J Digit Imaging (2012) 25:319–324
DOI 10.1007/s10278-011-9439-1

Real-Time Occupational Stress and Fatigue Measurement in Medical Imaging Practice

Elizabeth Krupinski · Bruce I. Reiner

Introduction

While occupational stress and fatigue have been well documented as a source of medical error, current medical practice largely leaves individual healthcare providers to their own device. With the exception of recently established rules limiting the number of hours physicians in training can work (to counteract errors related to sleep deprivation), few regulations exist addressing occupational stress and fatigue in medicine. In some clinical arenas, fatigue countermeasure programs have been studied, but have rarely been implemented on a larger scale [1]. Generally, an individual provider is expected to consistently operate at high levels of efficiency regardless of workload and task complexity.

In addition to physiologic, emotional, and cognitive fatigue encountered throughout medicine, radiologists are also subject to visual fatigue due to prolonged periods of time interpreting complex medical imaging datasets on computers. Collectively, these different types of fatigue have the potential to adversely affect radiologist performance and lead to medical errors. Effective counter-measures require some sort of objective analysis at the point of care, which can record fatigue indicators in a standardized fashion, correlate these measures with baseline data specific to the individual end-user and task being performed, and present real-time feedback for the purpose of performance improvement and education.

Current Stress and Fatigue Measurement Tools and Technologies

Sleep Quality

Although sleep quality is a readily accepted clinical construct, it actually represents a complex phenomenon which is difficult to define and objectively measure. The term “sleep quality” includes quantitative measures including sleep duration, latency, and number of arousals; along with subjective measures such as sleep depth and restfulness [2].

A number of tests are currently available to quantify sleepiness and sleep quality, including both subjective and objective measuring instruments. Subjective measures of sleep quality include the Stanford Sleepiness Scale [3], the Pittsburgh Sleep Quality Index [2], and the Epworth Sleepiness Scale [4].

Objective tests include the Psychomotor Vigilance Task [5] and Continuous Performance Test [6]. The former is a 10-min sustained attention test which is sensitive to sleepiness and alcohol, in which participants are asked to manually respond to scrolling numbers on a computer screen, with dependent variables including median reaction time and frequency of lapse. The latter is a 14-min computer vigilance task, requiring participants to manually respond to alphabetic letters, with dependent variables including errors of commission and omission.

The effects of sleep deprivation in medical practice have been well described among medical trainees with sleep

E. Krupinski
Department of Radiology, Psychology and the Arizona
Telemedicine Program, University of Arizona,
1609 N Warren, Bldg 211, Rm 112,
Tucson, AZ 85724, USA

B. I. Reiner (✉)
Department of Radiology,
Veterans Affairs Maryland Healthcare System,
10 North Greene Street,
Baltimore, MD 21201, USA
e-mail: breiner1@comcast.net

deprivation reported to negatively impact medical care delivery by adversely affecting neurobehavioral and work-related tasks, mood and affect, learning, commission of medical errors, and the health and overall well-being of providers [7, 8]. Another provider group of particular relevance in medical imaging is teleradiology providers, who frequently work during night time hours, and are therefore prone to sleep deprivation through disruption of normal circadian rhythms [9].

Physiologic and Emotional Fatigue

There are a number of surveys that have been developed to assess physiologic and emotional fatigue and/or stress in the workplace, but not all have been widely accepted or validated in different occupations. One exception is the Swedish Occupational Fatigue Inventory (SOFI) that was developed and validated to specifically measure perceived fatigue in work environments [10, 11]. The instrument consists of 20 expressions, evenly distributed on five 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, while 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 results from some recent studies on fatigue that used the SOFI in radiologists are summarized below. Although there are certainly other surveys that could be used in developing tools to monitor and counteract fatigue in radiologists, the SOFI is easy to implement on a computer and has already been shown to be useful in assessing fatigue states in radiologists. Thus, it seems to be a likely candidate for monitoring fatigue. It could perhaps been programmed to appear on a workstation every 4–5 h to be filled in by the radiologist. If certain thresholds in the scored variables are passed, then the computer could advise the radiologist to take a break.

There are objective measures of physiologic fatigue such as monitoring (absolute values and/or changes over time) heart rate, pulse, blood pressure, galvanic skin response, or response times on a reaction task (e.g., hold a ruler at the bottom with two fingers then let go and grasp as soon as possible—increased fatigue reduces reaction time and people either miss grasping it altogether or catch it closer to the top than if they were not fatigued). With modern remote monitoring devices, it may be rather simple for radiologists to wear them or to incorporate them into workstations (if they are willing to be monitored). A less invasive alternative may be to develop a reaction time “video game” that would appear periodically throughout the day. Radiologists would have to perform the designated task (e.g., catch × number of aliens in a set amount of time) in order to continue reading or take a break if they do not.

Visual Fatigue

Eyestrain is caused by additional work that oculomotor systems must perform to maintain accommodation (focus), convergence (single vision), and gaze (directing the fovea). Accommodative asthenopia is caused by strain of the ciliary muscles, whereas muscular asthenopia is caused by strain of the external ocular muscles. Both lead to physical symptoms: blurred or double vision, headaches, and pain in and around the eyes. Therefore, there is not a single potential cause of symptoms, and presence of a symptom may indicate any of several different malfunctions of oculomotor control. Different work environments stress the eyes in different ways. For example, inertial forces acting on the eyes of a pilot are not present in the radiology reading room. Likewise, perceptual activities, such as cine display of a computed tomography (CT) dataset under viewer control, are not found in other work environments. Several kinds of oculomotor fatigue have to be measured to determine what breakdowns in physiological mechanisms underlie the symptoms that radiologists experience.

One of the most obvious manifestations of fatigue with prolonged use of computer displays is eyestrain or asthenopia [12–16]. Some of the perceivable symptoms of visual fatigue are blurred vision, ocular pain, ocular swelling, headaches, and dry eyes [16]. Viewing distance [12], ambient lighting [17, 18], display resolution [19], *mental workload* [20], glare [21, 22], viewing angle [23, 24], and *length of continuous viewing time* [6] all have been shown to increase visual fatigue with video display terminal use. Sanchez-Roman et al. compared two groups of workers, one using computers and the other not [25] and found that 4 h at the computer was sufficient to produce asthenopia.

One of the main issues for establishing metrics for a fatigue tool is how to measure visual fatigue. Many of the general display-induced fatigue studies use subjective rating scales that are certainly useful, but need to be complemented by more objective measures. Spontaneous blink rate and duration have been proposed as measures of visual fatigue [26, 27]. Eye-tracking methods are quite useful for measuring blink rate, gaze, and other aspects of visual search that may be affected by fatigue, but these are rather impractical to implement in daily reading environments.

Other common objective measures of visual fatigue in are accommodation and vergence. Both accommodation and vergence are part of a neural control system designed to keep objects in the visual field focused. There are both negative and positive feedback loops involved in the system to control the amount of accommodation under different stimulus and environmental (e.g., ambient light) conditions, and for the most part they operate without voluntary control. Every person differs in the amount of accommodation triggered by changes in vergence and this amount changes as a function of vergence requirements imposed by external stimuli. In continual near-viewing situations, as when a radiologist interprets radiographic studies on a digital display, the vergence system may be strained leading to increased asthenopia.

In a series of recent studies by Krupinski and Berbaum, we measured visual accommodation and dark vergence as a function of how long (before and after a long day of reading clinical images) radiologists had been reading clinically and correlated those measures with diagnostic accuracy and subjective measures of fatigue. In the first study [28], we hypothesized that the current practice of radiology produces oculomotor fatigue that reduces diagnostic accuracy. We measured visual accommodation of radiologists before and after diagnostic viewing work using an autorefractor that is capable of make multiple measurements of accommodation per second. Three radiologists and three residents focused on a simple geometric (asterisk) target placed at near to far distances while accommodation was measured. The target distances varied from 20 to 183 cm from the eye. The data were collected prior to and after a day of digital diagnostic viewing. Results indicated that accommodation at near distances is significantly worse overall compared to far distances and is significantly worse after a day of digital reading at all distances. Because diagnostic image interpretation is performed at near-viewing distances, this inability to maintain focus on the image could impact diagnostic accuracy. As expected, younger residents had better accommodative accuracy than older radiologists.

In the next study, we measured diagnostic accuracy of fracture detection, visual accommodation, and subjective ratings of fatigue and visual strain before and after a day of clinical reading. Forty attending radiologists and radiology residents viewed 60 bone exams, half with fractures before and after a day of clinical reading. Visual accommodation was measured before and after each reading session. SOFI and oculomotor strain were collected. It was found that diagnostic accuracy was reduced significantly after a day of clinical reading, with an average receiver operating characteristic (ROC) area under the curve (AUC) of 0.885 for reading prior to a day of work and 0.852 for after ($p<0.05$). After a day of image interpretation, error in visual accommodation was greater ($p<0.01$) and subjective fatigue ratings were higher. We concluded that 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 [29].

Skeletal radiographic examinations commonly have images that are displayed statically, so the third study investigated whether diagnostic accuracy for detecting pulmonary nodules in CT of the chest displayed dynamically would be similarly affected by fatigue. Twenty-two radiologists and 22 residents were given two tests searching CT chest sequences for a solitary pulmonary nodule before and after a day of clinical reading. To measure search time, ten lung CT sequences, each containing 20 consecutive sections and a single nodule, were inspected using free search and navigation. To measure diagnostic accuracy, 100 CT sequences, each with 20 sections and half with nodules, were displayed at preset scrolling speed and duration. Accuracy was measured using ROC analysis. Visual strain was measured via dark vergence, an indicator of the ability to keep the eyes focused on the display. Diagnostic accuracy was reduced after a day of clinical reading ($p=0.0246$), but search time was not affected ($p>0.05$). After a day of reading, dark vergence was significantly larger and more variable ($p=0.0098$), reflecting higher levels of visual strain and SOFI were also higher. After their usual workday, radiologists experience increased fatigue and decreased diagnostic accuracy for detecting pulmonary nodules on CT. Effects of fatigue may be mitigated by active interaction with the display [30].

These three studies clearly demonstrate that there are objective measures of visual fatigue that correlate with subjective measures of fatigue and changes in diagnostic accuracy. The key to creating a fatigue countermeasure tool that would rely on these measures to warn radiologists of potentially increasing and negative levels of fatigue is to somehow incorporate versions of these devices either into the reading workstation itself or conveniently in the reading room environment in some other fashion.

Cognitive Fatigue

The literature on the cognitive deficits associated with fatigue has a long history and is still studied intensely today. Cognitive errors (i.e., poor decisions) are the most obvious form of deficit associated with fatigue. For example, Schellekens et al. [31] engaged subjects on two workdays, one with highly demanding information-processing tasks and the other with less-demanding tasks. *Errors increased significantly* after the demanding workday but not in the less demanding workday (both compared to morning or control performance). In addition to the increase in errors, the demanding workday led to shorter reaction times and investment of *less effort* in the assigned task (memory search for word pairs). Errors increased because subjects took shortcuts in their cognitive strategies. If radiologists take shortcuts as they become fatigued, the potential for diagnostic errors increases. Lorist et al. [32] examined the effects of mental fatigue on planning and preparation performance using the “switching paradigm”.

Subjects were either engaged in a set of tasks that did not change from trial to trial, or in a set of tasks that switched cognitive demands every second trial. With fatigue, preparation time declined and decision errors increased significantly in the cognitively demanding "switch" tasks. In radiology, each case may represent a new task that requires preparation and planning on how to approach its interpretation. Why else would clinical history affect the interpretation? [33–40]. If fatigue causes radiologists to take less time processing relevant information and preparing themselves cognitively to interpret a case, diagnostic errors will increase.

Outside distractors such as noise can exacerbate fatigue and error effects [41]. Radiology reading rooms are often hectic places where people talk, move around and create other distractions that are likely to compound fatigue effects in radiologists. Age may also influence fatigue and cognitive functioning to some degree in computer-based work. Czaja and Sharit [42] engaged subjects ranging from 25 to 70 years in age in a computer-based task and found that the older subjects tended to have more errors, longer response times, and greater subjective reports of fatigue and task difficulty. Horowitz et al. [43] studied visual search performance across the circadian cycle and found that reduced alertness (low points in the circadian cycle) resulted in longer reaction times and more errors. Fatigued subjects did not modify their search behaviors as a function of their fatigue (i.e., they should have slowed search down and been more careful with increased fatigue), resulting in "reckless" search and more errors. Since a main component of the radiological interpretation process is visual search, there is a very strong possibility that decrements in cognitive performance and increased error rates will occur as radiologists become fatigued.

Researchers in radiology have studied variation in diagnostic performance over the course of an ordinary professional workday [44, 45]. Gale et al. [45] found a significant morning to afternoon drop in sensitivity in the detection of pulmonary nodules in chest radiographs. Brogdon et al. [44] 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. More recently, reader accuracy at different times during the day has been examined with mixed results. Taylor-Phillips et al. [46] looked data from the UK Breast Screening Programme. There were nearly 200,000 cases and they attempted to relate accuracy to time of day and reading time and they found that recall rates varied with time of day but not in the same way for the individual readers. Some readers had lower recall rates in the afternoon, but others did not. Recall rates generally declined with increased reading time (i.e., recall rates were lower around lunch and the end of the day), but results varied considerably. The sample was rather noisy and they could not document anything significant beyond a possible trend. This study did not however directly examine fatigue or conduct a controlled study in which readers read a dedicated set of cases before and after a day of clinical reading. Al-s'adi et al. [47] found that breast lesion detection varies with time of day, but there was no particular time of day that had a significant effect. Radiologists were recruited at a national meeting to read mammograms during one of four reading times (7:00–10:00, 10:00–13:00, 13:00–16:00, 16:00–20:00). There were no significant differences in sensitivity, specificity or ROC AUC as a function of time of day. Limitations included the fact that readers only participated in a single session and they could choose the time of their participation, possibly choosing a time of higher performance or motivation.

One recent study found that working in teams may actually counteract some effects of decision fatigue [48]. Baranski et al. examined the effects of 30 h of sleep loss and long hours of cognitive work on performance in a distributed team decision-making environment. They had 16 teams each comprised of four members. Three members made threat assessments in a military surveillance task and then had to forward their decisions to the leader. The leader made the final decision on behalf of the team. They found that sleep loss had an antagonistic effect on team decision-making accuracy and decision time. However, it was mediated by being part of a team compared to performing the same task individually. They concluded that there was evidence of a "motivational gain" effect in the sleepy teams. For radiology, it suggests that reading alone in an isolated environment may be more detrimental than working in an environment where collaborations or discussions about cases can take place.

As noted previously, the choice of methods to determine whether radiologists are too cognitively fatigued to continue reading and should take a break can be determined via and number of existing validated decision assessment tools. The challenge is going to be implementing them into the routine workflow in such a way that they do not create more of a burden than intended and truly do measure fatigue in a valid and reliable fashion.

Conclusion

A number of existing measurement tools and technologies are currently available which can be applied to medical imaging for the assessment of occupational stress and fatigue with the goals of improving quality, safety, and clinical outcomes. In addition to physiologic, emotional, and cognitive fatigue which are encountered throughout all medical disciplines, medical imaging professionals are particularly susceptible to visual fatigue due to the data-intensive and image-centric nature of radiology.

Successful integration of these fatigue measurement tools into existing medical imaging and information systems technologies requires workflow optimization, interoperability, and data standardization. The ability to standardize occupational stress and fatigue data provides a mechanism for meta-analysis and creation of best-practice guidelines; related to technology usage, employee and patient safety, and clinical outcomes. The ultimate goal is to simultaneously improve quality and performance for all end-users, in a manner which can take into account individual end-users' attributes, needs, and preferences.

References

1. Scott LD, Hofmeister N: Implementing a fatigue countermeasures program for nurses. J Nurs adm 40:233–240, 2010
2. Buysse DJ, Reynolds III, CF, Monk TH, et al: The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 28:193–213, 1989
3. Hoddes E, Dement W, Zarcone V: Quantification of sleepiness: a new approach. Psychophysiology 10:431–436, 1973
4. Johns MW: A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 24:540–545, 1991
5. Dinges DF, Powell JW: Microcomputer analyses of performance on a portable simple visual RT task during sustained operations. Behav Res Method Instrum Compt 17:652–655, 1985
6. Epstein JN, Erkanli A, Connors CK, et al: Relations between continuous performance test measures and ADHD behaviors. J Abnorm Child Psychol 31:543–554, 2003
7. Owens JA: Sleep loss and fatigue in medical training. Curr Opin Pulm Med 7:411–418, 2001
8. Baldwin Jr, DC, Daugherty SR: Sleep deprivation and fatigue in medical residency training: results of a national survey of first and second-year residents. Sleep 27:217–223, 2004
9. Bradley WG: Offshore teleradiology. J Am Coll Radiol 1:244–248, 2004
10. Ahsberg E: Dimensions of fatigue in different workplace populations. Scand J Psych 41:231–241, 2000
11. Ahsberg E, Gamberale F, Gustafsson K: Perceived fatigue after mental work: an experimental evaluation of a fatigue inventory. Ergonomics 43:252–268, 2000
12. Ankrum DR. Viewing distance at computer workstations. Work-Place Ergonomics, 1996; Sept/Oct: 10–12
13. Murata A, Uetake A, Otsuka M, Takasaw Y: Proposal of an index to evaluate visual fatigue induced during visual display terminal tasks. Int J Hum-Comput Interact 13:305–321, 2001
14. Watten RG, Lie I: Time factors in VDT-induced myopia and visual fatigue: an experimental study. J Hum Ergol 21:13–20, 1992
15. Watten RG, Lie I, Birketvedt O: The influence of long-term visual near-work on accommodation and vergence: a field study. J Hum Ergol 23:27–39, 1994
16. Yunfang L, Wenjing W, Bingshuang H, Changji L, Chenglie Z: Visual strain and working capacity in computer operators. Homeost Heal Dis 40:27–29, 2000
17. Duffy VG, Chan AHS: Effects of virtual lighting on visual performance and eye fatigue. Hum Fact Ergon Manuf 12:193–209, 2002
18. Siegel EL, Reiner B: Image workflow. In: Dreyer KJ, Mehta A, Thrall JH Eds. PACS a guide to the digital revolution. Springer, New York, 2002, pp 161–190
19. Ziefle M: Effects of display resolution on visual performance. Hum Factors 40:554–568, 1998
20. Backs RW, Walrath LC: Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. Appl Ergon 23:243–254, 1992
21. Dillon A: Reading from paper versus screens: a critical review of the empirical literature. Ergonomics 35:1297–1326, 1992
22. Dillon TW, Emurian HH: Some factors affecting reports of visual fatigue resulting from use of a VDU. Comput Hum Behav 12:49–59, 1996
23. Sommerich CM, Joines SMB, Psihogios JP: Effects of computer monitor viewing angle and related factors on strain, performance, and preference outcomes. Hum Factors 43:39–55, 2001
24. Turville KL, Psihogios JP, Ulmer TR, Mirka GA: The effects of video display terminal height on the operator: a comparison of the 15″ and 40″ recommendation. Appl Ergon 29:239–246, 1998
25. Sanchez-Roman FR, Perez-Lucio C, Juarez-Ruiz C, Velez-Zamora NM, Jimenez-Villarruel M: Risk factors for asthenopia among computer terminal operators. Salud Publica Mex 38:189–196, 1996
26. Jachinski W, Bonkacker M, Alshuth E: Accommodation, convergence, pupil diameter and eye blinks at a CRT display flickering near fusion limit. Ergonomics 39:152–164, 1996
27. Kaneko K, Sakamoto K: Spontaneous blinks as a criterion of visual fatigue during prolonged work on visual display terminals. Percept Mot Ski 92:234–250, 2001
28. Krupinski EA, Berbaum KS: Measurement of visual strain in radiologists. Acad Rad 16:947–950, 2009
29. Krupinski EA, Berbaum KS, Caldwell RT, Schartz KM, Kim J: Long radiology workdays reduce detection and accommodation accuracy. J Am Coll Radiol 7:698–704, 2010
30. Krupinski EA, Berbaum KS, Caldwell RT, Schartz KM, Madsen MT, Kramer DJ. Do long radiology workdays impact nodule detection in dynamic CT interpretation? J Am Coll Radiol 2011 (in press)
31. Schellekens JM, Sijtsma GJ, Vegter E, Meijman TF: Immediate and delayed after-effects of long lasting mentally demanding work. Biological Psych 53:37–56, 2000
32. Lorist MM, Klein M, Nieuwenhuis S, DeJong R, Mulder G, Meijman TF: Mental fatigue and task control: planning and preparation. Psychophysiology 37:614–625, 2000
33. Berbaum KS, El-Khoury G, Franken EA, Kathol M, Montgomery W, Hesson W: The impact of clinical history on the detection of fractures. Radiology 168:507–511, 1988
34. Berbaum KS, Franken Jr, EA, Anderson KL, Dorfman DD, Erkonen WE, Farrar GP, Geraghty JJ, Gleason TJ, MacNaughton ME, Phillips ME, Renfrew DL, Walker CW, Whitten CG, Young DC: The influence of clinical history on visual search with single and multiple abnormalities. Investig Radiol 28:191–201, 1993
35. Berbaum KS, Franken EA, Dorfman DD, Barloon T, Ell SR, Lu CC, Smith W, Abu-Yousef MM: Tentative diagnoses facilitate the detection of diverse lesions in chest radiographs. Investig Radiol 2l:532–539, 1986
36. Berbaum KS, Franken EA, Dorfman DD, Barloon T: Influence of clinical history upon detection of nodules and other lesions. Investig Radiol 23:48–55, 1988. 60–61
37. Berbaum KS, Franken Jr, EA, Dorfman DD, Lueben K: Influence of clinical history on perception of abnormalities in pediatric radiographs. Acad Radiol 1:217–223, 1994
38. Berbaum KS, Franken Jr, EA, Dorfman DD, Miller EM, Caldwell R, Kuehn DM, Berbaum ML: Role of faulty visual search in the satisfaction of search effect in chest radiology. Acad Radiol 5:9–19, 1998

39. Berbaum KS, Franken EA, El-Khoury GY: Impact of clinical history on radiographic detection of fractures: a comparison of radiologists and orthopedists. Am J Roentgenol 153:1221–1224, 1989
40. Berbaum KS, Franken EA, Smith WL: The effect of comparison films upon resident interpretation of pediatric chest radiographs. Investig Radiol 20:124–128, 1985
41. Takahashi K, Sasaki T, Hosokawa T, Kurasaki M, Saito K: Combined effects of working environmental conditions in VDT work. Ergonomics 44:562–570, 2001
42. Czaja SJ, Sharit J: Age differences in the performance of computer-based work. Psychol Aging 8:59–67, 1993
43. Horowitz TS, Cade BE, Wolfe JM, Czeisler CA: Searching night and day: a dissociation of effects of circadian phase and time awake on visual selective attention and vigilance. Psych Science 14:549–557, 2003
44. Brogdon BG, Kelsey CA, Moseley RD: Effect of fatigue and alcohol on observer perception. Am J Roentgenol 130:971–974, 1978
45. 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. Elsevier, London, 1984
46. Taylor-Phillips S, Clarke A, Wallis M, Wheaton M, Duncan A, Gale AG. The time course of cancer detection performance. Proc SPIE Med Imag 7966:796605-1-8, 2011
47. Al-s'adi M, McEntee MF, Ryan E. Time of day does not affect radiologists' accuracy in breast lesion detection. Proc SPIE Med Imag 7966:796608-1-7, 2011
48. Baranski JV, Thompson MM, Lichacz FM, McCann C, Gil V, Pasto L, Pigeau RA: Effects of sleep loss on team decision making: motivational loss or motivational gain? J Hum Factors Ergon Soc 49:646–660, 2007