

## How Many Screens Does a CT Workstation Need?

D.V. Beard, B.M. Hemminger, K.M. Denelsbeck, and R.E. Johnston

A considerable number of prototype and commercial workstations have been developed during the last 10 years for electronic display of computed tomographic (CT) images during clinical interpretation. These CT workstations have varied widely in the number and size of monitors available for the display of the medical images ranging from a single  $1,024 \times 1,024$ -pixel monitor, to eight  $2,500 \times 2,000$ -pixel monitors. Image display times also have varied considerably, ranging from as fast as .11 seconds, to as slow as 26 seconds to fill a single monitor. No consensus has formed in the workstation community with regard to display area and response time requirements. To address this issue, we have constructed a time-motion model of CT interpretation. Model accuracy is experimentally verified with three workstations as well as with the film alternator. In general, CT interpretations with an electronic workstation become faster as display area increases and display time decreases. Results can be used by workstation designers and purchasers to roughly estimate differences in interpretation speeds among contending CT workstation designs.

Copyright © 1994 by W.B. Saunders Company

**KEYWORDS:** computed tomography (CT), image display and recording, diagnostic radiology observer performance, images, display, time and motion analysis.

**I**MPROVED OPERATIONAL efficiency of radiology departments potentially is available with a picture archiving and communication system (PACS). Possible improvements include reduced costs,<sup>1,2</sup> concurrent access to images, elimination of lost images, and cost-effective access to various image processing techniques.<sup>3,4</sup> The digital modalities, such as computed tomography (CT) and magnetic resonance imaging (MRI) are particularly well suited to PACS for two reasons: first, because the data is fundamentally digital, and second because the images, even at full resolution, are small enough to fit completely onto available monitors making unnecessary the time-consuming roaming-and-zooming of individual images. However, the potential benefits of PACS for the digital modalities are only available with a workstation facilitating fast and accurate interpretations.

CT workstations developed during the last 10 years vary widely in the number and size of monitors available for the display of the medical images. In general, CT interpretations using an electronic workstation become faster as display

area increases and display time decreases. But so does cost, and no consensus has formed in the workstation community as to how much display area a CT workstation needs, nor as to how fast those monitors must be filled with CT images before radiologists can use such a workstation to interpret a CT study as quickly as they can with film and alternator. Faster interpretations of reduced quality are not generally considered useful, so equivalent interpretation quality must be maintained while interpretation speed is reduced.

In this report, we present results from detailed time-motion models of CT interpretations. First, we provide background on CT interpretation, CT workstations in general, and three prototype workstations we have constructed in particular. Second, we provide background on time-motion models and present evidence as to their potential accuracy. Third, we describe the time motion models of CT interpretation that we have constructed and the experimental results that we have generated to verify their accuracy. Finally, we describe the relationship between display area, display speed, and interpretation speed, and list caveats that should be considered when using these results to estimate the relative interpretation times of CT workstations.

### CT WORKSTATIONS

A number of groups have evaluated workstations for CT and MRI interpretation.<sup>5-10</sup> In general, these studies show acceptable interpretation quality for CT and MRI, but unacceptably slow interpretation times. Johnston et al evaluated CT images for image quality using film light box and a three-screen  $1,024 \times 1,024$ -pixel display system.<sup>7</sup> They concluded that CT

---

*From the Departments of Radiology and Computer Science, University of North Carolina, Chapel Hill, NC.*

*This work was sponsored by National Institutes of Health Grant No. R01 CA44060.*

*Address reprint requests to David Beard, PhD, Department of Radiology, 509 Old Clinic Bldg, School of Medicine, UNC Chapel Hill, NC 27599-7510.*

*Copyright © 1994 by W.B. Saunders Company  
0897-1889/94/0702-0007\$03.00/0*

images, displayed on a good quality monitor with appropriate intensity window settings, will provide the image quality available on film. Foley et al conducted an observer experiment comparing film with a commercial workstation with two  $1,024 \times 1,024$  monitors and a 7.5-second to 26-second image display time.<sup>6</sup> Not only was the workstation considerably slower than the alternator, but interpretation quality was degraded even though Foley believed that the monitors provided acceptable image quality. Poor interpretation quality again may have been the result of strained human "working" or "short-term" memory,<sup>11</sup> caused by the very slow response times. Finally Brown et al evaluated a workstation with eight  $1,024 \times 1,024$  monitors for MRI knee studies.<sup>8</sup> As with Johnston, image quality was acceptable. However, workstation interpretation time was again twice as long as with film.

Recently, experiments have been conducted indicating that interpretations of electronically displayed CT images can be very fast if sufficient image display area and sufficient image display speed are provided. Straub et al<sup>10</sup> described a workstation using two  $2,000 \times 2,000$ -pixel monitors with a .1-second image display, as well as a second CT workstation that displayed a single CT image at a time using a superimposed or "cine" display. A receiver operating characteristic study not only showed equivalent accuracy of the workstation interpretations to the film interpretations, but also indicated equivalent speed. However, interpretation speed might have been different if a free-form dictated interpretation report, rather than a findings form, had been used. This is because the findings form may have acted as a working memory aid, affecting cognitive load, interpretation behavior, and thus possibly speed.

The cine results are particularly interesting because, if true, they would greatly reduce the required cost and size of a CT workstation. However, cine display is outside the scope of our time-motion models. Unlike film alternator and other workstations, cine does not require eye movements between images. Because our models do not factor in these eye movements, making time-motion comparisons between multiple-image display and cine display systems is impossible.

Over the last 8 years, we have developed and experimentally evaluated a number of CT workstation prototypes.<sup>12,14</sup> Experiments comparing film with the FilmPlane workstation (University of North Carolina [UNC], Chapel Hill) with a single  $1,024 \times 1,024$ -pixel monitor and a 1.5-second image display time<sup>12,13</sup> indicated that interpretation with such a workstation takes about twice as long as film. Although interpretation quality appeared acceptable, there also were clear signs of human working memory strain<sup>11</sup> on the part of the radiologists, such as the radiologists losing track of where they were in their interpretation.

However, as with Straub et al, much better results were obtained when sufficient image display area and speed were provided. Another workstation prototype, called FilmStrip (UNC),<sup>14</sup> uses a single  $2,048 \times 2,560$ -pixel monitor (Megascan Technology, Littleton, MA) with an image display time of .11 seconds. Because a good visual mental model<sup>15</sup> or metaphor is essential to a well-designed computer human interaction, FilmStrip uses the metaphor of a single-panel vertical alternator in which the images are arrayed in a long vertical "filmstrip" with "up" and "down" scroll buttons used to control movement. Twelve CT images were displayed simultaneously in a three-column by four-row filmstrip similar in organization and size to film. Experimental results<sup>14</sup> showed FilmStrip to be as fast as the film alternator for interpretation of single-CT chest studies with apparently identical accuracy.

However, cost is a problem with a clinical workstation based on the FilmStrip hardware; FilmStrip requires two  $2,048 \times 2,560$ -pixel monitors to compare two CT cases for a total hardware cost of well over \$90,000 plus software and marketing for a total rough price of at least \$150,000. Although  $2,048 \times 2,560$ -pixel monitor workstations may be required for computed radiography (CR) interpretation and, thus, eventually will be available for CT and MRI interpretation, we believe that a workstation designed for the digital modalities alone must have a greatly reduced price tag to be clinically viable.

Therefore, we constructed the FilmStripLet prototype to provide the interpretation speed of FilmStrip using low-cost commodity-priced hardware.<sup>16</sup> FilmStripLet, implemented using the X

windowing system and C++, runs on a SPARC 2 workstation (Sun Microsystems, Mountain View, CA) with 128 Mbyte of main memory and two  $900 \times 1,100$ -pixel, 8-bit framebuffers and monitors. These two monitors allow eight full-resolution CT images to be displayed simultaneously. FilmStripLet uses the same horizontal filmstrip mental model that we used with FilmStrip. Speed is of the essence. FilmStripLet, like FilmStrip, is able to display the next set of four images in less than .2 seconds. Although a SPARC 2 with 8-bit framebuffers cannot intensify window a 12-bit image and display it in this amount of time, it can very quickly transfer an 8-bit image onto the screen. Because the FilmStrip system showed the effectiveness of preset intensity windowing for CT workstations, we precompute the lung, soft tissue, and liver intensity windows for all the slices in a CT study and store these images in three large X-window pixmaps, allowing very rapid image display. Thus, although preset intensity windows were rapidly available, there was no provision for dynamic intensity windowing.

#### TIME-MOTION MODELS

Time-motion analysis, initially popularized by Gilbreth<sup>17</sup> was refined into a reasonable computer-human interaction tool by Card et al<sup>18,19</sup> with follow-up work by Roberts,<sup>20</sup> Kieras,<sup>21-23</sup> and others. Such time-motion models have been verified with a large number of controlled subject experiments.<sup>19,20,24</sup> To build our time-motion models, we used a variation of Card's GOMS model (Goals, Operators, Methods, and Selection rules). GOMS, and its precursor Key-stroke model, provide crude but acceptable accuracy for a number computer-human interaction situations.

Time-motion models estimate the completion time of a task by summing the completion times of the task's subtasks which are in turn recursively analyzed until atomic tasks such as button presses are encountered. Task times for these atomic tasks can either be located in a table, calculated based on human performance metrics,<sup>25</sup> estimated ad hoc, or determined experimentally by timing user interaction with small interface mockups. Typical generic atomic-task times include .2 seconds for a button press, .3 seconds per character for typing, 1.1 seconds to

point the cursor at a screen object with the mouse, and .5 seconds to move the hand from the keyboard to the mouse.<sup>19</sup>

A very simple example of a time-motion model may help the reader understand how GOMS models function. Suppose a workstation with a  $1,024 \times 1,024$ -pixel monitor, can display four full-resolution CT images at the same time, and has an interface consisting of two buttons for scrolling forwards and backwards through a study; each button press displays four new images. We can further assume an image display time of 3 seconds and a generic CT study with 40 slices. Given a single  $1,024 \times 1,024$ -pixel monitor and starting at the top, a user will have to scroll four times, on average, to locate a slice containing a particular anatomical object. To scroll once, the user must press the backwards or forwards button (.2 seconds) and wait for the system to respond (3 seconds), resulting in 3.2 seconds per scroll for a total of 12.8 seconds for the four-scroll search operation. What if the image display time was reduced to .1 seconds? This would then result in a scroll-time of .3 seconds and a task time of 1.2 seconds. If search tasks such as this were frequent during CT interpretation, then a difference of 11.6 seconds per search task could have a significant impact on total interpretation time. (Those of you familiar with time-motion analysis will note that I have simplified the discussion by ignoring delays caused by "mental pauses."<sup>19</sup>)

#### CT MODELS AND VERIFICATION

We developed a family of CT-interpretation GOMS time-motion models based on film alternator and three CT workstation prototypes. First, we developed models for workstation and film alternator interpretation of chest CT scans and verified the model's predictive abilities by comparing model output with experimentally-derived interpretation times. Then, we determined image manipulation time while systematically varying image display time and display area.

#### *CT Model Limitations*

Unfortunately, our GOMS time-motion models of CT interpretation have limitations. Not only is GOMS designed to model only the error-free performance of expert users, but

worse, GOMS models do not easily describe all aspects of decision support interactions such as medical image interpretation. This is because these time-motion models cannot account for the time the radiologists spend viewing the various images or interacting with the dictation machine; these models only account for the time the users spend actually pressing buttons and otherwise manipulating the interface. Image manipulation time is only one component of image interpretation time. Image manipulation time denotes the time to load and unload the films, to scroll the alternator or workstation, and to adjust intensity windows on the workstation. Image interpretation time is composed of image manipulation time as well as the time for the radiologist to view the images, to develop an interpretation, and to dictate a report. Note that a difference in image manipulation time between two systems may not result in exactly the same difference in image interpretation time because the human cognitive activities of image viewing, interpretation, and dictation may occur in parallel to computer image display and, to a certain extent, in parallel to human manipulation of the computer controls. Nevertheless, this effect is limited, and a sufficiently large difference in image manipulation time should indicate a corresponding difference in interpretation time. Thus, these models can be used to provide rough comparisons between candidate workstation designs allowing design decisions to be made before prototype implementation or system purchase.

There are a number of additional limitations to GOMS models. First, because we do not model radiologists moving their chairs back and forth, there is a limit to the number of monitors for which the model can account. Second, we do not include the time to move the head from looking at one monitor to looking at another, so a large number of monitors would produce further error. Third, because these models do not include the time for the radiologist's eyes to move from one CT image on a sheet of film to another, the model will not accurately predict the difference in task times of film alternator interpretation and interpretation using a superimposed or cine display.<sup>10</sup> A number of other caveats are listed below.

### *CT Model Construction*

Our initial time-motion model analyzed film alternator CT interpretation. First, we determined the average number of lung, soft tissue, and liver images in a set of 10 CT chest examinations, and created a CT chest generic examination having the average number of these categories of images. The 10 CT studies used to develop the generic examination were also used in the film alternator, FilmPlane, FilmStrip, and FilmStripLet experiments described below. Second, based on observation of about 100 CT chest examinations being interpreted using film alternator by more than 10 radiologists, we developed a generic scenario to describe a typical radiologist's behavior when interpreting the generic examination. In this generic scenario the radiologist takes the films out of the film folder, mounts them onto the alternator, and views the lung, soft tissue, and liver images in turn. The radiologist then views two critical regions with interesting findings, dictates a report while re-viewing these critical regions, unloads the films, and places them back into the image folder. We realize that many radiologists have their films mounted before interpretation, but the actual choice of scenario will not effect our model's predictive accuracy, only the actual interpretation time estimated for film and alternator.

On different occasions, the same modeler may develop significantly different GOMS models for the same CT interpretation method. To examine this, we developed a second GOMS model for each CT interpretation method. Thus 6 months after the first film alternator scenario and model was developed, the same modeler developed a second generic scenario from the same generic CT chest examination and then developed a second GOMS model of film alternator interpretation. The second scenario and model were developed without review of the first one. Estimated image manipulation times were 1.3 minutes for both models suggesting little inner-modeler variability.

After building the film alternator models, we modeled FilmStrip using the generic scenario used with the first film alternator model described above. Figure 1 shows a GOMS model of FilmStrip interpretation of the generic examination using the generic scenario. Six months

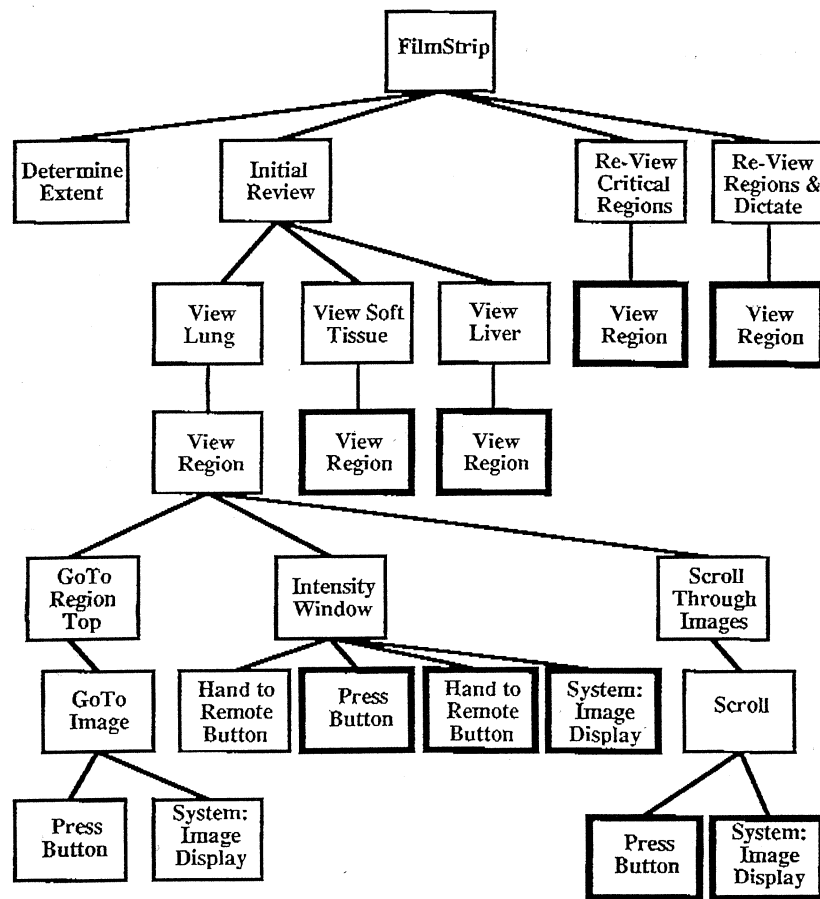


Fig 1. GOMS model of CT chest interpretation using the FilmStrip workstation and the first generic behavior scenario. Bold-bordered tasks are "reused" versions of other tasks.

later, a second FilmStrip model was developed by the same modeler with the generic scenario used with the second film alternator model described above. Estimated image manipulation times were .5 minutes for the first model and .4 minutes for the second. In a similar manner, two models of FilmStripLet interpretation were developed using the two generic scenarios. Estimated image manipulation times were .6 minutes for the first model, and .5 minutes for the second. Finally, we developed two models of our 5-year-old FilmPlane workstation with its single  $1,024 \times 1,024$ -pixel monitor and 1.5-second image display time. Estimated image manipulation times were 3.1 minutes for the first model, and 3 minutes for the second.

#### CT Model Verification

**Speed.** In 1989, we conducted a timing study to measure the interpretation speed of radiologists reading difficult cases using the FilmPlane workstation.<sup>12,13</sup> Average interpretation time was

8.2 minutes. In 1991 and 1992, we conducted a timing study to determine actual CT chest interpretation times for film alternator and FilmStrip.<sup>14</sup> Average interpretation time for the four radiologists using FilmStrip was 5.7 minutes. Average interpretation time using film alternator was 6.2 minutes including 1.2 minutes to load and unload the films. Finally in 1992 and 1993, we conducted a timing study with two radiologists interpreting chest CT cases using FilmStripLet. Average interpretation times for FilmStripLet was 5.9 minutes. The same 10 CT cases were used in all these timing studies and there was considerable overlap among the radiologists, so we expect these interpretation times to reflect actual speed differences between the various interpretation methods. Confidence intervals varied among these studies, but generally were  $\pm 50$  seconds around these interpretation-time means.

**Accuracy.** Dictated reports were used for accuracy measurement to provide a cognitive

task similar to an actual clinical interpretation. ROC studies typically use findings forms which allow very precise results and simple grading. However, these forms can act as a working memory aid and therefore may distort the timing (and possibly accuracy) results of a study. To determine interpretation accuracy, dictated reports were examined, a list of findings were generated, and the report was graded as either acceptable or unacceptable with an acceptable report having all findings relevant to the requisition, all critical findings, and no incorrect findings. No unacceptable reports were generated during these interpretations; this result was not unexpected because these were experienced board-certified radiologists on the faculty of a major university who were working under laboratory conditions, and who had been instructed to "work as quickly as possible while generating reports of clinically acceptable accuracy." Although these experiments did not have sufficient experimental power to determine equivalent accuracy between film and workstation, we believe that these results indicate that the timing results can be accepted as representative.

Table 1 summarizes these experimentally determined interpretation times as well as the model-estimated image manipulation times for film alternator and the three workstations as described above. Two models were constructed for each interpretation method by the same modeler 6 months apart to show model repeatability.

We offer two indications that these models provide accuracy which is similar to the experimental data (ie,  $\pm 50$  seconds). First, note that the model-predicted ordering of the interpretation methods (left to right, Table 1) is the same ordering as provided by the experimentally determined interpretation times. Thus, if we had used the model to predict whether one

**Table 1. Model-Estimated Image Manipulation Time and Experimentally Determined Interpretation Time (Minutes)**

Times	Film Strip	Film StripLet	Film Alternator	Film Plane
Estimated manipulation time				
Model 1	.5	.6	1.3	3.1
Model 2	.4	.5	1.3	3
Actual interpretation time	5.7	5.9	6.2	8.2

**Table 2. Average Predicted Difference Versus Actual Difference in Interpretation Time (Minutes)**

Interpretation Pair	Predicted Difference	Actual Difference	Model Error
Film Alternator/FilmPlane	1.8	2	.2
Film Alternator/FilmStrip	.9	.5	.4
Film Alternator/FilmStripLet	.8	.3	.5
FilmPlane/FilmStrip	2.6	2.5	.1
FilmPlane/FilmStripLet	2.5	2.3	.2
FilmStrip/FilmStripLet	.1	.2	.1

interpretation method would have been faster than another, the experimental results would have generated the same results.

A second indication of model accuracy is provided in Table 2. Given that the models measure image manipulation time while these subject experiments measured total interpretation time, we cannot directly compare model with experimental results. We can use the model results to estimate how much faster one interpretation method would be to another, and then use the difference in experimental results to measure the accuracy of that model's prediction.

Table 2 summarizes the average model-estimated and actual differences between the various interpretation methods. To use the first row of Film Alternator/FilmPlane as an example, predicted difference is the difference between the Film Alternator estimate of 1.3 minutes and 3.1 minutes—the average of the estimates for FilmPlane from Table 1. Actual difference is the difference between 6.2 minutes and 8.2 minutes—the actual interpretation times of film alternator and FilmPlane, respectively. Model error is the difference between the predicted difference and the actual difference. Average model error is .24 minutes. The maximum model error is .5 minutes, which is within the general confidence interval of  $\pm 50$  seconds for the interpretation times of the various workstation timing experiments.

## DISPLAY AREA AND TIME

### Model Results

Assuming that our workstation GOMS model is reasonably accurate in predicting image manipulation time, we can use it to estimate the difference in interpretation times for various CT workstation configurations, and thus deter-

mine the relative impact on interpretation time of image display area and image display time. Further, a hospital or radiology department PACS designer can extrapolate the results to provide a rough estimate of interpretation speed for a commercial workstation being considered for clinical use.

Table 3 shows CT interpretation times estimated with the models for various workstation configurations assuming interpretation of the generic CT case using the generic scenario. Estimates were generated by first calculating the difference in the model-estimated image manipulation time for a display area/display time configuration and the model estimate for FilmStrip, and then adding that difference to FilmStrip's experimentally determined interpretation time to estimate interpretation time for that configuration. Bold face entries indicate experimentally determined data points, with FilmPlane having one 1,024 × 1,024-pixel monitor (4 images) and 1.5 seconds, FilmStripLet having two 1,024 × 1,024-pixel monitors (8 images) and .1 seconds, and FilmStrip having one 2,560 × 2,048-pixel monitor (12 images) and .1 seconds.

Table 3 shows the importance of having at least two 1,024 × 1,024-pixel monitors (8 CT images) and a 1-second image display time or one 1,024 × 1,024-pixel monitors (4 CT images) and a .5-second display time. One way the reader could use Table 3 is in developing a very rough estimate of the potential interpretation times (of our generic task and CT case) for a given workstation configuration. These estimates might be useful to the CT workstation designer choosing between possible hardware configurations or the workstation purchaser trying to estimate how a particular commercial workstation might function in the clinic. Thus, someone could use Table 3 to roughly estimate

that a CT workstation with a single 1,024 × 1,024-pixel monitor and a .75-second image display time would require about 7.5 minutes to interpret the generic CT case.

### Caveats

A number of factors should be considered before applying the information in Table 3 to clinical workstations. First, there are large differences in the behavior and therefore the resulting interpretation times of different radiologists. Further, different cases or protocols will produce differing interpretation speeds. Thus, although our GOMS models do allow comparisons for our generic examination and generic scenario (ie, radiologist behavior), a given radiologist and case may generate very different results in the real world. Second, fewer displayed images and longer image display times will increase the cognitive load on radiologists, and possibly modify their behavior. For example, if too few images are displayed, radiologists might scroll back and forth several times over the same images trying to develop a mental model of the anatomy, or if monitors are too far apart, the radiologists may not use all available monitors. Third, our time-motion models and experimental results refer to interpretation of single CT chest studies. Comparison with a prior CT study or an examination of another modality would require additional display area. Fourth, our time and motion models assume the extremely simple three-button interface of FilmStrip. Selecting an object or menu item on the monitor takes about 1.1 seconds, moving the hand from the keyboard to the mouse can take from 1 to 2 seconds, and dynamic intensity windowing can take more than 10 seconds. Because 1 second here and 10 seconds there can easily evolve into a significant increase in interpretation time, we would expect workstations with complex interactions to result in much longer interpretation times. Fifth, because our models do not consider eye movement between simultaneously display images, we can not model the relative interpretation time of a cine display with that of film alternator simultaneously displaying multiple images. Further, interimage eye movement between monitors is likely to

**Table 3. Model-Estimated CT Interpretation Time for an Optimal Interaction Using the Generic Case and Generic Scenario**

Monitor Configuration (pixels)	Image Display Time Per Scroll			
	(1.5 sec)	(1 sec)	(.5 sec)	(.1 sec)
one 1,024 × 1,024	8.2 min	7.8 min	7.2 min	6.8 min
two 1,024 × 1,024	6.4 min	6.2 min	5.9 min	5.8 min
one 2,560 × 2,048	6.2 min	5.9 min	5.8 min	5.7 min

Film interpretation would require 6.2 minutes under similar conditions.

take more time than eye movement on the same monitor, so multiple monitor systems may be considerably slower than we have estimated.

It is also important to realize the data given in Table 3 is at best only accurate to within .5 minutes, and given our limited data and all the caveats listed above, we would not be surprised if actual results varied even more. Thus, caution should be used when extrapolating these results

to other systems. Nevertheless, rough estimates of relative system performance can be made, and should prove useful to the CT workstation designer or system purchaser if applied carefully.

#### ACKNOWLEDGMENT

We wish to thank Sridhar Seshadri for help in preparing this manuscript.

#### REFERENCES

1. Arenson RL, van der Voorde F, Stevens JF: Improved financial management of the radiology department with a microcosting system. *Radiology* 166:255-259, 1988
2. Beard DV, Parrish D, Stevenson D: A cost analysis of film image management and four PACS based on different network protocols. *J Digit Imaging* 3:108-118, 1990
3. Pizer SM: Psychovisual issues in the display of medical images, in Hoehne KH (ed): *Pictorial Information Systems in Medicine*. Berlin, Germany, Springer Verlag, 1985, pp 235-250
4. Pizer SM, Beard DV: Medical image workstations: State of science and technology. *J Digi Imaging* 2:185-193, 1989
5. Arenson RL, Chakraborty DP, Seshadri SB, et al: The digital imaging workstation. *Radiology* 176:303-315, 1990
6. Foley WD, Jacobson DR, Taylor AJ, et al: Display of CT studies on a two-screen electronic workstation versus a film panel alternator: Sensitivity and efficiency among radiologists. *Radiology* 174:769-773, 1990
7. Johnston RE, Yankaskas BC, Perry JR, et al: Agreement experiments: A method for quantitatively testing new medical image display approaches. *SPIE Medical Imaging IV: PACS System Design and Evaluation* 1234:621-630, 1990
8. Brown JJ, Malchow SC, Totty WG, et al: MR examination of the knee: Interpretation with multiscreen digital workstation vs hardcopy format. *AJR Am J Roentgenol* 157:81-85, 1991
9. Berbaum KS, Franken EA, Honda H, et al: Evaluation of a PACS workstation for assessment of body CT studies. *J Computer Assisted Tomography* 4:853-858, 1990
10. Straub WH, Gur D, Good WF, et al: Primary CT diagnosis of abdominal masses in a PACS environment. *Radiology* 1:739-743, 1991
11. Miller GA: The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psycholog Rev* 63, 81-97, 1956
12. Beard DV: Designing a radiology workstation: A focus on navigation during the interpretation task. *J Digit Imaging* 3:152-163, 1990
13. Beard DV, Perry R, Muller K, et al: Evaluation of total workstation CT interpretation quality: A single-screen pilot Study. *SPIE Medical Imaging V: PACS Design and Evaluation* 1446:52-58, 1991
14. Beard DV, Hemminger BM, Perry JR, et al: Single-screen workstation vs. film alternator for fast CT interpretation. *Radiology* 187:1-6, 1993
15. Young R: The machine inside the machines: User's models of pocket calculators. *Int J Man-Machine Studies* 15:51-85, 1981
16. Beard DV, Hemminger BM, Pisano ED, et al: CT Interpretations with a low cost workstation: A timing study. *JDI* 1993 (in press)
17. Gilbreth FB: *Motion Study*. New York, NY, Van Nostrand, 1911
18. Card SK, Moran TP, Newell A: The keystroke-level model for user performance time with interactive systems. *Communications of the ACM* 23:396-410, 1980
19. Card SK, Moran TP, Newell A: *The psychology of human-computer interaction*. Hillsdale, NJ, Lawrence Erlbaum Associates, 1983
20. Roberts, TL: *Evaluation of computer text editors*. Doctoral Thesis, Stanford, CA, Stanford University, 1979
21. Kieras DE, Polson PG: An approach to the formal analysis of user complexity. *Int J Man-Machine Studies* 2:365-394, 1985
22. Kieras DE, Bovair S: The acquisition of procedures from text: A production-system analysis of transfer of training. *J Memory Language* 2:507-524, 1986
23. Kieras DE: The role of cognitive simulation models in the development of advanced training and testing systems, in Frederiksen N, Glaser R, Lesgold A, et al (eds): *Diagnostic Monitoring of Skill and Knowledge Acquisition*, Hillsdale NJ, Erlbaum, 1990
24. Grey WD, John BE, Atwood ME: The precision of project ernestine or an overview of a validation of GOMS. *Proc CHI Monterey, CA, '92 Human Factors in Computing Systems*, New York, NY, ACM, 1992
25. Fitts PM: The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol* 47:381-391, 1954