Staying Out of Trouble

Philip Levitt

Published online: 16 February 2013
© Society for Imaging Informatics in Medicine 2013

Introduction

Nearly all present-day discussions of patient safety begin with the Institute of Medicine's 1999 report, *To Err is Human* [1]. It recommended "creating safety systems inside health care organizations through the implementation of safe practices at the delivery level. This level is the ultimate target of all the recommendations." Attempts at the realization of the proposals included giving aspirin to all heart attack patients on admission and increasing safeguards to prevent medication errors. "Time outs" are required in which a surgeon has to declare the type and site of an operation before he is passed a scalpel.

Medical quality assurance started long before systems error reduction came into prominence. The Flexner report in the early twentieth century focused on the competency of the physicians who were graduating from the nation's medical schools [2]. The report resulted in the closure of scores of inferior schools that were producing dangerously inept doctors and the enactment into law in all states of standards assuring a thorough grounding in the basic sciences and rigorous clinical exposure for medical students. Earnest Codman, a Boston academic surgeon of the same era, criticized his surgical colleagues' care of patients and proposed tabulating their surgical outcomes as well as his own and publicizing them. As a result, he found it necessary to leave the hospital where he practiced. Subsequently, he helped found the American College of Surgeons and its Hospital Standardization Program, the forerunner of the Joint Commission.

Throughout the twentieth century, accomplished senior doctors taught eager, competent residents. Physicians had to

go to hospital conferences and committee meetings that dealt with problem cases and to national specialty meetings. They read journals and attended to all the irritating details in their practices. They communicated with colleagues in a responsive and informative way. Privately, they struggled with and agonized over their bad outcomes, trying to prevent their recurrence.

Today, the goal of assuring patient safety has taken on a more abstract and global aspect. *To Err is Human* awakened the medical profession to the problem of harmful medical errors in a new way by treating them in the aggregate and projecting the number of patients affected in the entire country. Its highly publicized claim that 44,000–98,000 patients died each year as a consequence of medical care in the nation's 5,500 hospitals aroused the public like nothing else since the Flexner report. Its findings were immediately criticized. Sox and Woloshin took issue with the death figures because the calculations used in deriving them were neither published nor explained [3].

The primary source of *To Err is Human* was the Harvard Medical Practice Studies (HMPS), a series of papers that were based on the examination by specially trained physicians for harmful errors of over 30,000 randomly selected charts from hospitals in New York State [4, 5]. The study found that 3.7 % of the charts revealed disabling or deadly "adverse events" (AEs)—the term for bad outcomes of medical management.

The majority of patients who died of AEs in the HMPS were elderly and very sick. The authors of the HMPS anticipated this criticism with the example of a patient with diffusely metastasized carcinoma and a pleural effusion who died immediately after the faulty insertion of a chest tube. This episode was included among the deadly AEs. They also acknowledged that it was likely that some of the patients requested and obtained limited care not documented in the charts.

P. Levitt (✉)
Division of Neurosurgery, St Mary's Hospital, 901 45th Street,
West Palm Beach, Florida, USA
e-mail: doclevittmd@aol.com

Hayward and Hofer wrote in 2001 of the fatal adverse events at several VA hospitals using similar criteria and methods of chart review to the HMPS [6]. In their introduction they said of the 44,000–98,000 figures, “if these inferences are correct, the health care system is a public health menace of epidemic proportions.” They found that although the immediate cause of death was often an adverse event, many patients were so sick that they would have been cognitively impaired or died anyway within 3 months. When they took that into consideration, the number of lives of quality lost from adverse events shrank by a factor of 12. However, that did not make the care any better. Two subsequent meticulously performed studies of 2010, albeit with much smaller numbers of patients than in the HMPS, have borne out the estimates in *To Err is Human* [7, 8].

Radiologists’ errors were not included in the HMPS, but all the clinical specialists one would expect to be caring for patients in a general hospital were. These were analyzed by specialty with neurosurgery, vascular surgery, and cardiac surgery having the most frequent AEs. All other specialties were in a group at a lower level of mishaps. All specialties had about the same percentage of *negligent* errors.

The second and initially less known facet of *To Err is Human* was the recommendation that adverse events could be prevented by systems methods. This had been done effectively in the commercial airline industry. But the science behind this assumption needed to be questioned more rigorously. The second HMPS paper described 1,133 patients with disabling or deadly injuries caused by medical treatment or 3.7 %, in just over 30,000 admissions [5]. In comparable but not identical studies from England and Canada, the percentages were 10.8 and 7.5 %, respectively [9, 10].

In an analysis by the author of this paper, a different picture of the main source of adverse events emerged than that painted in *To Err is Human*. Table 7 of the second HMPS paper gives the incidence of 37 categories of errors among the 1,133 patients. Two categories, which gave rise to 61 % of the mistakes, *technical errors* (559) and *failure to use indicated tests* (134), are attributed by this author to the hands and thought processes of individual doctors, not flawed systems. This should have put into question the validity of approaching the problem of adverse events exclusively by correcting flawed systems; further, no subsequent data contradicting the above findings have been published. Most of the other 35 categories of error could be attributed to either individual or systems errors.

The authors of *To Err is Human* anticipated a reduction of 50 % in 5 years by following a systems approach. However, hope and reality collided in two reports, one by HHS and the other from Harvard published in 2010 [7, 8]. They showed that the number of deaths associated with adverse events in the decade after *To Err is Human* was not smaller

and may have increased. The Harvard study looked at hospitals in North Carolina exclusively because the hospitals of that state had an outstanding record of compliance with safety measures. The HHS study projected the number of deadly adverse events in Medicare patients at 180,000 per year, while the number of *preventable* deaths for all types of patients could be projected as 135,000 per year in the Harvard/North Carolina study. In summary, without vigorous intervention, well over a million deaths in the next decade will be due to preventable adverse events and most of them will be from the errors of individual doctors.

The Synergy

Ordering the correct diagnostic test, a problematic category in the HMPS, is the end result of several processes which are the burden of the clinician: history taking, a physical exam, mulling over the diagnostic possibilities, and knowing which test is most likely to confirm the diagnosis while doing the least harm. Radiologists are very well versed in the indications for imaging procedures. It is essential that they pass on their learning to their clinical colleagues.

Diagnostic radiology is the keystone of diagnosis in most specialties. What might have been found only at autopsy in previous decades is now routinely detected and explicated by digital imaging. This remarkable technical progress notwithstanding, a well considered completion of the imaging request form with the patient’s history and physical findings is essential and should include more than just the couple of words used to justify the imaging. The counterpart of an informative request, a clear, thorough, and promptly transmitted report requires that the radiologist compose a piece of writing with its reader clearly in mind, making every effort to convey a true diagnostic impact. Leslie, Jones, and Goddard did a controlled study on the effects of clinical information and its accuracy in interpreting CT scans of the head, chest, and abdomen [11]. An accurate history had a salutary effect on the interpretation, quite often having a major clinical impact. Inaccurate clinical information led to misinterpretations, particularly in chest and abdominal CT.

Direct discussion has benefits for the patient. It goes almost without saying that where the diagnosis remains obscure clinician and radiologist need to discuss the case and rethink the diagnosis, a quality improvement strategy enhanced by interpersonal rapport. When an urgent or unexpected diagnosis is made, direct communication between clinician and radiologist is mandatory. One study found that in urgent and/or complicated cases direct clinicoradiological communication resulted in a change of the clinical diagnosis in 50 % of cases and a change of treatment in 60 % [12].

Cognitive Processing by Doctors

The diagnostic process differs between two specialty blocs, those that are predominantly visually based, radiology and pathology, and clinical medicine which uses a variety of sensory modalities. The error rate among radiologists has been estimated to be lower than that among clinicians, 3–5 vs. 15 % [13]. That difference is attributed to the greater complexity of data accumulation and synthesis involved in clinical medicine. However, another estimate of the radiologic error rate is 30 % [14].

Hypothetico-deductive processing [15], using a differential diagnosis, is the way doctors are taught to evaluate a new patient. They make educated guesses after doing a history and physical and list several diseases in order of likelihood and order the tests that will best help to sort things out. It is a slow, analytical, conscious process. The clinician takes a corrective course of action if the initially favored diagnosis does not pan out. The method often cannot be applied easily to a rapidly deteriorating patient.

The pattern recognition technique is experience based and is best used by seasoned clinicians and radiologists. It is quick, operates on an unconscious level and usually produces the correct diagnosis. However, when the rapid impression is wrong, the patient may suffer an adverse event. Three closely allied and overlapping cognitive phenomena, common to radiologists and clinicians, are at play when pattern recognition fails:

1. Anchoring bias occurs when a doctor draws conclusions too rapidly by avoiding the tedious, time-consuming process of the differential diagnosis and other diagnostic aids such as check lists and computer-based diagnostic programs.
2. Confirmation bias is one in which the doctor ignores data that fails to confirm the initial diagnosis or points to another.
3. The availability heuristic is a term explaining how less common diagnoses are missed.

A retrospective study of primary physicians' serious diagnostic misses by Ely et al. concluded that the most frequent type of error was making a common benign diagnosis when the patient had an uncommon serious disease. The correct diagnosis had not initially occurred to the studied physicians 82 % of the time [16]. Four patients were originally diagnosed by their doctors with gall bladder disease. The final diagnoses in those cases were pancreatitis, ovarian cancer, pulmonary embolus, and herpes zoster.

The process by which the right answer abruptly and clearly comes to consciousness is not completely understood. A key recommendation of the doctors who missed a diagnosis was to broaden the differential diagnosis.

The pattern recognition and hypothetico-deductive methods are equally effective in arriving at the correct diagnosis. More importantly, when used together they are more effective than when used alone [17]. Diagnostic radiology is as pure a pattern recognition task as exists in medicine. Combining pattern recognition with a differential diagnosis is helpful in making the initial diagnosis and in making recommendations for additional studies. However, this creates a dilemma which needs to be resolved somehow in favor of the patient. How does one reach the correct diagnosis while dealing with the time constraint of reading a large cache of images and the financial limits imposed by reimbursement policies? For example, there are 38 possible interpretations of a pulmonary nodule under 4 cm in diameter on a chest X-ray [18]. Only experience and good fortune seem to help with that quandary.

Two additional diagnostic approaches [15] are:

1. Algorithms, preset pathways with criteria for diagnoses. They have been criticized for their inflexibility and for blocking independent thinking.
2. The worst-case scenario approach in which the doctor retains a list of life and limb threatening "cannot miss" disorders. This is epitomized by a sign that hangs in many obstetricians' offices: "IS THIS AN ECTOPIC?"; 3 % of the missed diagnoses in the above study of primary physicians' missed diagnoses were of the "cannot miss" type.

Types of Radiology Errors

Errors in observation include failing to scan the entire image and failing to detect an abnormality despite adequate scanning. Common interpretive errors include calling a structure normal when it is a lesion, being distracted by another finding in the field and missing the clinically relevant abnormality, the so-called satisfaction of search error, and consciously reading as normal equivocal findings in order to reduce the number of false positives [14]. In one study of radiologic errors, just under half were observational; a fourth was interpretive. Those based on inadequate knowledge occurred only 1 % of the time [19].

Observational and interpretive errors have been shown to become more frequent as a result of fatigue, interruptions, excessive workload, and even the time of day. Late morning and early afternoon are somehow linked to an increased chance of image reading mistakes [20].

The most common types of misses are fractures on skeletal X-rays and missing cancerous lesions on a variety of studies. The most often missed fracture sites are in the femur, the navicular, and the cervical spine [14]. The cancerous lesions most likely to be missed are carcinomas on barium enemas, breast cancer on mammograms, lung nodules on plain chest X-ray, and bone tumors on skeletal images.

Detecting and Tabulating Imaging Errors

Reducing errors requires detecting, amassing, and analyzing them. This is followed by planning and testing strategies for making them as infrequent as possible. It is essential to have standards for detecting imaging errors. The gold standard for radiology is the pathologic diagnosis. This is most readily available in diseases such as appendicitis and breast cancer. In other disorders where excision or biopsy are not carried out such as pneumonia or congestive heart failure, the detection of inter-reader discrepancies, a more subjective method, has become the standard investigational tool.

The RADPEER program of the ACR is geared towards quality assurance/peer review that consumes as little time as possible but still satisfies the institutional peer review requirements of board recertification [21]. It relies on the rereading of prior examinations of the same patient by a different radiologist during the course of interpreting new exams, thus fitting into the routine of busy radiologists. The arbiter of whether an error exists is a committee of the radiologists in the department.

One of the findings of RADPEER has been the discovery of a small number of outlier radiologists with higher discrepancy rates than their peers [21]. This is a very sensitive subject. It is at the heart of the individual vs. systems dualism suggested in the introduction. In order to foster physician compliance, this type of information cannot be used for economic retaliation or turf battles within a department. The goals of all of the data collection are education and finding the best way collectively and individually to reduce errors to the ultimate benefit of the patient.

However, there may be public pressure for the disclosure of such data. In Florida, for instance, there is a constitutional amendment that allows the legal discovery of peer review information. The supreme court of that state upheld the amendment by refusing to hear appeals from several hospitals wishing to withhold peer review information from plaintiffs' attorneys in malpractice lawsuits.

Conclusions

There are over 100,000 preventable deaths per year in our hospitals. The contribution of imaging errors to that total is not known. When they occur, they harm patients a significant minority of the time. There is no standard for acceptable levels of diagnostic accuracy. It is the goal of collecting the necessary data to set such a standard, establish rigorous peer review and reduce patient harms by reducing diagnostic errors. It is hoped that such measures will reduce the likelihood of lawsuits, hospital privilege restrictions, and licensure actions. Meticulous quality assurance measures could also have the highly desirable effect of reducing the

number of times that a radiologist will have to deal with the terrible feelings of having made a harmful error.

References

1. Kohn LT, Corrigan JM, Donaldson MS Eds: To err is human: building a safer health system. National Academy Press, Washington, DC, 1999
2. Mitka M: The Flexner report at the century mark a wake-up call for reforming medical education. *JAMA* 303:1465–1466, 2010
3. Sox HC, Woloshin S: How many deaths are due to medical error? Getting the number right. *Eff Clin Pract* 6:277–283, 2000
4. Brennan TA, Leape LL, et al: Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study. *New Engl J Med* 324:370–376, 1991
5. Leape LL, Brennan TA, et al: The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *New Engl J Med* 324:377–384, 1991
6. Hayward RA, Hofer TD: Estimating hospital deaths due to medical errors. Preventability is in the eye of the reviewer. *JAMA* 286:415–420, 2001
7. Department Of Health And Human Services Office Of Inspector General Adverse events in hospitals: national incidence among Medicare beneficiaries. November, 2010, Oei-06-09-00090
8. Landrigan CP, Perry GJ, et al: Trends in rates of patient harm resulting from medical care. *New Engl J Med* 363:2124–2134, 2010
9. Vincent C, Neale G, Woloshynowych M: Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 322(7285):517–519, 2001
10. Baker GR, Norton PG, et al: The Canadian adverse events study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 170(11):1678–1686, 2004
11. Leslie A, Jones AJ, Goddard PR: The influence of clinical information on the reporting of CT by radiologists. *Br J Radiol* 73:1052–1055, 2000
12. Dalla Palma L, Stacul F, et al: Relationships between radiologists and clinicians: results from three surveys. *Clin Radiol* 55:602–605, 2000
13. Berner ES, Graber ML: Overconfidence as a cause of diagnostic error in medicine. *Am J Med* 121(5A):S2–S23, 2008
14. Pinto A, Brunese L: Spectrum of diagnostic errors in radiology. *World J Radiol* 2(10):377–383, 2010
15. Sandhu H, Carpenter C: Clinical decision making: opening the black box of cognitive reasoning. *Ann Emerg Med* 48:715–719, 2006
16. Ely JW, Kaldjian LC, et al: Diagnostic errors in primary care: lessons learned. *JABFM* 25:87–97, 2012
17. Norman G: Dual processing and diagnostic errors. *Adv Health Sci Educ Theory Pract* 1:37–49, 2009
18. Felson B, Reeder MM: Gamuts in radiology: comprehensive lists of roentgen differential diagnosis, 2nd edition. Audiovisual Radiology of Cincinnati, Cincinnati, 1987
19. Renfrew DL, Franken EA, et al: Error in radiology: classification and lessons in 182 cases presented at a problem case conference. *Radiology* 183:145–150, 1992
20. Australian Patient Safety Foundation, Radiology Events Register. Final Report, 2009
21. Borgstede JP, Lewis RS, Bhargavan M, Sunshine JH: RADPEER quality assurance program: a multifacility study of interpretive disagreement rates. *J Am Coll Radiol* 1:59–65, 2004