

Leonard Berlin*

Radiologic errors, past, present and future

Man must strive, and striving he must err
Goethe, *Faust*, Part I [1].

Abstract: During the 10-year period beginning in 1949 with publication of five articles in two radiology journals and UKs *The Lancet*, a California radiologist named L.H. Garland almost single-handedly shocked the entire medical and especially the radiologic community. He focused their attention on the fact now known and accepted by all, but at that time not previously recognized and acknowledged only with great reluctance, that a substantial degree of observer error was prevalent in radiologic interpretation. In the more than half-century that followed, Garland's pioneering work has been affirmed and reaffirmed by numerous researchers. Retrospective studies disclosed then and still disclose today that diagnostic errors in radiologic interpretations of plain radiographic (as well as CT, MR, ultrasound, and radionuclide) images hover in the 30% range, not too dissimilar to the error rates in clinical medicine. Seventy percent of these errors are perceptual in nature, i.e., the radiologist does not "see" the abnormality on the imaging exam, perhaps due to poor conspicuity, satisfaction of search, or simply the "inexplicable psycho-visual phenomena of human perception." The remainder are cognitive errors: the radiologist sees an abnormality but fails to render a correct diagnoses by attaching the wrong significance to what is seen, perhaps due to inadequate knowledge, or an alliterative or judgmental error. Computer-assisted detection (CAD), a technology that for the past two decades has been utilized primarily in mammographic interpretation, increases sensitivity but at the same time decreases specificity; whether it reduces errors is debatable. Efforts to reduce diagnostic radiological errors continue, but the degree to which they will be successful remains to be determined.

Keywords: diagnostic error; radiological error; radiology.

*Corresponding author: Leonard Berlin, MD, FACR, Department of Radiology, Skokie Hospital, 9600 Gross Point Road, Skokie, IL 60076, USA, Phone: +847-933-6111, Fax: +847-933-6113, E-mail: lberlin@live.com; and Professor of Radiology, Rush University, University of Illinois, Chicago, IL, USA

The past

Nearly 65 years ago a California radiologist named L. Henry Garland shocked the medical community with publication in a radiology journal his article entitled, "On the Scientific Evaluation of Diagnostic Procedures" [2]. Summarizing investigations that revealed a "surprising" degree of inaccuracy in many non-radiologic clinical and laboratory as well as radiological tests in that article and several others that were published soon thereafter [3–5], Garland enumerated studies that found a 34% error rate in the diagnosis of myocardial infarction, an only 15% agreement among eight experienced internists in determining the presence of "the most simple signs" of emphysema when examining the chests of patients afflicted with that disease, a marked disparity in clinical evaluation of 1000 school children of the indications for tonsillectomy, an agreement rate of only 7% among five experienced pediatricians determining clinically whether children were suffering from malnutrition, a 20% error rate in the interpretation of electrocardiograms, a 28% error rate among 59 different hospital clinical laboratories in reporting the results of chemical analyses, and a 28% error rate among clinical laboratories in measuring the erythrocyte count.

Most of Garland's attention, however, was focused on radiologic errors. He found that experienced radiologists missed 30% of chest radiographs positive for radiologic evidence of disease, overread 2% of them that were actually negative for disease, and disagreed with themselves 20% of the time [5]. Garland commented:

Many clinicians continue to believe that their observations are accurate, and are unaware of the need . . . to reduce error. They feel . . . that Roentgen tests may be subject to faulty interpretation, but not careful "observation." Not only should clinicians recognize their own errors; they should admit them.

Garland went on to relate that when one of his friends, a well-known professor of radiology, learned that Garland's research disclosed that radiologists missed about one-third of roentgenologically positive films, the friend expressed the hope that Garland would discontinue his "investigations in this field because they were so morale-disturbing." When other radiologists were confronted with this data,

continued Garland, their usual reaction was, “Well, in my everyday work, this does not apply; I would do better than those busy investigators.” How wrong they were.

The present: error rates confirmed

Garland’s revelations about the incidence of errors in clinical medicine have been confirmed and expanded upon by subsequent researchers. Agreement among academic faculty physicians performing physical examination for spleen enlargement [6], liver enlargement [7], abdominal ascites [8], acute otitis media [9], and other assorted physical findings [10–12], has shown to be remarkably poor. Large autopsy studies have disclosed frequent clinical errors and misdiagnoses, with error rates as high as 47% [13, 14]. Errors ranging from 25% to 49% in pathologists’ interpretations of biopsy specimens, and a 24% error rate in laboratory results have also been reported [15, 16].

In the decades following Garland’s classic articles, a number of investigators replicated Garland’s findings relative to radiologic interpretations [17–26]. In a 1976 study at the University of Missouri, an error rate of 30% was reported among staff radiologists in their interpretation of chest radiographs, bone exams, gastrointestinal series, and special procedures [27]. Elsewhere researchers found that as many as 20% of colonic tumors were missed on lower gastrointestinal examinations [28]. Harvard University researchers [29] reported that radiologists disagreed on the interpretation of chest radiographs as much as 56% of the time. Additional studies conducted by researchers at major academic medical centers disclosed that from 26% to 90% of all lung carcinomas were missed by radiologists interpreting plain chest radiographs [30–32].

Numerous reports also documented similarly high error rates among the more recent “high-tech” modalities utilized in radiologic practice, such as sonography [33, 34], arteriography [35], MR angiography [36], MRI when evaluating lumbar disk herniation [37], MR when evaluating rotator cuff injury [38], MR when evaluating prostatic cancer [39], and radionuclide scans [40]. Similar error rates were also reported with chest CT scans harboring lung cancer [41, 42].

In a 2010 study in which three experienced radiologists who specialize in abdominal imaging initially reviewed 90 abdominal and pelvic CT examinations and then at a later date while blinded to previous interpretations were asked to reinterpret exams that had been read not only by themselves but by their colleagues, the interobserver discrepancy rate was 26%, the intraobserver

discrepancy rate 32% [43] – almost identical to Garland’s data 60 years earlier. Although these figures today are readily acknowledged, in the 1950s they were indeed astonishing to all radiologists.

Still other studies have confirmed a 35% error rate among radiologists interpreting radiologic studies obtained in patients who had undergone trauma [44–46]. Statistics disclosing inaccuracies in the interpretation of mammograms are startling [47–54]. A report from Yale University School of Medicine found that upon retrospective review of mammograms originally interpreted by experienced radiologists as normal, from 15% to 63% of breast carcinomas had been overlooked at initial readings [55]. A University of Arizona study found that in 75% of mammograms initially interpreted as normal, breast carcinomas could be seen on retrospective evaluation [56].

Retrospective experimental vs. “real-time” error rates

As explained by Garland, error rates can be calculated in two different ways, depending on the denominator used [5]:

If a series of 100 roentgenograms contains 10 positive and 90 negative films, and a reader misses three of the positive films and over reads two of the negative films, he may be regarded as having only a 5% error. On the other hand, since the series of 100 roentgenograms is being examined to detect patients with disease, the reader who misses three of the ten positive films has an error rate of 30%. Coupled with an over reading of two of the ninety negative films, the combined error rate in the example mentioned is about 32%.

In virtually all of the studies of error rates to which this article has thus far referred, the denominator consists of a preselected number of abnormal radiologic studies. Thus, if a radiologist participating in a research project is given 100 radiographs known to be abnormal and misses 30 of them, the error rate is obviously 30%. This should not be construed as indicating that radiologists commit an average 30% error rate in their everyday practices. Several studies have measured “real-time” error rates of radiologists by determining how many errors were committed among a large number of radiologic exams interpreted by a radiologist in a practice situation over a selected period of time – in other words, with a denominator including both normal and abnormal exams.

University of Texas researchers [57] reviewed imaging interpretations rendered in radiology departments of six community hospitals and found a 4.4% mean rate of

interpretation error. Other researchers [58] reviewed the performance of more than 250 radiologists who had interpreted more than 20,000 examinations as part of clinical testing of a performance improvement product, *Radpeer*. They found an all-case error rate of 3% to 3.5%. Still another group of researchers [59] reviewed the results of a quality improvement study conducted among 26 radiologists who read 6703 cases, and found an overall error rate of 3.48%. To summarize, if the denominator consists only of radiology studies that harbor abnormalities, the error rate averages 30%; if the denominator consists of an “everyday” mixture of abnormal and normal cases as is usually found in daily practices, the error rate averages 3.5%–4%.

It should be emphasized that none of the studies referred to in this article reflect the degree to which patient care is injured or otherwise jeopardized because of reader misinterpretation. “Extrapolation of reader error to medical care is complex” [29]. Although some radiologic errors may indeed result in serious injury and/or mismanagement of a patient, most are either corrected quickly or, fortunately, not clinically important, and thus exert no adverse effect on the health or management of the patient.

Causes of radiological errors: perceptual

Diagnostic errors in radiology may be perceptual or cognitive. Although it is not known exactly what percentage of diagnostic errors in radiology are due to perceptual misses, it has been estimated to be in the 60%–70% range [17, 23].

The failure to detect a radiologic abnormality is often attributed to the subtlety of the radiologic finding, or poor *conspicuity*, a term that is defined as the ratio between the contrast enhancement of the lesion relative to the surrounding tissues. While this definition may adequately explain how a truly subtle lesion can be missed, it is woefully inadequate to explain how an obvious abnormality can be missed. The phenomenon of simply not “seeing” initially an abnormality that is easily and clearly seen on a second look has never been explained to anyone’s satisfaction. Referred to as the “human factor” [2, 60], as the “foibles of human perception” [24], as an “irreducible necessary fallibility emanating from uncertainties inherent in medical predictions based on human observation and the laws of natural science” [13], and as a constant “inherent in all human activity” by Leape [61], the missing of an overt lesion remains as much a mystery and enigma today as it was 61 years ago.

Probably no event is more perplexing or frustrating to radiologists than the realization that they have committed a perceptual error – that they did not “see” on a radiologic imaging exam an abnormality that later is plainly evident. Despite the voluminous material that has been published in the radiologic literature on the subject of perceptual misses, we still do not know exactly why we miss obvious radiographic findings. We still do not know the answer to the question that all too often the erring radiologist asks in exasperation; “How and why did I not see that abnormality?”

One pioneering researcher in radiologic perception commented [60]:

So long as human beings are responsible for Roentgen interpretation, the process will be subject to the variability of human perception and “reader-error” due to the interpreter’s failure to perceive critical detail. The processes governing search behavior and mediating visual perceptual are correspondingly complicated, and our knowledge of them is fragmentary. Enough is known, however, to suggest that errors of perception are for the most not the result of carelessness or willful bias on the part of the radiologist, but rather a consequence of the physiologic processes of perception. Errors of perception are an unavoidable hazard of the “human condition.”

A British radiologist-researcher observed, “Although technology has made enormous progress in the last century, there is no evidence for similar improvement in the performance of the human eye and brain” [62].

Author Malcolm Gladwell in an article published in *The New Yorker* [63] made the following insightful observation regarding perceptual errors in radiologic interpretation:

The reason a radiologist is required to assume that the overwhelming number of ambiguous things are normal is that the overwhelming number of ambiguous things really are normal. Radiologists are, in a sense, a lot like baggage screeners at airports. The chances are that the dark mass in the middle of the suitcase isn’t a bomb, because you’ve seen a thousand dark masses like it in suitcases before, and none of those were bombs – and if you flag every suitcase with something ambiguous in it, no one would ever make his flight. But that doesn’t mean, of course, that it isn’t a bomb. All you have to go on is what it looks like on the X-ray screen – and the screen seldom gives you quite enough information.

Cognitive errors: alliterative and satisfaction of search

Whereas up to 70% of diagnostic radiologic errors are perceptual in nature – failing to “see” something on the

radiologic image – the remainder is due to *cognitive* errors or errors in judgement – i.e., attaching the wrong significance to a finding that is seen. It is impossible to delve into the minds of radiologists who have rendered erroneous conclusions, but one possible explanation referred to as “faulty reasoning” [17], is the radiologists’ failure to think of possibilities when interpreting radiographs: radiologists simultaneously combine perception of an abnormality with the notion of it, and the notion is often so strong that other features or information that might have modified the decision are rejected. Thus, “more things are missed through not being thought of, and so not looked for, than through not being known.”

Alliteration is defined as the occurrence in a phrase, or line of speech or writing, of two or more words having the same initial sound [64]. In the context of radiologic errors, the *alliterative* error results from the influence that one radiologist exerts on another: if one radiologist fails to detect an abnormality or attaches the wrong significance to an abnormality that is easily perceived, the chance that a subsequent radiologist will repeat the same error is increased. Alliterative errors occur because radiologists read the reports of previous examinations before or while reviewing the newly obtained radiologic studies and therefore are more apt to adopt the same opinion as that rendered previously by a colleague (or oneself) [65]. To what extent radiologists repeat the same errors as those committed by predecessor radiologists is not known with certainty, but they are not uncommon.

“Satisfaction of search” refers to the fact that the detection of one radiologic abnormality may interfere with the detection of additional abnormalities in the same examination. In other words, when viewing radiologic studies, there is a tendency to become “satisfied” after identifying the first abnormality that leads to a failure to search for additional findings [66]. In a study in which radiologists were shown in random order a number of exams in which one abnormality was present and the same number of exams in which two or three abnormalities were present, 75% of the abnormalities were reported when the examination contained one or two abnormalities. In examinations that contained three or more abnormalities however, only 41% were detected [67].

Computed assisted detection (CAD)

Over the past two decades, a multitude of articles discussing new technologies has appeared in the radiologic literature. In 1998, the US Food and Drug Administration

approved computed assisted detection (CAD) to assist radiologists in their interpretation of radiologic examinations. Since that time, CAD has been used primarily in mammography [68–70], where it has improved sensitivity by raising the level of the radiologist’s suspicion for breast cancer. CAD was initially shown to be of value in reducing radiologic errors and improving interpretation in mammography [71, 72]. Early studies disclosed a significant increase in breast cancer detection when CAD was utilized, but more recent studies have suggested that the value of CAD may have been overrated.

A study published in 2011 that examined records from 685,000 women who received more than 1.6 million mammograms from 1998 through 2006 disclosed that although CAD was associated with a statistically significant decrease in specificity, there was only a non-statistically significant increase in sensitivity but no statistically significant improvement in cancer detection rates [73]. It was also found that CAD increased a woman’s risk of being recalled unnecessarily for further testing. An accompanying editorial concluded that millions of women are being exposed to “a technology that may be more harmful than it is beneficial” [74]. Thus far, CAD has not been shown to reduce radiologic error.

Radiologic errors: the future

Awareness and understanding of all medical errors – clinical and radiologic – have grown rapidly in the past decade and a half. Nevertheless, it has been estimated that up to 80,000 US hospital deaths occur due to misdiagnosis annually, 5% of autopsies reveal lethal diagnostic errors for which a correct diagnosis would have averted death, and physician errors resulting in adverse events were more likely to be diagnostic than drug related [75]. In 2007, the Agency for Healthcare Research and Quality (AHRQ) announced special emphasis on funding diagnostic errors research. Health information technology, improved education, and increasing acknowledgment of diagnostic errors hold promise in error reduction [76], although such efforts still remain a goal rather than a reality. Diagnostic accuracy remains low [77]. Cognitive errors in radiology have been reduced through continuing medical education and by providing more complete patient history and clinical findings to the interpreting radiologist. Reducing perceptual errors, however, remains a challenge. A lament 44 years ago [60] that “the ultimate solution to the problem of ‘reader error’ is not yet clear,” still remains true today.

Fifty-four years ago, Garland exhorted future radiologists to continue “attempts at elucidation and correction of the factors involved” in causation of radiological errors. Urged on by this charge, radiologist and non-radiologist researchers today still pursue this goal, and will undoubtedly continue to do so for many years to come.

Conflict of interest statement: The author declares no conflict of interest.

Received September 10, 2013; accepted October 30, 2013

References

- Goethe, Faust, part I. Tripp KT editor. The international thesaurus of quotations. New York: Crowell Company, 1970:209.
- Garland LH. On the scientific evaluation of diagnostic procedures. *Radiology* 1949;52:309–28.
- Garland LH. On the reliability of roentgen survey procedures. *AJR* 1950;64:32–41.
- Garland LH, Miller ER, Zwerling HB, Harkness HT, Hinshaw HC, Shipman SJ, et al. Studies on the value of serial films in estimating the progress of pulmonary disease. *Radiology* 1952;58:161–77.
- Garland LH. Studies on the accuracy of diagnostic procedures. *AJR* 1959;82:25–38.
- Grover SA, Barkun AN, Sackett DL. Does this patient have splenomegaly? *J Am Med Assoc* 1993;270:2218–21.
- Naylor CD. Physical examination of the liver. *J Am Med Assoc* 1994;271:1859.
- Williams JW, Simel DL. Does this patient have ascites? *J Am Med Assoc* 1992;267:2645–8.
- Pichichero ME, Poole MD. Assessing diagnostic accuracy and tympanocentesis skills in the management of otitis media. *Arch Pediatr Adolesc Med* 2001;155:1137–42.
- Eliot DL, Hickam DH. Evaluation of physical examination skills; reliability of faculty observers and patient instructors. *J Am Med Assoc* 1987;258:3405–8.
- Sackett DL. A primer on the precision and accuracy of the clinical examination. *J Am Med Assoc* 1992;267:2638–44.
- Gruver RH, Freis ED. Study of diagnostic errors. *Ann Intern Med* 1957;47:108–20.
- Anderson RE, Hill RB, Key CR. The sensitivity and specificity of clinical diagnostics during five decades: toward an understanding of necessary fallibility. *J Am Med Assoc* 1989;261:1610–7.
- Roosen J, Frans E, Wilmer A, Knockaert DC, Bobbaers H. Comparison of premortem clinical diagnoses in critically ill patients and subsequent autopsy findings. *May Clin Proc* 2000;75:562–7.
- Landro L. Hospitals move to cut dangerous lab errors. *Wall Street Journal*, June 14, 2006: D1, D11.
- Plebani M. Errors in clinical laboratories or errors in laboratory medicine? *Clin Chem Lab Med* 2006;44:750–9.
- Smith MJ. Error and variation in diagnostic radiology. Springfield, IL: Thomas, 1967:4, 71, 73–74, 144–69.
- Stevenson CA. Accuracy of the X-ray report. *J Am Med Assoc* 1969;207:1140–1.
- Berlin L. Does the “missed” radiographic diagnosis constitute malpractice? *Radiology* 1977;123:523–7.
- Markus JB, Somers S, Franic SE, Moola C, Stevenson GW. Interobserver variation in the interpretation of abdominal radiographs. *Radiology* 1989;171:69–71.
- Berlin L. Reporting the “missed” radiologic diagnosis: medicolegal and ethical considerations. *Radiology* 1994;192:183–7.
- Berlin L. Errors in judgment. *AJR* 1996;166:1259–61.
- Berlin L. Perceptual errors. *AJR* 1996;167:587–90.
- Renfrew DL, Franken EA, Berbaum KS, Weigelt FH, Abu-Yousef MM. Error in radiology: classification and lessons in 182 cases presented at a problem case conference. *Radiology* 1992;183:145–50.
- Potchen EJ. Measuring observer performance in chest radiology: some experiences. *J Am Coll Radiol* 2006;3:423–32.
- Berlin L. Defending the “missed” radiographic diagnosis. *AJR* 2001;176:317–22.
- Lehr JL, Lodwick GS, Farrell C, Braaten O, Vitama P, Kolvisto EL. Direct measurement of the effect of film miniaturization on diagnostic accuracy. *Radiology* 1976;118:257–63.
- Cooley RN, Agnew CH, Rios G. Diagnostic accuracy of the barium enema study in carcinoma of the colon and rectum. *AJR* 1960;84:316–31.
- Herman PG, Gerson DE, Hessel SJ, Mayer BS, Watnick M, Blesser B, et al. Disagreement in chest roentgen interpretation. *Chest* 1975;68:278–82.
- Forrest JV, Friedman PJ. Radiologic errors in patients with lung cancer. *West J Med* 1981;134:485–90.
- Muhm JR, Miller WE, Fontana RS, Sanderson DR, Uhlenhapp MA. Lung cancer detected during a screening program using four-month chest radiographs. *Radiology* 1983;148:609–15.
- Austin JH, Romney BM, Goldsmith LS. Missed bronchogenic carcinoma: radiographic findings in 27 patients with a potentially respectable lesion evident in retrospect. *Radiology* 1992;182:115–22.
- James AE, Fleischer AC, Sacks GA, Greeson T. Ectopic pregnancy: a malpractice paradigm. *Radiology* 1986;160:411–3.
- Hertzberg BS, Kliewer MA, Paulson EK, Sheafor DH, Freed KS, Bowie JD, et al. PACS in sonography: accuracy of interpretation using film compared with monitor display. *AJR* 1999;173:1175–9.
- Manning WJ, Li W, Adelman RR. A preliminary report comparing magnetic resonance coronary angiography with conventional angiography. *N Engl J Med* 1993;328:828–32.
- Litt AW, Eidelman EM, Pinto RS, Riles TS, McLachlan JJ, Schwartzberg S, et al. Diagnosis of carotid artery stenosis: comparison of 2DFT time-of-flight MR angiography with contrast angiography in 50 patients. *AJNR* 1991;12:149–54.
- van Rijn JC, Klemetso N, Reitsma JB, Majoie CB, Huisman FJ, Pevl WC, et al. Observer variation in MRI evaluation of patients suspected of lumbar disk herniation. *AJR* 2005;184:299–303.

38. Robertson PL, Schweitzer ME, Mitchell DG, Schlesinger F, Epstein RE, Friedman BG, et al. Rotator cuff disorders: interobserver and intraobserver variation in diagnosis with MR imaging. *Radiology* 1995;194:831–5.
39. Schiebler ML, Yankaskas BC, Tempany BC, Spritzer CC, Rifkin MD, Pollack HM, et al. MR imaging in adenocarcinoma of the prostate; interobserver variation and efficacy for determining stage C disease. *AJR* 1992;158:559–62.
40. Sigal SI, Soufer R, Fetterman RC, Matterna JA, Wackers FJ. Reproducibility of quantitative planar thallium-201 scintigraphy: quantitative criteria for reversibility of myocardial perfusion defects. *J Nucl Med* 1991;32:759–65.
41. Gurney JW. Missed lung cancer at CT: imaging findings in nine patients. *Radiology* 1996;199:107–12.
42. White CS, Romney BM, Mason AC, Austin JH, Miller BH, Protopapas Z. Primary carcinoma of the lung overlooked at CT: analysis of findings in 14 patients. *Radiology* 1996;199:109–15.
43. Abujudeh HH, Boland GW, Kaewlai R, Rabinar P, Halpern EF, Gazelle GS, et al. Abdominal and pelvic computed tomography (CT) interpretation: discrepancy rates among experienced radiologists. *European Radiol* 2010;20:1952–7.
44. Janjua KJ, Sugrue M, Deane SA. Prospective evaluation of early missed injuries and the role of tertiary trauma survey. *J Trauma* 1998;44:1000–7.
45. FitzGerald R. Error in radiology. *Clin Radiol* 2001;56:938–46.
46. FitzGerald R. Radiological error: analysis, standard setting, targeted instruction and teamworking. *Eur Radiol* 2005;15:1760–7.
47. Sickles EA. Breast imaging: from 1965 to the present. *Radiology* 2000;215:1–16.
48. Berlin L. The missed breast cancer: perceptions and realities. *AJR* 1999;173:1161–7.
49. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast imaging reporting and data system: inter- and intraobserver variability in feature analysis and final assessment. *AJR* 2000;174:1769–77.
50. Baines CJ, McFarlane DV, Miller AB. Role of the reference radiologist: estimates of inter-observer agreement and potential delay in cancer detection in the National Breast Screening Study. *Invest Radiol* 1990;25.
51. Beam CA, Layde PM, Sullivan DC. Variability in the interpretation of screening mammograms by US radiologists. *Arch Intern Med* 1996;156:209–13.
52. Kerlikowske K, Grady D, Rubin S, Sandrock C, Ernster V. Efficacy of screening mammography: a meta-analysis. *J Am Med Assoc* 1995;273:149–54.
53. Mushlin AI, Kouides RW, Shapiro DE. Estimating the accuracy of screening mammography: a meta-analysis. *Am J Prev Med* 1998;14:143–53.
54. Kerlikowske K, Grady D, Barclay J, Ernster V, Frankel SD, Ominsky SH, et al. Variability and accuracy of mammography interpretation using the American College of Radiology Breast Imaging Reporting and Data System. *J Natl Cancer Inst* 1998;90:1801–9.
55. Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretations of mammograms. *N Engl J Med* 1994;331:1493–9.
56. Harvey JA, Fajardo LL, Innis CA. Previous mammograms in patients with impalpable breast carcinoma: retrospective vs blinded interpretation. *AJR* 1993;161:1167–72.
57. Siegle RL, Baram EM, Reuter SR, Clarke EA, Lancaster JL, McMahan CA. Rates of disagreement in imaging interpretation in a group of community hospitals. *Acad Radiol* 1998;5:148–54.
58. Borgstede JP, Lewis RS, Bhargavan M, Sunshine JH. RADPEER quality assurance program: a multifacility study of interpretive disagreement rates. *J Am Coll Radiol* 2004;1:59–65.
59. Soffa DJ, Lewis RS, Sunshine JH, Bhargavan M. Disagreement in interpretation: a method for the development of benchmarks for quality assurance in imaging. *J Am Coll Radiol* 2004;1:212–7.
60. Tuddenham WJ. Roentgen image perception – a personal survey of the problem. *Radiol Clin N Amer* 1969;7:499–501.
61. Leape LL. Error in medicine. *J Am Med Assoc* 1994;272:181–7.
62. Robinson PJ. Radiology's achilles' heel: error and variation in the interpretation of the roentgen image. *Br J Radiol* 1997;70:1085–98.
63. Gladwell M. The picture problem: mammography, air power, and the limits of looking. *The New Yorker* December 13, 2004
64. *The American Heritage Dictionary*, Second College Edition. Boston: Houghton Mifflin Co., 1985:95.
65. Berlin L. Alliterative errors. *Am J Roent* 2000;174:925–31.
66. Rogers LF. Keep looking: satisfaction of search. *Am J Roent* 2000;175:287.
67. Ashman CJ, Yu JS, Wolfman D. Satisfaction of search in osteoradiology. *Am J Roent* 2000;175:541–4.
68. Warren Burhenne LJ, Wood SA, D'Orsi, CJ, Ferg SA, Kopans DB, O'Shaughnessy KF, et al. Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology* 2000;215:554–62.
69. Freedman M, Osicka T. Reader variability: what we can learn from computer-aided detection experiments. *J Am Coll Radiol* 2006;3:446–55.
70. Khorasani R, Erickson BJ, Patriarche J. New opportunities in computer-aided diagnosis: change detection and characterization. *JACR* 2006;3:468–9.
71. Morton MJ, Whaley DH, Brandt KR, Amrami KK. Screening mammograms: interpretation with computer-aided detection – prospective evaluation. *Radiology* 2006;239:375–83.
72. Dean JC, Ilvento CC. Improved cancer detection using computer-aided detection with diagnostic and screening mammography: prospective study of 104 cancers. *Am J Roent* 2006;187:20–8.
73. Fenton JJ, Abraham L, Taplin SH, Geller BM, Carney PA, D'Orsi C, et al. Effectiveness of computer-aided detection in community mammography practice. *J Natl Cancer Inst* 2011;103:1–10.
74. Berry DA. Computer-assisted detection and screening mammography: where's the beef? *J Natl Cancer Inst* 2011;103:1139–40.
75. Newman-Toker DE, Pronovost PJ. Diagnostic errors – the next frontier for patient safety. *J Am Med Assoc* 2009;301:1060–2.
76. Lee CS, Nagy PG, Weaver SJ. Cognitive and system factors contributing to diagnostic errors in radiology. *Am J Roent* 2013;201:611–7.
77. Meyer AN, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests. A vignette study. *JAMA Intern Med*. doi: 10.1001/JAMAinternmed.2013.10081. Published online August 26, 2013.