# Long Radiology Workdays Reduce Detection and Accommodation Accuracy

Elizabeth A. Krupinski, PhD[a], Kevin S. Berbaum, PhD[b], Robert T. Caldwell, MFA[b],
Kevin M. Schartz, PhD[b], John Kim, MD[b]

**Purpose:** The aim of this study was to measure the diagnostic accuracy of fracture detection, visual accommodation, reading time, and subjective ratings of fatigue and visual strain before and after a day of clinical reading.

**Methods:** Forty attending radiologists and radiology residents viewed 60 deidentified, HIPAA-compliant bone examinations, half with fractures, once before any clinical reading (early) and once after a day of clinical reading (late). Reading time was recorded. Visual accommodation (the ability to maintain focus) was measured before and after each reading session. Subjective ratings of symptoms of fatigue and oculomotor strain were collected. The study was approved by local institutional review boards.

**Results:** Diagnostic accuracy was reduced significantly after a day of clinical reading, with average areas under the receiver operating characteristic curves of 0.885 for early reading and 0.852 for late reading ($P < .05$). After a day of image interpretation, visual accommodation was no more variable, though error in visual accommodation was greater ($P < .01$), and subjective ratings of fatigue were higher.

**Conclusions:** After a day of clinical reading, radiologists have reduced ability to focus, increased symptoms of fatigue and oculomotor strain, and reduced ability to detect fractures. Radiologists need to be aware of the effects of fatigue on diagnostic accuracy and take steps to mitigate these effects.

**Key Words:** Reader fatigue, observer performance, visual accommodation

*J Am Coll Radiol 2010;7:698-704. Copyright © 2010 American College of Radiology*

## INTRODUCTION

Radiology services, especially high-technology modalities [1], second opinion [2], and teleradiology [3], have increased significantly in recent years. Fewer radiologists now read more studies, each containing more images, in less time [4-8]. This increase in time spent viewing more images may increase strain on a radiologist's oculomotor system, resulting in eyestrain (known clinically as asthenopia) [9,10].

Although eyestrain has not been extensively studied in radiology, we have self-report data showing that radiologists report increasingly severe symptoms of eyestrain, including blurred vision and difficulty focusing, as they read more imaging studies [11]. These findings are corroborated by the self-report data of other radiology researchers [12,13]. Eyestrain occurs when the oculomotor systems must work to maintain accommodation, convergence, and direction of gaze. Visual accommodation is a common objective measure of visual strain or fatigue in studies of computer displays [14-17].

We recently collected accommodation data on 3 attending radiologists and 3 radiology residents before and after a day of clinical reading [18]. Errors in accommodation indicating increased visual strain and, as a consequence, a reduced ability to focus increased significantly after a day of clinical reading. Error was greater at close viewing distances such as those used by radiologists to interpret images. The inability to maintain focus on a diagnostic image could affect diagnostic accuracy. Therefore, the goal of the present study was to measure diagnostic accuracy before and after a day of diagnostic image interpretation and study corresponding changes in ac-

---

[a] Department of Radiology, University of Arizona, Tucson, Arizona.

[b] Department of Radiology, University of Iowa, Iowa City, Iowa.

Corresponding author and reprints: Elizabeth A. Krupinski, PhD, University of Arizona, Department of Radiology, 1609 N Warren, Building 211, Room 112, Tucson, AZ 85724; e-mail: krupinski@radiology.arizona.edu.

This work was supported in part by grant R01 EB004987 from the National Institute of Biomedical Imaging and Bioengineering (Bethesda, Md).

0091-2182/10/$36.00 • DOI 10.1016/j.jacr.2010.03.004

commodative response. We hypothesized that the accuracy of visual accommodation (reflecting visual strain) and detection accuracy for fractures would decrease after a day of clinical reading.

## METHODS

This study was approved by the institutional review boards at both the University of Arizona and the University of Iowa.

### Images

All images were stripped of patient identifiers to comply with HIPAA standards. We used skeletal images from earlier satisfaction of search studies [19]. There were 66 cases, each with 2 to 4 images. One case served as a demonstration to familiarize observers with the procedure and presentation software, 5 served as practice cases, and the remaining 60 were the test cases. Half of the cases had no fractures, and half had a single moderate to very subtle fracture. In some cases, the fractures were visible in multiple views. The study included wrist, hand, ankle, foot, long bones, and shoulder and rib examinations. The conspicuity of a fracture was rated (easy vs hard) by the frequency with which it had been detected in previous studies [19].

The 60 cases were presented in a randomized order for each observer. The first 30 cases, which had predominantly easy fractures, had a separate randomization than the second 30 cases, which had predominantly hard fractures. Cases were displayed using customized WorkstationJ software developed at the University of Iowa [20]. The software presented each case sequentially, with the first screen having the age and gender of the patient, thumbnails of all available views, and the toolbar. Observers were allowed to bring each image to full size for viewing and were allowed to adjust window and level settings using the mouse, hot keys, or select presets. The confidence of positive decisions was reported as definite, probable, possible, or suspicious, along with a percentage confidence rating (0%-100% in 10% intervals), with 100% indicating a high degree of confidence. Negative decisions did not require input and were recorded as such by default when the observer went to the next case. The program recorded total viewing time per case, which images were viewed and in what sequence, how long an image was displayed, how often the observer used window and level settings, and how often the observer used presets.

### Observers

Observers were attending radiologists and radiology residents at the University of Arizona and the University of Iowa. There were 10 attending radiologists and 10 radiology residents at each institution. Table 1 provides the gender, average age, months since last eye examination, dominant eye, percentage wearing corrective lenses, type of lenses worn, and type of vision disorder for the observers at both institutions. Table 2 provides information regarding at what time radiologists woke up on the day of the experiment; how many hours of sleep they had; how long they had been reading cases that day; the number of cases; what percentage had colds, allergies, and itchy or watery eyes; and what percentage had used eye drops that day.

### Procedure

Data were collected at two points in time for each observer: once in the morning (before any diagnostic read-

**Table 1.** Characteristics of participating University of Arizona and University of Iowa attending radiologists and residents

| | Attending Radiologists | | Residents | |
|---|---|---|---|---|
| Variable | Arizona | Iowa | Arizona | Iowa |
| Men/women | 7/3 | 10/0 | 9/1 | 9/1 |
| Average male age (y) | 44.43 ± 15.75 (range, 31-69) | 51.10 ± 12.06 (range, 31-71) | 31.44 ± 3.81 (range, 28-40) | 32.22 ± 4.63 (range, 28-42) |
| Average female age (y) | 42.00 ± 8.19 (range, 35-51) | — | 33 ± 0 (range, 0) | 35 ± 0 (range, 0) |
| Months since last eye examination | 25.90 ± 37.10 (range, 2-120) | 13.65 ± 12.73 (range, 0.5-36) | 29.40 ± 35.73 (range, 4-120) | 18.30 ± 18.67 (range, 4-60) |
| Dominant eye | 90% right | 57% right | 80% right | 80% right |
| Wear corrective lenses | 50% | 50% | 90% | 80% |
| Types of lenses | 50% glasses/contact lenses full-time; 50% readers | 100% glasses/contact lenses full-time | 60% glasses/contact lenses full-time; 40% computer glasses | 88% glasses/contact lenses full-time; 12% driving |
| Vision | 50% nearsighted; 17% farsighted; 33% presbyopia | 50% nearsighted; 12% farsighted; 12.5% astigmatism; 25% nearsighted with presbyopia | 100% nearsighted | 17% nearsighted; 17% astigmatism; 66% nearsighted with astigmatism |

**Table 2**. Data for attending radiologists and residents for the early and late sessions regarding sleep, case reading, and eye conditions on the days of the study

| | Attending Radiologists | | Residents | |
|---|---|---|---|---|
| Variable | Early | Late | Early | Late |
| Time up | 4-7:30 AM | 5-6:45 AM | 5-8:30 AM | 5-7:15 AM |
| Hours of sleep | 7.10 ± 0.66 (range, 6-8) | 6.88 ± 0.86 (range, 5-8) | 6.93 ± 0.80 (range, 6-8.5) | 6.48 ± 0.92 (range, 4-8) |
| Hours reading | 0.44 ± 0.79 (range, 0-3) | 6.48 ± 2.43 (range, 2-10) | 0.28 ± 0.70 (range, 0-2,5) | 7.73 ± 2.06 (range, 4-14) |
| Number of cases | 6.05 ± 11.21 (range, 0-40) | 70.55 ± 47.31 (range, 8-200) | 2.40 ± 6.96 (range, 0-30) | 27.45 ± 19.54 (range, 5-75) |
| Cold/allergies | 25% | 25% | 0% | 10% |
| Itchy/watery eyes | 28.57% | 0% | 37.50% | 0% |
| Used eye drops | 0% | 12.5% | 0% | 0% |

ing activity [early]) and once in the late afternoon (after a day of diagnostic reading [late]) on days they spent interpreting cases. Observers completed surveys regarding their current physical status (eg, how many hours of sleep they had, whether they had allergies) and the number of hours spent reading that day along with the type of images. They completed the Swedish Occupational Fatigue Inventory (SOFI), which was developed and validated to specifically measure perceived fatigue in work environments [21,22]. The instrument consists of 20 expressions, evenly distributed on 5 latent factors: lack of energy, physical exertion, physical discomfort, lack of motivation, and sleepiness. Physical exertion and physical discomfort are considered physical dimensions of fatigue, whereas lack of motivation and sleepiness are considered primarily mental factors. Lack of energy is a general factor reflecting both physical and mental aspects of fatigue. Lower scores indicate lower levels of perceived fatigue than higher scores. The SOFI does not measure visual fatigue, so it was complemented with the oculomotor strain subscale from the Simulator Sickness Questionnaire [23,24].

Visual accommodation (strain) was measured using the WAM-5500 Auto Refkeratometer (Grand Seiko, Hiroshima, Japan), which collects refractive measurements and pupil diameter measurements every 0.2 seconds. Two sets of measurements were made before and after each reading session. For each set, the observer first fixated an asterisk for 30 seconds and then fixated a 2 × 2 inch image of a finger fracture displayed on a liquid crystal display for 30 seconds while accommodation was measured. The asterisk is a standard target for the device. Our premise for using the fracture was that the image was similar to a real radiologic examination.

After an introduction and 5 practice cases, the observers viewed the series of skeletal images on a 3-megapixel liquid crystal display (at the University of Arizona: Dome C3i, Planar Systems, Inc, Beaverton, Ore; at the University of Iowa: National Display Systems, San Jose, California) that was calibrated to the Digital Imaging and Communications in Medicine Grayscale Standard Display Function [25]. Their task was to determine if a fracture was present, locate it with a cursor, and provide ratings of their decision confidence to be used in a receiver operating characteristic analysis of the data.

## RESULTS

### Diagnostic Accuracy

The area under the receiver operating characteristic curve was used to measure accuracy for detecting fractures [26,27]. The area under the curve was estimated for each observer in each experimental condition, and the average areas were compared using analysis of variance (ANOVA). Independent variables were institution (University of Arizona or University of Iowa), level of training (attending radiologist or resident), and the reading session time of day (early or late). A more complex ANOVA added session order (readers assigned to early first, then late, or to late first, then early) and case difficulty (first 30 with 15 easier fractures, second 30 with 15 harder fractures) as other independent variables.

There was a significant drop in detection accuracy for late vs early reading. The average areas under the curves were 0.885 for early and 0.852 for late readings, $F(1, 36) = 4.15$, $P = .049$. There were no other significant effects. The more complex ANOVA revealed that although attending radiologists and residents were about the same on easy cases, not surprisingly, residents were somewhat less accurate on hard cases. Supplemental analyses suggested that the reduction in accuracy for late reading was based on about the same increase in false-positives as the decrease in true-positives.

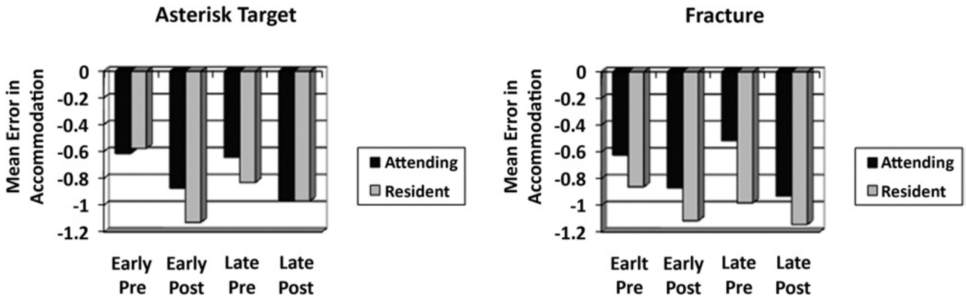

**Fig 1.** Error in accommodation for the asterisk (left) and fracture (right) targets for before and after measurements made early and late in the day.

## Reading Time

Total inspection time for interpreting the examinations was also analyzed. The ANOVA treated total inspection time as a dependent variable and included fracture status (no fracture or fracture), institution, fracture difficulty, training level, and cases as independent variables. Each examination took 52.1 seconds on average for early reading and 51.5 seconds for late reading. On average, each examination took radiologists 50.7 seconds and residents 52.8 seconds. The only main effect was a significantly greater reading time for normal examinations than examinations with fractures (56.7 vs 46.9 seconds), $F(1, 36) = 18.84$, $P = .0001$.

To determine whether search time was affected by time of day, we studied the time to report fractures for cases in which the fractures were detected in both the early and late sessions. There was no significant difference between early and late reading time to report the fractures for all examinations (37.0 vs 38.3 seconds), easier examinations (33.0 vs 34.0 seconds), or harder examinations (42.5 vs 44.2 seconds). For all examinations, the average response times were 42.8 seconds for early reading and 36.0 seconds for late reading when the early session occurred first. The average response times were 31.2 seconds for early reading and 40.0 seconds for late reading when the late session occurred first, $F(1, 32) = 20.84$, $P = .0001$. Similar results were obtained when easy and hard examinations were analyzed separately. These results suggest that responses in the second session were faster. This apparent practice effect is hardly surprising. The main finding was that when a fracture was found both early and late, the same amount of search time was required.

## Visual Strain Results

Accommodation measures (as a measure of visual strain) were taken every 0.2 seconds over a number of seconds. Medians were computed for each reader before and after the early and late reading sessions. An ANOVA was used to analyze the accommodation measures with the fracture and asterisk targets. For the fracture, there was significantly greater accommodative error after the workday (−1.16 diopters late vs −0.72 diopters early), $F(1, 29) = 27.01$, $P < .0001$. For the asterisk target, there was also significant main effect for session time of day (−1.04 diopters late vs −0.64 diopters early), $F(1, 34) = 22.005$, $P < .0001$. This suggests that readers were more myopic and experienced more visual strain after their workdays. Overall, there was no main effect for measures before and after the reading session or for level of training. A significant before vs after × attending radiologist versus resident interaction showed that although the attending radiologists tended to have less accommodative error after the reading session than before, residents tended to have more (Figure 1).

We further hypothesized that if readers had greater visual strain and thus more difficulty maintaining focus after visual work, their accommodation measures would be more variable. Analyses of variance on the standard deviations of the accommodation measurements were computed. For the fracture target, there were no significant main effects or 2-way interactions. There was a significant 3-way interaction of before vs after × attending radiologist vs resident × early vs late, $F(1, 34) = 4.35$, $P < .05$. For the asterisk target, residents' accommodation was significantly more variable than faculty members' (0.13 vs 0.17 diopters), $F(1, 29) = 4.72$, $P < .05$. There were no other significant main effects or 2-way interactions. The 3-way interaction of before versus after × attending radiologist vs resident × early versus late was again significant, $F(1, 29) = 8.12$, $P < .01$). Because the nature of the 3-way interactions was not consistent between the 2 targets (fracture and asterisk), nothing could be concluded beyond that the variability for residents was greater than for faculty members. Overall, we must conclude that variability of accommodation was unaffected by visual work in our experiment.

**Table 3.** Mean ± SD of the SOFI and SSQ survey ratings for attending radiologists and residents early and late in the day

| | Attending Radiologists | | Residents | |
|---|---|---|---|---|
| **Item** | **Early** | **Late** | **Early** | **Late** |
| Lack of energy | 1.60 ± 1.28 | 3.36 ± 2.62 | 2.20 ± 2.15 | 4.41 ± 2.54 |
| Physical discomfort | 1.51 ± 0.81 | 1.94 ± 1.31 | 1.58 ± 0.95 | 2.36 ± 1.59 |
| Sleepiness | 1.58 ± 1.50 | 2.78 ± 2.70 | 2.20 ± 2.06 | 3.84 ± 2.64 |
| Physical exertion | 1.23 ± 0.53 | 1.25 ± 0.63 | 1.20 ± 0.44 | 1.25 ± 0.47 |
| Lack of motivation | 2.01 ± 1.66 | 2.66 ± 2.27 | 2.38 ± 1.74 | 3.46 ± 2.31 |
| SSQ eye strain | 1.13 ± 0.22 | 1.55 ± 0.50 | 1.21 ± 0.35 | 1.66 ± 0.56 |

Note: SOFI = Swedish Occupational Fatigue Inventory; SSQ = Simulator Sickness Questionnaire.

## Fatigue Survey Results

The scores for each of the 5 SOFI factors were analyzed with an ANOVA with session (early vs late) and experience (attending radiologist vs resident) as independent variables. Average rating values for each factor are shown in Table 3.

For lack of energy, $F(1, 76) = 16.19$, $P = .0001$, physical discomfort, $F(1, 76) = 5.091$, $P = .0269$, and sleepiness, $F(1, 76) = 7.761$, $P = .0067$, there were statistically significant differences as a function of session but not experience. For physical exertion and for motivation, there were no statistically significant differences as a function of either session or experience. Additional analyses indicated that there were no statistically significant differences on any of the factors as a function of gender or site.

The scores from the 7 questions on the oculomotor strain subscale of the Simulator Sickness Questionnaire were averaged and analyzed with an ANOVA as a function of session and experience (see Table 2). As with the SOFI, low scores represent lower levels of perceived oculomotor strain. There was a statistically significant difference in rated symptoms of oculomotor strain as a function of session, $F(1, 75) = 20.39$, $P < .0001$, but not experience, $F(1, 75) = 0.99$, $P = .32$.

## DISCUSSION

### Diagnostic Accuracy

The results of this study suggest that because of increased visual strain as reflected in their lowered accommodation measures, radiologists' ability to focus on images was reduced, making them less accurate after a day reading diagnostic images. Several authors have studied variation in diagnostic performance over the course of an ordinary professional workday [28,29]. Gale et al [29] found a significant morning-to-afternoon drop in sensitivity in the detection of pulmonary nodules on chest radiographs. However, Brogdon et al [28] found no significant effect of fatigue on observer sensitivity or specificity between early and late reading of chest images with pseudonodules during an ordinary workday.

Our study demonstrated reduced diagnostic accuracy after the radiology workday, but the difference between accuracy before and after work was small, on the order of 4%. It seems that our sample of 40 readers reading 60 multiview examinations was just sufficient to detect this difference at the .05 significance level.

Christensen et al [30] compared performance after rest with performance after working a minimum of 15 consecutive hours and found no deterioration in performance with fatigue. Other researchers have studied the discordance between resident readings during night call with readings made by radiologists the next morning. As in Christensen et al's laboratory study, a lack of sleep is added to the fatigue that results from image interpretation work extending well beyond a clinical workday. An explanatory problem in these studies is that the nighttime readers are residents, whereas the next morning readers are faculty members, so the disparity may reflect training and experience rather than just fatigue and sleeplessness. The morning reading is treated as the gold standard, and the goal is often to determine the cost in diagnostic accuracy of using residents rather than radiologist readers at night.

Other experiments have evaluated ways for mitigating the detrimental effect of sleep loss. These "discrepancy" experiments are easy to perform, because although relatively large numbers of patient examinations are sampled, the truth of diagnosis is only followed or arbitrated when there is discordance between the night and daytime readings. In a variety of circumstances, the discordance rate and impact of "misses" is small [31-33]. However, in more complex examinations, the discordance rate can be substantially higher [34-36]. These studies in radiology and studies in other medical specialties usually explain the errors or discrepancies by pointing to the breakdown of cognitive functions that accompany sleep loss. For example, Harrigal and Erly [37] noted:

> There is an important relationship between sleep and the consolidation of procedural and declarative memory and learning. Twenty-four hours without sleep results in decreased achievement in cognitive tasks requiring critical thinking. One study revealed that after a single night without sleep there is a significant decline in the performance of tasks using inference and deduction. In addition to affecting performance of these higher cognitive tasks, there is a decreased perception of these deficits. The effects of sleep deprivation are most evident in higher cognitive functions of the prefrontal cortex including attention, judgment, memory, and problem solving. For radiologists these tasks are crucial to image interpretation and ultimately, patient care and safety.

Although there was a difference in the time of day between early and late reading sessions, sleep loss was not present in our experiment. This excludes factors that might explain a detection accuracy decrement that would need to be considered were sleep loss present.

## Visual Strain and Reading Time

When we began this experiment, we thought that although oculomotor fatigue or strain might reduce the ability to stay focused on an image, observers might compensate by taking more time. This did not happen. Accommodation accuracy was reduced, reading time was the same. Viewing time was unchanged late in the day, and time to report fractures was no different. No extra time was taken to achieve better accommodation during the fracture detection experiment. Perhaps examinations read at the end of a workday are interpreted under the burden of having the eyes focused further in front of the display screen than at the beginning of the workday. From our experiment, we cannot reach the conclusion that the reduction in detection accuracy is caused by the reduction in ability to keep the eyes focused on the display screen. Other neural mechanisms could be responsible for reduced detection accuracy. Further research is needed to establish a causal link between the myopia induced by a day of medical image interpretation and reduced diagnostic accuracy at the end of that workday.

Many radiologists work more hours than we studied. Even when sleep loss is not a factor, some radiologists work considerably longer on a given day than those in our study. Given that a small but significant reduction in detection accuracy was demonstrated for an average workday of about 8 hours, we suspect that more extended reading may expose the reader to greater decrements in accuracy.

An interesting question is why the average accommodation measurement for our readers was in front of the display screen (−0.6 diopters for the fracture target and −0.7 diopters for the asterisk target) at the beginning of the day. An explanation is that refraction using automatic refraction may differ from the method of interchangeable trial lenses used in an ophthalmologist's office. Autorefractors use only small portions of the eye's optic, and the technique is generally less refined. Moreover, there is reason to believe that autorefractors may measure "more myopic" than ophthalmologists, according to the American Academy of Ophthalmology [38]:

> So-called instrument myopia, the tendency to accommodate when looking into instruments, has caused major problems with automated refractors in the past. Various methods of fogging and automatic tracking have been developed to overcome this problem, with some success.

## Subjective Ratings of Fatigue

The symptom self-report scales indicate general fatigue with negative effects on visual, physical, cognitive, and emotional status. But if the current study could not establish a causal link, what further research could reveal causes? Eliminating other potential causes may require exhaustive study with isolating causes. Getting a definitive answer may require a true experimental manipulation of oculomotor control mechanisms rather than field observations. For example, accommodation might be experimentally stressed while treating detection as a dependant variable. It is hard to see how this could be done in radiologists or clinical reading.

## Limitations

A limitation of the present study was that only radiographic examinations were used. Computed tomographic and MRI examinations contain hundreds of images that must be scrolled through; this is potentially more fatiguing than reading static images. We are currently conducting a study of nodule detection on chest CT examinations, in which detection depends on the discrimination of different kinds of temporal modulation (2-D motion vs on and off with no change in position).

# CONCLUSION

After a day of clinical reading, radiologists have reduced ability to focus on displayed images, increased symptoms of fatigue and oculomotor strain, and reduced detection accuracy. Radiologists need to be aware of the effects of fatigue on diagnostic accuracy and take steps to mitigate these effects.

# ACKNOWLEDGMENT

We would like to thank the 40 attending radiologists and radiology residents in the departments of radiology at the University of Arizona and the University of Iowa for their time and participation in this study.

## REFERENCES

1. Bhargavan M, Sunshine JH. Utilization of radiology services in the United States: levels and trends in modalities, regions, and populations. Radiology 2005;234:824-32.
2. DiPiro PJ, vanSonnenberg E, Tumeh SS, Ros PR. Volume and impact of second-opinion consultations by radiologists at a tertiary care cancer center: data. Acad Radiol 2002;9:1430-3.
3. Ebbert TL, Meghea C, Iturbe S, Forman HP, Bhargavan M, Sunshine JH. The state of teleradiology in 2003 and changes since 1999. AJR Am J Roentgenol 2007;188:W103-12.
4. Sunshine JH, Maynard CD. Update on the diagnostic radiology employment market: findings through 2007-2008. J Am Coll Radiol 2008;5:827-33.
5. Lu Y, Zhao S, Chu PW, Arenson RL. An update survey of academic radiologists' clinical productivity. J Am Coll Radiol 2008;5:817-26.
6. Nakajima Y, Yamada K, Imamura K, Kobayashi K. Radiologist supply and workload: international comparison—working group of Japanese College of Radiology. Radiat Med 2008;26:455-65.
7. Mukerji N, Wallace D, Mitra D. Audit of the change in the on-call practices in neuroradiology and factors affecting it. BMC Med Imag 2006;6:13.
8. Meghea C, Sunshine JH. Determinants of radiologists' desired workloads. J Am Coll Radiol 2007;4:143-4.
9. Ebenholtz SM. Oculomotor systems and perception. New York: Cambridge University Press; 2001.
10. MacKenzie W. On asthenopia or weak-sightedness. Edinburgh J Med Surg 1843;60:73-103.
11. Krupinski EA, Kallergi M. Choosing a radiology workstation: technical and clinical considerations. Radiology 2007;242:671-82.
12. Vertinsky T, Forster B. Prevalence of eye strain among radiologists: influence of viewing variables on symptoms. AJR Am J Roentgenol 2005;184:681-6.
13. 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;232:762-6.
14. Andre JT, Owens DA. Predicting optimal accommodative performance from measures of the dark focus of accommodation. Human Factors 1999;41:139-45.
15. Hasebe S, Graf EW, Schor C. Fatigue reduces tonic accommodation. Ophthalmic Physiol Optics 2001;21:151-60.
16. Murata A, Uetake A, Otsuka M, Takasaw Y. Proposal of an index to evaluate visual fatigue induced during visual display terminal tasks. Int J Human Comput Interact 2001;13:305-21.
17. Watten RG, Lie I, Birketvedt. The influence of long-term visual near-work on accommodation and vergence: a field study. J Hum Ergol (Tokyo) 1994;23:27-39.
18. Krupinski EA, Berbaum KS. Measurement of visual strain in radiologists. Acad Radiol 2009;16:947-50.
19. Berbaum KS, El-Khoury GY, Ohashi K, et al. Satisfaction of search in multi-trauma patients: severity of detected fractures. Acad Radiol 2007;14:711-22.
20. Schartz K, Berbaum K, Caldwell B, Madsen M. Software to facilitate medical image perception and technology evaluation research. Presented at: Medical Image Perception Society XII Conference; Iowa City, Iowa; October 17-20, 2007.
21. Ahsberg E. Dimensions of fatigue in different workplace populations. Scand J Psych 2000;41:231-41.
22. Ahsberg E, Gamberale F, Gustafsson K. Perceived fatigue after mental work: an experimental evaluation of a fatigue inventory. Ergonomics 2000;43:252-68.
23. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psych 1993;3:203-20.
24. Kennedy RS, Lane NE, Lilienthal MG, Berbaum KS, Hettinger LJ. Profile analysis of simulator sickness symptoms: application to virtual environment systems. Presence 1992;1:295-301.
25. National Electrical Manufacturers Association. Digital Imaging and Communications in Medicine. Available at: http://medical.nema.org. Accessed July 2, 2009.
26. Metz CE, Pan X. "Proper" binormal ROC curves: theory and maximum-likelihood estimation. J Math Psych 1999;43:1-33.
27. Pan X, Metz CE. The "proper" binormal model: parametric ROC curve estimation with degenerate data. Acad Radiol 1997;4:380-9.
28. Brogdon BG, Kelsey CA, Moseley RD. Effect of fatigue and alcohol on observer perception. AJR Am J Roentgenol 1978;130:971-4.
29. Gale AG, Murray D, Millar K, Worthington BS. Circadian variation in radiology. In: Gale AG, Johnson F, eds. Theoretical and applied aspects of eye movement research. London: Elsevier Science; 1984.
30. Christensen EE, Dietz GW, Murry RC, Moore JG. The effect of fatigue on resident performance. Radiology 1977;125:103-5.
31. Lal NR, Murray UM, Eldevik OP, Desmond JS. Clinical consequences of misinterpretations of neuroradiologic CT scans by on-call radiology residents. AJNR Am J Neuroradiol 2000;21:124-9.
32. Le AH, Licurse A, Catanzano TM. Interpretation of head CT scans in the emergency department by fellows versus general staff non-neuroradiologists: a closer look at the effectiveness of a quality control program. Emerg Radiol 2007;14:311-6.
33. Ruchman RB, Jaeger J, Wiggins EF III, et al. Preliminary radiology resident interpretations versus final attending radiologist interpretations and the impact on patient care in a community hospital. AJR Am J Roentgenol 2007;189:523-6.
34. Cervini P, Bell CM, Roberts HC, Provost YL, Chung T-B, Paul NS. Radiology resident interpretation of on-call CT pulmonary angiograms. Acad Radiol 2008;15:556-62.
35. Meyer RE, Nickerson JP, Burbank HN, Alsofrom GF, Linnell GJ, Filippi CG. Discrepancy rates of on-call radiology residents' interpretations of CT angiography studies of the neck and circle of Willis. AJR Am J Roentgenol 2009;193:527-32.
36. Filippi CG, Schneider B, Burbank HN, Alsofrom GF, Linnell G, Ratkovits B. Discrepancy rates of radiology resident interpretations of on-call neuroradiology MR imaging studies. Radiology 2008;249:972-9.
37. Harrigal CL, Erly WK. On-call radiology: community standards and current trends. Semin Ultrasound CT MRI 2007;28:85-93.
38. American Academy of Ophthalmology. Clinical optics (section 3 of the basic and clinical science course 2005-2006). San Francisco, Calif: American Academy of Ophthalmology; 2005:312-5.