Missed and Delayed Diagnoses in the Emergency Department: A Study of Closed Malpractice Claims From 4 Liability Insurers

Allen Kachalia, MD, JD
Tejal K. Gandhi, MD, MPH
Ann Louise Puopolo, BSN, RN
Catherine Yoon, MS
Eric J. Thomas, MD, MPH
Richard Griffey, MD, MPH
Troyen A. Brennan, MD, JD
David M. Studdert, LLB, ScD

Study objectives: Diagnostic errors in the emergency department (ED) are an important patient safety concern, but little is known about their cause. We identify types and causes of missed or delayed diagnoses in the ED.

Methods: This is a review of 122 closed malpractice claims from 4 liability insurers in which patients had alleged a missed or delayed diagnosis in the ED. Trained physician reviewers examined the litigation files and the associated medical records to determine whether an adverse outcome because of a missed diagnosis had occurred, what breakdowns were involved in the missed diagnosis, and what factors contributed to it. Main outcome measures were missed diagnoses, process breakdowns, and contributing factors.

Results: A total of 79 claims (65%) involved missed ED diagnoses that harmed patients. Forty-eight percent of these missed diagnoses were associated with serious harm, and 39% resulted in death. The leading breakdowns in the diagnostic process were failure to order an appropriate diagnostic test (58% of errors), failure to perform an adequate medical history or physical examination (42%), incorrect interpretation of a diagnostic test (37%), and failure to order an appropriate consultation (33%). The leading contributing factors to the missed diagnoses were cognitive factors (96%), patient-related factors (34%), lack of appropriate supervision (30%), inadequate handoffs (24%), and excessive workload (23%). The median numbers of process breakdowns and contributing factors per missed diagnosis were 2 and 3, respectively.

Conclusion: Missed diagnoses in the ED have a complex cause. They are typically the result of multiple breakdowns in the diagnostic process and several contributing factors. [Ann Emerg Med. 2007;49:196-205.]

INTRODUCTION

Medical error continues to capture the attention of the medical profession, policymakers, and the public. Inpatient care has been the major focus of attention, but there is increasing recognition of the risks of iatrogenic harm in the outpatient setting, including the emergency department (ED). Diagnostic errors are of particular concern and throughout the last decade have become the most prevalent type of malpractice claim in the United States.

The ED is an especially challenging environment in which to consistently make accurate and timely diagnoses. Patients often present with high-acuity illness, elevating the stakes from the outset. Triage, consultations, admissions, discharge, and other steps in emergency care are operationally complex and must usually be executed under tight time constraints. Emergency physicians seldom have a continuous relationship with the

Editor’s Capsule Summary

What is already known on this topic
Mistaken or delayed diagnoses in the emergency department (ED) can produce adverse outcomes. The mechanisms by which they occur are poorly understood.

What question this study addressed
Seventy-nine closed malpractice claims involving mistaken or delayed ED diagnoses were intensively scrutinized to develop descriptions of the settings in which such failures occurred and to explore possible contributing factors.

What this study adds to our knowledge
Most diagnostic failures resulted from multiple contributing factors; they were not “single-point” failures. Diagnostic failures were spread over a variety of clinical conditions; in this collection of cases, there was no single clinical condition or group of conditions that could be targeted for amelioration.

How this might change clinical practice
At present, understanding of the origins of diagnostic failures in ways that might be practically addressed in practice is minimal. The traditional approach has been punishment and training. The multifactorial nature of diagnostic failures suggests that novel approaches are needed to improve performance.

patients they treat, and the continuous nature of an ED necessitates a perpetual cycle of shift changes and handoffs. Supervision needs are high because trainees with widely varying clinical backgrounds and skills participate in care delivery. These intrinsic pressures of emergency care are amplified by crowding and increasing utilization by uninsured patients.

Previous studies of missed diagnoses in the ED have focused on specific diagnoses or used epidemiologic methods to identify clinical risk factors. However, little is known about the system-of-care factors that lead to diagnostic errors. Medical malpractice claims files present a potentially valuable source of information. They often involve severe injuries; they represent a powerful catchment point for information on errors; and by drawing together documentation from both formal legal inquiries and confidential internal investigations, they present a substantially richer body of information about the antecedents of medical injury than the medical record alone. Several clinical areas, most notably anesthesiology, have made impressive use of malpractice claims file analysis.

We analyzed a sample of medical malpractice claims involving allegations of misdiagnosis in the ED. The study goal was to determine specifically where breakdowns in the diagnostic process occurred and what contributing factors (systems, cognitive, and patient-related) played a role in their occurrence. Such descriptive information may help to identify priority areas for interventions to enhance safety in the ED.

MATERIALS AND METHODS

Four malpractice insurance companies based in 3 regions in the United States (northeast, southwest, and west) participated in the study. Collectively, the insurers covered approximately 21,000 physicians, 46 acute care hospitals (20 academic and 26 nonacademic), and 390 outpatient facilities. Institutional review boards at the investigators’ institutions and at each review site approved the study.

Data were extracted from random samples of closed malpractice claims files at each insurer. The claims file is the repository of information accumulated by the insurer during the life of a claim. It captures a wide variety of data and usually includes the statement of the claim, interrogatories, depositions, and other litigation-related documents; reports of internal investigations; expert opinions from both sides; medical reports and records detailing the plaintiff’s pre- and postevent condition; and, while the claim is open, medical records pertaining to the episode of care at issue. We reacquired the relevant medical records from insured institutions for sampled claims.

Following previous studies, we defined a claim as a written demand for compensation for medical injury. Claims involving missed or delayed diagnosis were defined as those alleging an error in diagnosis or testing that resulted in a delay in appropriate treatment or a failure to act or follow up on diagnostic test results. We excluded allegations related to pregnancy or to care rendered solely in the inpatient setting.

We focused on the outpatient setting, including the ED, because of the prevalence of this type of claim and the perceived importance of outpatient diagnostic errors in patient safety research and medical malpractice policy.

We established a target number of claims for each site, which represented the insurer’s estimate of the total number of claims closed during the previous 5 years that would be eligible for inclusion in the sample. Working with staff at the insurers, we used administrative databases to generate lists of candidate claims and reviewed narrative summaries to confirm that they met the applicable category definition. We began with the most recently closed claims and moved back in time until the target number was reached. We anticipated that the claims of missed diagnosis would implicate care during long periods and multiple sites of care, with a majority of the sites being the primary care physician’s office. The principal location of alleged errors was not known before commencement of the claim file reviews.

We reviewed a total of 429 claims of missed or delayed diagnoses, 122 (28%) of which involved allegations of substandard diagnostic care in the ED. This analysis focuses on those ED claims.

Sampled claims files were reviewed at the insurers’ offices or insured facilities by physicians. We used board-certified attending physicians, fellows, or third-year residents in internal
medicine to conduct the reviews. Physician investigators (A.K., T.K.G., E.J.T.) trained the reviewers in the content of claims files, use of the study instruments, and confidentiality procedures in 1-day sessions at each site. The reviewers were also assisted by a detailed manual. Reviews took 1.4 hours per file on average. Resource constraints and the labor-intensiveness of reviews dictated use of 1 reviewer per file. However, to test the reliability of the review process, 42 of the 429 diagnostic claims were rereviewed by a second reviewer from the same group who was unaware of the first review.

A sequence of 4 instruments guided the review. Figure 1 overviews the review process and key determinations. For all claims, administrative details of the case (Administrative Screening Data Form) were recorded. Next, reviewers verified that the patient had sustained an adverse outcome (Outcome Assessment Form) and graded its severity on a 9-point severity scale, ranging from emotional injury only to death. To simplify presentation of our results, we grouped scores on this scale into 5 categories (emotional, minor, significant, major, and death).

To make the error determination, reviewers were led through questions that considered the potential role of a range of contributing factors (Human Factors Form) in causing the patient’s adverse outcome; these factors were selected according to a review of the patient safety literature and covered system-, clinician-, and patient-related factors. Reviewers then judged, in light of all available information and their decisions about contributing factors, whether the adverse outcome was due to diagnostic error. We used the Institute of Medicine’s definition of error, namely, “the failure of a planned action to be completed as intended (ie, error of execution) or the use of a wrong plan to achieve an aim (ie, error of planning).” The error determination was based on prevailing practice approaches and technology available when the alleged error occurred.

Reviewers recorded their judgment on a 6-point confidence scale, ranging from 1, “little or no evidence that adverse outcome resulted from error/errors” to 6, “virtually certain evidence that adverse outcome resulted from error/errors.” Claims that scored 4 (“more likely than not that adverse outcome resulted from error/errors; more than 50-50 but a close call”) or higher were classified as having an error.

Finally, for claims with errors, reviewers completed a form (Missed and Delayed Error Form) that collected additional clinical information about the missed diagnosis, including the setting and the importance of involved clinicians’ contributions (graded on 5-point Likert scale, ranging from “somewhat important” to “highly important”). This form also presented reviewers with a defined sequence of diagnostic steps (eg, medical history/physical examination, test ordering, creation of follow-up plan) and asked them to grade their confidence that a breakdown had occurred at each step (5-point Likert scale, ranging from “highly unlikely” to “highly likely”). If a breakdown was judged to have been at least “somewhat likely” (score of 3 or higher on the scale), reviewers were directed to consider a non–mutually exclusive list of reasons for the breakdown at this point in the diagnostic process. Reviewers were not blinded to the litigation outcomes, because it was logistically impossible to censor this information in the files. However, they were instructed to ignore this outcome and exercise independent clinical judgment in rendering injury and error determinations. Training sessions stressed that the study definition of error is not synonymous with the legal definition of negligence and that a mix of factors besides merit influences whether claims are paid during litigation.

A secondary review was conducted by a board-certified emergency physician (R.G.), who considered free-text summaries of all claims in which the original reviewer had...
determined that a missed diagnosis led to an adverse outcome. This special review was conducted because we became concerned that the original reviewers’ lack of specialty training in emergency medicine may have led to incorrect determinations as to what qualified as missed diagnoses. This retrospective check on the work of the original reviewers focused on the possibility of false positives; for cases judged not to involve adverse outcomes or error, there was insufficient information about the clinical circumstances to support a secondary review.

**Primary Data Analysis**

The hand-completed data forms were electronically entered and verified by a professional data entry vendor and sent to the Harvard School of Public Health for analysis. Additional checks and data cleaning were performed by study programmers to ensure the data set’s quality. Analyses were conducted with the SAS 8.2 (SAS Institute, Inc., Cary, NC) and Stata/SE 8.0 (Stata Corporation, College Station, TX) statistical software packages.

The primary unit of analysis is the entire sequence of care in a claim judged to involve diagnostic error that led to an adverse outcome. For ease of exposition, we hereafter refer to such sequences as “missed diagnoses.” We examined characteristics of the claims, patients, and injuries in our sample and the frequency of the various process breakdowns and contributing factors. \( \kappa \) Scores were used to measure interrater reliability of the injury and error determinations.\(^{31}\)

**RESULTS**

One hundred twenty-two of the claims alleged diagnostic error in the ED. The claims alleged injuries sustained between 1979 and 2001. All the claims were closed between 1984 and 2003. In 80% of the claims, the alleged diagnostic error occurred in 1990 or later, and in 46%, it occurred in 1994 or later.

In 3% (4/122) of the claims, no adverse outcome or change in the patient’s clinical course was evident. Thirty-two percent (39/122) of the claims contained an adverse outcome but no error. The remaining 65% (79/122) of the claims involved a diagnostic error that was linked to an identifiable adverse outcome. This group of 79 missed diagnoses is the focus of our analyses.

In 40 of the 42 claims that underwent 2 independent reviews, the original reviewers agreed about whether an adverse outcome had occurred (95% agreement). There was 72% agreement among the reviewers about whether an error had occurred (\( \kappa = 0.42, 95\% \) confidence interval \(-0.05 \text{ to } 0.66\)).

In the secondary review, the emergency physician reviewer agreed with 55 of the 79 original error determinations (70%). For the rest, he was not comfortable verifying the original determination without further information beyond the short narrative available to him. In no case did he determine that a mistake had been made with respect to the original error determination.

### Table 1. Key characteristics of 79 diagnostic errors in the ED.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>41</td>
<td>N/A</td>
</tr>
<tr>
<td>SD</td>
<td>20</td>
<td>N/A</td>
</tr>
<tr>
<td>Range</td>
<td>1–81</td>
<td>N/A</td>
</tr>
<tr>
<td>Female</td>
<td>35</td>
<td>44</td>
</tr>
<tr>
<td><strong>Health insurance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>28</td>
<td>58</td>
</tr>
<tr>
<td>Uninsured</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>Medicaid</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Medicare</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td><strong>Clinicians involved</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency physician</td>
<td>41</td>
<td>52</td>
</tr>
<tr>
<td>Primary care physician</td>
<td>22</td>
<td>28</td>
</tr>
<tr>
<td>Surgeon</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>Radiologist</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Nurse</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Trainees(^{\dagger})</td>
<td>44</td>
<td>56</td>
</tr>
<tr>
<td><strong>Adverse outcome</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatric/emotional only</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Minor physical</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Significant physical</td>
<td>26</td>
<td>33</td>
</tr>
<tr>
<td>Major physical</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Death</td>
<td>31</td>
<td>39</td>
</tr>
<tr>
<td><strong>Missed or delayed diagnosis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fracture</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>Infection</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Cancer</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Cerebral vascular disease</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Embolism</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Appendicitis</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Other abdominal disease</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Aneurysm</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Other cardiac disease</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Other(^{\dagger})</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td><strong>Characteristics of Study Subjects</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Patient’s health insurance was missing in 30 claims (38%). Percentages calculated using nonmissing observations as the denominator.

\( ^{\dagger}\)Percentages do not sum 100% because multiple providers were involved in some errors.

\( ^{\dagger}\)Resident, fellow, or intern.

\( ^{\dagger}\)The levels of injury represent groupings of the scores on the National Association of Insurance Commissioners’ 9-point severity scale: psychiatric/emotional only (1), minor (2 and 3), significant (4-6), major (7 and 8), and death (9).

\( ^{\dagger}\)Other than cancer.

\( N/A \), not applicable.

**Characteristics of Study Subjects**

The mean age of patients was 41 years (median 39 years), with 13% of missed diagnoses occurring in the care of children (Table 1). Of these missed diagnoses, 39% led to death and 48% involved adverse outcomes rated as significant (33%) or major (15%).

The types of missed diagnoses were diffuse. The leading types were fractures (19%), infections (15%), myocardial infarction (10%), and cancer (9%), which together accounted for more than half. Overall, the missed diagnoses tended to
Example 1. A 72-year-old woman with history of coronary angioplasty presents by ambulance for chest pain, nausea and vomiting, diarrhea, shortness of breath, and bilateral arm tingling. An EKG revealed new ST depressions that were not diagnosed; the ED physician did not compare with an old EKG. Cardiac enzyme tests were not ordered. The patient was discharged with a diagnosis of gastroenteritis. A cardiologist reviewed the EKG later and noted the abnormalities but did not immediately notify the ED. After the ED was notified, the patient was asked to return. The patient subsequently died from the myocardial infarction.

Example 2. A 44-year-old man with obesity and history of peptic ulcer disease presented to the ED with anorexia, epigastric pain, and blood-streaked emesis. He was tachycardic and mildly hypotensive. Blood was drawn for laboratory tests, and results were within normal limits. The patient was given ketorolac and then discharged. The patient sustained a cardiac arrest 2 days later. He was found to have had a perforated duodenal ulcer.

Example 3. A 9-year-old girl presented to the ED 4 times in 1 week with fever, sore throat, and back and abdominal pain. During the first visit, a throat culture was taken and later was positive for streptococcus, but it was not specifically reported to anyone. During the next 2 visits, the culture results were not reviewed, and the patient was diagnosed with a viral syndrome. She presented for her fourth ED visit with paralysis and loss of bladder control. After some delay in evaluation during that visit, she was found to have an epidural abscess.

Example 4. A 30-year-old man presented to ED with jaw pain after trauma. He was evaluated by a first-year surgical house officer, who ordered a CT scan for head trauma. No facial radiographs or CTs were ordered. The patient was diagnosed with a jaw fracture 3 weeks later.

Example 5. A 39-year-old man presented to ED with jaw pain 3 weeks after trauma. He was evaluated by a first-year surgical house officer, who ordered a CT scan for head trauma. No facial radiographs or CTs were ordered. The patient was diagnosed with a jaw fracture 3 weeks later.

Example 6. A 55-year-old man with hypertension presented to the ED with 6-7 days of back and flank pain. He was sent for an intravenous pyelogram (IVP) to evaluate for renal stones. Because of an exceedingly busy ED, his systolic blood pressure of 70 was not diagnosed until after the IVP. He was then taken to the operating room for emergency aortic abdominal aneurysm rupture more than 12 hours after he presented.

Figure 2. Clinical examples of identified diagnostic errors.

Table 2. Diagnostic steps and frequency of breakdowns at each step.

<table>
<thead>
<tr>
<th>Step</th>
<th>No.</th>
<th>%*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient notes problem and seeks care</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Provider performs medical history and physical examination</td>
<td>33</td>
<td>42</td>
</tr>
<tr>
<td>Provider orders appropriate tests</td>
<td>46</td>
<td>58</td>
</tr>
<tr>
<td>Ordered tests performed in a timely manner</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Ordered tests performed correctly</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Test results transmitted to and received by the provider</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Test results transmitted to and received by the patient</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Interpretation of test results</td>
<td>29</td>
<td>37</td>
</tr>
<tr>
<td>Provider orders consultation (or referral)†</td>
<td>26</td>
<td>33</td>
</tr>
<tr>
<td>Requested consultation (or referral) occurs†</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Creation of proper follow-up plan</td>
<td>21</td>
<td>27</td>
</tr>
<tr>
<td>Patient adherence with plan</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

*Calculated as a percentage of 79 claims with identified errors.
†A failure to order a consultation in the ED includes the failure to order an immediate consultation in the ED, the failure to order an appropriate outpatient subspecialty referral, and the failure of trainees to consult with more senior physicians.

Missed and Delayed Diagnoses in the Emergency Department

Table 2. Diagnostic steps and frequency of breakdowns at each step.

<table>
<thead>
<tr>
<th>Step</th>
<th>No.</th>
<th>%*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient notes problem and seeks care</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Provider performs medical history and physical examination</td>
<td>33</td>
<td>42</td>
</tr>
<tr>
<td>Provider orders appropriate tests</td>
<td>46</td>
<td>58</td>
</tr>
<tr>
<td>Ordered tests performed in a timely manner</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Ordered tests performed correctly</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Test results transmitted to and received by the provider</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Test results transmitted to and received by the patient</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Interpretation of test results</td>
<td>29</td>
<td>37</td>
</tr>
<tr>
<td>Provider orders consultation (or referral)†</td>
<td>26</td>
<td>33</td>
</tr>
<tr>
<td>Requested consultation (or referral) occurs†</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Creation of proper follow-up plan</td>
<td>21</td>
<td>27</td>
</tr>
<tr>
<td>Patient adherence with plan</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

*Calculated as a percentage of 79 claims with identified errors.
†A failure to order a consultation in the ED includes the failure to order an immediate consultation in the ED, the failure to order an appropriate outpatient subspecialty referral, and the failure of trainees to consult with more senior physicians.
diagnoses involved at least 1 of these “cognitive” factors. The other leading contributing factors were patient-related factors (34%), lack of appropriate supervision (30%), inadequate handoff (24%), and excessive workload (23%).

Missed diagnoses frequently involved multiple breakdowns, contributing factors, and contributing clinicians (Figure 3). The median number of process breakdowns per missed diagnosis was 2; 81% had at least 2 process breakdowns, 38% had at least 3, and 13% had at least 4, which left only 15 missed diagnoses that involved a breakdown at only 1 point in the care process (“single-point breakdowns”). A majority (8/15) of these single-point breakdowns were failures to order tests.

The median number of contributing factors per missed diagnosis was 3; 85% had at least 2 contributing factors, 61% had at least 3, and 39% had at least 4. Hence, although cognitive factors contributed to virtually all missed diagnoses,
they operated alone in only one third of cases; in the remaining two thirds, they were present alongside other contributing factors. With respect to contributing clinicians, two thirds (52/79) of the missed diagnoses involved more than 1 clinician, 27% of the errors had at least 3, and 10% had at least 4.

The factors that contributed to missed diagnoses involving trainees differed from those that led to missed diagnoses without trainees in several ways. The trainee events were more likely to involve inadequate supervision (55% versus 0%), patient-related factors (48% versus 17%), and excessive workload (36% versus 6%).

LIMITATIONS

The use of malpractice claims for addressing patient safety has limitations. First, severe injuries are probably overrepresented because they are more likely to trigger litigation. Second, certain breakdowns or contributing factors may not have been discernible in claims file review, even though they played a role; to the extent that this occurred, the prevalence findings for such estimates will be lower bounds, and the multifactorial causality we observed probably understates the true complexity of missed diagnoses. Alternatively, reviewers may have overstated the role of some breakdowns or contributing factors, particularly cognitive factors in circumstances in which the reasons for the adverse outcome were difficult to explain.

Third, certain factors or breakdowns that lead to litigated missed diagnoses cases may differ systematically from the factors or breakdowns that lead to nonlitigated ones, although we know of no reason why they would. To the extent that the profile of missed diagnoses in malpractice claims diverges from that of missed diagnoses more broadly, some of the targets for intervention our findings highlight may have a disproportionately large impact on the type of events that prompt claims.

Our study had several other limitations. EDs in teaching hospitals are overrepresented in our sample, so the missed diagnoses we identified may not be generalizable to every ED. Reviewers were general internists, not emergency physicians. The lack of this specialist perspective may have led to some misidentification of missed diagnoses, process breakdowns, and contributing factors (both false positives and false negatives).

Finally, reviewers’ judgments about the appropriateness of care is likely to have been influenced by hindsight bias. One possible version of this bias is knowledge of the litigation outcome, which may have encouraged findings of errors in paid claims and vice versa. A second version relates to the presence of adverse outcomes, especially severe ones, which may have prompted inferences that care was inappropriate. A third, more general version of hindsight bias may have stemmed from reviewers’ recognition that the data source was a malpractice claim file.

DISCUSSION

By reviewing malpractice claims files related to care in the ED, this study identified 79 missed diagnoses that varied widely in type, often involved acute illnesses, and frequently resulted in severe injury. The cause of these events was complex, with the majority involving multiple breakdowns in the diagnostic process, several contributing factors, and more than 1 provider. The most common breakdown points were test ordering and interpretation, performance of the medical history and physical examination, and initiation of consultations. Cognitive factors contributed to almost every missed diagnosis, but they usually acted in concert with other types of factors, particularly supervision, handoffs, workload levels, and patient-related factors.

Most research into diagnosis errors in the ED has been aimed at specific diagnoses, particularly myocardial infarctions, fractures, and infections. Studies have identified problems in diagnostic steps such as history taking, interpretation of ECGs and radiographs, and clinical decisionmaking. Previous research has also identified gaps in clinician knowledge, mistakes in judgment, fatigue, and poor handoffs as factors in missed diagnoses. Our findings echo these results.

From a prevention perspective, targeting interventions at selected diagnoses with high potential for being overlooked may produce significant gains. Such an approach is well suited to the use of explicit clinical algorithms designed by appropriate experts. However, serious diagnostic errors—at least those evident in malpractice claims data—cover a broad range of diagnoses; the top 3 diagnoses accounted for less than half of the misses identified in our study sample. Such dispersion suggests that error prevention strategies aimed at specific diagnoses, even the most prevalent ones, will leave a large number of diagnostic problems unaddressed. Furthermore, designing and implementing interventions diagnosis by diagnosis may be logistically impractical; for example, diagnostic algorithms may not prospectively capture all atypical presentations.

As the patient safety movement has gathered momentum during the past decade, experts have increasingly pushed for attention to causes of error at the systems level, noting that this approach has the potential for cross-cutting gains. We sought to adopt this perspective in examining the heterogeneous group of missed and delayed diagnoses found among malpractice claims.

Breakdowns in the diagnostic steps that require active clinician decisionmaking—specifically, conducting patient medical histories and physical examinations, ordering and interpreting tests, ordering consultations, and creating follow-up plans—were common, occurring in all but 2 of the missed diagnoses (97%). Failure to order appropriate tests was the most frequent breakdown. Appropriate test ordering, like other steps that involve active decisionmaking, requires 2 key ingredients: the availability of the right information on which to base the decision and correct application of cognitive skills to this
information. We found the latter component to be particularly problematic. The same was generally true with the other active decisionmaking steps.

There have been many attempts to implement clinical decision support systems for diagnosis and treatment determinations. 47–50 Although evaluations have demonstrated benefits in the selection of treatments, 51–55 evidence of the value in the diagnostic area is more mixed. 48,50,56–59 Workflow impediments and efficiency concerns appear to have obstructed adoption or effectiveness. 48,56–58 This study’s findings emphasize the importance of continuing to press for successful implementation of decision support systems.

Process breakdowns at steps in which clinician decisionmaking played no, or a relatively minor, role—for example, the proper performance of ordered tests, transmission of test results, and follow-through on requested consultations—were much less common. Nonetheless, breakdowns in some such “passive” steps are still cause for concern. In 1 in 6 missed diagnoses, for example, test results did not reach the correct clinicians.

The problem of transmission of critical test results in clinical practice is now well recognized. 60,61 Breakdown in transmission is especially troubling in the ED setting, where tests are generally ordered for immediate review. EDs should therefore pay close attention to their communication policies and procedures for critical test results, including strategies such as direct communication of findings between radiology or laboratories and the ordering providers.

Problems in test result transmission also highlight the stresses on continuity of care—shift changes, multiple providers per patient, and discharge back to the primary care provider—that are, to some degree, essential features of emergency treatment. Not surprisingly, handoff breakdowns were present in almost a quarter of the missed diagnoses we identified. There is a growing awareness of the implications of discontinuities in care; our findings underscore the fact that the diagnostic process in the ED is also affected. 62 Strategies being promoted nationwide to improve these discontinuity problems through standardized handoff procedures and communications 63 should not ignore the ED.

Excessive workload in the ED was identified as a contributing factor in almost a quarter of the cases. This is consistent with reports that EDs today are frequently crowded. 11,12,63,64 The significant harm associated with excessive workload levels suggests that whereas the traditional image of hurried providers rapidly triaging and treating patients in the ED may provide for an exciting work environment, it also poses a threat to patient safety.

ED staff in academic medical centers face the dual challenge of providing high-quality care and providing excellent training to medical students and house officers from multiple specialties. Supervision of residents has drawn increasing attention today, especially in light of the “80-hour” workweek in teaching hospitals. 65 The prevalence of supervision problems in the errors we identified suggests that ongoing efforts in this area may pay dividends in patient safety. Close oversight of trainees’ diagnostic evaluations in the ED may bring particularly large returns at times of high workload or when complicating patient factors such as atypical presentation and complicated medical histories are involved. Of course, this strategy poses major implementation challenges because times of hectic workload are often exactly the moments when supervision is most difficult.

Awareness of the potential severity of patient injury caused by missed diagnoses in the ED should motivate efforts to avoid these breakdowns. Our findings highlight potential opportunities for achieving this. In terms of prioritization, the lower proportion of missed diagnoses that involved breakdowns in more passive cognitive points in the diagnostic process suggests that, even though interventions targeted at these steps will improve patient safety, better support for active cognitive processes has the potential to avert more harm. Addressing problems with particular contributing factors—such as handoffs, supervision, and workload—may also prove to be high-yield strategies.

Cost and feasibility remain key challenges for mounting interventions. Challenges associated with measures to guard against cognitive errors have proved to be especially formidable. 66,67 Cross-matching frequent process breakdowns with underlying contributing factors may help. For example, automatically double-checking clinician interpretations of test results might be useful in reducing breakdowns in this step. Enhanced staffing during periods of heavy workload might improve supervision (attending involvement) of trainee physicians. However, even targeted interventions may face serious implementation problems and may introduce new safety concerns. 68–70 The net effects of such interventions should thus be carefully evaluated.

In conclusion, missed diagnoses in the ED can have severe consequences and are a major patient safety concern. They are also complex, involving multiple process breakdowns and contributing factors. Prevalent breakdown points and contributing factors represent targets of opportunity in preventing missed diagnoses. Use of medical malpractice claims data to unravel their cause and identify such targets has limitations but may offer valuable insights.

The authors thank Allison Nagy and Ilina Chaudhari for their invaluable assistance with data collection.

Supervising editor: Robert L. Wears, MD, MS

Authors contributions: TAB and DMS conceived the study and obtained funding. AK, TKG, and DMS designed this analysis. TKG, ALP, EJT, and DMS designed the data collection instruments. AK, TKG, ALP, EJT, and DMS trained physician reviewers. TAB and DMS undertook recruitment of participating sites. ALP, CY, and DMS managed the data collection, including quality control. RG contributed to quality control efforts after the data had been collected. AK, TKG, CY, and DMS analyzed the data. AK, TKG, and DMS drafted the article, and all authors contributed substantially to revision of its intellectual content. AK, TKG, and DMS take responsibility for the paper as a whole.
Funding and support: This study was supported by grants from the Agency for Healthcare Research and Quality (HS011886-03) and the Harvard Risk Management Foundation. Dr. Studdert was also supported by grant KO2HS11285 from the Agency for Healthcare Research and Quality.

Publication dates: Received for publication November 13, 2005. Revisions received March 15, 2006, and June 7, 2006. Accepted for publication June 26, 2006. Available online December 1, 2006.

Reprints not available from the authors.

Address for correspondence: David M. Studdert, LLB, ScD, Department of Health Policy and Management, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115; 617-432-5209, fax 617-432-4494.

REFERENCES

26. Cheney FW. The American Society of Anesthesiologists Closed Claims Project: what have we learned, how has it affected practice, and how will it affect practice in the future? Anesthesiology. 1999;91:552-556.