



Published in final edited form as:

*J Am Coll Radiol.* 2018 December ; 15(12): 1709–1716. doi:10.1016/j.jacr.2017.12.019.

## The Effects of Fatigue from Overnight Shifts on Radiology Search Patterns and Diagnostic Performance

Tarek N. Hanna, MD, Matthew E. Zygmunt, MD, Ryan Peterson, MD, David Theriot, MD, Haris Shekhani, MD, Jamlik-Omari Johnson, MD, and Elizabeth A. Krupinski, PhD

Emory University, Division of Emergency and Trauma Imaging, Department of Radiology and Imaging Sciences, Emory University Midtown Hospital, 550 Peachtree Road, Atlanta GA 30308

### Abstract

**Purpose**—Assess the effect of overnight shifts (ONS) on radiologist fatigue, visual search pattern, and diagnostic performance.

**Materials and Methods**—This experimental study was approved by the IRB. 12 radiologists (5 faculty and 7 residents) each completed two sessions: one during a normal work-day (“not fatigued”) and another in the morning following an ONS (“fatigued”). Each radiologist completed the Swedish Occupational Fatigue Inventory (SOFI). During each session radiologists viewed 20 bone radiographs consisting of normal and abnormal studies. Viewing time, diagnostic confidence, and eye-tracking data were recorded.

**Results**—SOFI results demonstrated worsening in all 5 variables (lack of energy, physical exertion, physical discomfort, lack of motivation, sleepiness) following ONS ( $p < 0.01$ ). Overall, participants demonstrated worse diagnostic performance in the fatigued vs not fatigued state ( $p < 0.05$ ). Total viewing time per case was longer when fatigued ( $35.9 \pm 25.8$  seconds) than not fatigued ( $24.8 \pm 16.3$ ) ( $p < 0.0001$ ). Total viewing time per case was longer for residents ( $p < 0.05$ ). Mean total fixations generated during the search increased 60% during fatigued sessions ( $p < 0.0001$ ). Mean time to first fixate on the fracture increased 34% during fatigued sessions ( $p < 0.0001$ ), and was longer for residents ( $p < 0.01$ ). Dwell times associated with true and false positive decisions increased, while those with false negatives decreased.

**Conclusion**—Following ONS, radiologists were more fatigued with worse diagnostic performance, a 45% increase in view time per-case, a 60% increase in total gaze fixations, and a 34% increase in time to fixate on the fracture. The effects of fatigue were more pronounced in residents.

---

Corresponding Author: Tarek N. Hanna, MD, Division of Emergency and Trauma Imaging, Department of Radiology and Imaging Sciences, Emory University Midtown Hospital, 550 Peachtree Road, Atlanta GA 30308, Fax: 404-686-5709, Phone: 404-451-9825, tarek.hanna@emory.edu.

Data originating from this study was presented at both the American Society of Emergency Radiology and the Radiologic Society of North America (2017 Annual Meetings).

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Introduction

Increasing evening and overnight diagnostic radiology volumes coupled with the growing demand for contemporaneous radiology interpretations has fueled an increase in atypical diagnostic radiology schedules (1–3). In addition, a substantial portion of afterhours work is emergent requiring shorter radiology report turnaround times to decrease patient length of stay and improve overall emergency department (ED) flow (4). In the absence of attending-level coverage at academic medical centers, residents provide afterhours interpretations with discrepancy rates approaching the interobserver error rates of attending radiologists (5). Whether staffed by attendings or residents, there is a paucity of research into the effects of overnight shift work on diagnostic radiologists (6).

Physiologic effects of overnight work result principally from disturbances of the normal human sleep-wake cycle and the accompanying disruption of the circadian rhythm (7). Metabolic syndrome, obesity, diabetes, adverse effects on cardiovascular health, and even increased rates of cancer are longer term impacts of overnight work (6, 8). In prior research of radiology residents, Bruni *et al.* documented that major discrepancies were more frequently made in the final four hours of a night float shift (9). While other investigations have been unable to reproduce these findings (2), a number of recent studies have demonstrated that diagnostic accuracy is impaired after just eight hours of clinical image interpretation (10–14). Importantly, diagnostic radiology lacks the alerting effects of ambulation, light, and human interaction which are prevalent in specialties like surgery or emergency medicine. These environmental differences may exaggerate fatigue-related errors in overnight radiology (8, 15, 16).

Although prior studies on fatigue in radiology interpretation (10–14) have examined diagnostic accuracy and the role of changes in vision as the day progresses (i.e., reduced ability to accommodate or focus after a long day of work) (10, 11), none to date have examined whether the way radiologists examine images changes as they become fatigued or whether these changes could in part explain how or why errors are made. Therefore, the purpose of our study was to employ eye-tracking technology to assess the effect of overnight shifts (ONS) on radiologist fatigue, visual search pattern, and diagnostic performance.

## Materials and Methods

Institutional review board approval for this experimental study was obtained, and was compliant with the Health Insurance Portability and Accountability Act (HIPAA). Each subject provided written informed consent prior to the first eye-tracking session. The resident radiologists were compensated \$100 for participation in the study. The faculty received no compensation.

## Participants and Shifts

A total of 12 radiologists (5 faculty and 7 residents) each completed two eye-tracking sessions in a counter-balanced fashion. One session occurred near the end of a normal 9-hour work-day (shift: 8 AM – 5 PM, session at 4 PM; “not fatigued”) and another in the morning following a 9-hour overnight shift (shift: 10 PM – 7 AM, session at 8 AM;

“fatigued”). Both the faculty and residents were participating in overnight shift work as part of their normal clinical duties.

### **Fatigue Measure: Swedish Occupational Fatigue Inventory Survey**

Radiologists completed the Swedish Occupational Fatigue Inventory (SOFI) survey at the beginning of each session. SOFI is a robust survey tool which captures both objective (physical) and subjective (mental) fatigue, has substantial validation in prior literature (17, 18) and has been used in radiology fatigue studies (11, 12). It was developed specifically for measuring work-related fatigue (17, 18). SOFI consists of 25 expressions which are divided into five latent subscales: lack of energy, physical exertion, physical discomfort, lack of motivation, and sleepiness (18). The physical exertion and physical discomfort subscales are considered physical fatigue factors, while the lack of motivation and sleepiness subscales are considered the mental fatigue factors. The lack of energy subscale correlates with both physical and mental fatigue, and with the four other subscales, suggesting that it may be a common factor in all human fatigue.

### **Study Design**

During each of the two sessions radiologists viewed 20 bone radiographs consisting of normal and abnormal exams. There was no limit on the amount of time radiologists could spend viewing each case. These bone radiographs have been used in previous studies, and we therefore have data on how difficult the fractures are to detect (11, 19). Fifteen of the cases had a single moderate to subtle fracture, and five had no fracture. In some cases, the fractures were visible in multiple views. The study included wrist, hand, ankle, foot, long bones, shoulder and rib examinations. Viewing time, diagnostic confidence, and eye-tracking search data were recorded for each radiograph for each participant. Diagnostic confidence was recorded by asking the participant to rate their level of confidence (high, medium, low) after reporting an interpretation for each case. They were asked to identify the type and location of all fractures reported.

### **Eye-Tracking & Image Display Apparatus**

A Tobii T60 XL Eye Tracker (Tobii Technology, Inc. Stockholm, Sweden) was used to conduct the study. The T60 has the eye-tracking sensors integrated into a high resolution (1920 × 1200 pixels: 16:10 aspect ratio) 24-inch TFT wide screen monitor with an NVIDIA Quadro 4000 video driver. It was calibrated to the DICOM Grayscale Standard Display Function (GSDF). It uses bright/dark pupil tracking with a sampling rate of 60 Hz and has a manufacturer reported accuracy of 0.4 deg. Participants were seated approximately 40 cm from the display with their eye level approximately at the center. No chin rest was used but the chair was placed in a set position and participants were instructed to maintain a constant but relaxed position throughout the study. Ambient room lights were set to 45 lux.

Prior to the start of data collection, participants were calibrated using the Tobii calibration software which successively presents five calibration dots on a plain, light gray background, and the participant focused his/her visual attention on the center of the dot. The eye tracker measures characteristics of the eyes and uses an internal, physiological 3D eye model to calculate gaze data. All participants calibrated successfully and were observed (by

periodically rechecking calibration) to maintain calibration prior to viewing each radiograph throughout the sessions.

## Data Analysis

The results from the Swedish Occupational Fatigue Inventory (SOFI) were analyzed using an Analysis of Variance (ANOVA) with fatigued versus not fatigued and faculty versus residents as independent variables and the five latent SOFI variables as the dependent variables. Correlation coefficients were calculated between age and the five SOFI variables.

The eye tracking data were analyzed with respect to certain key parameters including total viewing time, time to first fixate the fractures, fixation durations, and dwell times associated with true and false, positive and negative decisions. For each, a repeated measures Analysis of Variance (ANOVA) was used with fatigue versus non-fatigued as the within subjects measure and resident versus faculty as the between subjects measure. The system samples eye positions every 1/60 of a second to generate raw x-, y- coordinate eye-position data. Fixations are formed by grouping x- and y- coordinates of the raw data using a running mean distance calculation having a 0.58 radius threshold. Dwell time can be calculated for each fixation, summed across fixations, then associated with a given region of interest or location in the stimulus image (k) and the type of decision (true or false, positive and negative) associated with a given location.

The eye tracking data were analyzed with respect to the key parameters noted above. For each one a repeated measures Analysis of Variance (ANOVA), with Protected Least Squares Difference post hoc tests as appropriate, was used with fatigue vs non-fatigued as the within subjects measure and resident vs faculty as the between subjects measure. Confidence data were analyzed using the OB-DBM MRMC ROC (Obuchowski-Rockette, Dorfman-Berbaum-Metz Multi-Reader Multi-Case Receiver Operating Characteristic) technique (20–22). Data are presented as mean  $\pm$  standard deviation. Significance was set at  $p < 0.05$ .

## Results

There were 12 participants in the study, 5 faculty (41.7%) and 7 residents (58.3%); 1 female (8.3%) and 11 males (91.7%). Mean age of the faculty was  $37.80 \pm 2.86$  years (range = 36–43) and mean age of the residents was  $30.29 \pm 2.13$  years (range = 28–33).

SOFI results demonstrated worsening on all 5 variables (lack of energy, physical exertion, physical discomfort, lack of motivation, sleepiness) following ONS ( $p < 0.01$ ) (Table 1).

There was a significant difference overall between fatigued and not fatigued diagnostic performance ( $p = 0.031$ ). The individual area under the curve (AUC) values are presented in Table 2. Overall faculty performed higher in both fatigued and non-fatigued conditions than the residents, but the difference was not significant ( $p = 0.225$ ). Across all participants ( $n = 12$ ) and images ( $n = 5$  no fracture, 15 fracture), there were 108 true positive (TP), 72 false negative (FN), 7 false positive (FP) and 53 true negative (TN) decisions following ONS vs 124 TP, 56 FN, 4 FP and 56 TN when not fatigued.

Analysis of the eye tracking data regarding total viewing time per case was performed. There was a significant difference as a function of fatigue ( $p < 0.0001$ ) and as a function of faculty versus resident ( $p = 0.018$ ). Overall total viewing time was longer when fatigued ( $35.9 \pm 25.8$  seconds) than when not fatigued ( $24.8 \pm 16.3$ ); and the effect was more pronounced with residents (mean fatigued = 39.5; not fatigued = 25.1) than faculty (mean fatigued = 30.9; not fatigued = 24.2) (Figure 1A).

The second parameter for analysis of the eye tracking data was total number of fixations generated during the search of each image (Figure 2, Figure 1B). There was a significant difference ( $p < 0.0001$ ) when comparing fatigue versus non-fatigue states, and when comparing faculty versus resident ( $p = 0.009$ ). Overall more fixations were generated when readers were fatigued ( $104.6 \pm 68.1$ ) than not ( $65.4 \pm 34.5$ ); the effect was more pronounced for residents (mean fatigued = 118.2; not fatigued = 61.9) than faculty (mean fatigued = 85.5; not fatigued = 70.2).

The average duration of the individual fixations (Figure 1C) differed as a function of fatigue ( $p = 0.0005$ ) but not resident vs faculty ( $p = 0.2795$ ). Overall, mean fixation duration when fatigued was greater (mean =  $0.34 \pm 0.06$ ; faculty = 0.33, residents = 0.34) than when not fatigued (mean =  $0.32 \pm 0.08$ ; faculty = 0.31, residents = 0.32).

Time to first fixate on the fracture target (Figure 1D) differed significantly ( $p < 0.0001$ ) as a function of fatigue and resident versus faculty ( $p = 0.0069$ ). It took longer to first fixate the fractures when fatigued (mean =  $14.9 \pm 7.02$  seconds) than not fatigued (mean =  $11.1 \pm 5.7$  seconds) and the effect was more pronounced in residents (mean fatigued = 15.5; not fatigued = 11.8) than faculty (mean fatigued = 14.0; not fatigued = 10.2).

Dwell times associated with the fractures and decisions made (TP, FN, FP) were analyzed. Data were analyzed for each decision using an unpaired t-test. The average dwell time associated with the TP decisions was significantly longer ( $t = 2.26$ ,  $p = 0.025$ ) when fatigued (mean = 839.34 msec + 419.15) than when not (mean = 715.63 msec + 410.69). The average dwell time associated with the FP decisions did not differ significantly ( $t = 0.96$ ,  $p = 0.362$ ) when fatigued (mean = 1065.57 msec + 355.91) than when not (mean = 872.25 msec + 235.67). The average dwell time associated with the FN decisions was significantly shorter ( $t = 2.06$ ,  $p = 0.042$ ) when fatigued (mean = 547.03 msec + 375.64) than when not (mean = 689.30 msec + 403.33).

## Discussion

Following overnight shifts, radiologists were more fatigued, demonstrated worse diagnostic performance with decreased confidence, demonstrated a 45% increase in view time per case with a 60% increase in gaze fixations, and a 34% increase in time to fixate on the fracture. It also took longer to render a positive (true or false) decision when fatigued, although due to the relatively low FP rate, it did not reach statistical significance for these decisions but did for TPs. Interestingly, the time associated with FNs were shorter when fatigued compared to not fatigued. Although further research is needed to explore these differences, our results suggest that a fatigued radiologist takes *more time* to either process the fracture information

or decide whether there is enough evidence to report the possible features detected as an actual fracture. It also takes more time to render the *wrong* decision about a fracture being present (false positive). On the other hand, fatigued radiologists spend *less time* on true fractures that attract less attention and thus end up being not reported (FNs).

Although small, our study demonstrates the significant alterations in fundamental performance that can occur following overnight shift work. Previous radiology and non-radiology studies have shown that fatigue from overnight shift work can impact an individual's performance (2, 7, 23, 24). We corroborate and quantify these alterations with eye-tracking technology and provide insight into the specific mechanistic changes that occur when a fatigued radiologist views individual studies. With the increase in afterhours radiology volumes and the commensurate increase overnight radiology staffing, this issue becomes more important for schedule optimization and error minimization (1–3). Perceptual errors (the failure to detect an abnormality) account for the majority of errors in radiology and can result from faulty visual search patterns (25, 26). The authors believe that once we objectively document fatigue-related changes in search and recognition, we can introduce environmental changes to reverse these effects, such as bright lights, required ambulation, or even periodic breaks. Further study in this area is necessary.

Since the first eye-tracking study in medical imaging was conducted in 1963 by Llewellyn-Thomas and Lansdown (27), multiple studies have used eye-tracking technology to examine a variety of parameters in visual search including the effect of display parameters on diagnostic accuracy (28), the relation of search patterns to error occurrence (29, 30), and the variation between novice and expert viewers (31). A recent systematic review on visual search patterns relating to diagnostic expertise identified a total of 22 relevant articles published from 1994 to 2015 and describes a 2-part global-focal model of visual search (32). The global-focal search model distinguishes between a relatively fast impression that signals abnormalities in two-dimensional images and then a slower more refined visual search pattern requiring more attention and focused mental effort. Studies providing evidence for the global search pattern found that expert radiologists can correctly identify abnormal images in a fraction of a second (33). Eye tracking parameters are related to experience, with expertise resulting in decreased total view time, decreased time to first fixate on the abnormality, and decreased total fixations (32). In our study, following overnight shift work both more experienced readers (faculty) and less experienced readers (residents) demonstrated an increased number and duration of fixations, as well as an increased time to fixate on the fracture. In effect, in their fatigued states, faculty had eye tracking parameters more characteristic of non-fatigued residents.

Total view time per case increased 45% in our study following overnight shifts. Interestingly, Burling *et al.* found that radiologists spend less time interpreting CT colonography examinations as they near the end of a work day, with the last five cases interpreted on average 29% faster than the first five cases (34). This increase in interpretive speed at the end of a shift indicates reduced thoroughness at the end of a long reading period, possibly secondary to decreased image coverage/gaze volume. Prior work (35) suggests that the fatigued radiologist may neglect a portion of the image with resultant search errors. Unlike radiologists near the end of the work day in the Burling et al study, our study showed that



radiologists spent *more* time on each case; this difference may reflect inherent differences in night-shift fatigue (sleepiness) versus end-of-day fatigue (impatience), or could be the result of our laboratory setting (subjects spent longer on cases because they knew they were being monitored). Despite this 45% increase in time-per study, the diagnostic accuracy of our subjects decreased. However, in a real world reading room setting, radiologists may be forced to continue reading at a quicker pace despite fatigue. We theorize that this could further decrease diagnostic accuracy. This is an area for further investigation.

Resident physicians have less cumulative radiologic work experience compared to attending physicians and the effects of fatigue were more pronounced in residents. One study showed that toward the end of a day shift, a decrease in detection rate was greater among residents than attending radiologists (11). Studies have also shown that experienced radiologists spend less time inspecting images than do those with less clinical experience (10). The majority of evening and weekend radiology imaging interpretations at academic institutions are provided by residents or fellows (36). Almost every radiology residency program has overnight call for their residents, ranging from 24 hour call shifts to consecutive 10–14 hour night float shifts (37). Thus, compared to attending physicians, resident physicians spend more time viewing images and likely experience a greater strain on their oculomotor system, resulting in asthenopia or eyestrain (11). We theorize that the deteriorated performance by residents following ONS work is due to cumulative lack of experience, coupled with greater impact of fatigue and eye strain.

Our eye-tracking data clearly demonstrate that a number of key parameters such as total time, number of fixations, fixation duration and time to first fixate the fracture are all impacted when radiologists are fatigued. It also seems (Figure 2) that the search patterns at a more global level seem to be changed in a more qualitative way. Readers seem to be less organized in their search strategies and engage in more “idling” (simply staring at a location in the image) more when fatigued. It is more difficult, however, to quantify these types of differences as opposed to the more discrete search and fixation patterns that we did analyze. This type of characterization could be an interesting avenue for research in the future.

## Limitations

Our study has limitations. First, we sampled a relatively small group of radiologists with relatively uniform demographic characteristics predominately young ( 43 years old) and male (11/12 individuals). There is known gender, age, and personality variability in coping with overnight shift work, and we may fail to capture such variation. Second, we examined our radiologists in an observed laboratory setting. This may have influenced their behavior in terms of image review – resulting in a longer examination of studies (although the longer time did not translate into better performance) than may have occurred if the subjects’ performance was not being monitored. We only looked at radiograph fracture cases, and the effect of fatigue on eye tracking patterns of cross sectional imaging or other radiographic pathology may differ.

## Acknowledgments

Support for this study originated from NIH/NIBIB 5R01EB004987-09 (PI: Dr. Elizabeth Krupinski). The authors declare no conflict of interest. There was no direct monetary industry funding for this project.

## References

1. Carroll T. Trends in on-call workload in an academic medical center radiology department 1998–2002. *Academic Radiology*. 2003; 10:1312–20. [PubMed: 14626306]
2. Hanna TN, Loehfelm T, Khosa F, Rohatgi S, Johnson JO. Overnight shift work: factors contributing to diagnostic discrepancies. *Emerg Radiol*. 2015
3. Hanna TN, Shekhani H, Lamoureux C, Mar H, Nicola R, Sliker C, et al. Emergency Radiology Practice Patterns: Shifts, Schedules, and Job Satisfaction. *J Am Coll Radiol*. 2017; 14(3):345–52. [PubMed: 27927590]
4. Towbin AJ, Iyer SB, Brown J, Varadarajan K, Perry LA, Larson DB. Practice policy and quality initiatives: decreasing variability in turnaround time for radiographic studies from the emergency department. *Radiographics : a review publication of the Radiological Society of North America, Inc*. 2013; 33(2):361–71.
5. Ruchman RB, Jaeger J, Wiggins EF 3rd, Seinfeld S, Thakral V, Bolla S, et al. Preliminary radiology resident interpretations versus final attending radiologist interpretations and the impact on patient care in a community hospital. *AJR American journal of roentgenology*. 2007; 189(3):523–6. [PubMed: 17715095]
6. Rohatgi S, Hanna TN, Sliker CW, Abbott RM, Nicola R. After-Hours Radiology: Challenges and Strategies for the Radiologist. *AJR American journal of roentgenology*. 2015; 205(5):956–61. [PubMed: 26496543]
7. Berger F. Impact of Shift Work on the Health and Safety of Nurses and Patients. *Clinical journal of oncology nursing*. 2005; 10(4):465–72.
8. Smith M, Eastman C. Shift work: health, performance and safety problems, traditional countermeasures and innovative management strategies to reduce circadian misalignment. *Nature and Science of Sleep*. 2012; 4:111–32.
9. Bruni S, Bartlett E, Yu E. Factors Involved in Discrepant Preliminary Radiology Resident Interpretations of Neuroradiological Imaging Studies: A Retrospective Analysis. *AJR: American Journal of Roentgenology*. 2012; 198:1367–74. [PubMed: 22623550]
10. Krupinski EA, Berbaum KS. Measurement of visual strain in radiologists. *Acad Radiol*. 2009; 16(8):947–50. [PubMed: 19406673]
11. 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. [PubMed: 20816631]
12. Krupinski EA, Berbaum KS, Caldwell RT, Schartz KM, Madsen MT, Kramer DJ. Do long radiology workdays affect nodule detection in dynamic CT interpretation? *J Am Coll Radiol*. 2012; 9(3):191–8. [PubMed: 22386166]
13. Krupinski EA, Berbaum KS, Schartz KM, Caldwell RT, Madsen MT. The Impact of Fatigue on Satisfaction of Search in Chest Radiography. *Acad Radiol*. 2017
14. Taylor-Phillips S, Elze MC, Krupinski EA, Dennick K, Gale AG, Clarke A, et al. Retrospective review of the drop in observer detection performance over time in lesion-enriched experimental studies. *J Digit Imaging*. 2015; 28(1):32–40. [PubMed: 25005866]
15. Postnova SRP, Postnov DD. Adaption to Shift Work: Physiologically Based Modeling of the Effects of Lighting and Shifts' Start Time. *PLoS one*. 2013; 8(1)
16. Pallesen S, Bjorvatn B, Mageroy N, Saksvik I, Waage S, Moen B. Measures to counteract the negative effects of night work. *Scand J Work, Environ Health*. 2010; 36(2):109–20. [PubMed: 20011984]
17. EA, FG. Perceived fatigue during physical work: an experimental evaluation of a fatigue inventory. *International Journal of Industrial Ergonomics*. 1998; 21:14.



18. EA, FG, AK. Perceived quality of fatigue during different occupational tasks. Development of a questionnaire. *International Journal of Industrial Ergonomics*. 1997; 20:14.
19. Krupinski EA, Lund PJ. Differences in time to interpretation for evaluation of bone radiographs with monitor and film viewing. *Acad Radiol*. 1997; 4(3):177–82. [PubMed: 9084774]
20. Hillis SL, Obuchowski NA, Schartz KM, Berbaum KS. A comparison of the Dorfman-Berbaum-Metz and Obuchowski-Rockette methods for receiver operating characteristic (ROC) data. *Stat Med*. 2005; 24(10):1579–607. [PubMed: 15685718]
21. Hillis SL. A comparison of denominator degrees of freedom methods for multiple observer ROC analysis. *Stat Med*. 2007; 26(3):596–619. [PubMed: 16538699]
22. Hillis SL, Berbaum KS, Metz CE. Recent developments in the Dorfman-Berbaum-Metz procedure for multireader ROC study analysis. *Acad Radiol*. 2008; 15(5):647–61. [PubMed: 18423323]
23. Costa G. Shift work and occupational medicine: an overview. *Occupational Medicine*. 2003; 53(2): 83–8. [PubMed: 12637591]
24. Wright K, Bogan R, Wyatt J. Shift work and the assessment and management of shift work disorder (SWD). *Sleep Med Rev*. 2013; 17(1):41–54. [PubMed: 22560640]
25. Donald JJ, Barnard SA. Common patterns in 558 diagnostic radiology errors. *J Med Imaging Radiat Oncol*. 2012; 56(2):173–8. [PubMed: 22498190]
26. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol*. 1978; 13(3):175–81. [PubMed: 711391]
27. Thomas EL, Lansdown EL. Visual Search Patterns of Radiologists in Training. *Radiology*. 1963; 81:288–91. [PubMed: 14053771]
28. Krupinski EA, Roehrig H. The influence of a perceptually linearized display on observer performance and visual search. *Acad Radiol*. 2000; 7(1):8–13. [PubMed: 10645452]
29. Kundel HL, Nodine CF, Krupinski EA. Searching for lung nodules. Visual dwell indicates locations of false-positive and false-negative decisions. *Invest Radiol*. 1989; 24(6):472–8. [PubMed: 2521130]
30. Nodine CF, Mello-Thoms C, Kundel HL, Weinstein SP. Time course of perception and decision making during mammographic interpretation. *AJR American journal of roentgenology*. 2002; 179(4):917–23. [PubMed: 12239037]
31. Nodine CF, Kundel HL, Lauver SC, Toto LC. Nature of expertise in searching mammograms for breast masses. *Acad Radiol*. 1996; 3(12):1000–6. [PubMed: 9017014]
32. van der Gijp A, Ravesloot CJ, Jarodzka H, van der Schaaf MF, van der Schaaf IC, van Schaik JPJ, et al. How visual search relates to visual diagnostic performance: a narrative systematic review of eye-tracking research in radiology. *Adv Health Sci Educ Theory Pract*. 2017; 22(3):765–87. [PubMed: 27436353]
33. Oestmann JW, Greene R, Kushner DC, Bourgooin PM, Linetsky L, Llewellyn HJ. Lung lesions: correlation between viewing time and detection. *Radiology*. 1988; 166(2):451–3. [PubMed: 3336720]
34. Burling D, Halligan S, Altman DG, Atkin W, Bartram C, Fenlon H, et al. CT colonography interpretation times: effect of reader experience, fatigue, and scan findings in a multi-centre setting. *Eur Radiol*. 2006; 16(8):1745–9. [PubMed: 16636802]
35. Roge J, Pebayle T, El Hannachi S, Muzet A. Effect of sleep deprivation and driving duration on the useful visual field in younger and older subjects during simulator driving. *Vision Res*. 2003; 43(13):1465–72. [PubMed: 12767314]
36. Hoffmann JC, Singh A, Mittal S, Peterkin Y, Flug J. Results of a National Radiology Attending Physician Survey: The Effects of In-House Late and Overnight Attending Coverage on Radiology Resident Training. *Curr Probl Diagn Radiol*. 2016; 45(5):304–11. [PubMed: 27020255]
37. Ruutiainen AT, Durand DJ, Scanlon MH, Itri JN. Increased error rates in preliminary reports issued by radiology residents working more than 10 consecutive hours overnight. *Acad Radiol*. 2013; 20(3):305–11. [PubMed: 23452475]

**Sentence Summary**

Following overnight shifts, radiologists were more fatigued, demonstrated worse diagnostic performance with decreased confidence, demonstrated a 45% increase in view time per case with a 60% increase in gaze fixations, and a 34% increase in time to fixate on the fracture.

Author Manuscript

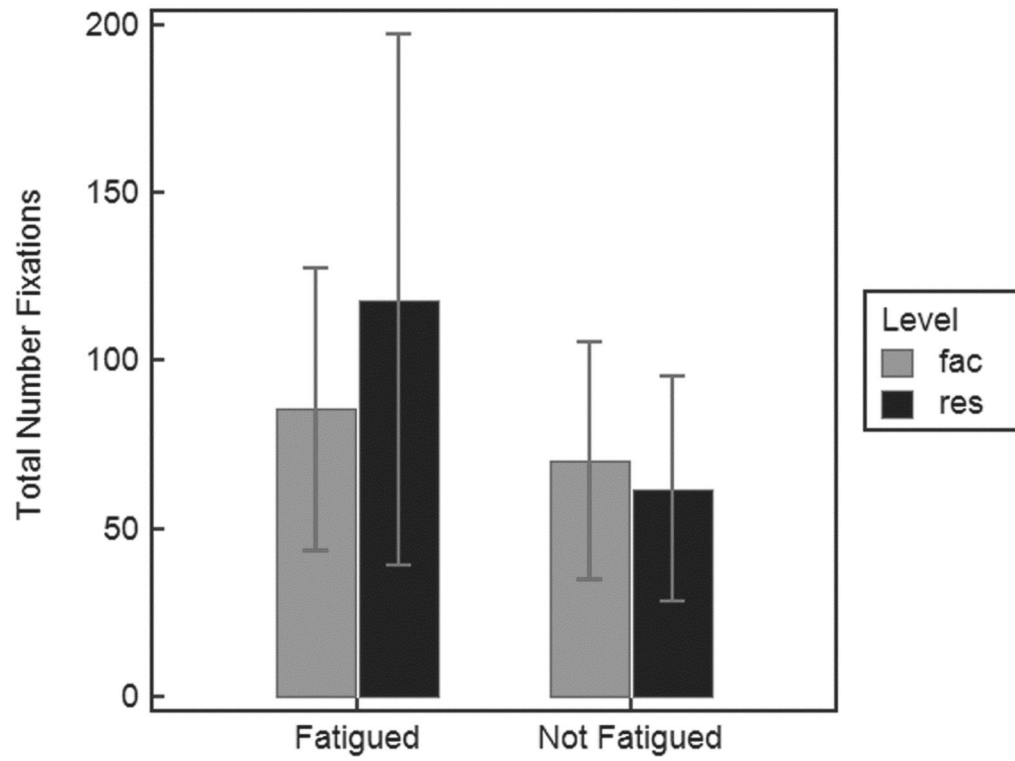
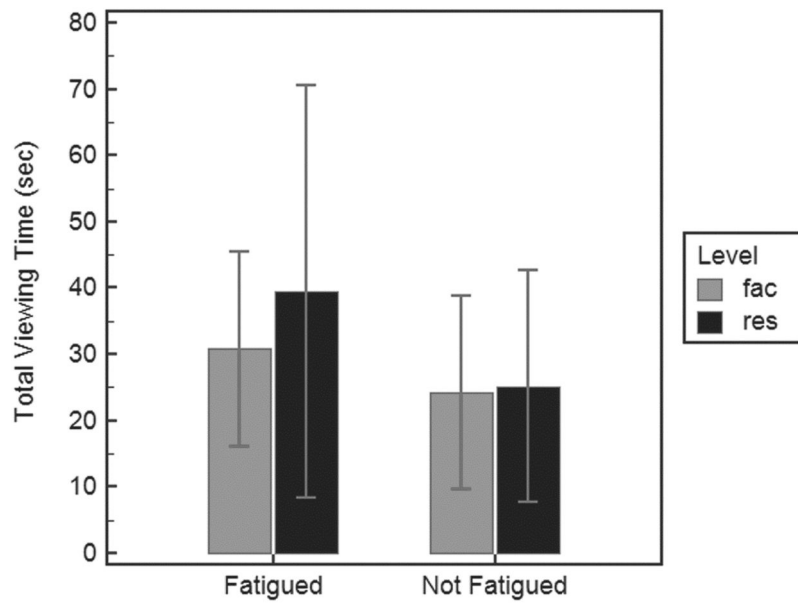
Author Manuscript

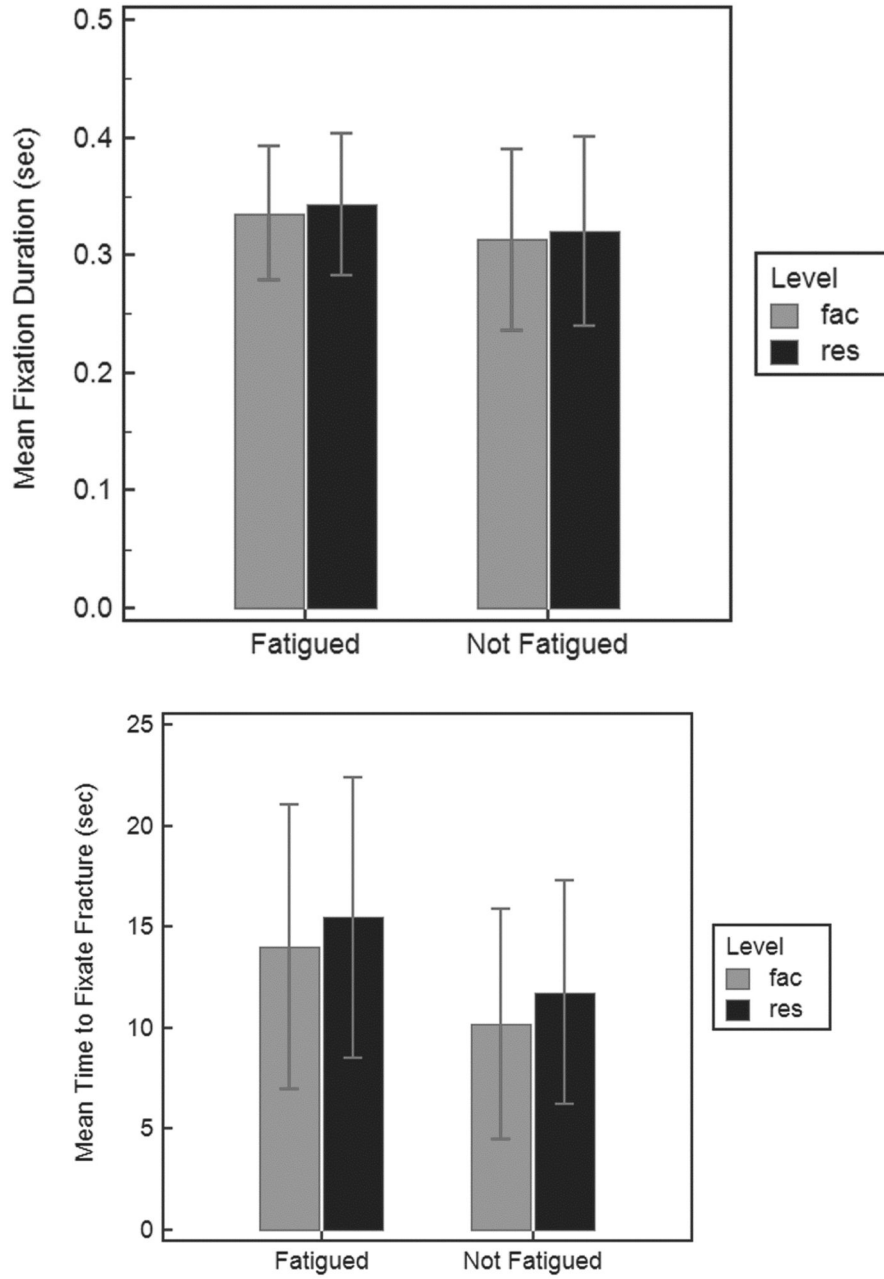
Author Manuscript

Author Manuscript

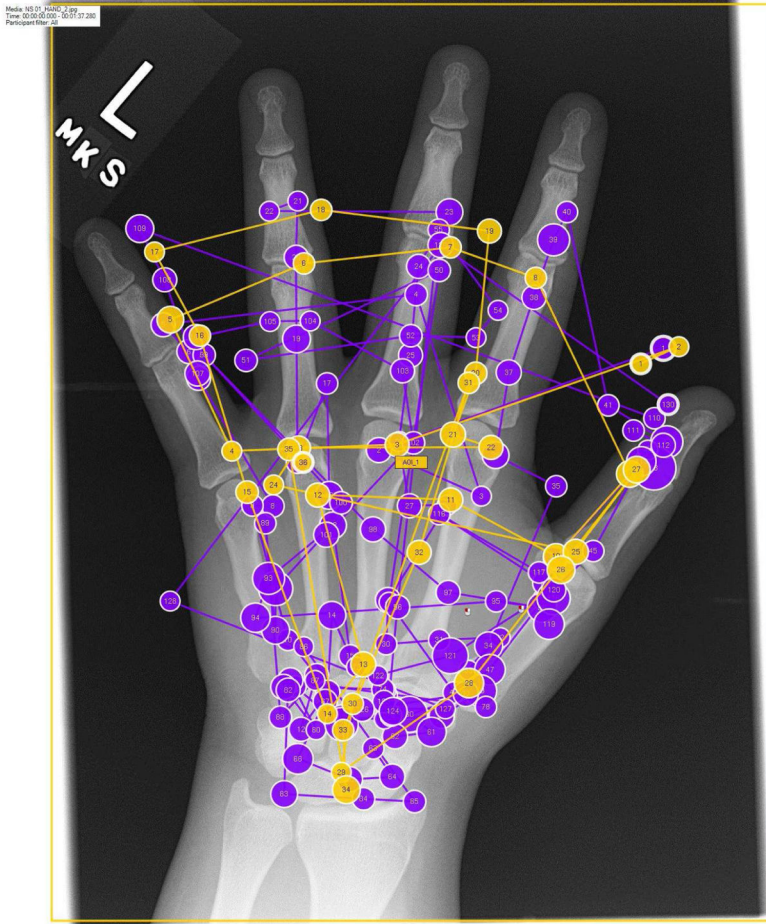
### Take Home Points

- Following overnight shift work, both faculty and resident radiologists demonstrate a lack of energy, increased physical discomfort, lack of motivation, increased sleepiness, and limited tolerance for physical exertion (all  $p < 0.01$ ).
- Compared with interpretations near the end of a daytime shift, following overnight shift work, radiologists demonstrate significantly worse interpretive performance, with an increase in both false positives and false negatives, and mean area under the ROC curve decreasing by 13%.
- Following overnight shift work, mean total viewing time per case increased by 45% ( $p < 0.0001$ ). Total viewing time per case was longer for residents ( $p < 0.05$ ), and the effect of fatigue on lengthening per-case view time was more pronounced with residents.
- Fatigue from overnight shift work resulted in substantial alterations in qualitative and quantitative metrics of visual search patterns: total gaze fixations generated during the search of each image increased 60% and time to first fixate on the fracture increased 34%.

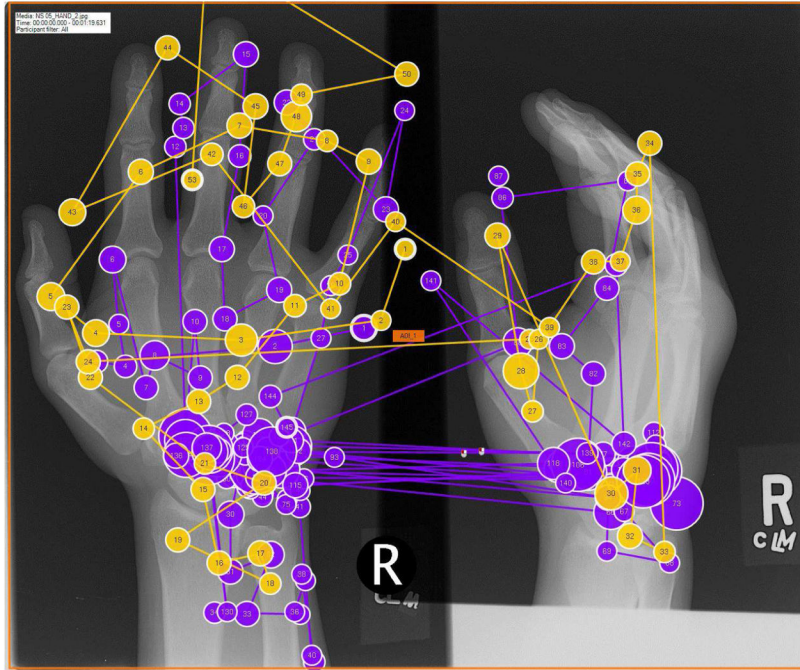




**Figure 1.** Graphical depiction of the changes in the visual search patterns of radiologists near the end of a normal work-day (4 PM during a 9hr 8AM–5PM shift, “Not Fatigued” Session) vs following an overnight shift (8 AM following a 9hr 10PM–7AM shift, “Fatigued” Session). **A)** Mean (with SD bars) total viewing time for faculty (fac) and residents (res) when fatigued vs not fatigued. **B)** Mean (with SD bars) total number of fixations generated for faculty (fac) and residents (res) when fatigued vs not fatigued. **C)** Average duration (with SD bars) of the individual fixations for faculty (fac) and residents (res) when fatigued vs not fatigued. **D)** Average time (with SD bars) to first fixate the fracture targets for faculty (fac) and residents (res) when fatigued vs not fatigued.







**Figure 2.** Visual search patterns of radiologists captured by eye-tracking technology near the end of a normal work-day (4 PM during a 9hr 8AM–5PM shift, “Not Fatigued” Session) vs following an overnight shift (8 AM following a 9hr 10PM–7AM shift, “Fatigued” Session). **A)** Example of one of the normal test images with the search pattern of a single reader in fatigued (purple) and non-fatigued (yellow) states superimposed. Each circle represents a gaze fixation (location where foveal vision is directed). The size of the circle represents fixation dwell time with larger circles reflecting increased dwell times. The lines between the fixations represent saccades, which are rapid eye movements between fixations. Notice the significant increase in gaze fixations in the fatigued state. **B)** Example of one of the abnormal test images (carpal fracture, which is obscured by the eye-tracking pattern). Notice the increase in gaze fixations and dwell time in the carpus during the fatigued state.

**Table 1**

Results from the Swedish Occupational Fatigue Inventory for faculty and residents.

	Faculty		Resident		ANOVA Results			Correlation with Age
	Not Fatigued Faculty	Fatigued Faculty	Not Fatigued Resident	Fatigued Resident	Fatigued vs Not Fatigued	Faculty vs Resident		
<b>Lack of Energy</b>	2.45	5.50	1.21	6.04	F = 30.10	F = 0.238	r = 0.25	
	(1.02)	(1.85)	(1.11)	(2.40)	P < 0.0001	P = 0.631		
<b>Physical Exertion</b>	0.10	1.85	0.14	0.96	F = 11.39	F = 1.22	r = 0.17	
	(0.14)	(1.21)	(0.38)	(1.30)	P = 0.003	P = 0.282		
<b>Physical Discomfort</b>	1.00	2.90	0.32	2.61	F = 9.94	F = 0.54	r = 0.34	
	(1.02)	(2.31)	(0.37)	(2.05)	P = 0.005	P = 0.473		
<b>Lack Motivation</b>	0.85	4.45	0.71	4.14	F = 27.36	F = 0.12	r = 0.07	
	(0.52)	(1.74)	(0.77)	(2.45)	P < 0.0001	P = 0.745		
<b>Sleepiness</b>	1.35	4.70	0.82	5.12	F = 18.49	F = 0.005	r = 0.10	
	(0.95)	(3.23)	(0.73)	(2.69)	P = 0.0003	P = 0.946		

**Table 2**

Results for faculty (top) and residents (bottom) when fatigued vs not fatigued.

Reader	Faculty/Resident	Fatigued	Not Fatigued
1	Faculty	0.907	0.999
2	Faculty	0.870	0.999
3	Faculty	0.935	0.907
4	Faculty	0.678	0.942
5	Faculty	0.927	0.936
<b>Mean</b>	<b>Faculty</b>	<b>0.863</b>	<b>0.957</b>
6	Resident	0.683	0.741
7	Resident	0.859	0.997
8	Resident	0.558	0.800
9	Resident	0.866	0.994
10	Resident	0.621	0.845
11	Resident	0.901	0.996
12	Resident	0.867	0.949
<b>Mean</b>	<b>Resident</b>	<b>0.765</b>	<b>0.903</b>
<b>Mean</b>	<b>Overall</b>	<b>0.806</b>	<b>0.926</b>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript