and that different cognitive pills might be needed for our different ills (Keren 1990). Cognitive biases reside mostly in the intuitive mode of decision-making, and their power should not be underestimated. We should be aware that simplistic approaches toward debiasing are unlikely to be effective. Except in cases of significant affective arousal, we cannot expect that one-shot interventions will work (Lilienfeld et al. 2009) or that one particular type of intervention will be sufficient; many biases have multiple determinants, and it is unlikely that there is a “one-to-one mapping of causes to bias or
of bias to cure” (Larrick 2004). It seems certain that successful debiasing will inevitably require repeated training using a variety of strategies – the initial cognitive inoculation should be followed by repeated booster shots and will probably require lifelong maintenance.

Nevertheless, we can get on with improving those conditions known to be associated with cognitive failure – cognitive overload, fatigue and sleep deprivation/sleep debt. All of this work should be strongly embedded in the clinical context of patient safety. Additional strategies such as using checklists (Ely et al. 2011), forcing functions, algorithms, cognitive tutoring systems (Crowley et al. 2007) and other clinical decision support all require further research and development. As Thomas and Brennan (2010) note, the first chapter on diagnostic adverse events is now written: we know they are there and how important they are. The next chapter should focus on understanding the cognitive roots of diagnostic failure, its prevention and the improvement of patient care.

References


Ely J, Graber M, and P. Croskerry. 2011. “Checklists to Reduce Diagnostic Errors.” Academic Medicine 86:307-313. (This paper inspired the Perioperative Interactive Education (PIE) group at the Department of Anesthesia at Toronto General Hospital to develop an interactive Diagnostic Checklist website at http://pie.med.utoronto.ca/dc.)


Perspectives on Diagnostic Failure and Patient Safety  Pat Croskerry


About the Author  Pat Croskerry, MD, PhD, is a clinical consultant in patient safety, professor in Emergency Medicine, and director of the Critical Thinking Program in the Division of Medical Education at Dalhousie University in Halifax, Nova Scotia. He can be reached by e-mail at: croskerry@eastlink.ca.

Brought to you by:

jobs.longwoods.com

“Except in cases of significant affective arousal, we cannot expect that one-shot interventions will be sufficient.” See page 54.

56 Healthcare Quarterly Vol.15 Special Issue 2012