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