Diagnostic Error in Medicine

Analysis of 583 Physician-Reported Errors

Gordon D. Schiff, MD; Omar Hasan, MD; Seijeoung Kim, RN, PhD; Richard Abrams, MD; Karen Cosby, MD; Bruce L. Lambert, PhD; Arthur S. Elstein, PhD; Scott Hasler, MD; Martin L. Kabongo, MD; Nela Krosnjar; Richard Odwazny, MBA; Mary F. Wisniewski, RN; Robert A. McNutt, MD

Background: Missed or delayed diagnoses are a common but understudied area in patient safety research. To better understand the types, causes, and prevention of such errors, we surveyed clinicians to solicit perceived cases of missed and delayed diagnoses.

Methods: A 6-item written survey was administered at 20 grand rounds presentations across the United States and by mail at 2 collaborating institutions. Respondents were asked to report 3 cases of diagnostic errors and to describe their perceived causes, seriousness, and frequency.

Results: A total of 669 cases were reported by 310 clinicians from 22 institutions. After cases without diagnostic errors or lacking sufficient details were excluded, 583 remained. Of these, 162 errors (28%) were rated as major, 241 (41%) as moderate, and 180 (31%) as minor or insignificant. The most common missed or delayed diagnoses were pulmonary embolism (26 cases [4.5% of total]), drug reactions or overdose (26 cases [4.5%]), lung cancer (23 cases [3.9%]), colorectal cancer (19 cases [3.3%]), acute coronary syndrome (18 cases [3.1%]), breast cancer (18 cases [3.1%]), and stroke (15 cases [2.6%]). Errors occurred most frequently in the testing phase (failure to order, report, and follow-up laboratory results) (44%), followed by clinician assessment errors (failure to consider and overweighing competing diagnosis) (32%), history taking (10%), physical examination (10%), and referral or consultation errors and delays (3%).

Conclusions: Physicians readily recalled multiple cases of diagnostic errors and were willing to share their experiences. Using a new taxonomy tool and aggregating cases by diagnosis and error type revealed patterns of diagnostic failures that suggested areas for improvement. Systematic solicitation and analysis of such errors can identify potential preventive strategies.

Arch Intern Med. 2009;169(20):1881-1887

Author Affiliations:
Departments of Medicine (Drs Schiff and Kim and Mss Krosnjar and Wisniewski) and Emergency Medicine (Dr Cosby), Cook County Hospital, Chicago, Illinois; Division of General Medicine and Primary Care, Brigham and Women's Hospital, Boston, Massachusetts (Drs Schiff and Hasan); Department of Medicine, Rush University, Chicago (Drs Schiff, Abrams, Hasler, and McNutt and Mr Odwazny); Departments of Health Policy and Administration (Dr Kim) and Medical Education (Dr Elstein) and College of Pharmacy (Dr Lambert), University of Illinois at Chicago; and Department of Family and Preventive Medicine, University of California, San Diego (Dr Kabongo).

E RRORS RELATED TO DELAYED or missed diagnoses are a frequent and underappreciated cause of patient injury.[1-5] While the exact prevalence of diagnostic error remains unknown, data from autopsy series spanning several decades conservatively and

consistently reveal error rates of 10% to 15%.[6-9] Such errors are also the leading cause of medical malpractice litigation, accounting for twice as many alleged and settled cases as medication errors.[10,11] Despite this, few studies have examined diagnostic errors in detail, in part because of the challenges in reliably identifying and analyzing them.[3,12]

Given the inherently unpredictable and varying disease presentations and the limitations of diagnostic testing, clinicians are justifiably reluctant, and often defensive, about judging their own or colleagues' potentially missed or delayed diagnoses. Under these circumstances, one approach for identifying and understanding this type of error is to anonymously survey clinicians about cases in which they personally committed or observed what they considered to be diagnostic errors. As part of a multiyear

Agency for Healthcare Research and Quality–funded patient safety grant, the Diagnostic Error Evaluation and Research (DEER) project, we collected such reports by surveying physicians from multiple hospitals across several states.

The objectives of this study were 3-fold: to identify commonly missed diagnoses, to delineate recognizable patterns and

Figure 1. Diagnostic Error Evaluation and Research clinician survey questionnaire. LOS indicates length of stay; pt, patient.

themes from reported cases, and to apply the DEER taxonomy tool for analyzing cases. The development and pilot testing of this tool have been described previously, and the tool has been used in other research studies on diagnostic error.[3,13]

METHODS

DEFINITION OF DIAGNOSTIC ERROR

We defined diagnostic error as any mistake or failure in the diagnostic process leading to a misdiagnosis, a missed diagnosis, or a delayed diagnosis. This definition could include any failure in timely access to care; elicitation or interpretation of symptoms, signs, or laboratory results; formulation and weighing of differential diagnosis; and timely follow-up and specialty referral or evaluation.[3]

SURVEY DESCRIPTION

After pilot testing with 10 physicians with iterative revisions, we anonymously administered a 6-item survey (**Figure 1**) to solicit reports and descriptions of perceived diagnostic errors from physicians. Participating physicians were asked to review the definition in the preceding paragraph and then to describe 3 clinically significant diagnostic errors that they have seen or committed. Respondents were then asked to describe the (correct) diagnosis that should have been made, the error or failure that occurred, and the factors that contributed to the occurrence of the error.

For each error identified in this way, we asked respondents to rate its clinical impact or outcome: (1) no impact (no impact at all); (2) minor (patient inconvenience or dissatisfaction); (3) moderate (short-term morbidity, increased length of stay, need for higher level of care, or invasive procedure); or (4) major (death, permanent disability, or near life-threatening event). We asked respondents to assess how often they had seen this type of error: (1) rarely (1 or 2 cases seen); (2) infrequently (1 case seen every few years); (3) occasionally (a few cases seen each year); or (4) commonly (several cases seen each month).

SURVEY ADMINISTRATION

We surveyed a convenience sample of physicians, including general internists, medical specialists, and emergency physicians in 2 ways: either by mail, through institutional (internal) mailings to internists and emergency physicians at 2 participating academic medical centers, or by distributing questionnaires to be filled out during medical grand rounds presentations on the topic of diagnosis errors, given at 20 community and smaller teaching hospitals across the United States. For mailed surveys, physicians were given unlimited time to respond, while for grand rounds presentations, 10 to 12 minutes were allotted for survey completion. We did not have access to patients' medical records; therefore, case descriptions relied solely on information provided without attempt to verify or clarify the reported information from other sources such as the reported cases' medical records.

ANALYSIS

All cases were entered into a Microsoft Access database by a research assistant to ensure the anonymity of respondents whose handwriting might be recognizable. Diagnoses were coded and grouped, and descriptive statistics were analyzed. Cases were analyzed by final diagnosis, contributing factors, who made the error (self, others, or both), clinical impact or outcome (none, minor, moderate, or major), frequency (rare, infrequent, occasional, or common), and demographics of the reporting physician (specialty and years in practice).

Study investigators reviewed case descriptions, and 2 physician investigators (G.D.S. and O.H.) discussed ambiguous cases to determine whether or not a diagnostic error had occurred. Cases that described only a medication error, did not contain sufficient information to analyze reliably, or were judged not to contain an identifiable diagnostic error were excluded from this report of our study. We applied the DEER taxonomy tool to classify included cases according to the location and type of error in the diagnostic process. Classification was a 2-stage process that first involved identifying broadly where (ie, at what step in the diagnostic process) a problem had occurred, followed by a determination of what that problem was (ie, what went wrong in that step).

We relied entirely on respondents' descriptions in the "What went wrong?" and "Why did it happen?" sections of the survey to assign DEER category and refrained from making judgments about reporting physicians' accuracy and judgment. In instances in which multiple breakdowns in the diagnostic process were responsible for the eventual diagnostic failure, the step judged to have made the largest contribution to failure was used to assign a primary DEER category, while others were coded as secondary or tertiary. A random 10% sample of 60 cases was independently reviewed by 2 study investigators (G.D.S. and O.H.) to assess the interrater reliability of the DEER classification.

Differences in proportions were evaluated with the χ^2 test using SAS version 9.1 (SAS Institute Inc, Cary, North Carolina). No identifiable patient information was submitted with the survey responses, and participating physicians gave their consent for study inclusion. This study was approved by the institutional review board of Cook County Hospital and Rush University, Chicago, Illinois.
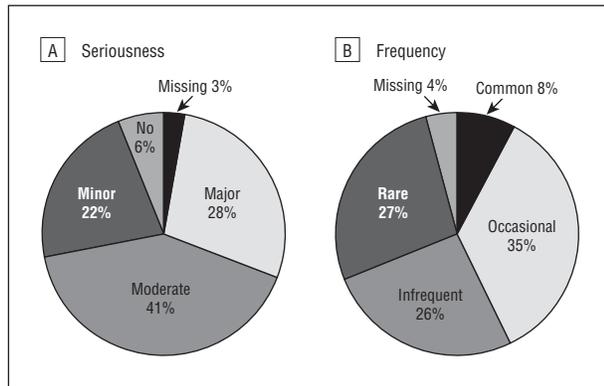
Figure 2. Perceived seriousness (A) and frequency (B) of the reported diagnosis error as rated by the physician reporting the error.

Table 1. Most Frequently Missed Diagnoses Among 583 Physician-Reported Cases of Diagnostic Error

Diagnosis	No. (%)
Pulmonary embolism	26 (4.5)
Drug reaction or overdose	26 (4.5)
Lung cancer	23 (3.9)
Colorectal cancer	19 (3.3)
Acute coronary syndrome	18 (3.1)
Breast cancer	18 (3.1)
Stroke, including hemorrhage	15 (2.6)
Congestive heart failure	13 (2.2)
Fracture, various types	13 (2.2)
Abscess, various locations	11 (1.9)
Pneumonia, including type	10 (1.7)
Aortic aneurysm/dissection	9 (1.5)
Appendicitis	9 (1.5)
Depression	9 (1.5)
Diabetes mellitus	8 (1.4)
Tuberculosis	8 (1.4)
Anemia	6 (1.0)
Bacteremia	6 (1.0)
Metastatic cancer	6 (1.0)
Spinal cord compression	6 (1.0)

RESULTS

A total of 669 cases were reported by 310 survey respondents, 243 (36%) via mail and 426 (64%) during grand rounds presentations. Response rates varied by survey method, ranging from a low of 10% to 20% for mailed surveys in certain medical subspecialty divisions to 70% to 90% for selected general medicine divisions and grand rounds participants. Each respondent was asked to report up to 3 cases: 1% reported 4 cases, 47% reported 3 cases, and 19% reported 2 cases, with the remaining 33% reporting 1 case each. After medication errors (14 cases) and reports lacking sufficient information or a clear description of a diagnostic error (72 cases) were excluded, 583 cases, reported by 283 respondents from 22 institutions in 6 states, were selected for detailed analysis.

Of the 283 respondents, 47% identified themselves as primary care physicians, 22% as specialty physicians, and 11% as other, while 20% did not provide relevant information. Respondents had been in practice an average of 9 years (median, 6 years), with 75% having practiced medicine for 15 years or less. Primary care physicians reported a wide range of common as well as rare diagnoses, whereas specialists typically reported cases limited to their specialty.

Of the 583 reported errors, 30% directly involved the reporting physician, 68% were witnessed being committed by others (without direct involvement of the reporting physician), and 2% were missing data on who made the error. Twenty-eight percent of the total reported errors were rated by the reporter to be major in severity, 41% moderate, and 22% minor, with the remainder being considered to have unknown or no patient impact (**Figure 2**A). The severity of the errors directly involving the reporting physician was graded as major in 29% of cases, moderate in 39%, and minor in 26%. In comparison, 28% of errors committed by others were graded as major, 44% as moderate, and 21% as minor (χ^2 P value for differences in proportions, .63).

Among all errors, only 8% were considered common, while 35% were considered occasional, 26% infrequent, and 27% rare; error frequency was missing for 4% of reported cases (Figure 2B). Of all the errors graded major in severity, only 5% were considered common, while 25% were considered occasional, 30% infrequent, and 40% rare. In comparison, 10% of all errors graded moderate in severity were
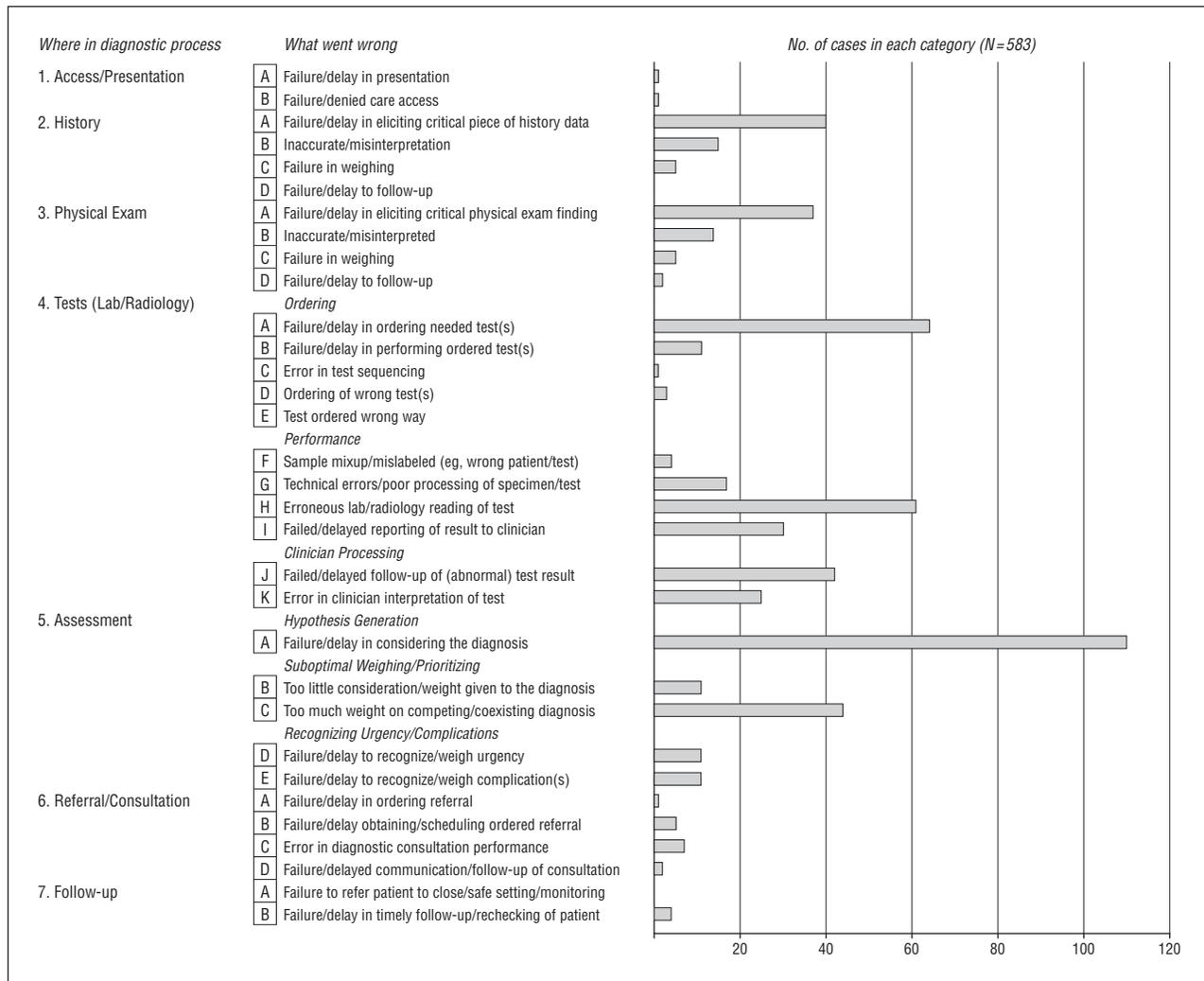
considered common, 40% occasional, and 28% infrequent. Similarly, 10% of all errors graded minor in severity were considered common, 42% occasional, and 24% infrequent (χ^2 P value for differences in proportions, <.01), suggesting that respondents perceived more serious errors to be significantly less common.

Reported diagnostic errors took place in diverse clinical settings, from inpatient to emergency and outpatient care, and encompassed a variety of acute as well as slowly evolving medical conditions (**Table 1**). Pulmonary embolism and drug reactions (including overdose and poisoning) were the 2 most commonly missed diagnoses (4.5% each), followed closely by lung cancer (3.9%) and colorectal cancer (3.3%). Cancers originating in the lung, colon, or breast accounted for 10.3% of diagnostic errors. Acute coronary syndrome (including acute myocardial infarction) and stroke (including intracranial hemorrhage) together accounted for 5.7% of cases. Surgical emergencies, including bone fracture, abscess, aortic aneurysm or dissection, acute appendicitis, and spinal cord compression, represented 8.2% of cases. All types of cancer together constituted the largest disease category, with 118 reported cases (20.2%).

Using the DEER taxonomy tool to determine where the diagnostic process broke down (**Figure 3**), we found that laboratory and radiology testing (including test ordering, performance, and clinician processing) accounted for the largest proportion of errors (44%), followed by clinician assessment (32%) (including hypothesis generation, weighing or prioritizing, and recognizing urgency or complications). In terms of identifying the specific process failure that occurred, failure or delay in considering the diagnosis (5A in Figure 3) accounted for the largest number of diagnostic failures (19%), followed by failure or delay in ordering needed tests (4A in Figure 3) and erroneous laboratory or radiology reading of tests (4H in Figure 3) in almost equal frequency (11%). Interrater reliability of the

Where in diagnostic process	What went wrong	No. of cases in each category (N=583)
1. Access/Presentation	A Failure/delay in presentation	
	B Failure/denied care access	
2. History	A Failure/delay in eliciting critical piece of history data	
	B Inaccurate/misinterpretation	
	C Failure in weighing	
	D Failure/delay to follow-up	
3. Physical Exam	A Failure/delay in eliciting critical physical exam finding	
	B Inaccurate/misinterpreted	
	C Failure in weighing	
	D Failure/delay to follow-up	
4. Tests (Lab/Radiology)	*Ordering*	
	A Failure/delay in ordering needed test(s)	
	B Failure/delay in performing ordered test(s)	
	C Error in test sequencing	
	D Ordering of wrong test(s)	
	E Test ordered wrong way	
	Performance	
	F Sample mixup/mislabeled (eg, wrong patient/test)	
	G Technical errors/poor processing of specimen/test	
	H Erroneous lab/radiology reading of test	
	I Failed/delayed reporting of result to clinician	
	Clinician Processing	
	J Failed/delayed follow-up of (abnormal) test result	
	K Error in clinician interpretation of test	
5. Assessment	*Hypothesis Generation*	
	A Failure/delay in considering the diagnosis	
	Suboptimal Weighing/Prioritizing	
	B Too little consideration/weight given to the diagnosis	
	C Too much weight on competing/coexisting diagnosis	
	Recognizing Urgency/Complications	
	D Failure/delay to recognize/weigh urgency	
	E Failure/delay to recognize/weigh complication(s)	
6. Referral/Consultation	A Failure/delay in ordering referral	
	B Failure/delay obtaining/scheduling ordered referral	
	C Error in diagnostic consultation performance	
	D Failure/delayed communication/follow-up of consultation	
7. Follow-up	A Failure to refer patient to close/safe setting/monitoring	
	B Failure/delay in timely follow-up/rechecking of patient	

Figure 3. Classification of diagnostic errors in 583 physician-reported cases using the Diagnostic Error Evaluation and Research project tool to localize where in the diagnostic process error occurred.

DEER classification for 60 randomly selected cases was moderate, yielding a κ statistic of 0.58 for agreement on where the diagnostic process broke down.

In the case of pulmonary embolism and drug reactions (including overdose and poisoning), failure or delay in considering the diagnosis (5A in Figure 3) accounted for 46% of errors (same percentage for each). In contrast, failure or delay in ordering needed tests (4A in Figure 3 [15%]) and erroneous laboratory or radiology reading of tests (4H in Figure 3 [14%]) were the 2 leading causes of diagnostic error among the aggregated cancer cases (n=118).

A subset of cases (n=179 [31% of total]) involved a secondary diagnostic process error that contributed to the eventual diagnostic failure. We found that the following types of errors clustered together (Figure 3):

4A and 5A (7 cases): Failure or delay in ordering needed tests and failure or delay in considering the diagnosis.

5A and 5C (6 cases): Failure or delay in considering the diagnosis and too much weight on a competing or coexisting diagnosis.

3A and 5C (5 cases): Failure or delay in eliciting critical physical examination findings and too much weight on a competing or coexisting diagnosis.

2A and 3A (4 cases) and 2A and 4A (4 cases): Failure or delay in eliciting a critical piece of history data, associated with failure or delay in eliciting critical physical examination findings or failure or delay in ordering needed tests.

In a subgroup analysis of major diagnostic errors (n=162), 43% were related to clinician assessment and 42% to laboratory and radiology testing. Of these 162 major errors, almost one-fourth (24%) were the consequence of a failure or delay in considering the diagnosis (5A in Figure 3), with failure or delay in ordering needed tests (4A in Figure 3) and placing too much weight on a competing or coexisting diagnosis (5C in Figure 3) in a tie for second place (12% each). Nine percent of these cases were attributable to failed or delayed follow-up of an abnormal test result. The most common diagnoses in this group were lung cancer (6.2%), pulmonary embolism (6.2%), poisoning or overdose (5.6%), stroke (4.9%), acute coronary syndrome (4.3%), aortic aneurysm or dis-

Table 2. Illustrative Case Descriptions and Diagnostic Error Evaluation and Research (DEER) Categorizations

Correct Diagnosis	What Went Wrong?	Why Did It Happen?	DEER Category
Opioid intoxication	Not considered in initial differential diagnosis as potential cause in patient presenting with history of medication overdose; not given appropriate treatment initially, delay	Failure to produce extensive list of differential diagnoses	5A
Renal failure	Two patients with same name had serum chemistry values interchanged	Improper labeling of samples by nursing staff	4F
Depression with suicide attempt–valproic acid	Only 1 level of valproic acid level drawn; always obtain 2 levels to appraise trend	Test not performed	4A
Pulmonary embolism	Patient had acute shortness of breath along with chronic history of pneumonia and lung abscess; no fever, no leukocytosis, 2- to 3 day history of immobility; tachycardia/tachypnea ignored	Symptoms and signs were interpreted as recurrence of pneumonia	5C
Acute appendicitis	Incomplete history and physical examination; radiologist did not find any findings on the CT consistent with appendicitis	Patient was unable to give clear history, 1 element of this misdiagnosis; also, the radiologist did not have skills to read CTs, his specialty was ultrasonography	2B primary and 4H secondary
Incarcerated hernia	Discrepancy between 2 radiology readings, 1 official and 1 by a "moonlighter"	Expert opinion of the 2 individual radiologists; however, clinical evidence and final agreement did not delay the surgical management of the case	4H
Myelodysplastic syndrome	Patient had anemia for a long time that was never worked up	Other medical conditions were more acute and needed to be taken care of, so symptoms were assumed to be due to iron deficiency	5D
Small cell lung cancer	Delayed biopsy; no report available	Computer system change	4I
Thyroid cancer	Patient had thyroid nodule; medical resident performed thyroid function tests only; results were normal; no follow-up.	Failed to order needed biopsy	4A
Colon cancer	Patient assumed to have hemorrhoidal bleeding	High patient volume; no other diagnosis sought	5A

Abbreviation: CT, computed tomogram.

section (4.3%), colorectal cancer (3.7%), and pneumonia (3.1%); the remaining diagnoses each represented less than 3%. Illustrative case descriptions, using the respondents own language, and our DEER categorizations are presented in **Table 2**.

COMMENT

Using in-person and local institutional solicitation, we collected and analyzed 583 cases of physician-identified diagnostic errors. While this convenience sample should not be considered perfectly representative of the universe and pattern of diagnostic errors arising in clinical practice, it does represent the largest reported case series of diagnostic errors to date and affords valuable insights into the types of errors that physicians are committing and witnessing. Physicians readily recalled and were willing to share cases of diagnostic error with (of 3 cases requested) an average of 2.2 errors reported by each respondent. This readiness suggests that diagnostic error is not unusual in clinical practice, and actively soliciting such cases represents an opportunity for tapping into a hidden cache of medical errors that are not generally collected by existing error surveillance and reporting systems.[14-16]

The most frequently missed or delayed diagnoses reported by our survey respondents mirrored those reported from previous analyses of large malpractice claims databases, with cancer being the leading category, followed by cases of pulmonary embolism, acute coronary syndrome, stroke, and infections.[10,17-19] Unlike the mal-

practice setting, our collection of anonymous cases and descriptions of "what went wrong" and "why did it happen" afforded those committing an error the opportunity to candidly share (or, in the words of several respondents, "to confess") errors in a blame-free context. As expected, specialist physicians generally reported cases in their specialty areas, a factor that likely biased rates of certain diagnoses in favor of sampled specialties but, at the same time, was useful in illustrating referred cases whose diagnoses were missed or delayed by the referring primary care physicians.

Our physician case reviewers found the DEER taxonomy tool intuitive and easy to use, with resulting moderate interrater reliability. This tool has now been applied to classifying diagnostic errors in other settings with cases in which access to the medical record details was available, unlike our cases.[13] When defined broadly to include failures in any stage of the (laboratory and radiology) testing process, including test ordering, specimen processing, test performance, interpretation, and follow-up, testing emerged as the step with the greatest number of reported process failures. There were slightly fewer failures in the clinician assessment stage, often referred to as cognitive errors, which included errors in considering a diagnosis (hypothesis generation) and giving appropriate weight to the various differential diagnoses. There were also sizable numbers of diagnostic errors that we designated as errors in recognizing the severity of a patient's illness.

Given the overlapping and inseparable nature of our 2 leading categories—testing and assessment—we found that

clearly differentiating them can be difficult and perhaps of questionable value in shedding insight into the error. If a critical test was not ordered because a diagnosis was not considered, eg, failure to consider anemia or hypokalemia, how can one classify or even separate patients whose diagnosis was ultimately or incidentally found on a hemogram or an electrolyte panel but was not apparently considered before the test? In each case, we attempted to identify a primary cause, but we believe that one of the important insights to emerge from our review involves the overlapping and clustering of certain patterns of errors, patterns that may be useful to consider when designing error reduction and prevention strategies.

In addition to the above-mentioned limitations related to selection bias, self-reporting, and ambiguities in classification, several additional limitations also need to be noted. In-person respondents were given only a brief opportunity to recall errors, creating the potential for a variety of recall biases.[20,21] The study lacked the ability to independently review medical records of reported cases and thus had to rely exclusively on respondents' accounts of events. While we potentially benefited from their first-hand knowledge, these self-reported error descriptions lack the objectivity and thoroughness that an independent root cause analysis would likely yield. Physicians' ratings of error seriousness were likewise subjective and likely commingled judgments about the seriousness of the outcome or diagnosis as well as the error itself. Finally, we observed considerable variability in the quality and details of case descriptions provided by survey respondents, varying from detailed descriptions of the error and its potential multifactorial causes to only a few words describing the circumstances that were considered responsible for the diagnostic error.

What can we learn from these cases beyond simply reinforcing awareness that such errors are occurring? While we are unable to quantify incidence rates, the respondents' cases paint a varied picture of heterogeneous types and causes of diagnosis errors. The DEER investigators identified a number of ways that such reporting could be of value. First, there are the potential benefits of the exercise itself, compelling physicians to recall and reflect on personal cases. Practicing such introspection is a key attribute of reflective physicians who systematically examine and learn from cases in which things go wrong.[22-24] Creating more systematic approaches, checklists, or automated decision support to aid in recalling, learning from, and sharing such cases has the potential to help others to avoid repeating similar errors.[3,25] Second, aggregating cases by diagnosis or diagnostic category allowed us to discern patterns of failures that were not otherwise apparent. Inadequate follow-up of abnormal imaging studies emerged as a leading process error, although other issues, such as erroneously attributing a worrisome symptom to another existing diagnosis, were also repeatedly reported.[26] Certainly, ensuring reliable follow-up of abnormal test results represents a "low hanging fruit" ripe for improvement of 1 error-prone process identified by our survey respondents and the medical literature.[27-29] Third, such reports permit us to look beyond individual diagnoses to identify cross-cutting generic factors that contribute to diagnostic errors. However, the question of whether diagnosis-specific vs more generic types of improvement efforts will be more productive remains an important unanswered question. Fourth, we found that our case reports highlight the complexities, relationships, and dialectics between different aspects of diagnostic decision making.[3,12] While diagnostic errors have traditionally been dichotomized into so-called cognitive vs system errors,[30-32] we believe that these cases demonstrate the ways in which these 2 realms overlap. For example, the missed appendicitis case (Table 2) had commingled system and cognitive failures.[3] Finally, collecting and reporting diagnostic errors gives continued visibility to a ubiquitous but often less overt type of medical error. With the decline of autopsies over the past half century,[33-35] we are more often literally and figuratively burying our mistakes.[8] Highlighting diagnostic error cases can help remind leaders of health care institutions of their responsibility to foster conditions that will better address and minimize the occurrence and consequences of errors that might otherwise have remained hidden, residing largely in the private memories of individual clinicians rather than becoming institutional knowledge for learning and improvement.

Correspondence: Gordon D. Schiff, MD, Division of General Medicine and Primary Care, Brigham and Women's Hospital, 1620 Tremont St, Third Floor, Boston, MA 02120 (gschiff@partners.org).
Author Contributions: *Study concept and design:* Schiff, Kim, Abrams, Cosby, Lambert, Elstein, Odwazny, Wisniewski, and McNutt. *Acquisition of data:* Schiff, Kim, Abrams, Cosby, Kabongo, Krosnjar, Wisniewski, and McNutt. *Analysis and interpretation of data:* Schiff, Hasan, Abrams, Cosby, Lambert, Elstein, Hasler, Krosnjar, Odwazny, and McNutt. *Drafting of the manuscript:* Schiff, Hasan, Kim, and McNutt. *Critical revision of the manuscript for important intellectual content:* Schiff, Hasan, Abrams, Cosby, Lambert, Elstein, Hasler, Kabongo, Krosnjar, Odwazny, Wisniewski, and McNutt. *Statistical analysis:* Hasan, Lambert, Odwazny, and McNutt. *Obtained funding:* Schiff, Kim, Cosby, and McNutt. *Administrative, technical, and material support:* Schiff, Hasan, Kim, Abrams, Cosby, Kabongo, Krosnjar, Odwazny, and Wisniewski. *Study supervision:* Schiff, Kim, and McNutt. *Additional conceptual foundations:* Lambert. *Cognitive psychology:* Elstein.

us cases of diagnostic errors they had committed or observed.

REFERENCES

1. 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(6):377-384.
2. Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care*. 2000;38(3):261-271.
3. Schiff GD, Kim S, Abrams R, et al. Diagnosing diagnostic errors: lessons from a multi-institutional collaborative project. In: *Advances in Patient Safety: From Research to Implementation*. Vol 2. Agency for Healthcare Research and Quality Web site. www.ahrq.gov/qual/advances. Accessed July 10, 2009.
4. Graber M. Diagnostic errors in medicine: a case of neglect. *Jt Comm J Qual Patient Saf*. 2005;31(2):106-113.
5. Newman-Toker DE, Pronovost PJ. Diagnostic errors—the next frontier for patient safety. *JAMA*. 2009;301(10):1060-1062.
6. Goldman L, Sayson R, Robbins S, Cohn LH, Bettmann M, Weisberg M. The value of the autopsy in three medical eras. *N Engl J Med*. 1983;308(17):1000-1005.
7. Sonderegger-Iseli K, Burger S, Muntwyler J, Salomon F. Diagnostic errors in three medical eras: a necropsy study. *Lancet*. 2000;355(9220):2027-2031.
8. Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy detected diagnostic errors over time. *JAMA*. 2003;289(21):2849-2856.
9. Kirch W, Schafii C. Misdiagnosis at a university hospital in 4 medical eras. *Medicine (Baltimore)*. 1996;75(1):29-40.
10. Phillips RL Jr, Bartholomew LA, Dovey SM, Fryer GE Jr, Miyoshi TJ, Green LA. Learning from malpractice claims about negligent, adverse events in primary care in the United States. *Qual Saf Health Care*. 2004;13(2):121-126.
11. CRICO/RMF forum. CRICO/RMF Web site. www.rmf.harvard.edu/high-risk areas/diagnosis/index.aspx. Accessed August 14, 2009.
12. McNutt R, Abrams R, Hasler S. Diagnosing diagnostic mistakes: case and commentary. emergency medicine, spotlight case. Agency for Healthcare Research and Quality Web site. http://www.webmm.ahrq.gov/case.aspx?caseID=95. Accessed August 14, 2009.
13. Singh H, Thomas EJ, Khan MM, Petersen LA. Identifying diagnostic errors in primary care using an electronic screening algorithm. *Arch Intern Med*. 2007; 167(3):302-308.
14. Milch CE, Salem DN, Pauker SG, Lundquist TG, Kumar S, Chen J. Voluntary electronic reporting of medical errors and adverse events: an analysis of 92,547 reports from 26 acute care hospitals. *J Gen Intern Med*. 2006;21(2):165-170.
15. Rowin EJ, Lucier D, Pauker SG, Kumar S, Chen J, Salem DN. Does error and adverse event reporting by physicians and nurses differ? *Jt Comm J Qual Patient Saf*. 2008;34(9):537-545.
16. Phillips RL, Dovey SM, Graham D, Elder NC, Hickner JM. Learning from different lenses: reports of medical errors in primary care by clinicians, staff and patients. *J Patient Saf*. 2006;2:140-146.
17. Gandhi TK, Kachalia A, Thomas EJ, et al. Missed and delayed diagnoses in the ambulatory setting: a study of closed malpractice claims. *Ann Intern Med*. 2006; 145(7):488-496.
18. Kachalia A, Gandhi TK, Puopolo AL, et al. Missed and delayed diagnoses in the emergency department: a study of closed malpractice claims from 4 liability insurers. *Ann Emerg Med*. 2007;49(2):196-205.
19. Singh H, Thomas EJ, Petersen LA, Studdert DM. Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. *Arch Intern Med*. 2007; 167(19):2030-2036.
20. Fischhoff B. Hindsight not equal to foresight: the effect of outcome knowledge on judgment under uncertainty. *Qual Saf Health Care*. 2003;12(4):304-312.
21. Neugebauer R, Ng S. Differential recall as a source of bias in epidemiologic research. *J Clin Epidemiol*. 1990;43(12):1337-1341.
22. Epstein RM. Mindful practice. *JAMA*. 1999;282(9):833-839.
23. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med*. 2003;78(8):775-780.
24. Mamede S, Schmidt HG, Rikers R. Diagnostic errors and reflective practice in medicine. *J Eval Clin Pract*. 2007;13(1):138-145.
25. Schiff GD. Minimizing diagnostic error: the importance of follow-up and feedback. *Am J Med*. 2008;121(5)(suppl):S38-S42.
26. Kostopoulou O, Delaney BC, Munro CW. Diagnostic difficulty and error in primary care—a systematic review. *Fam Pract*. 2008;25(6):400-413.
27. Schiff GD, ed. *Getting Results: Reliably Communicating and Acting on Critical Test Results*. Oakbrook Terrace, IL: JCHAO; 2006.
28. Wahls T, Haugen T, Cram P. The continuing problem of missed test results in an integrated health system with an advanced electronic medical record. *Jt Comm J Qual Patient Saf*. 2007;33(8):485-492.
29. Nepple KG, Joudi FN, Hillis SL, Wahls TL. Prevalence of delayed clinician response to elevated prostate-specific antigen values. *Mayo Clin Proc*. 2008; 83(4):439-445.
30. Kassirer JP, Kopelman RI. Cognitive errors in diagnosis: instantiation, classification, and consequences. *Am J Med*. 1989;86(4):433-441.
31. Graber M, Gordon R, Franklin N. Reducing diagnostic errors in medicine: what's the goal? *Acad Med*. 2002;77(10):981-992.
32. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med*. 2005;165(13):1493-1499.
33. Roberts WC. The autopsy: its decline and a suggestion for its revival. *N Engl J Med*. 1978;299(7):332-338.
34. Hill RB. The current status of autopsies in medical care in the USA. *Qual Assur Health Care*. 1993;5(4):309-313.
35. Xiao J, Krueger GR, Buja LM, Covinsky M. The impact of declining clinical autopsy: need for revised healthcare policy. *Am J Med Sci*. 2009;337(1):41-46.