

A Study of Radiologists Viewing Multiple Computed Tomography Examinations Using an Eyetracking Device

David V. Beard, R. Eugene Johnston, Osamu Toki, and Claire Wilcox

Understanding the scan patterns radiologists use to view medical images is critical to the design of image viewing devices. In this study, an eyetracker, a device for recording eye and head movement, was used to determine the scan patterns during the interpretation of single and multiple computed tomographic (CT) examinations presented on a four-over-four viewbox. CT examinations were used because they represent complex viewing situations. In two separate studies, radiologists viewed patient folders containing single or multiple CT chest examinations and dictated a report. Eye movement was recorded with an eyetracker and video camera. After mounting the films in order, radiologists generally started with a sequential scan through the entire examination, followed by careful viewing of two to four clusters of three to six images, followed by dictation. These results indicate that a well designed radiology workstation should provide an image index, sufficient display area to simultaneously view 10 or more images, random and sequential movement through the examination, image comparison, and image marking.

© 1990 by W.B. Saunders Company.

KEY WORDS: human factors, computed tomography interpretation, radiology workstations.

PREVIOUS INSIGHT into the radiologist's interpretation process has been gained from experiments that recorded eye movements during image interpretation.¹⁻⁴ Much of this research focused on determining the source of interpretation errors. Experienced radiologists employ a variable, though basically circumferential scan pattern when reading radiographs.⁵ Scan patterns develop with experience,⁶ are affected by prior knowledge,¹ and deviate from textbook recommendations.² While some studies have suggested that misreadings (false-negatives, which range around 30%,⁷⁻⁹ may occur because large

areas of film are not viewed foveally,¹⁰ or because there is nonuniform coverage of the film.¹¹ Eyetracking experiments indicate that only about 30% of missed lung nodules can be attributed to the lesion not having been foveally viewed.¹²

In 1986, the University of North Carolina developed a fingerprinting study to identify the visual scan patterns (fixation location and fixation duration) used by different radiologists when reading chest CT image sets^{3,4}. In other words, how often, in what order, and how long was each image viewed? The accuracy level of fixation location was a single image (CT SLICE), not locations within the image. Five radiologists each read three CT examinations using a four-over-four viewbox. Each examination was accompanied by its requisition form, and all its films could simultaneously fit onto the viewbox. The experimenter sat next to the radiologist and recorded times for certain activities: information accessed, placement of the films on the viewboxes, and the visual scan path as indicated by the radiologist. Radiologists were instructed to read and interpret an examination as naturally as possible, except always to point to the image under regard, and to generate verbal descriptions.

Despite individual differences, basic similarities were observed among radiologists. Usually, they started a reading session by sequentially scanning the images while placing them on the viewboxes. In a second phase, radiologists concentrated (fixated) on several small groups or clusters of specific images that were suspect. Finally, during a dictation (or reporting) phase, they scanned the images in a nonsequential manner fixating on those groups (clusters) of images on which they were reporting. A cluster was operationally defined as a group of proximate images connected by repeated back-and-forth fixation scan paths.

There were two possible sources of error in the fingerprinting approach: first, there was no information on where the radiologist actually looked, only where he indicated he was looking. Second, the radiologist's scan path might have been somewhat modified by this task.

Sponsored by National Cancer Institute ROI CA 44060 and POI CA 47982.

From the Department of Radiology and the Department of Computer Science, University of North Carolina, Chapel Hill.

Address reprint requests to David Beard, Department of Radiology, CB# 7510 Old Clinic Bldg, University of North Carolina, Chapel Hill, School of Medicine, Chapel Hill, NC 27599-7510.

© 1990 by W.B. Saunders Company.

0897-1889/90/0304-0012\$03.00/0

We repeated the fingerpointing study using an eyetracker, a device for measuring a subject's eye movement while he or she is performing a visual task. In this study, radiologists read patient folders containing both single and multiple CT examinations. As with the fingerpointing study, our objectives were to determine the following: what scan patterns are used, how images in different examinations are cross-compared, the size of clusters, how the films are managed on the viewbox space, and how much time is spent manipulating them.

We first describe the eyetracker study of single CT chest examinations. A second eyetracker study is then presented in which the radiologists interpret patient folders containing multiple CT examinations. Finally, we briefly describe the implications this work has on the design of viable radiology workstations.

SINGLE CT CHEST EXAMINATIONS

Patient folders containing single CT chest examinations were chosen for the initial eyetracker study. CT chest examinations are representative of many complex image manipulation and viewing tasks performed by radiologists.

Equipment and Observers

To determine eye and head movements as the radiologists conducted their reading of medical images, we used an eyetracking device, the Eye Mark Recorder Model V (EMR-V) (NAC, Inc)

(Fig 1). The system has two main components, the head-goggle unit, and the camera controller unit with remote control. The goggle unit is a head mounted unit that contains the eyetracking optics and electronics. In operation an infrared light-emitting diode light source (950 nm wavelength), which is not sensed by the eye, projects a spot of light onto the subject's cornea. The spot is reflected from the cornea along an optical path, detected by a metal-oxide semiconductor (MOS) image device (video camera), and sent to the camera controller for processing. Besides the MOS image devices for each eye, there is a third "field unit" that observes the central portion of the subject's field of view. The camera controller provides power to the LEDs and video cameras, and superimposes the eye position indicator spots onto the video image of the subject's field of view, which is in turn available to a video monitor and video cassette recorder. The unit's field of view is 60° horizontal and 45° vertical with a manufacturer's quoted accuracy of 0.6°.

Radiologists need to have freedom to move their heads and bodies in any fashion during the reading process. While many eyetrackers require fixed head position, the EMR-V allows free movement, and was chosen in order to minimize any intrusion on the radiologist's normal image reading process. The cost of allowing this freedom was a sacrifice in tracking accuracy and necessitated periodic calibration checks and occa-



Fig 1. Interpreting a CT scan with an eyetracker.

sional recalibration during a reading session. Using our calibration charts, we determined that horizontal eye-spot position accuracy was within 1° ($\pm 1/3$ in at a 20 in viewing distance) but in the vertical direction, as the eye moved toward the vertical limits, the accuracy degraded to about 3° (± 1 in at a 20 in viewing distance). Since a single CT image slice (15 images on a 14 in by 17 in sheet of film) is 3 in by 4 in, at an average viewing distance of 20 in we could resolve which image slice was being viewed, but could not accurately determine the location within a given image slice that was being viewed.

We taped each reading session with a video camera that recorded audio and time with a built-in stopwatch. These records helped in data analysis by providing timing information and voice clues for what the radiologist was doing when. We chose the four-over-four viewbox to simplify tracking which image was being viewed as it typifies the standard display space available for viewing films. Four board-certified radiologists and one senior radiology resident participated in the study. All subjects had experience with chest CT examinations. None wore eyeglasses, but one had contact lenses. In our experience, contact lenses do not affect the use of the eyetracker. Each radiologist read an initial examination wearing the eyetracker to become comfortable with the device. All subjects reported that the eyetracker did not affect the quality of their interpretation or their operational methods, but that the eyetracker occasionally began to feel heavy toward the end of a reading session. The latter occurred when the calibration procedure was unusually long.

Procedure

The test material consisted of three different patient cases each having a single chest CT examination. Each imageset contained about 45 image slices all of which were recorded on film with an intensity window selected for viewing the mediastinum (three sheets of film). Another two sheets of film contained a subset of about 20 images that were recorded with an intensity window selected for viewing lung. The referring physician's requisition was provided. The radiologists were instructed to read the chest CT examination, dictate their interpretations using the same professional standards they would use

in the clinic, and end the study by replacing the films in the patient folder. Their report was dictated via a microphone onto the Eyetracker VCR.

Data Collection and Analysis

The objectives of this experimental work were to determine radiologists' film handling habits, interimage search patterns, and how they used text information, CT scout views, and screen space. We also wanted to know the size and number of image clusters they viewed. Cluster size, ie how many single images make up the cluster, is important in determining the minimum amount of display area a radiology workstation must have for simultaneous display of the smallest working group of images.

Visual scan patterns were determined manually. The tape was reviewed by two experimenters on a playback machine and the scan path documented at 0.5 second intervals. The playback machine allowed frame by frame viewing if more frequent intervals were needed for analysis. From the video tapes, the experimenters documented the eye path of the subject noting when the radiologist paused (fixated) on an image. A fixation was defined as a visual concentration of 1.0 or more seconds on a given image. A cluster was defined by repeated (more than three) rapid, back and forth viewings of two or more proximate images with no eye movement onto intervening noncluster images. We documented how many images there were in a cluster and how many clusters there were in a CT examination. We recorded the amount of time taken for each phase of the reading session.

Results

Film handling and organization. Lung and mediastinum (MS) intensity-windowed images were printed in two separate sequential series, and this separation was maintained by the radiologists when loading films onto the viewboxes. Generally, the radiologists placed MS images on the bottom viewboxes, with the lung images either placed in a single pile on one upper viewbox, or spread out on the upper viewbox. One radiologist only used the bottom viewboxes, using the left three viewboxes for the MS images, and piling the lung images on the right-most box. For the other radiologists, lung images on the

upper viewboxes were viewed by looking up at them, by removing them and viewing them in hand, by swapping them with MS images on the lower boxes, or by standing up. Radiologists constantly changed their viewing position during the interpretation, often moving close to an image (zooming in) for a detailed look at the image, or moving away from the viewbox (zooming out) for a broader view. They would move left or right from the center to better view images on the outer viewboxes. Three radiologists occasionally stood up to view images on the upper viewboxes.

Scan patterns. The following scan pattern was used by all but one radiologist during interpretation of these single CT examinations. First, images were removed from the patient folder, sorted, and selected films were placed onto the viewbox. Images were often viewed as they were being loaded onto the viewboxes. Second, during what we call an overview phase, a rapid scan was performed over all the images. Sometimes this overview search was by organ, but in general, a systematic sequential scan pattern was initially used to view the series of slices in the CT examination, from the first to last image. These radiologists first viewed the images intensity-windowed for the mediastinum, followed by those intensity-windowed for the lungs. Comparison between images of different intensity windowing occurred only once by one radiologist. Third, during the detail phase, two to six clusters of image slices showing important radiology findings were reviewed. These clusters typically contained 3 or 6 images, with occasionally as many as 9, depending on the local anatomy and radiology finding under consideration (Fig 2). Cluster size seemed to follow the row and column layout of the images on the films. From the video

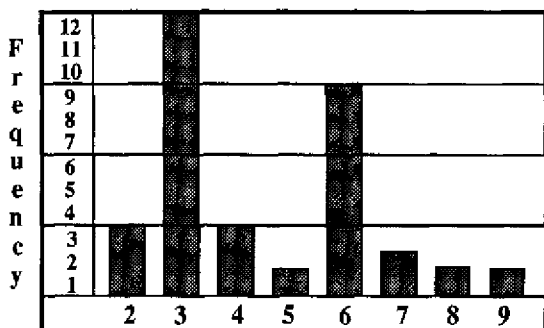


Fig 2. Histogram cluster for all studies and radiologists.

Table 1. Mean Times for Single Examination Interpretation

Case no.	Mean Times (min:sec)			
	#1	#2	#3	Mean
Film manipulation*	0:47	0:38	0:30	0:38
Film viewing/dictation	5:23	5:45	5:02	5:15
Total Interpretation	6:10	6:23	5:32	5:53

*Some data points are missing due to eyetracker calibration loss. Does not include time to unload films from viewboxes.

recordings, we determined that the radiologists averaged less than 0.2 seconds when moving eyes from one image to another, with typically a two to five second pause at each image. Fourth, during the dictation phase an interpretation report was generated, often while the images continued to be viewed. Fifth, the films were removed from the viewbox and replaced back into the patient folder.

A great deal of time was spent locating and nonsequentially accessing these small clusters of images showing important radiological findings. When asked how much of the patient's folder must be simultaneously viewed, radiologists typically answer "all of it."^{3,4} The above results contradict this, indicating that a maximum of nine images are viewed in rapid sequence demanding simultaneous display. But radiologists are correct in feeling they need most of the patient folder displayed on the viewbox. The simultaneous display of the complete contents of a patient folder appeared to be used as a visual image index, allowing the radiologist to locate quickly any particular image, with rapid eye and head movement. This visual image index was not perfect. Some radiologists occasionally used their finger to mark the location of critical image clusters to speed subsequent access. The finger-pointing method could not show this because one hand was busy pointing out the eye's location, and the other manipulating films on the viewboxes.

Timing. Table 1 summarizes several categories of mean times for interpretations. Film viewing-dictation time is the total time radiologists spent actually viewing images, rather than moving films. Film manipulation time is the total time radiologists spent moving and manipulating films, including time to load the images onto the viewboxes, and time for radiologists to move images from one location on the viewbox to another location. The time to unload the films

and place them back into the patient's folder was not timed, but took about 10 seconds per case. Sometimes radiologists would hold images in their hands for viewing, making it difficult to separate image-viewing time from film-handling time. If radiologists held films to light sources, and their eyes moved over the images on the films, we considered it viewing time rather than film-movement time. Total Interpretation time is from the moment radiologists picked up patient folders and started loading films until they replaced films into the folders. Sometimes partial times are missing due to eyetracker calibration loss.

MULTIPLE CT EXAMINATIONS

Method

A second study was conducted with three patient folders, each containing more than one CT examination with the number of films exceeding the available viewing space. The intent of this study was to document how radiologists used the available viewing area, how images from the different examinations were compared with one another, and how the films were arranged and organized and the times taken to carry out the various steps in the reading process. The first patient folder contained two abdominal CT examinations conducted 2 weeks apart. There were seven sheets of film with a total of 84 images. The second patient folder had three abdominal CT examinations with a total of 12 sheets of film and 135 images. Finally, the third patient folder had three abdominal CT examinations and one MR examination (obtained after the first CT examination) with a total of 18 sheets of film and 244 images. Radiologists were presented with the patient folder, the requisition for the most recent examination, and the radiologist report for the previous examination. The remainder of the experimental protocol remained the same as for the single CT examination study.

Radiologists compared images from the various examinations to evaluate changes that occurred over time. So, we observed the size of image clusters and how they were cross compared. The analysis of the video records, both from the video camera and the eyetracker, were manually reviewed, scan patterns were documented, and the number of fixations, image cluster size, and timing were obtained the same

as for the single CT examination. Additional information was recorded regarding how radiologists sorted and organized films on the viewbox and what they did with the films not occupying viewbox space.

Results

Film handling and organization. Radiologists began the interpretation by reading the requisition form and, in most cases, the previous interpretation reports, and put the films into an order of their liking. Images were usually quickly scanned during the sorting process. Images were piled by intensity window within an examination, with different piles for each examination. Piles were located either on the table or on the upper viewboxes. Radiologists appeared to use both the image appearance as well as the examination date printed on each image for sorting purposes. In two cases, there was far more film than viewbox space, so radiologists kept the films in the piles until needed. This resulted in a great deal of film movement during the interpretation, most likely, requiring a great deal of attention. Further, access to images in the piles was much slower than access to images displayed on the viewbox. Radiologists could no longer use their spatial knowledge of the viewbox organization to quickly index to the required image. Rather, they had to slowly move through a pile, looking at each film in turn until the desired one was located. The bottom panel (or row of four viewboxes) was generally used to hold the current examination, with the upper panel—and sometimes the outer viewboxes on the bottom panel—used to hold films being compared. As with the single examination interpretation, films were read while being held, and the radiologists moved to the left and right, and sometimes stood up in order to better view the images. One radiologist organized each examination, along vertical rather than horizontal lines.

Scan patterns. After reading the requisition form, possibly the previous interpretation reports, and sorting the films, the radiologist began scanning the images. The initial overview phase observed with single CT examination was still evident, but more diverse. In different orders, all radiologists sequentially viewed the current examination, selectively viewed the previous examination (most likely focusing on radiologic findings

identified by the requisition and previous interpretation report) and looked at any older images, (this varied with the radiologist and the case). This overview phase was again followed by a more complex detail phase with a focus on small clusters. A final dictation phase was again present with the same review of critical clusters. Radiologists repeatedly (about 3 to 5 times) accessed related text information such as requisitions, interpretation reports, and text information such as examination step size written on the films, before and during the interpretation. This text information appeared to be very important to interpretation strategy for decision making.

Two methods of pair-wise cluster comparison were observed. In one, the radiologists rapidly moved their eyes back and forth many times between a single image in one study and an image in another study with a maximum of a two second fixation per image. With the second method, clusters from two studies were compared with rapid fixations on each image within a cluster and far fewer movements back and forth between clusters. The current and most recent examinations were the most likely to be cross compared. As before, clusters for single examination viewing typically ranged from two to six images with a maximum of nine. However, when clusters from different examinations were compared, the individual clusters ranged from two to four images for a total of four to eight images normally requiring simultaneous display.

When comparing clusters, the films from the different examinations were placed as close together as possible. Sometimes the films were placed above and below each other, sometimes side by side, and often one film remained on the viewbox with the other hand-held close by.

Timing. Table 2 contains the times for manipulating the films and the total times for the interpretations. Blanks occurred in several sessions with case three due to the eyetracker losing calibration. Both film manipulation time and total time were gathered. Film viewing-dictation time is the difference between the total and film manipulation times. We show total time for individual radiologists reading individual cases to allow the reader to ascertain the degree of individual difference among radiologists as well as the time differences between cases.

As with most human factors studies, the differences among response times for individual radiologists is almost as much as that between the different cases. Even considering the individual variations, the total film manipulation time and time per film roughly, but dramatically, increases with the number of films for a case. This observation matches what we have observed under clinical conditions and, we believe, reflects the additional amount of film sorting and searching a larger number of films entails. We feel this is because, while examinations older than the current and previous are always viewed, they are not examined in detail. Therefore after the first

Table 2. Times for Multiple CT Interpretation

Study	Mean Times (min:sec)						Mean
	A	B	C	D	E	F	
#1 7 Films							
Film Manipulation	1:20	1:50	1:40	2:00	1:40	1:10	1:40
Viewing/Dictation	5:40	5:10	11:00	3:40	6:40	4:30	6:10
Total	7:00	7:00	12:40	5:40	8:20	5:40	7:50
#2 13 Films							
Film Manipulation	2:40	3:00	3:50	2:50	2:20	2:00	2:50
Viewing/Dictation	8:20	6:40	12:30	5:10	10:10	5:50	8:10
Total	11:00	9:40	16:20	8:00	12:30	7:50	11:00
#3 18 Films							
Film Manipulation	5:10	—	—	7:10	—	5:50	6:00
Viewing/Dictation	5:50	—	—	8:00	—	7:40	7:10
Total	11:00	—	—	15:15	—	13:30	13:10
Mean film manipulation	3:00	2:30	2:40	4:00	1:40	3:00	3:00
Mean viewing/dictation	6:40	5:50	11:50	5:40	8:30	6:00	7:10
Mean Total	9:40	8:20	14:30	9:40	10:10	9:00	10:10

NOTES. Film viewing/dictation = total time - film manipulation. All times are rounded to the nearest 10 seconds.

two CT examinations, additional ones add only a small amount of interpretation time. Note also that the total time spent on film viewing-dictation does not change much between these three multiple examination cases, though viewing-dictation time is greater than that for interpretations of single CT examinations.

DISCUSSION

The above results validate and expand upon the conclusions from the fingerprinting method described in the introduction. Radiologists sequentially overview the images, then examine critical clusters in detail, and finally dictate a report, again reviewing the clusters. While all the images arrayed on the viewbox appear to serve as a visual image index, typically only two to four small clusters consisting of from three to six images from the same examination are simultaneously viewed, and only clusters from different examinations of from one to four images are pair-wise compared. The small cluster sizes, particularly in cross-comparison, may be related to the film image layout, the physical restraints of keeping the images close together when comparing them, or attention-memory limitations.

The visual image index (all the images displayed on the viewbox) works well for image access, with radiologists generally moving their eyes directly to the target images (though they occasionally use their finger to mark a location for future reference). However, this visual image index fails when there are more films than viewbox space, as was often the case with multiple CT examinations. There appears to be a great deal of manipulation of films for larger cases. Films must either be left in piles on nearby tables, or stacked on several viewboxes. A great deal of time may be wasted looking for particular films, removing and replacing films on the viewboxes.

We would expect somewhat improved film handling with an alternator, but it by no means solves the film handling problem. Our observation of radiologists using an alternator indicates two problems: first, radiologists sometimes lose track of the location of particular examinations on the alternator's moving strip. Second, the seek time on the alternator, that is, the average time to locate a needed image, can be quite long.

The fact that little comparison, as measured by the eyetracker records, was made between different intensity windowed images, such as those intensity windowed for lung versus bone, was a surprise to both the radiologists and the experimenters. The radiologists thought that they did physically make that comparison. In one of our previous experiments with a prototype electronic workstation¹³ some radiologists placed the lung windowed image adjacent to the mediastinum-windowed image and thought this was useful, but there was no indication that it was an important or necessary feature. Of course if that comparison can be conveniently and rapidly made with an electronic workstation, the radiologists habits may change. This qualification, of course, must be applied to all the results we obtained from studying the radiologist using the traditional display device that they have been trained on that has imposed limitations on the reading process.

IMPLICATIONS FOR RADIOLOGY WORKSTATION DESIGN

Electronic workstations are the essential means for physicians to view electronically stored medical images.^{14,15} However, producing an acceptable image-display radiology workstation is difficult. The typical viewbox array can easily display the equivalent of 32 high-resolution (1024²) display monitors that the physician can quickly access using techniques (the movement of eyes and head) that have been practiced for an entire professional career. Even the best workstations barely have this much real memory, let alone display space. Current workstations begin to show their advantage only with large patient folders that overflow viewboxes.

From our eyetracking research, we conclude that a viable radiology workstation for single and multiple CT examinations at least needs the following: an image index for rapid access to patient folders, examinations, and images; sequential and arbitrary viewing of images in an examination; access to all current patient text information; sufficient area to simultaneously display a minimum of nine full-resolution CT images; function for comparing image clusters from different examinations; function for marking critical images and clusters; and function for inten-

sity windowing an image slice. Rapid response time, simple hand motions, the ability to quickly measure anatomy, and an easy to understand interaction are also clearly needed.

ACKNOWLEDGMENTS

We wish to thank Drs Bell, Kennedy, Klymenko, Mauro, Mikhail, Pizer, and Schiebler for their assistance in this research and in reviewing this manuscript.

REFERENCES

1. Kundel HL, Wright DJ: The influence of prior knowledge on visual search strategies during the viewing of chest radiographs. *Radiology* 93:315-320, 1969
2. Gale AS, Worthington BS: The utility of scanning strategies in radiology, eye movements and psychological functions. In: Groner R, Menz C, Fisher DF, et al (eds): *International Views*. Erlbaum, Hillsdale, NJ, 1983
3. Rogers DC, Johnston RE, Hemminger BM, et al: Development of and experience with a prototype medical image display. Presented at the Farwest Image Perception Conference, University of New Mexico, Department of Radiology, July 17-18, 1986
4. Beard DV, Pizer SM, Rogers DC, et al: A prototype single-screen PACs console development using human computer interaction techniques. *SPIE Medical Imaging*, 1987; 767:646-653, 1987
5. Carmody DP, Nodine CF, Kundel HL: Global and segmented search for lung nodules of different edge gradients. *Invest Radiol* 15:224-233, 1980
6. Kundel HL, LaFollette PS: Visual search patterns and experience with radiological images. *Radiology* 103:523-528, 1972
7. Yerushalmy J: The statistical assessment of the variability in observer perception and description of roentgenographic pulmonary shadows. *Radiology Clinics of North America* 1:381-392, 1969
8. Smith MJ: *Error and Variation in Diagnostic Radiology*. Thomas, Springfield, IL, 1967, p 148
9. Tuddenham WJ: Visual search, image organization, and reader error in roentgen diagnosis: Studies of the psychophysiology of roentgen image perception. *Radiology* 78:694-704, 1962
10. Llewellyn-Thomas E: Search behavior. *Radiology Clinics of North America* 1:403-417, 1969
11. Tuddenham WJ, Calvert WP: Visual search patterns in roentgen diagnosis. *Radiology* 76:225-256, 1961
12. Kundel HL, Nodine CF, Carmody D: Visual scanning, pattern recognition and decision making in pulmonary nodule detection. *Invest Radiol* 13:175-181, 1978
13. Johnston RE, Rogers DC, Hemminger BM, et al: A multiscreen multiple image display console. *SPIE Med Imag* 626:447-450, 1986
14. Beard DV: Designing a radiology workstation: Focus on navigation during the interpretation task. *J Digit Imag*, 3:152-163, 1990
15. Pizer SM, Beard DV: Medical image workstation: State of science and technology. *J Digit Imag* 2:185-193, 1989