

Review

Michael A. Bruno*

256 Shades of gray: uncertainty and diagnostic error in radiology

DOI 10.1515/dx-2017-0006

Received February 6, 2017; accepted May 8, 2017; previously published online July 28, 2017

Abstract: Radiologists practice in an environment of extraordinarily high uncertainty, which results partly from the high variability of the physical and technical aspects of imaging, partly from the inherent limitations in the diagnostic power of the various imaging modalities, and partly from the complex visual-perceptual and cognitive processes involved in image interpretation. This paper reviews the high level of uncertainty inherent to the process of radiological imaging and image interpretation vis-à-vis the issue of radiological interpretive error, in order to highlight the considerable degree of overlap that exists between these. The scope of radiological error, its many potential causes and various error-reduction strategies in radiology are also reviewed.

Keywords: cognitive bias; error; payment models; perception; perceptual error; performance improvement; statistical variation; uncertainty; working memory.

Introduction

Radiologists practice in an environment of extraordinarily high uncertainty. While they know a great deal about the appearance of disease processes on radiographs and scans, they are almost always given incomplete information as to the patient's clinical presentation which prompted the imaging study they are asked to interpret. There is also a great deal of variability inherent in the imaging process itself: the images from which radiologists extract their information are acquired using a wide range of equipment types and physical techniques, and are in turn processed and reconstructed digitally in myriad different ways that vary by equipment vendor,

settings and technologist preferences. The information contained within the images is “encoded” in often very subtle differences in audio- or radiofrequency signal or X-ray attenuation that can be highly variable from patient to patient, with a broad range of “normal” that usually overlaps the even broader range of “abnormal”. Radiologists must therefore continually make mental adjustments of their expectations to account for unknowable degree of variation that enters the diagnostic imaging process even before the images are ever viewed – variation in anatomy from patient to patient, in technique and positioning, image noise and digital image manipulation, patient size and body composition, motion, and the variable effects of contrast-enhancing materials (such as iodinated contrast or Gadolinium-based contrast) that may be observed between patients in both health and disease, as well as a potentially bewildering variety of physical and biological artifacts. Compounding this, there are few, truly pathognomonic appearances of disease on imaging. To the contrary, many very distinct disease states can appear essentially identical to each other, or at least share a common imaging appearance (see Figure 1). Finally, the practice of diagnostic radiology itself is a fallible, human endeavor, one involving complex perceptual, neuro-physiological and cognitive processes employed under a range of circumstances and practice settings.

In this paper I wish to first review the high level of uncertainty inherent to the process of radiological interpretation vs. what might constitute a radiological interpretive error – in order to highlight the considerable degree of overlap that exists, before addressing the scope of error, potential causes of error and possible error-reduction strategies in radiology.

The radiologist's task begins with visual detection of (often very subtle) variations in electronically displayed shades of gray in an image or set of images, which form the basis of discrimination. Although the human eye is capable of discriminating approximately 500 shades of gray, digital images are typically constructed with eight bits per pixel, which yields 256 shades of gray. These grayscale variations are the “findings” from which the presence or absence of disease is inferred. The radiologist

*Corresponding author: Michael A. Bruno, MS, MD, FACR, Penn State Health/Milton S. Hershey Medical Center and The Penn State College of Medicine, 500 University Drive, Mail Code H-066, Hershey, PA 17033, USA, E-mail: mbruno@pennstatehealth.psu.edu

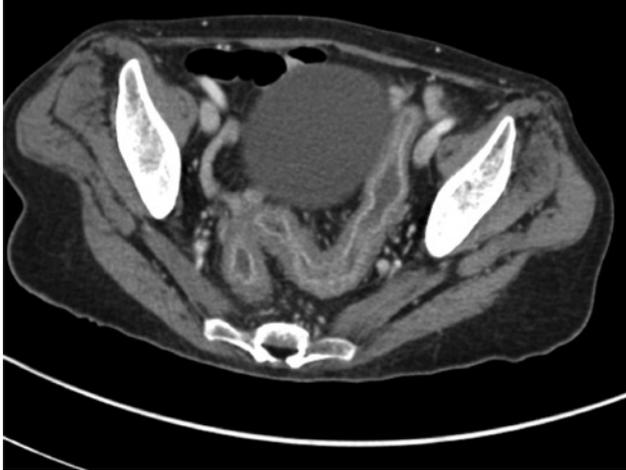


Figure 1: CT image of a patient showing a region of abnormal-appearing bowel.

The appearance is not pathognomonic for any particular disease; a reasonable differential diagnosis for this CT appearance includes ischemia, infection, and inflammatory bowel disease. Although they share a common appearance on imaging, these three distinct disease entities have little in common with each other in terms of etiology, treatment or prognosis.

must find the ones which are meaningful, ignore the ones that are not, and ultimately decide which might have diagnostic significance for a particular patient – and, if so, what. The basis of those discriminations is both dependent upon the neurophysiologic aspects of visual perception in the individual radiologist combined with their experience and training. This training, most commonly requiring 6 years after medical school, is a period of time in which radiologists-in-training will likely have viewed well over 60,000 medical imaging studies under the tutelage of more senior radiologists. Despite decades-long efforts aimed at standardizing the radiology residency curriculum and national certification testing, it is inevitable that individual radiologists’ training may or may not have included certain topics, and that their understanding of the importance of particular appearances seen on images will be variably influenced by the opinions and beliefs held by the senior radiologists who trained them. Further, each radiologist will inevitably bring into the interpretive process their own variable opinions and biases as to the likelihood of various disease entities being present within the population of patients whom they serve in their communities. Clearly, such a complex process (visual image interpretation) is inevitably subject to wide variation, both between and even within individual operators.

Finally, radiologists and other physicians rarely have the luxury of time; like other physicians radiologists are generally performing their duties under time-pressure due

to the urgency of the clinical questions they must address, the number of patients needing care, etc. In the world of medicine, inaction is rarely an option, despite the high degree of uncertainty that prevails. It is generally not possible to control most of the key variables in the diagnostic process, nor to have complete information about any given patient under a doctor’s care – which is why all physicians of all specialties generally must make decisions based on incomplete information.

It should not be surprising, therefore, that both the process and the performance outcomes of radiological diagnoses are subject to a great deal of variability. For example, in a 2010 study performed at the Massachusetts General Hospital by Abujudeh et al. a number of leading subspecialist radiologists reinterpreted a series of CT studies of the abdomen and pelvis which had been previously interpreted by their group, including cases they themselves had interpreted as well as those interpreted by one of their close colleagues. In these “second-look” image reinterpretations, which were performed blinded to patient outcomes, these noted expert radiologists disagreed with each other on more than 30% of the CT studies, and disagreed with their own original readings more than 25% of the time [1]!

Strictly speaking, however, the existence of inter- and intra-observer variation of this magnitude is not in itself proof that one or another radiologist has made an interpretive error; to the contrary, it merely illustrates the degree to which high levels of uncertainty and inherent variability in the process limits the conclusions of imaging tests.

Radiologic ‘error’ vs. radiologic uncertainty

To the extent that a defining reference standard may exist to constitute a “correct” diagnostic interpretation – such as an autopsy result or objective findings at surgery – we can potentially define a radiological error as having occurred when a radiologist’s diagnostic impression differs substantively from the objective, final diagnosis at surgery or autopsy. However, these sorts of “gold standards” are rarely available, and so historically radiologists have defined “error” as individual deviation from a defined (most commonly, ad hoc) peer consensus [2, 3]. Typically such a consensus is reached in retrospect once the final outcome is known and thus does not fully take the degree of prospective uncertainty into account. Further, with regard to reference standards, studies have shown considerable

The diagnostic process

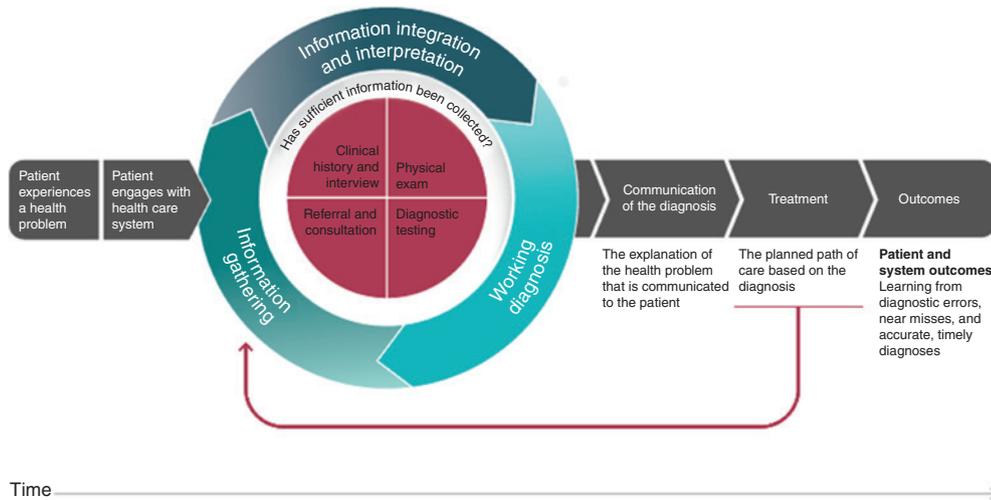


Figure 2: Conceptual model of the diagnostic process, from Ref. [6].

variability in pathologists' final diagnoses from surgical specimens and at autopsy when different pathologists re-examine the same specimens, calling into question the validity of the pathological diagnosis or even autopsy as an objective standard [4, 5]. Accordingly, we may reasonably conclude that a significant fraction of discrepant diagnoses, which are commonly deemed to represent radiologists' errors, more likely represent manifestations of the underlying uncertainty in which radiologists practice.

Radiology and diagnostic error

Diagnostic error in medicine is known to be a common occurrence, and a major cause of patient harm. The National Academies of Medicine 2015 report *Improving Diagnosis in Health Care* [6], provided a conceptual model of the diagnostic process as an iterative one (Figure 2). It provided a definition as to what constitutes diagnostic error, provided an estimate of the prevalence and potential impact of diagnostic error to involve 5% of US adults seeking outpatient care in any given year – such that every American is likely to experience one diagnostic error in their lifetime – and prescribed a national course of action to reduce both the incidence and impacts of diagnostic errors in practice.

The extent to which incorrect radiological interpretations contribute to the overall problem of diagnostic error is uncertain; given the high degree of reliance of physicians on imaging studies to establish a diagnosis, the

radiological contribution to the overall problem of diagnostic error is likely to be substantial [2]. In a recent audit of medical errors resulting patient harms reported in 2016 to the Pennsylvania Patient Safety Authority under the State's Act 13 reporting mandate, for example, the fraction of cases attributable to radiology was slightly <5%. Approximately 4.7% of Act 13 reports to the PA Patient Safety Authority in CY 2015 were attributed to radiology (unpublished data; personal communication). Combined with the Institute of Medicine's estimate of 12 million diagnostic errors occurring annually, and the estimate that there are approximately 30,000 practicing radiologists in the US [7] this would yield a very rough estimate of 19 or more diagnostic errors attributable to each practicing radiologist, on average, every year.

But there is reason to believe that the actual "error rate" of radiologists in practice is much higher than that, with each radiologist experiencing perhaps as many as three or four errors per day on average, based on multiple studies dating back to the work of Radiologist Leo Henry Garland in 1949. In fact, no reported measurements of radiologist errors has ever fallen below 4% under any circumstances, and were in some cases substantially higher [8–12]. Now, given a crude estimate that the average radiologist workload slightly exceeds 100 case interpretations per day, or 18,000 in a typical year [13], a figure which is based on older data and is likely underestimated, one can extrapolate a much larger lower boundary for average annual errors for each full-time radiologist each year, which is a number close to 700. Taking a different approach by working from a fairly recent estimate that there are nearly 500 million

radiologic imaging examinations performed in radiological practices within the US [14], multiplied by the empirical 4% minimal error rate yields an estimate of nearly 20 million radiologist errors per year divided by the 30,000 practicing radiologists, yields an average error figure also close to 700 errors per radiologist per year. I believe the discrepancy between the two estimates of radiologist errors (19 errors per radiologist per year based on extrapolation from the fraction of reported error/harm cases in the database of one state vs. an estimate of 700 errors per radiologist per year calculated from empirical estimates) lends further support to the conclusion that the vast majority of radiologist errors go undetected or do not result in patient harm.

From past studies of the types of errors that radiologists are prone to make, there is reason to believe that most radiological errors are errors of omission – most of them with statistically little chance of having a direct or immediate consequence to a patient’s final outcome. This conclusion seems likely based on existing evidence, although it is impossible to verify experimentally. It also suggests the need for future research on developing the means to detect radiologists’ errors more effectively and ideally in real time.

Types of radiologist error

Most radiological errors have been shown to be perceptual, i.e. errors that occur when a radiologist simply fails to perceive a finding on an image [15]; sometimes such findings are subtle, but they are often readily apparent in retrospect (Figure 3). This phenomenon accounts for between 60 and 80% of all radiologist interpretive errors in most series dating back several decades [8, 9, 15–19]. The reasons for this type of error are not well established, but are believed to be complex, and have, thus far, proved intractable [2].

There is also a well-described phenomenon known as “satisfaction of search”, first described in 1990, in which a radiologist fails to detect a second abnormality, apparently because of prematurely ceasing to search the images after detecting a “satisfying” finding, perhaps one that explains the patient’s clinical symptoms or is “satisfying” to the radiologist in some other way [20, 21]. In the recent study by Kim and Mansfield, this phenomenon was associated with a large number of radiologist interpretive errors, approximately half as many as due to perceptual errors [15]. Ashman et al. analyzing errors in the interpretation of bone radiographs, noted that the detection rate for the first finding in a study that contained multiple abnormalities was about 78%, but the second and third findings were each discovered approximately 40% of the time [22]. Most radiologists will report having succumbed



Figure 3: An example of a perceptual error. AP radiograph of the chest reveals a swallowed coin within the esophagus. This finding was missed – twice – by a skilled and experienced subspecialist radiologist. From Ref. [2].

to this type of error at some time during their career, although a more recent study by Berbaum, Krupinski et al. failed to detect this type of error in their study population, which suggests that its prevalence may possibly have been previously over-estimated [23].

Other types of radiologist interpretive errors have been described that may be more amenable to intervention, such as gaps in knowledge, the presence of cognitive biases, the failure to detect findings located in unexpected places on the images, or simply a misunderstanding of the pre-test probability of disease, which in turn may be due to misleading or erroneous clinical information having been provided to the interpreting radiologist by the patient’s doctor. In addition to perceptual errors and satisfaction of search, Kim and Mansfield also described 10 additional types/categories of radiologist errors, each accounting for a small fraction of the errors detected in their series. Of note, they observed that radiologists’ errors, from whatever apparent cause, tended to be repeated or propagated on subsequent radiological examinations of the same patient, either by themselves or by other radiologists who interpreted those subsequent studies [15].

Strategies for error reduction and harms prevention – “Sutton’s law”

Considering the immense complexity of the radiologist’s perceptual task, the high degree of variability in the

underlying processes involved to extract information by physical measures from the patient's body via imaging and ultimately transform these into the radiologists' final impression, and the extremely high level of uncertainty that prevails in every step of this process, it would appear that the occurrence of error in radiological testing is most likely inevitable, and that we would be justified in thinking that the relatively high incidence of radiological error in practice may indeed be irremediable. It makes sense, however to focus limited research resources and effort on those areas which are more likely to have the greatest benefit. When the notorious American bank robber Willie Sutton (1901–1980) was asked why he chose to rob banks for his career, he replied simply, "Because that's where the money is". The quote was ultimately the source of 'Sutton's law' in medicine, which, simply stated, is that one should focus effort where success is more likely (i.e. 'go where the money is').

Perceptual errors – wherein a radiologist simply fails to observe a finding which is sometimes subtle but is often readily apparent in retrospect – have been shown to be, far and away, the most common type of radiologist error [2, 10, 15]. Accordingly, research aimed toward understanding the underlying psychophysical processes of perception, whether entering the radiologists conscious awareness or not, and the role of working memory and other "mental" issues is needed, if there is to be any hope of significantly reducing the overall prevalence of radiological error.

Sub-conscious detection

Careful research performed very recently has unequivocally shown that radiologists can in fact detect abnormalities without having any conscious awareness that they have done so. For example, a recent study by Evans, Haygood et al. revealed that even a half-second glimpse of a mammogram yields a higher than chance probability that the radiologist will be able to guess whether a cancer is present in the patient's opposite (disease-free) breast [24]! Indeed, it has been long known based on eye-tracking studies that radiologists are often capable of identifying an abnormality on an image in a mere half-second, and presumably often do. It is thought that in such instances abnormalities may be initially detected in the peripheral visual field, leading to a very rapid reorientation of gaze to the area of interest [25]. Such ultrafast detection, of course, precedes any conscious awareness of the finding or thoughtful consideration of its meaning. Since the highest visual acuity is in the central (foveal)

vision, finding an abnormality by visual search requires moving the eyes around an image in order to aim the central visual field onto each of many areas of interest, one after another. Often, these eye movements are performed in a more or less habitual way by each radiologist, guided by such factors as personal habit, training and experience, sometimes influenced by the clinical question posed by the referring physician or other heightened pretest suspicion of disease, or even the radiologist's own clinical knowledge of the range of abnormalities being searched for. The pattern of eye movements performed has been studied for years using eye-tracking technologies, and is generally believed to be fairly consistent for each radiologist, and has become known as the radiologists' "search pattern". Studies of this type have shown that search patterns "deteriorate" due to radiologist fatigue, distraction or interruption [26], which suggests that the monitoring of eye movements could potentially provide an avenue to reveal otherwise difficult-to-detect perceptual errors, or at least reveal when there is heightened risk of such errors.

For example, in a 2014 study by Mallett et al. the "dwell time" of a radiologists gaze fixation in certain areas of a radiographic image seemed to be related to the presence of abnormalities in those areas of the images, whether or not the abnormality was consciously detected [27]; in fact, it was observed that the subject radiologists more frequently returned to areas of an image which contained (consciously undetected) abnormalities! Poorly understood attributes of working memory may also play an important role in these perceptual errors. One can hypothesize that some of these abnormalities that are detected visually, but perhaps below the level of consciousness, are somehow "rejected" and thus not stored in working memory sufficiently to be included in the radiologists' final impression. Any or all of these testable hypotheses could yield workable strategies – using existing technologies – to either prevent or reduce perceptual error or alternatively to provide "triggers" for early detection of errors before patient harm results.

A history of failed remediation: pitfalls vs. perception

It has been suggested that the triumph of the sciences has been resilience in the face of repeated failure [28], and more precisely a pervasive cultural willingness within the scientific community to admit, scrutinize and carefully analyze past failures and learn from them. Based on this reasoning, careful analysis of failed prior attempts to

reduce radiologists' errors may well be expected to prove both illustrative and useful.

In the more than six decades since Garland's initial work on radiologists' errors was published, there have been concerted and sophisticated efforts to address the problem within the radiology community. Most of this effort has been educational; following the hypothesis that gaps in the knowledge base of individual practitioners is what leads to failures in perception. A common mantra of radiologists – one sees what one knows – is correctly attributed to the German philosopher Goethe and is commonly recited in connection with continuing medical education (CME). The intensive education of radiologists that begins in their residency training continues more or less unabated after residency, with ongoing knowledge updates and re-training of practicing radiologists, enforced by Maintenance of Certification and State licensure requirements via various types of CME. Whether held in vast hotel ballrooms, university lecture halls, or in print journal articles or textbooks – and now increasingly via interactive electronic media – a massive ongoing educational effort and industry has developed around the knowledge-gap hypothesis. If one reviews the learning topics and titles from these educational sessions, the most common recurring keyword in CME is “pitfall”. Radiologists are constantly admonished to avoid one or another such “pitfalls”, or gaps in their knowledge leading to errors, and often are instructed how to do so in great detail by expert teachers. Unfortunately, it has become clear that this educational strategy to eliminate radiologist error, while not entirely without theoretical merit, has been pursued quite vigorously for decades across the entire spectrum of diagnostic radiology, but has ultimately failed to make any measurable difference [10].

Cognitive “de-biasing” and other promising but unproven strategies

A great many diagnosis errors appear to involve faulty or biased cognitive processes – lapses of logic, generally unconscious and emotionally-driven, which interfere with clear, objective thinking – but that can actually be adaptive under certain circumstances, including those under which human beings evolved. There is a large and growing psychology literature on this subject, with at least 150 such cognitive and affective biases having been described [29, 30]. Along with the search-satisficing tendency noted above, three of these are felt to be the most

common, and thus account for the greatest proportion impact on diagnostic accuracy. These are:

- “Anchoring bias”, in which one's initial conclusion is never called into question by the discovery of contrary information but instead the new information is either discounted or modified to match the prevailing hypothesis,
- “Confirmation bias”, a related mental process in which there is a tendency to search for confirmatory evidence to support one's hypothesis and discount or ignore contrary evidence, and
- “Availability bias”, in which a diagnosis is considered more likely when one has recently seen a case, especially if the diagnosis was missed in the prior instance.

It is interesting to note that these biases are also the most commonly encountered in studies of clinical decision-making involving internists [31].

For radiologists, the manifestations of cognitive bias may be easily appreciated in cases where an abnormality is detected perceptually, but the significance of a perceived finding is mistaken. But unconscious bias may also play a role even in cases that appear on the surface to be primarily perceptual errors (i.e. failure to perceive). In these cases the abnormality in question may be missed at least partially because of underlying bias in the radiologist's expectations of which findings are likely to be encountered, or what types of findings will be searched for within an imaging study.

Accordingly, strategies for “cognitive de-biasing” and related meta-cognitive interventions (i.e. thinking about thinking) have been advocated to remediate these common types of radiologist error [32–34], especially for the most common of the known cognitive biases that may impact clinical reasoning and information-gathering.

Although evidence that ‘de-biasing’ can reduce cognitive error is mixed, recent reviews have been favorable [35] and clearly this is an area that requires more research. A convincing counter-argument is that the various cognitive biases that undermine diagnostic accuracy may be hardwired, and thus refractory to conscious efforts to prevent them. However, with training it may be easier and more practical to detect these biases than to prevent them [36, 37] providing an opportunity to intervene before there is harm.

Another promising idea which has not yet made a difference in this sphere is computer-aided detection (CAD) technology. The basic idea is straightforward: a software algorithm is designed to identify suspicious features on images and bring these to the attention of the radiologist, who might otherwise have missed them. While

widely used in the current practice of mammography due to governmental payment incentives, it is commonly accepted among radiologists that CAD has no real impact in reducing the overall risk that significant abnormalities will be missed, an opinion which is supported by experimental evidence [38, 39]. Clearly, however, the underlying technology and algorithmic sophistication, especially with the advent of computer “deep learning”, is continually improving and so there is reason to believe that CAD still has promise for improving radiologist performance in the future [40]. It is certainly an area of very active research within technology companies worldwide. Likewise, applications of technologies based on eye-tracking technology, as discussed above, may also aid radiologists by providing subtle gaze direction or real-time feedback to radiologists about areas on an image where their visual “dwell time” suggests that an unrecognized abnormality may be present [41, 42].

Practice quality improvement (PQI), systems and failsafe strategies

If radiologist errors, especially perceptual errors, are indeed as inevitable and intractable as they seem, then developing the means to improve the early detection and self-correction of errors is of paramount importance for prevention of patient harm. It makes sense to invest effort into developing checks and balances to reduce the potential harm of errors after the fact, as well as to develop “trigger tools” to facilitate early detection of errors soon after they occur and hopefully before any irreparable harm is done.

Systems-based practice quality improvement strategies designed to better manage the radiologists’ workload, reduce fatigue and burnout, better optimize the speed of radiologists’ work (not too fast, but also not too slow), improve overall ergonomics and reduce interruptions and distractions should all have a positive effect [43, 44].

The importance of fostering an open and blameless culture within medicine, in order to foster rapid “near miss” and error-discovery and promote systems-based learning cannot be over-emphasized. The relevant science strongly supports the conclusion that the current prevailing organizational culture of assigning individual blame to radiologists for interpretive errors, a longstanding approach which itself arose from the traditional “blameful” culture of medicine (and which fuels the medical-malpractice industry), is rarely, if ever appropriate or helpful [45]. Changing this culture has been a primary, though

elusive goal of the quality and safety movement since at least the time of the 1999 IOM report, *To Err is Human* [46]. It remains an almost quixotic quest, a continual struggle against entrenched interests and habits of thought that extend well beyond medicine [47].

Where the money is – or is not

Double-reading of imaging studies, wherein each study is independently interpreted by two observers blinded to each other’s impressions would be expected to dramatically reduce the risk of random perceptual errors reaching the patient. This approach was first suggested by Garland in 1949 [48], has been tried sporadically, most commonly in the practice of mammography, and generally with good results [49]. Unfortunately, widespread double-reading is essentially impossible to implement in our nation’s healthcare system, owing primarily to the enormous number of radiological studies performed in the US vs. the limited interpretive capacity of the 30,000 or so practicing radiologists available to do the interpretive work, not to mention the currently insurmountable financial barriers to widespread implementation of double-reading, since the cost of second readings are generally not underwritten by health insurers. Further, since perceptual errors are random, a strategy of selective or sporadic double-reading of cases deemed to be particularly ‘difficult’ or otherwise ‘worthy’ of the added labor and expense is unlikely to make a difference. As the traditional fee-for-service payment model in healthcare is progressively replaced by capitated or bundled prospective-payment “value” payment models, however, the cost issue may become less of an absolute barrier in the years ahead. Matching the increased demand for radiologist services to the manpower will probably be more limiting.

Conclusions

In conclusion, when one considers the degree of uncertainty involved in the process of radiological interpretation, the incredible complexity of the radiologists’ neurocognitive tasks, the variability of technical process which radiologists must account for and the resultant inevitability of error, it is remarkable that radiologists perform as well as they do! It is certainly a testament to the diligence, training and professional expertise of these individual physicians. But until such time as the radiologists’ task is taken over by robots, the practice of radiology will

remain a human endeavor, and thus can never be made error-free. Future research, therefore, must be focused on developing methods to facilitate early error detection for prevention of harm. Basic research into understanding the underlying processes of human visual perception as applied to the radiologist's task are likely to yield new understandings that, in turn, may lead to workable strategies to substantially improve radiologists' performance. In the meantime, it would be foolish to pass over or ignore the low-hanging fruit represented by applying well-understood quality improvement strategies for systems-based learning and improvement, including fostering an open and blameless culture throughout medicine, and paying attention to common-sense systems/management interventions that serve to reduce the effects of fatigue, interruptions and distractions, and physician burnout.

Matching the financial incentives in our healthcare system to foster, rather than interfere with, the desired outcomes of healthcare will best satisfy both the letter and the spirit of Sutton's law.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Abujudeh HH, Boland GW, Kaewlai R, Rabiner P, Halpern EF, Gazelle GS, et al. Abdominal and pelvic computed tomography (CT) interpretation: discrepancy rates among experienced radiologists. *Eur Radiol* 2010;20:1952–7.
2. Bruno MA, Walker EA, Abujudeh HH. Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction. *RadioGraphics* 2015;35:1668–76.
3. Waite S, Scott J, Gale B, Fuchs T, Kolla S, Reede D. Interpretive error in radiology [review]. *AJR Am J Roentgenol* 2017;208:739–49.
4. Hill RB. The current status of autopsies in Medical Care in the USA. *Int J Qual Health Care* 1993;5:309–13.
5. Swapp RE, Aubry MC, Salomão DR, Cheville JC. Outside case review of surgical pathology for referred patients: the impact on patient care. *Arch Pathol Lab Med* 2013;137:233–40.
6. Balogh EP, Miller BT, Ball JR, editors. Board on Health Care Services, Institute of Medicine. Improving diagnosis in healthcare. Washington, DC: The National Academy of Sciences, The National Academies Press, 2015.
7. Hughes DR. Commentary: "How Many Radiologists? It depends on who you ask!" Website of the Harvey Niemen Health Policy Institute © 2017 American College of Radiology. <http://www.nei-manhpi.org/commentary/how-many-radiologists-it-depends-on-who-you-ask/>. Accessed: 6 Jan 2017.
8. Garland LH. Studies on the accuracy of diagnostic procedures. *Am J Roentgenol Radium Ther Nuc Med* 1959;82:25–38.
9. Kundel HL. Perception errors in chest radiography. *Semin Respir Med* 1989;10:203–10.
10. Berlin L. Radiologic errors, past, present and future. *Diagnosis* 2014;1:79–84.
11. Berlin L. Accuracy of diagnostic procedures: has it improved over the past five decades? *AJR Am J Roentgenol* 2007;188:1173–8.
12. Pinto A, Brunese L. Spectrum of diagnostic errors in radiology. *World J Radiol* 2010;2:377–83.
13. Bhargavan M, Kaye AH, Forman HP, Sunshine JH. Workload of radiologists in United States in 2006–2007 and Trends since 1991–1992. *Radiology* 2009;252:458–67.
14. Mettler F, Bhargavan M, Faulkner K, Gilley DB, Gray JE, Ibbott GS, et al. Radiologic and nuclear medicine studies in the United States and worldwide: frequency, radiation dose, and comparison with other radiation sources – 1950 to 2007. *Radiology* 2009;253:520–31.
15. Kim YW, Mansfield LT. Fool me twice: delayed diagnosis in radiology with emphasis on perpetuated errors. *AJR Am J Roentgenol* 2014;202:465–70.
16. Funaki B, Szynski G, Rosenblum JD. Significant on-call misses by radiology residents interpreting CT studies: perception vs. cognition. *Emerg Radiol* 1997;4:290–4.
17. Samei E, Krupinski E. Medical Image perception. In: Samei E, Krupinski E, editors. *The Handbook of medical image perception and techniques*. Cambridge, England: Cambridge University Press, 2010.
18. Donald JJ, Barnard SA. Common patterns in 558 diagnostic radiology errors. *J Med Imaging Radiat Oncol* 2012;56:173–8.
19. Quekel LG, Kessels AG, Goel R, van Engelshoven JM. Miss rate of lung cancer on the chest radiograph in clinical practice. *Chest* 1999;115:720–4.
20. Fleck MS, Samei E, Mitroff SR. Generalized 'satisfaction of search': adverse influences on dual-target search accuracy. *J Exper Psychol Appl* 2019;16:60–71.
21. Bernbaum KS, Schartz M, Caldwell RT, Madsen MT, Thompson BH, Mullan BF, et al. Satisfaction of search from detection of pulmonary nodules in CT of the chest. *Acad Radiol* 2013; 20:194–6.
22. Ashman CJ, Uy JS, Wolfman D. Satisfaction of search in osteoradiology. *AJR Am J Roentgenol* 2000;175:541–4.
23. Bernbaum KS, Krupinski EA, Schartz KM, Caldwell RT, Madsen MT, Hur S, et al. Satisfaction of search in chest radiography. *Acad Radiol* 2015;11:1457–66.
24. Evans KK, Haygood TM, Cooper J, Culpan AM, Wolfe JM. A half-second glimpse often lets radiologists identify breast cancer cases even when viewing the mammogram of the opposite breast. *Proc Natl Acad Sci USA* 2016;113:10292–7.
25. Kundel HL, Nodine CF. Interpreting chest radiographs without visual search. *Radiology* 1975;116:527–32.
26. Krupinski EA. Current perspectives in medical image perception. *Atten Percept Psychophys* 2010;72:1205–17.
27. Mallett S, Phillips P, Fanshawe TR, Helbren E, Boone D, Gale A, et al. Tracking eye gaze during interpretation of endoluminal

- three-dimensional CT colonography: visual perception of experienced and inexperienced readers. *Radiology* 2014;273:783–92.
28. Firestone S. *Failure: why science is so successful*. New York: Oxford University Press, 2016.
 29. Benson B. Cognitive bias cheat sheet. Better Humans blog. <http://betterhumans.coach.me/cognitive-bias-cheat-sheet-55a472476b18#.s2j8qandz>. Accessed: January 2017.
 30. “Cognitive Bias List,” https://en.wikipedia.org/wiki/List_of_cognitive_biases Wikipedia. Accessed: May 2017.
 31. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Int Med* 2005;165:1493–9.
 32. Crosskerry P, Singhal G, Mamede S. Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ Qual Safety* 2013;22(Suppl 2):ii58–64.
 33. Crosskerry P, Singhal G, Mamede S. Cognitive debiasing 2: impediments to and strategies for change. *BMJ Qual Safety* 2013;22ii:65–72.
 34. Crosskerry P. Clinical cognition and diagnostic error: applications of a dual-process model of reasoning. *Adv Health Sci Educ Theory Pract* 2009;14(Suppl 1):27–35.
 35. Lambe K, O’Reilly G, Kelly B, Curristan S. Dual-process cognitive interventions to enhance diagnostic reasoning: a systematic review. *BMJ Qual Safety* 2016;25:808–20.
 36. Dror I. A novel approach to minimize error in the medical domain: cognitive neuroscientific insights into training. *Medical Teacher* 2011;33:34–8.
 37. Graber ML, Kissam S, Payne VL, Meyer AN, Sorensen A, Lenfestey N, et al. Cognitive interventions to reduce diagnostic error: a narrative review. *BMJ Qual Safety* 2012;32:535–57.
 38. Fenton JJ, Taplin SH, Carney PA, Abraham L, Sickles EA, D’Orsi C, et al. Influence of computer-aided detection on performance of screening mammography. *N Engl J Med* 2007;356:1399–409.
 39. Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL, et al. “Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med* 2015;175:1828–37.
 40. Castellino RA. Computer aided detection (CAD): an overview. *Cancer Imaging* 2005;5:17–9.
 41. Bailey R, Mcnamara A, Sudarsanam N, Grimm, C. Subtle gaze direction. *ACM Trans Graph* 2001;3:1–2.
 42. Latif N, Gehmacher A, Castelhana MS, Munhall KG. The art of gaze guidance. *J Exp Psychol Human Percept Performance* 2014;40:33–9.
 43. Rohatgi S, Hanna TN, Sliker CW, Abbott RM, Nicola R. After-hours radiology: challenges and strategies for the radiologist. *AJR Am J Roentgenol* 2015;205:956–61.
 44. Krupinski EA, Bernbaum KS, Caldwell RT, Scharz KM, Kim J. Long radiology workdays reduce detection and accommodation accuracy. *J Am Coll Radiol* 2010;7:698–704.
 45. Larson DB, Donnelly LF, Podberesky DJ, Podberesky DJ, Merrow AC, Sharpe RE Jr, et al. Peer feedback, learning and improvement: answering the call of the institute of medicine report on diagnostic error. *Radiology* 2917;283:231–41.
 46. Kohn LT, Corrigan JM, Donaldson MS, editors. *Committee on Quality of Health Care in America, Institute of Medicine. To Err is Human: building a safer health system*. Washington, DC: National Academy Press, 2000.
 47. Abujudeh HH. Just culture.’ Is radiology ready? *J Am Coll Radiol* 2015;12:4–5.
 48. Garland LH. On the scientific evaluation of diagnostic procedures. *Radiology* 1949;52:309–28.
 49. Onega T, Aiello-Bowles EJ, Miglioretti DL, Carney PA, Geller BM, Yankaskas BC, et al. Radiologists’ perceptions of computer-aided detection vs. double-reading for mammography interpretation. *Acta Radiol* 2010;17:1217–26.