Pilot Study on the Effects of a Computer-Based Medical Image System

Shu-Chieh Wu
Department of Psychology
The Ohio State University
Columbus, Ohio

Jack W. Smith, M.D. Ph.D.
Division of Medical Informatics
Department of Pathology
The Ohio State University
Columbus, Ohio

J. Edward Swan II
Department of Computer
and Information Science
The Ohio State University
Columbus, Ohio

Current medical imaging systems are developed for the purpose of data management. Evaluations of these systems are usually done by assessing users' subjective appreciation rather than objectively gauging performance influence. The present report discusses the evaluation of a medical image presentation system prototype utilizing a cognitive approach. Experimental results showed hypothesized performance improvement attributed to advanced presentation techniques. However, this improvement was almost inadvertently masked by users' previous strategies and interactions with new technology. Overall these data demonstrate the potential benefit of implementing such a system in actual practice as well as provide an example of applying the cognitive approach in evaluating the usability of medical systems.

INTRODUCTION

Central to diagnosis and treatment of many patients is the accurate and efficient interpretation of medical images. To improve the accessibility of medical images in view of their excessive data volume, efforts have been made to develop digital imaging techniques [2] as well as computer-based medical imaging system for efficient data acquisition and management [3-7].

The present paper reports the preliminary evaluation of a computer-based medical image presentation system prototype. This evaluation was intended to address the issues of interface design to support medical diagnosis, an important but mostly neglected property in current medical imaging systems. In addition, the evaluation effort followed a cognitive approach [1], going beyond conventional subjectivity-prone methodology and gauging performance impact based on objective criteria. Results of the evaluation provide feedback on not only the potential of such a system in actual practice but also general methodology in evaluating medical systems.

Overview of Current State of Radiology Medical Imaging Systems

Two families of medical imaging systems, Picture Archiving and Communication Systems (PACS) and Radiology Information Systems (RIS), are currently being used in practice. PACS emerged to address the dilemma of the amount of mass storage required for digital medical images and efficient retrieval of medical images. These systems are implemented as distributed networked computers to enable comprehensive access to images stored centrally [3]. PACS accompanied with digital imaging systems also promise the economy of archiving space as well as the efficiency of image utilization [4].

RIS are known for handling radiology-related data, such as the scheduling of patient examination and billing information [5]. Data collected in these systems have potential use for evaluating work flow and quality of data management in the radiology department [6]. In addition, efforts have been made to combine RIS and PACS into an unified radiology image and information system [7].

Motivation

PACS and RIS are mainly developed for supporting efficient data management and image utilization. According to our review, most of them provide minimal image viewing capacity [3-5]; some of them in addition implement simple interfaces supporting manipulation of images [4,5]. Overall little attention has been paid to the design of such an interface.

However, the interface has been a major problem in medical image diagnosis, most evident from traditional examining processes using a lightbox. Medical image diagnosis is a highly sophisticated decision task involving specialized domain knowledge of human anatomy and radiology. Information on which decisions are based is noisy and variable in terms of the generating procedures. Performance of such a task takes place under environmental distractions and time pressure. On top of these, a presentation tool like the lightbox further demands irrelevant sorting, searching, and arrangement of negatives to fit its limited viewing space. Given the task requirements, limitations imposed by the viewing equipment as well as the information sources shape physician diagnostic behavior in ways which may significantly increase cognitive demands.

Among the medical imaging systems reviewed above, most provide panel viewing capacity imitating that already available from lightboxes [4]. Only one of them provides more flexibility in image viewing by presenting, in addition, stacked images [5]. Efforts for improving the viewing interface and enhancing structural information in the image are needed.

One recent study directly addresses the issue of interface design of PACS workstations [8]. The viewing interface in this study was designed under an alternator-filmstrip metaphor, which treats images as sequential frames in a filmstrip. The filmstrip was further augmented with a pictorial directory of all images associated with a patient. The user could select an image from the pictorial directory and scan sequentially away from the selected image in both directions, four at a time.

The pictorial directory was used as a long shot to provide an overview of all the images [8]. Along with the simultaneous display of successive images, it was expected to help establish the spatial representation that described the relationship between these images by enhancing visual momentum across individual images [9]. The efficacy of such an interface was evaluated by providing either one, two, or four concurrent displays as well as manipulating the availability of a pictorial directory. Results in terms of interpretation time with this program showed superior performance with the two-image format and no particular benefit of having a pictorial directory.

The filmstrip metaphor was proposed in view of the diagnostic advantage of large display space provided by conventional film alternators [8]. Therefore, it was implemented in the interface program in a way which resembled the physical properties of an film alternator. However, the ultimate goal of any data presentation method in the context of medical image diagnosis is to facilitate the reconstruction of analog representations of human anatomical structures in the head of the user. It is conceivable that the reconstruction may be perceived even easier if it is done externally and presented to the user. In addition, from a surgical perspective, such a presentation technique if successfully developed would better represent visually actual relationships of anatomical structures and could be further elaborated into a planning tool. Therefore, we developed a computer-based image presentation system which attempted to realize the idea of a well-reconstructed external representation.

The Computer-Based Medical Image Presentation System Prototype

The system prototype was implemented on a Silicon Graphics Indy Workstation, with an R4400 processor running at 174 MHz, 64 Mb of RAM, and running IRIX 5.3. The system had a 24 bit, 1280 x 1204 monitor (physical size measured 13.75w x 11h) with 76 MHz refresh rate.

The system prototype was developed for presenting CT and MRI images for diagnosis preceding head and neck surgeries of benign and malignant cranial base tumors. One major innovation of this system was that individual sliced images taken from the same orientation were organized into continuous movies with topographically related images placed in sequence and displayed as animations.

Movies of eight sources were provided in the system (see Figure 1). Four of them were made from those also available as lightbox negatives, including CT axial, MRI axial, MRI coronal, and MRI sagital cuts. One new source was generated by combining bony windows from CT and tissue information from MRI, and also displayed as continuous movies. The other three were 3-D reconstructions of axial cuts of CT, MRI, and combined CT-MRI images with major structures such as skin, bones, tissues, ventricles, tumors, arteries, etc. hand segmented by false coloring. For each case an animation was created in which the 3-D volumetric head was rotated and cut away in a way which best seemed to correspond to the possible surgical perspectives.

These information sources were communicated to users through a window-based mouse-control interface. Each of the eight sources was contained in a single display window. On the right half of the bottom bar of each window were function buttons supporting visual stream operations typical with a home VCR, such as Play, Step Forward, Step Backward, and Stop. The left half of the bottom bar was designed into a slider which users could drag to view sequential images in a self-paced fashion.

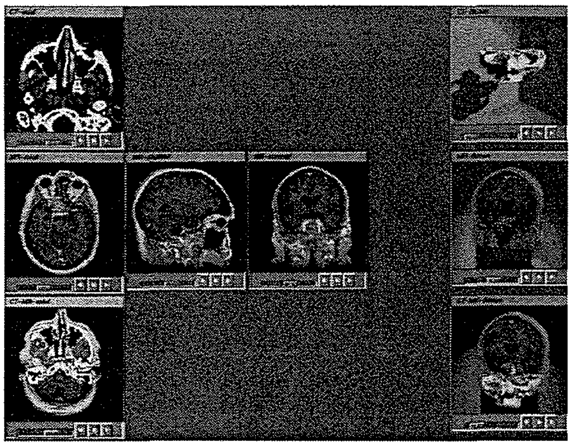

Figure 1. The system prototype screen layout

Figure 1 gives a snapshot of the screen layout. From left to right, the top row displayed CT axial cuts and 3-D reconstruction. The middle showed MRI axial, coronal, sagital cuts and 3-D reconstruction. Combined CT-MRI images and the corresponding 3-D reconstruction were shown on the bottom.

Hypotheses

With such an interface supporting convenient image manipulation and viewing as well as the availability of structurally richer information, we expected the system to improve the quality of diagnosis in the following ways compared to traditional panel-driven tools like lightboxes. First, we hypothesized there would be a decrease in time spent on the whole diagnostic process because operations like sorting have been done in the process of preparing the computerized images. Moreover, we hypothesized that by making all sources of information available at one location the system would reduce the cognitive demands associated with having limited viewing space on a lightbox and thus lead to better performance in terms of diagnostic accuracy.

Evaluation of Medical Imaging Systems

Evaluation of the usability of medical imaging systems has mostly involved interviews and questionnaires of user ratings [3]. These measures are easy to administer but inevitably imprecise, unreliable, and subject to personal biases [1]. In addition, it has been shown that they do not always correspond with performance measured objectively [8]. To overcome the limitations of conventional techniques, Kushniruk and Patel [1] propose instead a cognitive approach, which incorporates video-based data collection and protocol analysis of think-aloud behavior to assess the usability of medical computing systems. Video recordings can fully capture interactions between humans and computers; concurrent think-aloud protocols have been show to correspond to ongoing cognitive processes [10].

Although its strong backgrounds from studies of social sciences and cognitive psychology suggests theoretical plausibility, the utility of the cognitive approach in evaluating medical imaging systems has yet to be tested empirically. Therefore, we adopted it to evaluate the system prototype in discussion. Such an attempt would reveal the usability of this system prototype and also provide feedback on the application of this approach.

METHODS

Participants. Four surgeons from the Arthur G. James Cancer Hospital volunteered in this experiment. Two of them specialized in neurosurgery and two in head and neck surgery. All of them were attending physicians with experience ranging from 2 to 25 years.

Apparatus. One S. & S. 4-panel portable lightbox was used in the experiment, as well as the system prototype introduced previously. Activities in the experiment were recorded through a video camera connected to a 3/4" tape deck. Auditory signals were recorded through a lapel microphone worn by participants and connected to the audio channel of the tape deck.

Materials. Two cases labeled B and 7464 were used in this study as stimulus material to be diagnosed. Case B was a 76 year old white female, diagnosed as left ethmoid cancer with intracranial extension. Case 7464 was a 54 year old male with sphenoid sinus carcinoma along with blindness.

Design. Participants completed two sessions of diagnoses on both cases using the lightbox and the system prototype. The experimental design counterbalanced the use of tools and the stimulus material so that one participant diagnosed case B with the lightbox and case 7464 with the system prototype in the first session and the same two cases with swapped tools in the second session. To avoid any memory from the previous session an interval of at least two weeks between sessions was enforced.

Procedure. In each session, participants diagnosed two cases using both the lightbox and system prototype. During the process they were asked to think aloud, verbalizing their thoughts. Several standard thinking aloud practice problems were given to participants to ensure that they understood this procedure. Participants also received a tutorial and practiced on a movie depiction of a 3-D brain reconstruction on how to use the system prototype. All activities were videotaped and recorded. All mouse events were recorded by the workstation.

RESULTS

Interpretation Time

Interpretation time results of individual participants are summarized in Table 1. Extensive variances were observed across participants using the lightbox, ranging from 395 to 1381 seconds. Even larger variances were seen in the system prototype condition, ranging from 178 to 1675 seconds. At the overall level (the Total column), it seemed to take participants more time to diagnose with the system prototype.

However, comparisons made strictly to interpretation time on conventional CT and MRI information (the Subtotal column) excluding the time on CT-MRI combined images and all 3-D reconstructions (the New column) revealed a different picture: Participants generally spent less time diagnosing cases using the system prototype. Furthermore, these decreases in time are not accompanied by declines in diagnostic accuracy, as will be presented later.

Breaking down interpretation time by component operation categories clarifies the nature of these decreases. In Table 1, Sort measures time taken to sort and search negatives when using the lightbox, and to overview images by fast playing the movies when using the system prototype. Exam, or examination, measures time taken to interpret images. Cross-Ex, or cross-examination, measures time spent on another sources of information in the middle of examining a primary source to that being currently attended. Results showed that time on sorting was largely reduced with the system prototype. Proportion of time allocated to true examination was largely elevated (statistics not shown in the table).

Diagnostic Accuracy

Evaluation of diagnostic accuracy was based on comparisons made between participants' verbal protocols and radiologists' original reports on the two cases. From the reports we derived case codings of anatomical structures involving tumor tissues and

Table 1. Interpretation time results (seconds)

		Conventional CT-MRI Images					
		Sort	Exam	Cross-Ex	**Subtotal**	New	**Total**
Case B							
# 3	Lightbox	135	421	0	**556**	0	**556**
	Prototype	52	315	56	**423**	440	**863**
# 4	Lightbox	164	540	7	**711**	0	**711**
	Prototype	No data due to technical error					
# 5	Lightbox	139	477	0	**616**	0	**616**
	Prototype	63	619	17	**699**	326	**1025**
# 6	Lightbox	119	276	0	**395**	0	**395**
	Prototype	0	64	0	**64**	222	**286**
Case 7464							
# 3	Lightbox	262	424	0	**686**	0	**686**
	Prototype	70	809	0	**879**	796	**1675**
# 4	Lightbox	388	831	162	**1381**	0	**1381**
	Prototype	64	329	20	**413**	620	**1033**
# 5	Lightbox	210	798	0	**1008**	0	**1008**
	Prototype	0	317	17	**334**	389	**723**
# 6	Lightbox	614	490	0	**1104**	0	**1104**
	Prototype	18	81	0	**99**	79	**178**

then looked for matches in participants' diagnostic statements. Results are summarized in Table 2. The prototype neither enhanced nor degraded the participants' performance.

DISCUSSION

The system prototype was designed for the purpose of supporting medical image diagnosis in part by providing structurally rich information. It was a realization of the attempt to facilitate the

Table 2. Accuracy results in numbers of diagnostic statements matching radiologists' reports

		TruePositive & TrueNegative (CT&MRI)	TruePositive & TrueNegative (New)	**Total**
Case B				
# 3	Lightbox	16	0	**16**
	Prototype	11	3	**14**
# 4	Lightbox	11	0	**11**
	Prototype	No data due to technical error		
# 5	Lightbox	16	0	**16**
	Prototype	15	0	**15**
# 6	Lightbox	4	0	**4**
	Prototype	5	1	**6**
Case 7464				
# 3	Lightbox	11	0	**11**
	Prototype	11	0	**11**
# 4	Lightbox	10	0	**10**
	Prototype	8	0	**8**
# 5	Lightbox	9	0	**9**
	Prototype	7	3	**10**
# 6	Lightbox	9	0	**9**
	Prototype	4	0	**4**

reconstruction of spatial relationships contained in separate images by providing external complete representations. Augmented with a convenient image manipulation interface, this system prototype was expected to expedite diagnostic processes and improve accuracy. As the results showed, due to savings on sorting and searching, participants spent less time diagnosing sources available to both tools using the system prototype, and this time decrease did not compromise accuracy. However, time on reviewing new sources delayed overall performance with the system prototype.

By and large, the present study showed promising advantages of applying such a system in actual practice. In addition, the complication of having performance enhancement inadvertently masked by the interpretation of new information sources raised general methodological issues concerning the application of the cognitive approach in medical system evaluations.

First, investigators should be aware of possible task process changes when introducing new systems. In this study the efficiency of having a convenient interface and structurally enriched information was masked by the interpretation of new sources. In future studies there should be stricter controls over the amount of information provided in both tools to provide a better comparison.

Second, users' tendency to apply customized strategies to novel situations may hinder adequate evaluation of new systems. As observed in this study participants adhered to their fixed sequences of reviewing films (CT first, MRI later). As a result, new sources were often reviewed last in the process. Such a bias may account for the relatively few diagnostic statements made during reviewing those new sources. In our case, an adequate evaluation may be achieved by presenting each source singly.

CONCLUSION

A medical image presentation system was developed for the purpose of providing structurally richer information and a convenient interface. Results of this study demonstrated the potential of such a system in facilitating medical image diagnosis. Furthermore, as an evaluation experiment on the cognitive approach, lessons learned from this study will prompt investigators to pay special attention to the dynamics of humans interacting with new systems in addition to taking advantages of advanced recording techniques for comprehensive data collection.

Acknowledgments

Our thanks to Greg Wiet, M.D. and David Schuller, M.D. for inviting our participation in this project, to Don Stredney for advice and direction, and to Naeem Shareef for implementing the system. We also thank Gerald Disimone and Dennis Sessanna for assistance in running the experiments, Ferdi Sheepers for developing the CT-MRI merging system, and Debra Parker for transcribing the dialogue.

References

1. Kushniruk AW, Patel VL. Cognitive computer-based video analysis: Its application in assessing the usability of medical systems. *MEDINFO 95 Proceedings*, 1995; 1566-1569.
2. MacMahon H, Doi K. Digital chest radiography. *Imaging Strategies in Pulmonary Disease*, 1991; 12: 19-32.
3. Siegel E, Brown A. Preliminary impacts of PACS technology on radiology department operations. *Proceedings of AMIA*, 1994; 917-921.
4. Yachida M, Maeda T, Nishimoto H, Yoshida S, Tohyama Y. PACS in Kochi medical school hospital. *Computer Methods & Programs in Biomedicine*, 1991; 36: 85-88.
5. Ratib O, Ligier Y, Scherrer JR. Digital image management and communication and medicine. *Computerized Medical Imaging & Graphics*, 1994; 18 (2): 73-84.
6. Crabbe JP, Frank CL, Nye WW. Improving report turnaround time: An integrated method using data from a radiology information system. *AJR,* 1994; 163: 1503-1507.
7. Wimmer MS, Blaine GJ, Jost RG, Moore SM, Studt JW. A distributed approach to integrated inquiry and display for radiology. *Journal of Medical Systems*, 1991; 15(4): 299-309.
8. Gee JC, Barfield W, Haynor DR, Kim Y. User interface design for medical imaging workstations: Image display and processing. *Interacting with Computers*, 1993; 5(3): 279-294.
9. Woods DD. Visual momentum: A concept to improve the cognitive coupling of person and computer. *Int J Man-Machine Studies*, 1984; 21: 229-244.
10. Ericsson KA, Simon HA. *Protocol Analysis*, 1993. Cambridge, MA: MIT Press.