Radiology

Mandating Limits on Workload, Duty, and Speed in Radiology

Robert Alexander, PhD • Stephen Waite, MD* • Michael A. Bruno, MD, MS • Elizabeth A. Krupinski, PhD • Leonard Berlin, MD • Stephen Macknik, PhD • Susana Martinez-Conde, PhD*

From the Departments of Ophthalmology (R.A., S.M., S.M.C.), Radiology (S.W.), Neurology (S.M., S.M.C.), and Physiology & Pharmacology (S.M., S.M.C.), SUNY Downstate Health Sciences University, 450 Clarkson Ave, Brooklyn, NY 11203; Department of Radiology, Penn State Milton S. Hershey Medical Center, Hershey, Pa (M.A.B.); Department of Radiology and Imaging Sciences, Emory University, Atlanta, Ga (E.A.K.); and Department of Radiology, Rush University Medical College and University of Illinois, Chicago, Ill (L.B.). Received October 25, 2021; revision requested November 29; revision received January 30, 2022; accepted February 3. **Address correspondence to** S.W. (email: *Stephen.Waite@downstate.edu*).

Study supported by the New York State Empire Innovator Program.S.M.C. and S.L.M. supported by the National Science Foundation (grant number 1734887) and the National Institutes of Health (grant number R01EY031971); and S.M.C., S.L.M., and S.W. supported by the National Institutes of Health (grant number R01CA258021).

* R.A. and S.W. contributed equally to this work.

Conflicts of interest are listed at the end of this article.

Radiology 2022; 304:274–282 • https://doi.org/10.1148/radiol.212631 • Content codes: ED SQ

Research has not yet quantified the effects of workload or duty hours on the accuracy of radiologists. With the exception of a brief reduction in imaging studies during the 2020 peak of the COVID-19 pandemic, the workload of radiologists in the United States has seen relentless growth in recent years. One concern is that this increased demand could lead to reduced accuracy. Behavioral studies in species ranging from insects to humans have shown that decision speed is inversely correlated to decision accuracy. A potential solution is to institute workload and duty limits to optimize radiologist performance and patient safety. The concern, however, is that any prescribed mandated limits would be arbitrary and thus no more advantageous than allowing radiologists to self-regulate. Specific studies have been proposed to determine whether limits reduce error, and if so, to provide a principled basis for such limits. This could determine the precise susceptibility of individual radiologists to medical error as a function of speed during image viewing, the maximum number of studies that could be read during a work shift, and the appropriate shift duration as a function of time of day. Before principled recommendations for restrictions are made, however, it is important to understand how radiologists function both optimally and at the margins of adequate performance. This study examines the relationship between interpretation speed and error rates in radiology, the potential influence of artificial intelligence on reading speed and error rates, and the possible outcomes of imposed limits on both caseload and duty hours. This review concludes that the scientific evidence needed to make meaningful rules is lacking and notes that regulating workloads without scientific principles can be more harmful than not regulating at all.

Faulty detection—the failure to identify salient findings—is the most consequential source of interpretive error in radiology (1–4), as well as the most common reason for malpractice lawsuits initiated against radiologists, underlying 78% of cases (5). So-called reckless reading lawsuits, which allege that missed findings resulted from insufficient time spent viewing images, have become more common (6). One often-discussed malpractice case from 2020 alleged that a subdural hematoma in a 64-year-old man was missed because the radiologist spent a total of 6 minutes and 27 seconds reading a CT examination of the head and cervical spine (as determined by subpoenaed keystroke data). The plaintiff's attorney argued that the average viewing time of each image was one-half second per image, and therefore the radiologist was lax (ie, not sufficiently careful), thus engendering a larger settlement (7–9).

Is it medical negligence to interpret an image quickly? If so, then it follows that a radiologist should spend a certain minimal amount of time reviewing each image from a patient's study, and that anything less would be negligent (ie, below the standard of care) (10). Legally, this could mean that every patient requires a specific reading process, regardless of the case. Such potential legal outcomes have prompted extensive discussions regarding how a radiologist's speed and workload relates to accuracy (11–14). One recent article (12) advocated implementing limits on both shift duration and caseload. Radiologists should operate under the principle "first do no harm" (12); however, we do not agree that the available evidence supports mandated work limits in radiology.

That is not to say there is no minimum necessary viewing time to render a meaningful diagnosis; however, to our knowledge, nobody has yet determined the minimum. We argue that any minimum should be established through the principles of data-driven visual processing research, rather than through an arbitrary administrative fiat or legal process. Studies may find that viewing some images for many minutes is not sufficient to identify particularly difficult anomalies, whereas other images may contain such obvious anomalies that a viewing time of even a fraction of a second is sufficient for accurate determination. Therefore, it is possible that the range of appropriate viewing times is so wide that any meaningful application of limits is meaningless.

Although it is possible that long and off-hours work shifts may also lead to poor performance and serve as a source of medical error, to our knowledge, no research to date has provided the necessary evidence to set work duration or shift limits for any individual radiologist. Before any duration or workload limits are enforced, it is important to acknowledge that any significant change to the medical system would require careful design supported by principled research findings to be practical, sustainable, and not inadvertently cause patient harm.

Abbreviations

AI = artificial intelligence, RVU = relative value unit, 2D = two-dimensional, 3D = three-dimensional

Summary

Although it is almost certain that increased workloads, faster speeds, and prolonged shifts will at some point decrease interpretive accuracy, there is currently insufficient evidence to determine appropriate workload, speed, or duty limits.

Essentials

- So-called reckless reading lawsuits, which allege that missed findings result from insufficient time spent viewing images, have become more common.
- Scene processing in less than a second does not indicate carelessness or recklessness but, in some settings, can be characteristic of competent or expert visual processing.
- Evidence indicates that reading each image of a cross-sectional examination in less than a second defines the current standard of care in radiology.
- Although long shifts and so-called off-hours work may sometimes lead to poor performance and serve as a source of medical error, to the authors' knowledge, no research to date has provided the evidence necessary to establish appropriate limits for individual radiologists.

Our review aims to discuss policies that are arbitrary and possibly harmful, especially those driven by legal argument rather than scientific research. We review the evidence for the relationship between speed and accuracy in radiology as it relates to the potential effects of imposed limits on caseloads, work shift durations, and variations in performance as a function of time of day. We also review the potential effects of artificial intelligence (AI) on speed, caseload, and fatigue. Finally, we outline research studies that must be conducted before principled policy making.

The Science of Radiologists' Perceptual Expertise

Like any other perceptual skill, the ability to detect radiologic abnormalities improves with training: radiologists become faster and more accurate as they gain experience (15–18) (see Alexander et al [1] for a review). According to most models of radiologic search, expert radiologists know how normal images should look because of repeated viewing of many varied examples and have learned to process information from the whole image in a rapid initial stage of processing. Obvious abnormalities (especially those closer to the point of gaze) are detected rapidly (19). In the context of controlled experiments, an experienced reader can detect whether radiographs or other static two-dimensional (2D) images are normal or abnormal, with above-chance accuracy, even when the images are presented for less than half a second (20–25). Radiologists can rapidly identify whether mammography in one breast belongs to a woman who has cancer in the opposite breast (26) and predict abnormal mammography years before localized signs of cancer are visible (27). This skill may also extend to three-dimensional (3D) images: Treviño et al (28) found that radiologists can rapidly detect abnormalities during volumetric interpretation. Radiologists were able to discriminate between normal and abnormal (ie, containing lesions) stacks of 26 T2-weighted images from prostate MRI at above-chance levels when presented as a brief movie, with image sections shown in as little as 48 msec per section. It is important to note that more interpretation time is needed in actual clinical practice because subtle abnormalities require more time to be detected and much better than above-chance performance is necessary. These studies, however, highlight the perceptual abilities of trained radiologists.

Fast and accurate visual processing is not unique to radiologists. In baseball, the time from the pitch until the bat strikes the ball can be less than 0.5 seconds (29). Despite this brief viewing time, batters not only decide whether to swing the bat, but also continuously adjust their swing in response to precise visual information about the trajectory of the ball. Expert batters discriminate between different types of pitches (eg, curve ball vs fastball) in under 200 msec (30,31). Similarly, car drivers reposition their feet from brake to gas pedal in under 1000 msec when responding to a red traffic signal turning green, not only detecting the visual change but also producing the necessary motor actions (32). Moreover, drivers can accurately identify upcoming hazards (eg, pedestrian stepping into road) in under 250 msec (33). Indeed, in many contexts, scene processing in less than a second does not indicate carelessness or recklessness. Rather, it is consistent with efficient and successful processing of stimuli, such as when experts perform a trained oculomotor task.

Of course, not all image reading can or should occur at a high speed. Subtle abnormalities may require direct visual fixation for identification. Localization may also require some additional processing: whereas very brief viewing times may allow an expert radiologist to detect the presence of an abnormality, some additional time may be necessary to identify the placement of that abnormality in the image (34). During typical image evaluation, there is a second stage of processing in which potentially critical image regions are foveated: the radiologist looks directly at one region at a time (thus directing their eye's fovea at those regions, as opposed to viewing the region by using their peripheral vision). This is accomplished by switching gaze from one region to the next by saccadic eye movements (35–39). The targets of saccadic eye movements are not random; expert radiologists may direct their gaze to clinically relevant regions that they initially identified with their peripheral vision, during the first pass review (1,40). One consistent finding is that expert radiologists find abnormalities faster than novice residents, perhaps partly because of requiring fewer eye movements to foveate potential abnormalities (41,42). In this case, because expert radiologists also make fewer errors than residents, enhanced speed is correlated with reduced medical error, demonstrating the lack of a monotonic relationship.

Once a radiologist looks directly at an abnormality, the radiologist must keep looking at it long enough to recognize its features (eg, at least 500–1000 msec depending on the modality) or identification may fail (43,44). Many abnormalities are foveated but never reported, perhaps because of insufficient viewing times or deciding incorrectly that the detected features do not represent an abnormality, although other factors may be involved (4,45).

Caveats of Volumetric Imaging

Although numerous studies have examined search strategies with 2D medical images, relatively little is known about search strategies in 3D volumetric images such as CT and MRI (46). During stack mode viewing, radiologists simulate motion by scrolling through sequential images and searching for lesions that suddenly stand out from the background (46) without stopping to look at each individual image. This is similar to how one views a video clip or movie (9).

Thus, the fundamental characteristics of 2D search are qualitatively different from those in volumetric search, and the temporal image viewing parameters may differ between 2D and 3D image interpretation. The vast amount of 3D data that radiologists must scrutinize effectively prevents exhaustive foveation of each image region within a CT stack (47,48). Because of the way 3D image stacks are searched—individual structures can be followed as if moving while scrolling axially up or down between image planes—some image regions may only be seen peripherally. In CT and MRI search, radiologists often do not directly foveate on much of the examination. As such, peripheral vision is especially important. Unlike 2D imaging, the entire 3D image set is never visible at one time. Consequently, it is not possible to derive a complete perceptual gist that informs simultaneous comparisons of image perturbations throughout an entire 3D image stack.

Because the dependence on peripheral vision may differ between 2D and 3D image viewing, the detectability of certain lesions on 2D and 3D images may also differ (49). To our knowledge, it is not known whether search expertise generalizes from 2D to 3D. This is important because, to our knowledge, there is essentially no data suggesting minimum interpretation times for 3D volumetric scans, either on the level of the entire study or for a specific image.

Drew et al (42) identified two different global strategies that radiologists adopt during nodule detection tasks on chest CT images. So-called scanners search each section before moving to the next depth. So-called drillers hold their eyes relatively still in the x and y planes, limiting search to a single lung quadrant while quickly scrolling through sections in the z-axis (42,50). It is not known if one strategy is universally better than the other, or if they each have different dependencies or specific advantages in specific subspecialties. In real life, any optimal strategy for 3D imaging interpretation is likely to be modality and region specific (51). Rather than demonstrating a clear preference for drilling versus scanning (42), radiologists both drill and scan during interpretation of digital breast tomosynthesis (52). In addition, strategies may change during image interpretation. For example, drilling may be ideal for pulmonary nodule detection but not for examination of the mediastinum on the same study.

Researchers have examined scrolling data obtained during the interpretation of cross-sectional imaging to understand the development of radiologic expertise in volumetric imaging (53,54). van Montfort et al (55) found that during a 5-year residency period, radiology trainees decreased the percentage of time spent on full runs (scrolling through more than 50% of CT scan sections) and increased the percentage of time spent on task-relevant areas (sections with the abnormality present). These results are consistent with visual expertise theories suggesting that greater expertise affords residents the ability to form a global impression of a study more quickly, allowing them to strategically ignore irrelevant areas. However, it should be noted that neither the percentage of time spent on full runs nor on relevant sections predicted diagnostic accuracy (55).

In short, our current understanding of the relationship between search patterns, speed, and accuracy during volumetric interpretation is rudimentary. Without a refined understanding of how radiologists develop expertise during volumetric interpretation, how search patterns relate to performance, and how these patterns differ from searches of 2D images, it is not principled to conclude that image viewing durations of any particular duration are substandard.

Minimal Interpretation Time per Image

The so-called lax speed per image allegation referenced in the 2020 litigation example (9) has a number of logical flaws. Noting them is important.

Imaging studies usually have multiple series such as axial, coronal, and sagittal series in addition to series with dedicated window-level settings (56) and convolution algorithms and/or kernels (eg, bone and soft tissue) (57). These series are routinely provided to the interpreting radiologist without consideration of their usefulness in answering the posited clinical question. Although we are not aware of any specific studies, experience suggests that many of these series are often ignored and only viewed for problem solving.

In addition, the raw data obtained during volumetric imaging can be reconstructed to images of varying thicknesses, from submillimeter to 10 mm (58). To use the number of images reviewed per unit time as a metric of thoroughness would imply that a radiologist looking at series of 1-mm-thick sections should spend five times longer than a radiologist looking at series of 5-mm-thick sections. There is, however, no logical reason or data to assert that this practice would increase accuracy.

The fact that many of the provided series are ignored, and the unprincipled premise that thinner sections should be evaluated proportionately longer than thick sections, are important theoretical problems related to the use of sections per unit time as a meaningful metric.

In addition, there is reason to believe that reading images from a cross-sectional study in less than a second is the current standard of care. A quantitative understanding of the workload of modern radiologists would therefore be important to review.

The resource-based relative value scale in use today was adopted for the Medicare payment system in 1992. Relative value units (RVUs) are the basic components of the scale that describes and quantifies the work and resource costs needed to provide physician services. Therefore, RVU form the basis of physician fees by Medicare and other payers and are often used to measure physician productivity (59,60).

According to Muroff and Berlin (13), private practice radiologists generate 13 000–15 000 RVUs per year. Given 261 working (nonweekend) days in 2021 and assuming 8 weeks of vacation, radiologists would have to generate 63 RVUs per day to achieve 14000 RVUs per year. Assuming that a CT scan of the abdomen and pelvis generates 1.82 RVUs (61), this amounts to approximately 34 scans per day. In 2010, McDonald et al (62) reported that the average CT examination contained 679 images. By using this figure for the average number of images per examination and a 7-hour workday (assuming a standard 8-hour workday including a 1-hour lunch break), this leaves 1.1 seconds for radiologists to look at each image. This presumes no breaks, interruptions, consultations, or conferences; this is an unrealistic scenario. Yu et al (63) recently reported that on-call radiologists receive an average of 72 phone calls during a typical 12-hour overnight shift, with an average total phone time of 108 minutes. If we allow 90 minutes for interruptions, breaks, consultations, and conferences, then radiologists are left with less than 1 second (0.86 sec) to read each image. Many studies (eg, CT angiography) generate more than 1000 images per study, further decreasing reading time per section. Modern imaging habitually requires radiologists to evaluate thousands of images during cross-sectional interpretation (64).

We acknowledge that these calculations rely on estimations and assumptions; however, the available evidence indicates that reading an image in a cross-sectional study in less than a second on average is the current standard of care in radiology, so implying that such behavior is negligent fails logic.

It is important to note that the American College of Radiology does not currently have a practice parameter that addresses minimum interpretation speed per image (9). This omission is warranted given the current state of knowledge and, therefore, any standard or practice parameter in this regard would not be scientifically justified.

Will Caseload Reduction Decrease Error?

Although caseloads vary widely across radiology practices (65–69), the numbers of studies and images that radiologists are required to interpret have been generally increasing. In one study, the number of images requiring interpretation each minute of every workday for staff radiologists grew from 2.9 in 1999 to 16.1 in 2010, an increase of over 550% (62). An exception is the overall workload reduction that occurred in 2020 during the COVID-19 pandemic (70).

Unfortunately, high caseloads have been found to be associated with increased interpretive error (71,72). Specifically, it is often assumed that increasing case volume directly results in less reading time per study, thus increasing error (12). In some cases of extreme time pressure, this is unequivocally true: reduced viewing time can lead to error (20). Cancers are more likely to be missed on chest radiographs with viewing times of 1 second or less versus 4 seconds or more (20). However, this pattern is not linear: In some experimental settings, performance with 4 seconds is similar to unlimited viewing time (20). A further nonlinearity relates to visibility of the specific disease. Obvious cancers on radiographs are noticed almost all of the time, even when viewed for only 250 msec, whereas subtle cancers are sometimes missed even with unlimited viewing time (20). Therefore, the available evidence suggests that appropriate reading times for plain film imaging may range from 250–4000 msec per image (where 4000 msec is equivalent to unlimited viewing time). Such a mandated range would serve no useful purpose, especially because data obtained in conditions of extreme time pressure are derived from laboratory experiments that do not reflect realistic clinical reading scenarios.

Sokolovskaya et al (73) found that when attending radiologists read studies twice as quickly as their own baseline (based on self-reported time), their rate of major misses increased by 166% (3.2 vs 1.2 average misses in 12 studies). However, as others have noted (74), this study tested only five radiologists (including one who had fewer misses at the faster speed), and it is unclear whether the findings might be replicated at a larger scale (75). Because changes in performance were relative to individual baselines, it is sensible that any attempts to establish viewing-time mandates should account for individual differences.

Hanna et al (71) found that shifts with errors (defined as discrepancies between preliminary and final reports that affect patient care) had an average of 13 examinations per hour ± 6.1 (SD), whereas shifts without errors had an average of 11 examinations per hour ± 6.8, suggesting that an approximately 16% reduction in reading speed decreased error. Moreover, in shifts where at least one error was made, higher error rates were associated with larger volumes of examinations (71). However, this study did not establish the precise conditions that were more likely to result in error. For instance, examination volume and reading speed might be linked in aggregate across a practice, although not necessarily involving each individual radiologist.

As in other search tasks (36,39), the speed of radiologic search and the accuracy of interpretation may vary widely with task complexity (ie, the difficulty in accurately perceiving or interpreting the image) (76). Artificially slowing radiologists' natural reading speed may reduce patient access to radiologic analysis (because radiologists will necessarily read fewer images per shift), but it is not clear that it would consistently decrease error rates. In a laboratory setting, Wolfe et al (77) found that slowing observer responses in a nonradiologic low-prevalence target task had no apparent effect on error rates. Instead, artificially slowing down radiologists' reading times could cause them to second-guess correct findings or seek new and incorrect interpretations, thus leading to new errors (19). Christensen et al (19) found that residents' observations toward the end of an image search were more likely to result in false-positive findings than in true-positive findings.

Would Shift Duration Limits Reduce Fatigue and Error?

Fatigue, weariness, and depleted mental energy is prevalent in radiology. Roughly half of radiologists report at least some degree of fatigue or burnout (78), and that fatigue and other aspects of burnout are increasing over time (79). Although fatigue often coexists with sleepiness, only half of surveyed radiologists report never or rarely being asked to read images when sleep deprived (79). Instead, 36.0% report doing this sometimes, 13.5% report doing this frequently, and 1.9% report doing this always.

Although increased shift duration increases fatigue, which can affect performance, one study (80) found that simulated surgical tasks were performed at a comparable level before and after trauma residents worked 24-hour shifts. This is despite fatigue being both established at a physiologic level and reported subjectively by the participants. This research did not, however, assess analytical abilities such as those used in radiologic interpretation. Regarding radiologic tasks, one study from 1977—when practices differed substantially from today—by Christensen et al (81) found no impact on performance after 15-hour workdays, though the results were confounded with differences in experience. In 2010, Krupinski et al (82) investigated the effect of fatigue on detection of easy- versus hard-to-detect bone fractures on plain film images. The authors found that after a day of diagnostic interpretation, readers had several issues compared with before the onset of diagnostic reading: asthenopia (induced myopia or nearsightedness because of long hours of reading images at close distances on computer monitors), more subjective fatigue, and more visual strain. Moreover, detection accuracy was lower for image reading in the late versus early parts of the shift. Subsequently, the authors investigated fatigue in CT scan interpretation during nodule detection and found that after a day of reading, radiologists reported increased visual strain and exhibited lower accuracy (83). In a study of almost 3 million cases from a teleradiology practice, Hanna et al (71) found that errors were most frequent around 9 hours into a shift. In a simulation-based study of critical care radiology, Sistrom et al (84) found decreasing resident performance throughout an 8-hour shift, reaching significance at 6 hours.

Even if fatigue (see Waite et al [85] for review) is a likely source of error in radiology, its effects may be ameliorated without arbitrary time restrictions (86). For example, hourly breaks decrease eye strain in radiologists (87). Other measures to reduce fatigue focus on optimizing the ergonomic design of the reading environment, such as by increasing ambient lighting and eliminating glare (86,88). Time dedicated to mentoring, practice building, continuing medical education, and reading physical journals may also help decrease error by providing not only a break from image viewing, but also by enhancing task-relevant knowledge and capabilities. Ensuring that other colleagues are present at the end of long shifts could reduce error by enabling fatigued radiologists to obtain consultations and second readings.

Off-hour Shifts

Although hospital-based radiology is a 24-hours-per-day, 7-days-per-week endeavor, work performed in the evening, overnight, weekends, and holidays is collectively considered off-hours (89).

Patel et al (89) found that most board-eligible and certified fellows made more interpretation errors in body CT examinations at night than during the day, with the highest error rate occurring in the second half of the night shift. Importantly, these work assignments were well within Accreditation Council for Graduate Medical Education guidelines to mitigate fatigue (including duty hour standards, requirements for educating residents and faculty about recognizing and responding to signs of fatigue and sleep deprivation, and programs to adopt fatigue mitigation strategies such as naps). This diminished diagnostic performance, despite fatigue mitigation efforts and relatively light caseloads (found in night vs daytime shifts), led the authors to suggest that circadian misalignment may have been a contributor (89). This study suggests that radiologists of all levels, and not only trainees, are susceptible to increased errors and diminished performance when working off hours (90).

These findings notwithstanding, Hanna et al (91) noted that residents exhibited decreased diagnostic discrepancies with increased consecutive shifts. This suggests that trainees may either acclimate to night work schedules, improve in accuracy because of enhanced perceptual learning (practice), or both. In addition, training during off-hour shifts may have potential benefits to the robustness of clinical performance in difficult circumstances. We propose that future research studies disentangle and measure the respective and possibly conflicting contributions of fatigue and perceptual learning to radiologic performance.

Individual Variability in Performance

Most of the studies described were primarily concerned with average group performance with respect to accuracy and interpretation times. However, both intra- and interradiologist performance can be highly variable (72). Some radiologists may be fast without compromising performance and some that are slower may not be more accurate (13). Indeed, the number of eye movements made, the locations they target, and visual attention deployment are highly variable among experts. Wen et al (92) found that certain saliency models performed better than others regarding how well they agreed with individual radiologists' eye positions during interpretation of chest radiography, CT, and PET scans. This implies that different radiologists may rely on different kinds of image information (eg, intensity, orientation, edges) during a visual search. If so, then there may be more than one optimal image analysis strategy. Research is required to determine the relative advantages of different strategic approaches to visual search with different imaging modalities and task conditions.

In addition, alignment between radiology subspecialty and case mix (eg, thoracic vs neuroradiologists reading chest studies) may be important in determining a radiologist's optimal interpretive speed.

There is ongoing research to develop educational and practical interventions to enhance radiologists' perceptual and decision-making skills (1–3,93). Importantly, any mandated limits may be rendered inappropriate whenever new training is adopted, as accurate performance may be achieved more quickly than before the training, or if new methods require different approaches to managing daily workloads.

Other Sources of Variability

Other sources of variability, such as practice setting, can play a role in interpretive error rates. Is the radiologist reading studies from an outpatient center with fewer sick patients than a tertiary cancer center? Are they reading complex multitrauma cases? These will be central questions to address before setting limits on practice. Further, environmental distractions (eg, interruptions or presence of trainees) may also play a role (94–96). There are too many unknowns to define, justify, or defend work-duration limits. Without consideration of variation among radiologists and subspecialties, it may be impossible to establish principled guidelines for what image viewing durations optimize accuracy.

Results of Mandated Resident Duty Hour Limits

The Libby Zion malpractice case (97) provided the impetus to reform resident work hours and supervision. A state commission, headed by Bertrand Bell, MD, was formed to address systemic problems in residency training (98), and in 1989, New York State 405 (Bell Commission) Workforce Regulations were enacted limiting residents to 80-hour work weeks (averaged over a 4-week period) and on-call shifts to no more often than every 3rd night. On July 1, 2003, the Accreditation Council for Graduate Medical Education, or ACGME, adopted resident duty hour standards for all ACGME-accredited residency programs (99).

The motivation behind regulating work hours was the growing recognition that sleep deprivation can result in poorer resident performance. The expectation was that limits would have a positive effect on patient care outcomes and resident quality-of-life measures (100,101). However, despite the extensive scientific evidence linking fatigue and impaired cognitive performance, little empirical data were used to guide the design of duty hour regulations (101). Indeed, in a letter to the *Journal of the American Medical Association* in 2007, Bell (102) reported that the 80-hour rule was arbitrarily developed "on my porch" by using "informal reasoning."

Although shortened work hours for residents improved their quality of life, encouraged better sleep, and caused less fatigue, a meta-analysis of duty hour restrictions did not demonstrate a uniform benefit to patient safety (100,101). Although any specific effects on radiology trainees are unknown, more broadly, critics have suggested that duty hour restrictions result in less continuity of coverage and abridged clinical exposure, resulting in impaired physician training and patient care (103).

Practical Ramifications of Workload, Speed, and Duty Hour Restrictions

Some have suggested shift limits of 8–10 hours (12) to ameliorate fatigue. The practical negative ramifications of any workflow or duty hours restrictions should not be underestimated, however. Even in situations where radiologists are not monetarily incentivized to read more studies, appropriate patient care in modern practice mandates a certain level of productivity. Interpretation limits and shift length limits are mutually inconsistent. Provided a certain number of cases, if radiologists slow their reading times, their workdays will necessarily lengthen. Studies can be left unread, but that is not a practical alternative. How unprincipled rules may affect the ability of radiologists to manage clinical workload and the impact on patient care are important considerations.

Attending-level radiologists working off-hours make more errors during night than during day assignments (89). The logical solution to this problem would therefore be to either not read studies overnight or, as Bruno suggested, to employ double reading in the "cool light of morning by a fresh radiologist" (90). In a systemic review of double reading, Geijer and Geijer (104) found that the rate of discrepancy ranged from insignificant to over 22% depending on the study setting. In particular, double reading by a subspecialist often led to high rates of changed reports. Unfortunately, double reading in the United States, despite its long-recognized benefits in reducing interpretative error (105), is not routinely practiced because it is time-consuming, requires additional manpower, and the second read is not reimbursed (106). Double-reading as a routine strategy will require an economic shift in our medical system to absorb increased expense and workload (because radiologists would have to read both daytime cases and cases from the previous night).

Some centers have reported successful use of limited, or targeted, double-reading of certain high-risk types of radiology studies, despite the high cost. Whereas two radiologists would be equally subject to perceptual error, it is unlikely that both readers, working independently, would miss the same abnormality, assuming such errors are random. A strategy of delayed double reading is not optimal in all settings and will not solve the problem of delayed care when an overnight error that affects patient treatment is not recognized until hours later.

Finally, it is important to note that 10 of the 32 fellows (31%) in the study by Patel et al (89) had fewer errors at night than during the day, indicating that consideration of individual differences may be an optimal approach.

Potential Roles of AI

AI based on machine learning and paired with computer vision technology has the potential to serve as a second reader in real time. AI and machine learning algorithms that are either in development phase or currently available may be sufficiently accurate at detecting abnormalities to augment human radiologists, thereby providing a safety net that improves accuracy (107).

Indeed, some studies (108) surmised that convolutional neural networks, which can produce a type of machine learning called deep learning, might help radiologists overcome perceptual or cognitive biases and other human limitations such as fatigue. Coppola et al (109) suggested that AI could "alleviate radiologists' traditional work burden," reducing the impact of increasing caseloads by offering "…new tools for quantitative analysis and image interpretation…saving time and effort during fatiguing and/or repetitive tasks." For example, Lexa and Jha (107) suggested that AI could do some of radiologists' "mundane tasks of daily labor such as measuring lymph nodes and lung nodules." In addition, AI might be used to improve the training of radiologists—that is, AI-empowered education—by personalizing learning to maximize expertise acquisition, leading to improvements in radiologists' accuracy (2,110). To our knowledge, however, no definitive proof exists that the use of AI directly reduces fatigue, and its effect on the caseloads of radiologists is unclear.

In a point-counterpoint series, Lexa and Jha (107) discussed a hypothetical scenario in which AI can do 50% of the work of radiologists. In this scenario, they note that to minimize costs, corporate and other managed environments are likely to reason that they need fewer radiologists—perhaps more than half as many, but certainly fewer than before AI. The notion that AI can do the work of radiologists might therefore lead to an ironic scenario in which radiologists with AI support have increased caseloads secondary to reduction of the employed radiologist workforce.

It is important to note that although there has already been considerable research and development devoted to AI and machine learning image classifier systems in radiology, progress has been slower than anticipated. Specific and narrow applications for AI have achieved performance levels comparable with those of humans (111). More importantly, studies that examine the impact of AI tools on radiologists' decisions and the ultimate effects of AI tools on patient care and outcomes are lacking. Early research assessing AI tools seems to parallel many of the studies conducted previously with computer-aided detection and computer-aided diagnosis tools. Therefore, the impact of AI tools on reader performance may vary because of a host of variables, including image type, disease type and severity, reader experience, and even the way in which the computer-aided detection (not AI) prompts are presented to the observer.

Considerable research and development efforts regarding AI and machine learning tools are underway and this technology is promising. Combining AI and radiologist assessment can improve accuracy compared with human interpretation alone and is a feasible solution to directly addressing errors of interpretation (112), including errors from fatigue, high caseloads, and off-hours shifts.

Future Studies

From a fundamental science perspective, further studies of how radiologist performance changes with different workloads, expertise, fatigue, and time of day are crucial to understand variability between and within readers. Does a radiologist work at the same speed and level of accuracy late on a Friday afternoon and early on a Monday morning, after a weekend off or after coming back from a week of vacation? The issue of assessing radiologists' productivity is not simple (113,114).

Research can help measure and optimize the accuracy of individual radiologists in several ways. One approach is to derive utility curves of the cost-benefit tradeoff for accuracy versus reading speed for each radiologist, in combination with a battery of measures of oculomotor and decision-making performance. These results could be used to optimize caseload and case type for individual radiologists. This research will be needed at a considerable scale to assess the relationship between speed and accuracy for large numbers of radiologists across multiple practices and specialties before any evidence-based recommendations can be made regarding maximum caseload or minimum viewing times.

It is important to note that depending on the experimental design, any conclusions drawn may be context specific. For example, experimentally derived workload and duty limits may not generalize to all situations (115). Reasonable standards will need to be established separately for different image modalities (high- vs low-complexity images) and clinical contexts (eg, isolated teleradiology vs in-person reading rooms with other radiologists).

Other studies could focus on peripheral factors that impact performance. For example, research might also be aimed at improving the environment in which radiologists perform their tasks. Environmental distractions, such as interruptions, can decrease radiologists' accuracy (94–96), and future studies could test potential interventions that may reduce these distractions (or otherwise minimize their effects on radiologists).

Conclusion

Whether examined at a macro scale (number of studies per day) (6) or a micro scale (images per unit time), what was true more than 20 years ago remains true today: It is unknown how many examinations or images radiologists can review in any period while maintaining accuracy. Whereas we agree that regulation may ultimately be required, arbitrary regulations that have no scientific basis are potentially more harmful than not regulating at all. Making rules without reliable scientific evidence is unprincipled and may fail to address the underlying problem (eg, the 80-hour resident work-hour rule has not resulted in decreased medical error), but it can create additional unforeseen problems such as an unacceptable backlog of unread images. Unprincipled regulations can worsen performance for radiologists who perform at their peak while near the margins of normal performance parameters, resulting in inadvertent exacerbation of medical error and compromised patient care.

Disclosures of conflicts of interest: **R.A.** No relevant relationships. **S.W.** No relevant relationships. **M.A.B.** Royalties from Oxford University Press and Elsevier. **E.A.K.** No relevant relationships. **L.B.** No relevant relationships. **S.M.** No relevant relationships. **S.M.C.** No relevant relationships.

References

1. Alexander RG, Waite S, Macknik SL, Martinez-Conde S. What do radiologists look for? Advances and limitations of perceptual learning in radiologic search. J Vis 2020;20(10):17.
2. Waite S, Farooq Z, Grigorian A, et al. A Review of Perceptual Expertise in Radiology-How it develops, How we can test it, and Why humans still matter in the era of Artificial Intelligence. Acad Radiol 2020;27(1):26–38.
3. Waite S, Grigorian A, Alexander RG, et al. Analysis of Perceptual Expertise in Radiology - Current Knowledge and a New Perspective. Front Hum Neurosci 2019;13:213. [Published correction appears in Front Hum Neurosci 2019;13:272.]
4. Waite S, Scott J, Gale B, Fuchs T, Kolla S, Reede D. Interpretive Error in Radiology. AJR Am J Roentgenol 2017;208(4):739–749.
5. Baker SR, Patel RH, Yang L, Lelkes VM, Castro A 3rd. Malpractice suits in chest radiology: an evaluation of the histories of 8265 radiologists. J Thorac Imaging 2013;28(6):388–391.
6. Berlin L. Liability of interpreting too many radiographs. AJR Am J Roentgenol 2000;175(1):17–22.
7. Lean R. Pictures worth $2 million: South Florida lawyers leverage CT scan to negotiate settlement. South Florida Daily Business Review, 2020.
8. Stempniak M. 2M settlement after subpoena of radiologist's keystrokes finds lax CT reading. Radiology Business, 2020.
9. Raskin MM. The "Lax Radiologist": Educating a jury as to how scans are read is key to a juror's understanding of how radiologists practice. ACR Bulletin October 2020. https://www.acr.org/Practice-Management-Quality-Informatics/ACR-Bulletin/Articles/October-2020/The-Lax-Radiologist. Accessed July 19, 2021.
10. Waite S, Scott J, Kolla S, Bruno MA. The Role of the Expert Witness in Radiology: Challenges and Strategies for Overcoming Them. J Am Coll Radiol 2021;18(2):318–323.
11. Berlin L. Medicolegal-Malpractice and Ethical Issues in Radiology. Faster Radiologic Interpretation, Errors, and Malpractice: An Unavoidable Triad? AJR Am J Roentgenol 2018;210(2):W92–W93.
12. Lexa FJ. Duty hour limits for radiologists: It's about time. J Am Coll Radiol 2021;18(1 Pt B):208–210.
13. Muroff LR, Berlin L. Speed versus interpretation accuracy: current thoughts and literature review. AJR Am J Roentgenol 2019;213(3):490–492.
14. Muroff LR, Berlin L. Reply to "The Speed-Accuracy Trade-Off". AJR Am J Roentgenol 2019;213(6):W300.
15. Nodine CF, Kundel HL, Lauver SC, Toto LC. Nature of expertise in searching mammograms for breast masses. Acad Radiol 1996;3(12):1000–1006.
16. Krupinski EA. Visual scanning patterns of radiologists searching mammograms. Acad Radiol 1996;3(2):137–144.
17. Kundel HL, Nodine CF, Conant EF, Weinstein SP. Holistic component of image perception in mammogram interpretation: gaze-tracking study. Radiology 2007;242(2):396–402.

18. Wood G, Knapp KM, Rock B, Cousens C, Roobottom C, Wilson MR. Visual expertise in detecting and diagnosing skeletal fractures. Skeletal Radiol 2013;42(2):165–172.
19. Christensen EE, Murry RC, Holland K, Reynolds J, Landay MJ, Moore JG. The effect of search time on perception. Radiology 1981;138(2):361–365.
20. Oestmann JW, Greene R, Kushner DC, Bourgouin PM, Linetsky L, Llewellyn HJ. Lung lesions: correlation between viewing time and detection. Radiology 1988;166(2):451–453.
21. Evans KK, Georgian-Smith D, Tambouret R, Birdwell RL, Wolfe JM. The gist of the abnormal: above-chance medical decision making in the blink of an eye. Psychon Bull Rev 2013;20(6):1170–1175.
22. Drew T, Evans K, Võ ML, Jacobson FL, Wolfe JM. Informatics in radiology: what can you see in a single glance and how might this guide visual search in medical images? RadioGraphics 2013;33(1):263–274.
23. Chin M, et al. Gist Perception and Holistic Processing in Rapidly Presented Mammograms. J Vis 2018;18(10):391.
24. Gandomkar Z, Ekpo EU, Lewis SJ, et al. Does the strength of the gist signal predict the difficulty of breast cancer detection in usual presentation and reporting mechanisms? In: Nishikawa RM, Samuelson FW, eds. Proceedings of SPIE: medical imaging 2019—image perception, observer performance, and technology assessment. Vol 10952. Bellingham, Wash: International Society for Optics and Photonics, 2019; 1095203.
25. Kundel HL, Nodine CF. Interpreting chest radiographs without visual search. Radiology 1975;116(3):527–532.
26. 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 U S A 2016;113(37):10292–10297.
27. Evans KK, Culpan AM, Wolfe JM. Detecting the “gist” of breast cancer in mammograms three years before localized signs of cancer are visible. Br J Radiol 2019;92(1099):20190136.
28. Treviño M, Turkbey B, Wood BJ, et al. Rapid perceptual processing in two- and three-dimensional prostate images. J Med Imaging (Bellingham) 2020;7(2):022406.
29. Ijiri T, Shinya M, Nakazawa K. Interpersonal variability in timing strategy and temporal accuracy in rapid interception task with variable time-to-contact. J Sports Sci 2015;33(4):381–390.
30. Burroughs WA. Visual simulation training of baseball batters. Int J Sport Psychol 1984;15(2):117–126. https://psycnet.apa.org/record/1985-16394-001.
31. De Lucia PR, Cochran EL. Perceptual information for batting can be extracted throughout a ball’s trajectory. Percept Mot Skills 1985;61(1):143–150.
32. Droździel P, et al. Drivers’ reaction time research in the conditions in the real traffic. Open Eng 2020;10(1):35–47.
33. Wolfe B, Seppelt B, Mehler B, Reimer B, Rosenholtz R. Rapid holistic perception and evasion of road hazards. J Exp Psychol Gen 2020;149(3):490–500.
34. Evans KK, Birdwell RL, Georgian-Smith D, Wolfe JM. Discrimination and Localization of Abnormalities in Mammograms from a Global Signal. In: Proceedings of the Radiological Society of North America 2010 Scientific Assembly and Annual Meeting. Oak Brook, Ill: Radiological Society of North America, 2010.
35. Wolfe JM. Guided Search 2.0 A revised model of visual search. Psychon Bull Rev 1994;1(2):202–238.
36. Alexander RG, Nahvi RJ, Zelinsky GJ. Specifying the precision of guiding features for visual search. J Exp Psychol Hum Percept Perform 2019;45(9):1248–1264.
37. Alexander RG, Zelinsky GJ. Effects of part-based similarity on visual search: the Frankenbear experiment. Vision Res 2012;54:20–30.
38. Alexander RG, Schmidt J, Zelinsky GJ. Are summary statistics enough? Evidence for the importance of shape in guiding visual search. Vis Cogn 2014;22(3-4):595–609.
39. Alexander RG, Zelinsky GJ. Occluded information is restored at preview but not during visual search. J Vis 2018;18(11):4.
40. Kundel HL. Visual search and lung nodule detection on CT scans. Radiology 2015;274(1):14–16.
41. Manning D, et al. How do radiologists do it? The influence of experience and training on searching for chest nodules. Radiography 2006;12(2):134–142.
42. Drew T, Vo ML, Olwal A, Jacobson F, Seltzer SE, Wolfe JM. Scanners and drillers: characterizing expert visual search through volumetric images. J Vis 2013;13(10):3.
43. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. Invest Radiol 1978;13(3):175–181.
44. Kundel HL. Perception errors in chest radiography. Semin Respir Crit Care Med 1989;10(3):203–210.
45. 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(6):1668–1676.
46. Nakashima R, Komori Y, Maeda E, Yoshikawa T, Yokosawa K. Temporal Characteristics of Radiologists’ and Novices’ Lesion Detection in Viewing Medical Images Presented Rapidly and Sequentially. Front Psychol 2016;7:1553.
47. Eckstein MP, Lago MA, Abbey CK. The role of extra-foveal processing in 3D imaging. Proc SPIE Int Soc Opt Eng 2017;10136:101360E.
48. Miller WT Jr, Marinari LA, Barbosa E Jr, et al. Small pulmonary artery defects are not reliable indicators of pulmonary embolism. Ann Am Thorac Soc 2015;12(7):1022–1029.
49. Eckstein MP, Lago MA, Abbey CK. The role of extra-foveal processing in 3D imaging. In: Kupinski MA, Nishikawa RM, eds. Proceedings of SPIE: medical imaging 2017—image perception, observer performance, and technology assessment. Vol 10136. Bellingham, Wash: International Society for Optics and Photonics, 2017; 101360E.
50. Kelahan LC, Fong A, Blumenthal J, Kandaswamy S, Ratwani RM, Filice RW. The Radiologist’s Gaze: Mapping Three-Dimensional Visual Search in Computed Tomography of the Abdomen and Pelvis. J Digit Imaging 2019;32(2):234–240.
51. Kliewer MA, Hartung M, Green CS. The Search Patterns of Abdominal Imaging Subspecialists for Abdominal Computed Tomography: Toward a Foundational Pattern for New Radiology Residents. J Clin Imaging Sci 2021;11:1.
52. Aizenman A, Drew T, Ehinger KA, Georgian-Smith D, Wolfe JM. Comparing search patterns in digital breast tomosynthesis and full-field digital mammography: an eye tracking study. J Med Imaging (Bellingham) 2017;4(4):1–10.
53. Venjakob AC, Mello-Thoms CR. Review of prospects and challenges of eye tracking in volumetric imaging. J Med Imaging (Bellingham) 2016; 3(1):011002.
54. den Boer L, van der Schaaf MF, Vincken KL, Mol CP, Stuijfzand BG, van der Gijp A. Volumetric image interpretation in radiology: scroll behavior and cognitive processes. Adv Health Sci Educ Theory Pract 2018;23(4):783–802.
55. van Montfort D, Kok E, Vincken K, et al. Expertise development in volumetric image interpretation of radiology residents: what do longitudinal scroll data reveal? Adv Health Sci Educ Theory Pract 2021;26(2):437–466.
56. Mabotuwana T, Qian Y, Sevenster M. Using image references in radiology reports to support enhanced report-to-image navigation. In: AMIA Annual Symposium Proceedings. Bethesda, Md: American Medical Informatics Association, 2013.
57. Cruickshank A, Bell D. Kernel (image reconstruction for CT). Radiopaedia.org. Published December 16, 2019. Accessed May 19, 2022.
58. Sandrasegaran K, Rydberg J, Tann M, Hawes DR, Kopecky KK, Maglinte DD. Benefits of routine use of coronal and sagittal reformations in multi-slice CT examination of the abdomen and pelvis. Clin Radiol 2007;62(4):340–347.
59. AAPC. What Are Relative Value Units (RVUs)? https://www.aapc.com/practice-management/rvus.aspx. Last reviewed 2020. Accessed May 19, 2022.
60. Coffta S. Understanding the Value of RVUs in Radiology. In: HAP Radiology Billing and Coding Blog. Media, Pa: Healthcare Administrative Partners, 2018.
61. Sebro R. Leveraging the electronic health record to evaluate the validity of the current RVU system for radiologists. Clin Imaging 2021;78:286–292.
62. McDonald RJ, Schwartz KM, Eckel LJ, et al. The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. Acad Radiol 2015;22(9):1191–1198.
63. Yu JPJ, Kansagra AP, Mongan J. The radiologist’s workflow environment: evaluation of disruptors and potential implications. J Am Coll Radiol 2014;11(6):589–593.
64. Andriole KP, Wolfe JM, Khorasani R, et al. Optimizing analysis, visualization, and navigation of large image data sets: one 5000-section CT scan can ruin your whole day. Radiology 2011;259(2):346–362.
65. Kirschner LB. AHRA (American Healthcare Radiology Administrators) survey. Staff utilization: Part I. Radiol Manage 1989;11(3):55–67.
66. Sunshine JH, Bansal S. Operational characteristics of radiology groups in the United States in 1992. Radiology 1994;193(3):613–618.
67. Conoley PM, Vernon SW. Productivity of radiologists: estimates based on analysis of relative value units. AJR Am J Roentgenol 1991;157(6):1337–1340.
68. 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(2):458–467.
69. Sunshine JH, Burkhardt JH. Radiology groups’ workload in relative value units and factors affecting it. Radiology 2000;214(3):815–822.

70. Shi J, Giess CS, Martin T, et al. Radiology workload changes during the COVID-19 pandemic: implications for staff redeployment. Acad Radiol 2021;28(1):1–7.
71. Hanna TN, Lamoureux C, Krupinski EA, Weber S, Johnson JO. Effect of shift, schedule, and volume on interpretive accuracy: a retrospective analysis of 2.9 million radiologic examinations. Radiology 2018;287(1):205–212.
72. Bechtold RE, Chen MY, Ott DJ, et al. Interpretation of abdominal CT: analysis of errors and their causes. J Comput Assist Tomogr 1997;21(5):681–685.
73. Sokolovskaya E, Shinde T, Ruchman RB, et al. The effect of faster reporting speed for imaging studies on the number of misses and interpretation errors: a pilot study. J Am Coll Radiol 2015;12(7):683–688.
74. Berlin L. Faster Reporting Speed and Interpretation Errors: Conjecture, Evidence, and Malpractice Implications. J Am Coll Radiol 2015;12(9):894–896.
75. Ruchman RB, Shinde T. Reply to "Faster radiologic interpretation, errors, and malpractice: an unavoidable triad?". AJR Am J Roentgenol 2018;211(3):W186.
76. Alexander RG, Yazdanie F, Waite S, et al. Visual Illusions in Radiology: Untrue Perceptions in Medical Images and Their Implications for Diagnostic Accuracy. Front Neurosci 2021;15(554):629469.
77. Wolfe JM, Horowitz TS, Van Wert MJ, Kenner NM, Place SS, Kibbi N. Low target prevalence is a stubborn source of errors in visual search tasks. J Exp Psychol Gen 2007;136(4):623–638.
78. Shanafelt TD, Hasan O, Dyrbye LN, et al. Changes in burnout and satisfaction with work-life balance in physicians and the general US working population between 2011 and 2014. Mayo Clin Proc 2015;90(12):1600–1613 [Published correction appears in Mayo Clin Proc 2016;91(2):276.].
79. Chen JY, Lexa FJ. Baseline survey of the neuroradiology work environment in the United States with reported trends in clinical work, nonclinical work, perceptions of trainees, and burnout metrics. AJNR Am J Neuroradiol 2017;38(7):1284–1291.
80. Di Stasi LL, McCamy MB, Macknik SL, et al. Saccadic eye movement metrics reflect surgical residents' fatigue. Ann Surg 2014;259(4):824–829.
81. Christensen EE, Dietz GW, Murry RC, Moore JG. The effect of fatigue on resident performance. Radiology 1977;125(1):103–105.
82. Krupinski EA, Berbaum KS, Caldwell RT, Schartz KM, Kim J. Long radiology workdays reduce detection and accommodation accuracy. J Am Coll Radiol 2010;7(9):698–704.
83. Krupinski E, Reiner BI. Real-time occupational stress and fatigue measurement in medical imaging practice. J Digit Imaging 2012;25(3):319–324.
84. Sistrom CL, Slater RM, Rajderkar DA, Grajo JR, Rees JH, Mancuso AA. Full Resolution Simulation for Evaluation of Critical Care Imaging Interpretation; Part 1: Fixed Effects Identify Influences of Exam, Specialty, Fatigue, and Training on Resident Performance. Acad Radiol 2020;27(7):1006–1015.
85. Waite S, Kolla S, Jeudy J, et al. Tired in the Reading Room: The Influence of Fatigue in Radiology. J Am Coll Radiol 2017;14(2):191–197.
86. Stec N, Arje D, Moody AR, Krupinski EA, Tyrrell PN. A Systematic Review of Fatigue in Radiology: Is It a Problem? AJR Am J Roentgenol 2018;210(4):799–806.
87. Vertinsky T, Forster B. Prevalence of eye strain among radiologists: influence of viewing variables on symptoms. AJR Am J Roentgenol 2005;184(2):681–686.
88. Goo JM, Choi JY, Im JG, et al. Effect of monitor luminance and ambient light on observer performance in soft-copy reading of digital chest radiographs. Radiology 2004;232(3):762–766.
89. Patel AG, Pizzitola VJ, Johnson CD, Zhang N, Patel MD. Radiologists Make More Errors Interpreting Off-Hours Body CT Studies during Overnight Assignments as Compared with Daytime Assignments. Radiology 2020;297(2):374–379.
90. Bruno MA. Radiology Errors across the Diurnal Cycle. Radiology 2020;297(2):380–381.
91. Hanna TN, Loehfelm T, Khosa F, Rohatgi S, Johnson JO. Overnight shift work: factors contributing to diagnostic discrepancies. Emerg Radiol 2016;23(1):41–47.
92. Wen G, Rodriguez-Niño B, Pecen FY, Vining DJ, Garg N, Markey MK. Comparative study of computational visual attention models on two-dimensional medical images. J Med Imaging (Bellingham) 2017;4(2):025503.
93. Ekpo EU, Alakhras M, Brennan P. Errors in Mammography Cannot be Solved Through Technology Alone. Asian Pac J Cancer Prev 2018;19(2):291–301.
94. Williams LH, Drew T. Distraction in diagnostic radiology: How is search through volumetric medical images affected by interruptions? Cogn Res Princ Implic 2017;2(1):12.
95. Bell LTO, James R, Rosa JA, Pollentine A, Pettet G, McCoubrie P. Reducing interruptions during duty radiology shifts, assessment of its benefits and review of factors affecting the radiology working environment. Clin Radiol 2018;73(8):759.e19–759.e25.
96. Krupinski EA, MacKinnon L, Hasselbach K, Taljanovic M. Evaluating RVUs as a measure of workload for use in assessing fatigue. In: Mello-Thoms CR, Kupinski MA, eds. Proceedings of SPIE: medical imaging 2015—image perception, observer performance, and technology assessment. Vol 9416. Bellingham, Wash: International Society for Optics and Photonics, 2015; 94161A.
97. Kramer M. Sleep loss in resident physicians: the cause of medical errors? Front Neurol 2010;1:128.
98. Wallack MK, Chao L. Resident work hours: the evolution of a revolution. Arch Surg 2001;136(12):1426–1431; discussion 1432.
99. Chahal HS. Work hour regulations and training of residents. J Oral Maxillofac Surg 2007;65(1):154–155.
100. Harris JD, Staheli G, LeClere L, Andersone D, McCormick F. What effects have resident work-hour changes had on education, quality of life, and safety? A systematic review. Clin Orthop Relat Res 2015;473(5):1600–1608.
101. Volpp KG, Rosen AK, Rosenbaum PR, et al. Mortality among hospitalized Medicare beneficiaries in the first 2 years following ACGME resident duty hour reform. JAMA 2007;298(9):975–983.
102. Bell BM. Resident duty hour reform and mortality in hospitalized patients. JAMA 2007;298(24):2865–2866; author reply 2866–2867.
103. Millard WB. For whom the bell commission tolls: unintended effects of limiting residents' hours. Ann Emerg Med 2009;54(4):A25–A29.
104. Geijer H, Geijer M. Added value of double reading in diagnostic radiology,a systematic review. Insights Imaging 2018;9(3):287–301.
105. Garland LH. On the scientific evaluation of diagnostic procedures. Radiology 1949;52(3):309–328.
106. Onega T, Aiello Bowles EJ, Miglioretti DL, et al. Radiologists' perceptions of computer aided detection versus double reading for mammography interpretation. Acad Radiol 2010;17(10):1217–1226.
107. Lexa FJ, Jha S. Artificial Intelligence for Image Interpretation: Counterpoint-The Radiologist's Incremental Foe. AJR Am J Roentgenol 2021;217(3):558–559.
108. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med 2018;15(11):e1002686.
109. Coppola F, Faggioni L, Gabelloni M, et al. Human, all too human? An all-around appraisal of the "artificial intelligence revolution" in medical imaging. Front Psychol 2021;12:710982.
110. Duong MT, Rauschecker AM, Rudie JD, et al. Artificial intelligence for precision education in radiology. Br J Radiol 2019;92(1103):20190389.
111. Kelly B, Judge C, Bollard SM, et al. Radiology artificial intelligence, a systematic evaluation of methods (RAISE): a systematic review protocol. Insights Imaging 2020;11(1):133.
112. Miles RC, Lehman CD. Artificial Intelligence for Image Interpretation: Point-The Radiologist's Potential Friend. AJR Am J Roentgenol 2021;217(3):556–557.
113. Duszak R Jr, Muroff LR. Measuring and managing radiologist productivity, part 1: clinical metrics and benchmarks. J Am Coll Radiol 2010;7(6):452–458.
114. Duszak R Jr, Muroff LR. Measuring and managing radiologist productivity, part 2: beyond the clinical numbers. J Am Coll Radiol 2010;7(7):482–489.
115. Holleman GA, Hooge ITC, Kemner C, Hessels RS. The 'Real-World Approach' and Its Problems: A Critique of the Term Ecological Validity. Front Psychol 2020;11:721.