Pilot Study on the Effects of 3-D Reconstruction of Medical Images on Physician Diagnostic Decision

Diagnosis using medical images like CT or MRI as a prelude to head and neck surgery is a highly sophisticated decision task. Traditionally the task is performed using slice images sorted on a lightbox, from which only limited viewing space is available. As a result, the space limitations imposed by the viewing equipment and the limited structural information available on each separate film shape physicians diagnostic behavior in ways which may significantly increase the cognitive demands of the task.

The present pilot study was done for the purpose of evaluating the usefulness of organizing 2-D images as movies and providing 3-D reconstruction of such imaging information as an aid for human performance. A window-based computer interface was developed for presenting these images. In this interface, images from the same source are organized into continuous movies with topographically related images placed in sequence. We hypothesized that this would decrease the need for potential large-scale sorting and searching evident using a lighbox. All images regarding a patient are presented concurrently by the interface and are readily available, supporting navigation from one information source to another. In addition, a new source of 2-D information was provided by the computer system combining bony windows from CT and tissue information from MRI. We also reconstructed 3-D volumetric images from two dimensional CT, MRI and combined this information to mimic actual human anatomy by segmenting and coloring the images by major anatomical structures.

From these changes in the presentation of information, we hypothesized there would be a decrease in time spent on the whole diagnostic process, because some imaging sorting had been done in the process of preparing the computerized images. We also hyopthesized that the new presentation would reduce the cognitive demands imposed by having limited viewing space on the lightbox, and thus lead to better performance in terms of diagnostic accuracy.

Preliminary analysis at a gross level of participant behavior on cases presented via lightbox or computer system did not show any significant advantage of using the computer interface to do diagnosis either in terms of time and accuracy. Physicians were generally slowed down by their fascination and unfamiliarity with the technology. They also maintained old habits of viewing and interpreting 2-D information to determine structural relationships. Accuracy did not improve either even with reduced sorting of images and the extra information sources of CT-MR combined and 3D reconstructed images. However, fine-grained result analysis based exclusively on the time participants spent using the computer system to interpret traditional 2-D CT and MRI information as continuous movies revealed expected changes in behavior. Physicians actually spent much less time on the computer interpreting movies of these sources, due to the reduction in time to sort and examine, especially the former. A larger proportion of time was then spent on examination with the computer, contrasting with the time spent using the lightbox.

Several lessons learned from the current pilot study will guide future efforts toward elucidating the conditions under which such computer systems will improve and alter diagnostic performance and how such evaluations can be used to suggest alternative interface designs. First, in future studies there should be stricter controls on the amount of information provided on both media for a valid comparison. New information sources under evaluation like the 3D reconstructions need to be singled out to overcome physicians' well-trained diagnostic strategies tied to the presentation of 2-D information in a fixed sequence. Second, more structured diagnostic tasks will need to be used in order to better highlight the performance difference of interest. In addition, the present data suggest alternatives in interface design, such as synchronization of information sources. There is also need to future explore issues concerning the temporal mapping of spatial information and its effect on temporal versus spatial scanning.

A more global difficulty with such studies and our current design methodology also emerges, our insufficient understanding of the process physicians use in image interpretation and its relationship to time on task as well as success or failure. The current technology driven approach to computer system design in this area is efficient in immediate application but is not grounded in an adequate understanding of what will be required to better support diagnosis. Therefore, we would argue for design approach driven by such a process model as the best hope of realizing the potential of novel 2-D information presentation and 3-D reconstruction. In this approach models of how diagnosis is currently done , what difficulties are currently involved and how new imaging presentations alter this process would be developed. We could then use these models as a major constraint on investigations and evaluations of alternative system designs.

Pilot Study on the Effects of 3-D Reconstruction of Medical Images on Physician Diagnostic Decision

LKBMS Technical Report #95-000

Introduction

Diagnosis on medical images like CT or MRI for head and neck surgery is a highly sophisticated decision task. It involves extensive specialized domain knowledge regarding human anatomy and radiology. Information on which decisions are based is noisy and variable in terms of the procedures to generate it. Physicians perform the task under environmental distractions, time pressure, and the tension that a patient's health is at stake.

Traditionally the task is done with a lightbox. Because of the limited viewing space available from one lightbox, people who perform the task spend time to sort the scans, choose scans they want, and then finally arrange them on the lightbox. The cycle occurs each time they move on to the next set of scans. Findings from one set of scans need to be memorized to be carried over. As a result, the limitations imposed by the equipment has shaped the diagnostic behavior in way which may further exacerbate the cognitive demands of the task.

The present pilot study was done for the purpose of aiding human performance in the medical image diagnostic task. We developed a window-based computer interface to present images for physicians to do diagnosis. In this interface, same source topographically related images are organized into continuous movies, accompanied with function buttons associated with common operations that can be used to manipulate movie streams, like Play, StepForward, StepBackward, etc. There is no need for large scale physical sorting and searching. Furthermore, all images regarding a patient are presented concurrently and readily available. People who use the interface can easily switch from one information source to another and have a better chance of viewing all of them at the same time. In addition, a new source of information is provided combining bony windows from CT and tissue information from MRI. We also reconstructed 3D volumetric images from two dimensional CT, MRI and combined this information to mimic actual human anatomy by segmenting and coloring the images by major anatomical structures.

By such innovation in presenting information, we anticipated the following things would happen. First, the new technology should decrease the time spent on the whole diagnostic process, because some sorting has been done in the process of preparing the computerized images. Second, the new technology would reduce the cognitive demands imposed by having limited viewing space on the lightbox, and thus lead to better performance in terms of accuracy improvement. Finally, because of the availability and accessibility of new information, the original one-line straightforward diagnostic process now may evolve into some other forms, which in turn may give feedback on how to improve the interface and shed light on the medical diagnostic behavior.

In this study, participants diagnosed two postoperative cases using both lightbox and computer, and were required to think aloud during the process. All activities were videotaped, along with their verbal protocols recorded. In the condition where they used the computer interface, key strokes and mouse movements were also tracked. Verbal protocols were later transcribed into texts to be compared with radiology reports to derive diagnostic accuracy results. Gross timings were taken from the videotape regarding several predefined operation categories to provide information on time allocation.

Preliminary analysis at a gross level did not show any significant advantage of using the computer interface to do diagnosis either in terms of time and accuracy. Physicians were generally slowed down by their fascination with the technology. Accuracy did not improve much either even with the extra information sources of CT-MR combined and 3D reconstructed images. However, fine-grained result analysis revealed expected patterns of behaviors, which also indicate promising further investigations. Following discussions regarding issues arising from the present study and topics to be explored later, we argue for a problem driven approach for analyzing this problem domain and generating design ideas.

Method

Participants

Five surgeons from the James Cancer Hospital volunteered in this experiment. Two of the participants were neurosurgeons. The other two were otolaryngologists. All of them were male.

Apparatus

The computer facility

The computer prototype was presented on a Silicon Graphics INDY UNIX graphics workstation. It was installed with 64MB memory. A monitor sized 13.75 (width) x 11 (height) provided 1280 x 1024 pixels resolution.

The computer prototype

Physical properties. The computer prototype provides nine sources of information regarding each case, including CT Axial, CT Reconstruction (abbreviated as Recon in the rest of the report), MR Axial, MR Sagital, MR Coronal, MR Recon, CT-MR Axial, CT-MR Recon. They are arranged on the screen as Figure 1 shows, with axial views occupying the very left column, followed by sagital and coronal views, and 3D reconstructions. CT is presented on the top, then followed by MRI and CT-MR combination scans. Table 1 summarizes the sizes of windows.

Functionality and operations. The prototype was designed into a window layout (Figure 2). On the header of a window described what information source is contained in that particular window. Function buttons that support operations lie in the bottom of the panel. For information sources that have left-right orientation property, letters L and R were put at the corresponding corner to designate the orientation of the information.[1]

The bottom of a panel, or the widget, can be in three modes, each of which supports different operations. The default mode is *NormalMode*, which contains the *Slider*, *StepBackward*, *StepForward*, and *Play* buttons. The *Slider* supports a mouse dragging interface for the user to navigate back and forth across scans. The *StepBackward* and *StepForward* buttons support mouse-clicking frame-by-frame presentation. The *Play* button supports one-time movie-style presentation.

The *Slider* supports two modes of operations. Users can click on the slider, hold down the left or middle mouse button, and drag the slider to go through scans frame by frame. This is the *DragMode*, which shows up in blue background color. On the other hand, users can click above the slider widget and get into the *FollowMode*, in which users can move through scans frame by frame just by moving the mouse left or right while it is over the slider widget, without depressing any mouse button. In this mode the slider background is brown. Users can toggle between the *DragMode* and *FollowMode* by clicking above the widget.

When the user clicks the *Play* button, the widget gets into the *PlayMode*, in which the *Stop* and *Pause* buttons are visible. If the user clicks the Pause button during movie playing, the widget then gets into the *PauseMode*, in which the *Stop* and *Continue* buttons are visible.

The functions of all buttons in the prototype are represented by conventional symbols that are generally used on VCR's.

Materials

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

Lightbox films. For case B, 11 CT image films, including axial bony windows, axial soft tissue window, axial fine cuts, and coronal bony window reconstruction were

---

[1] Because of errors in the material preparation process, some images were left right flipped regarding the actual images. Participant 3 who was run first on the computer pointed out this problem. Later the labels were removed from the prototype. All participants thereafter received unlabelled information and assumed orientation of the scans by their own familiar conventions.

available for diagnosis. MR image for case B films included axial T1 enhanced, axial T1 not enhanced, axial T2, sagital T1, coronal T1 enhanced, and axial T1 and T2 from some other unknown settings. Films available for case 7464 included 14 CT films and 12 MR films. CT films included axial bony window, axial soft tissue window and coronal bony window reconstruction. MR films included axial T1 not enhanced, sagital T1, coronal T1 enhanced, and axial T1 and T2 from some other unknown settings. Details regarding the size and number of each kind of scans are provided in Table 2.

Computer animation. (To be completed by Ed) ✓

Design

A two-factor complete Within-Subject design was used in this experiment. One factor is case. Two case materials were used, case B and case 7464. The other factor is the diagnostic media, which were the traditional lightbox and the computer virtual reality prototype. Participants completed two sessions on the two cases using both lightbox and computer. Odd numbered participants did case B on lightbox and case 7464 on computer in their first sessions, and did same cases on the other medium in their second sessions. Even numbered participants were assigned the opposite manipulation. Participant 7 was assigned as even numbered condition. Intervals between sessions varied from 19 to 63 days. A detailed summary of condition assignment is provided in Appendix A.

Procedure

In each experimental session, the participants were asked to diagnose a postoperative case on the case's preoperative scans. All practice and diagnostic activities were videotaped and recorded. All key stroke activities were recorded as well if the computer prototype was being used for diagnosis.

The experiment was conducted in two conference rooms; one was for the lightbox and the other was for the computer. Both rooms provided a quiet environment without any special noise prevention technology.

Before the experiment session, participants were debriefed on the purpose of this experiment. In the very first session for every participant, the experimenter provided the participant with instructions of thinking aloud and practiced him on several tasks, including one simple arithmetic problem and two anagrams (see Appendix B for procedure description).

In the experimental session using the computer prototype, participants were instructed on how to use the computer prototype. An movie depiction of a 3D brain reconstruction served as practice material. Participants tried out all possible operations on it until they felt comfortable with the prototype. Then the experimenter brought up the complete window layout of the computer prototype and explained the containment of each windows to the participant.

In the experimental session in which the lightbox was used, participants were given two envelops of unsorted scans, CT and MRI.[2] After the experimenter read the case scenario, the participant moved on to the diagnostic task if he did not have any further question or nothing else about the case the experimenter could provide. The participant was allowed to proceed the task in any order. The session ended when the participant gave a verbal summary about the case diagnosis or informed the experimenter that he had got an idea about the case. Sometimes the experimenter asked the participant to give a verbal summary about the diagnosis if the latter had not explicitly done so. The whole session took about 10 to 20 minutes per case.

In the experimental session in which the computer prototype was used, participants started the task after the experimenter had given the case scenario. They were able to freely perform the task in any order. The session ended when the participant summed up his diagnosis or he informed the experimenter about reaching a conclusion. Sometimes the experimenter prompted the participant to give a verbal summary about the diagnosis if he had not explicitly done so. The time spent in a session ranged from 20 to 60 minutes.

After the experimental session using the computer, participants were asked to comment on the prototype and completed a questionnaire about their training and experience on their specialty and computer technology, and their appreciation of the prototype.

The experimenter was present during the whole experiment session to give instructions, answer questions regarding the case scenario, and remind the participant to think aloud if the latter remained silent for any period of time. A video recording technician was also present to monitor and adjust the voice level. Sometimes there were one or two observers in the room as well. All except the experimenter were out of the sight of the participant and remained silent.

Results

Four of the five participants completed both sessions using the computer prototype and lightbox on both cases. Participant 7 only did one session using the computer prototype on case B and using the lightbox on case 7464.

Preprocessing of data

Videotapes were transferred to VHS format videotapes for analysis.

---

[2] Except Participant 4 and Participant 7 received sorted and numbered CT and MRI scans in their second and first sessions using the lightbox, respectively.

General impression

Participants were assumed to perform diagnosis on the lightbox in their usual manner. Obvious carryover effects from extensive experience in using the lightbox were observed on the computer condition as the participants still followed the examining sequence of CT first and MRI second, then viewed the new information.[3]

Gross timing on diagnostic processes

Gross timings were taken from the videotapes for the following operation categories.

Sort: Coded when explicit sorting and searching behavior was observed in conditions where the lightbox was used; also coded when image animation was played or scanned through quickly in conditions where the computer prototype was used. [4]

Exam: Coded when explicit diagnostic behavior was observed in experimental conditions using both media.[5]

Cross: Coded when participants referred to previously examined other information sources.

Total: Total time spent on the particular information sources.

Each of these operation categories can be associated with three possible information sources.

CT: Computed Tomography, including all available views and cuts for the particular case.[6]

MR: Magnetic Resonance Imaging, including all available views and cuts for the particular case.

EX: (only applicable to the computer prototype) All information sources that are not available on the lightbox, including CT Reconstruction, MR Reconstruction, CT-MR Combined, and CT-MR Combined Reconstruction.

Gross timing results are summarized in Table 3. One thing immediately emerges from the results is the extensive variance of reaction times across participants. Diagnoses

[3] Two exceptions, though. One participant explained that he started out with MRI because he did not know there were CT films as well. In the other case the participant chose to start on MRI first but explained that was only because MRI films were put on top of the CT films.

[4] In conditions where the lightbox was used, sorting time sometimes mixed with diagnostic examination. Therefore, the estimation of sorting time on the lightbox would sometimes be overestimated.

[5] In conditions where the computer prototype was used, examining behavior would always involve more or less degree of sorting and searching. Therefore, the estimation of time spent on examination on the computer prototype would mostly be overestimated.

[6] Although information available on both media did not match exactly (see Table 2), all image scans that came from original CT and MRI were categorized as such, regardless of their correspondence.

done on the lightbox took around 395 (Case B, Participant 6) to 1381 seconds (Case 7464, Participant 3). Even larger variances were seen in conditions using the computer prototype, ranging from 178 (Case 7464, Participant 6) to 1675 seconds (Case 7464, Participant 3). At this level, it seems to take participants more time to diagnose cases with the new technology.

Further comparisons between the lightbox and computer are made strictly to the use of CT and MR information in order to remove any effects possibly caused by having more information on the computer in the form of 3D reconstructed and CT-MR combined images. These times are illustrated in Table 3 and Figure 3. The results now show that generally participants spent less time on the computer, due to the reduction in time to sort and examine, especially the former. Examining time varies from participants to participants and from case to case, showing a slight trend of decreasing on the computer as well. To anticipate, these decreases in time are not accompanied by a decrease in diagnostic accuracy, as will be presented later.

Figure 4 shows the same data of Figure 3 but presented in terms of time allocation percentages. Results show that the computer prototype has changed the task performance by reallocating time among subtasks. The proportion of time spent on sorting images is significantly reduced if the computer was being used.

Diagnostic accuracy[7]

In addition to diagnosis time, we evaluated participants' diagnostic performance in terms of accuracy. The evaluation used radiologists' original reports on the two cases as sources of criteria (see Appendix C). We derived some case codings of anatomical structures that involved tumor tissues based on the radiologists' findings (see Appendix D) and compared those codings against participants' verbal reports.

Table 4 summarizes the comparison results. What appear in the table are frequencies of participants making the same statements regarding the involvement of certain anatomical structures. The *True Positive & True Negative (same source)* column contains frequency counts for which participants made the same statement on the same source of information as radiologists did. The *True Positive & True Negative (different source)* column contains frequencies for which participants made the same statement as radiologists but from a different information sources. An example would be that the participant made a statement based on CT, while the same statement made by radiologists was based on MR. The *True Positive & True Negative (EX)* describes the frequencies that participants made statements based on extra information sources which were only available on the computer. Therefore cells of this kind but corresponding to data collected from the lightbox would show zero frequency. The *False Positive & False Negative* column contains instances in which participants made opposite statements than radiologists regardless of information sources. The last column shows the *Total*

[7] Participant 6 was given wrong MRI films in his second session using the lightbox. Therefore his diagnosis for those films was excluded from accuracy analysis.

frequencies, including*False Positive & False Negative*, over the *Total Statements* made by radiologists. There is no significant enhancement in diagnostic accuracy observed in Table 4 comparing subject performance on the lighbox and computer.

Discussion

The computer prototype was design for displaying CT and MRI medical images in an animated fashion, plus there were combined CT and MRI and reconstructed 3D volumetric images. It was expected that this would lead to better diagnoses in terms of both speed and accuracy, since the prototype was expected to make the cognitive task of interpretation easier. In particular the reduction in physical sorting usually demanded by viewing films on a lightbox and the ability in the computer to provide the same information in several different layouts was predicted to improve performance. As the results showed, at the very surface level, participants' performance did not seem to confirm these hypotheses. They spent more time doing diagnosis and the accuracy level seemed to stay the same.

However, a finer-grained more focused analysis revealed different results. When the time spent by participants on the extra combined and reconstructed information sources was excluded from the analysis, we saw a reduction of sorting time and an increased proportion of time spent on examination in conditions where the computer prototype was used. No significant accuracy improvement was found, however. It should be noted that several participants made positive statements about how 3D reconstruction made them *see* CT and MRI information better, although our current analysis does not provide support for this. In addition to changes in performance, participants demonstrated creative behavioral patterns of using the computer prototype on their second encounters. Several participants started with the information gathered from the reconstruction and navigated across other standard information sources CT, MRI in non-traditional sequences.

By and large, the present study turns out to be a promising failure. It is a failure because overall diagnoses took longer with no significant accuracy increment. However, it is promising because more detailed analysis of component behaviors actually moved in the expected direction. Future work that should be done to more clearly show the differences can be divided into three areas: experimental control, research approach, and problem domain. The following is a discussion of issues regarding these three areas.

Experimental control: What could be done

Several things concerning experimental control probably contributed directly to the difficulties in deriving verifiable comparisons of time between the use of lightbox and computer. As mentioned repeatedly, participants were given unequal amounts and different sources of information in the two conditions. As a result, true performance improvement was attributable to the inclusion of extra information sources. Furthermore, even after practicing on an unrelated case, participants still showed signs of unfamiliarity and novelty with the computer prototype. Several of them did not have any experience using computers. It was not uncommon to find them verbalizing their fascination with the technology along with the diagnostic process in their verbal protocols. And there were instances where they accidentally hit the wrong keys and then spent time to realize what they should have done. In the future training to a higher level of familiarity would be required.

Another problem which might have decreased performance in the computer condition was underspecification of the experimental task. Though all participants are experts on diagnosis with the lightbox, they seemed to have uncertainty about what the experimenter wanted them to do in an experiment like this which involves the use of new technology. They might have gotten the wrong impression that they were involved in an usability experiment so that they needed to explore the whole thing before stopping. Ill-specified task requirement may have caused them to spend more time than necessary to make sure they met the experimenter's expectation.

On the other hand, diagnostic styles of the participants could have confounded accuracy measurement. All participants are experienced medical practitioners and have built their own pattern of performing diagnosis. They clearly carried the same strategy over to the new technology. Therefore, the accurate statements made on 3D reconstruction or combined data may have been underestimated because these information sources were most of the time approached later in the diagnostic process after interpretation has taken place using traditional CT and MRI data.

Research approach: What else can be explored

In addition to insufficiencies in the current study methodology, the results also suggest other possibilities. One recurring event in the above discussion is the unequal information given to the participants from the two media in the present study. One future direction would definitely be to provide exactly the same amount and sources of information. Since the combined CT and MRI images and the 3D reconstructions are the particular focus of this project, it would be interesting to see how well physicians do diagnosis with exclusively these new information sources.

Another possible alternative is to investigate the enhanced component behavior as observed in the present study by designing a series of structured specific tasks that directly approach the question at the level carryout. An example would be to ask participants to find the epicenter of a tumor using both media and compare performances.

There are other things worth exploring in the interface design. Participants sometimes demonstrated intentions to match the anatomical location on different kinds of films. With a lightbox, they set the interested scans aside, usually on the lightbox. The same goal can be accomplished on the computer by simply clicking on the relevant information source, playing through it, and stopping at the spot. If such cross reference is actually a significant part of the diagnostic process, the computer prototype can be redesigned to have a mode in which all information sources are automatically synchronized by their corresponding anatomical structures. In so doing, diagnostic time would be expected to be reduced even more and accuracy increased due to decreased cognitive overload in maintaining correspondence of the images mentally.

Although the computer prototype was intended to help surgeons diagnose images, it actually changes the way users acquiring information in the process. Consecutive images on the lighbox are concurrently distributed in several spatial locations. However, images are organized to occupy only one location on the computer, which requires users to navigate through a temporal dimension to get the whole picture. From the lightbox to computer, users shift from dealing with spatial mapping of information to temporal mapping. The effect of this change in representation of the task is an unanswered issue which strongly deserve further investigation.

Problem domain

As we went along figuring out how technology changes the way physicians diagnose cases, we were confronted with our ignorance of how they have been doing diagnosis using images in the first place. There are very few, if any, studies which directly addresses this problem domain. We started out from a technology driven approach, in which we have a system to evaluate and then inspect how things are changed by the system. A problem with this approach is that we always see the question through the system, which brings in its own biases. It would be helpful if we could step back and take a problem driven approach, in which we look at how diagnosis is really done, what difficulty is involved in a task like this, and then think of how to support it through the images.

Finale

Computer based image processing technology has become one prominent approach in medical decision support, but enhancing performance using these technologies may not be that straightforward. Little attention has been paid to critical questions. Why do current approaches work and why do they fail? What we presented in this report was an instance in which technology failed to improve performance but shows promise. We have argued for a more problem driven approach for investigating system design as the best hope of realizing this potential.

Appendix A
Experimental Condition Assignment Summary

	Session 1		Session 2	
	Lightbox	Computer	Lightbox	Computer
Participant 3	B 7/15/95	7464 7/15/95	7464 8/3/95	B 8/3/95
Participant 4	7464 7/15/95	B* 7/15/95	B 9/16/95	7464 9/16/95
Participant 5	B 7/15/95	7464 7/15/95	7464 8/26/95	B 8/26/95
Participant 6	7464 7/15/95	B 7/15/95	B 8/5/95	7464 8/5/95
Participant 7	7464 9/23/95	B 8/26/95		

* incomplete

Appendix B D
Procedure of Think Aloud Procedure

(To be completed by Ed)

~~Appdex~~
Appendix B
Sub information form

Appendix C
Consent Form

Appendix C
Radiologists' Case Reports

(Printouts)

Appendix D
Correspondence between Radiologist Findings and Case Coding

Case B: CT

Radiologist's Findings	CT FINDINGS
Epicenter	
• opacification of the frontal sinus	• CT opacification of the frontal sinuses
• frontal sinus expanded	
• opacification of the left maxillary sinus	• CT the left maxillary sinus
• opacification of the left ethmoid air cells	• CT the left ethmoid air cells (expanded w/ walls destroyed)
• opacification of the left nasal cavity	• CT the left nasal cavity
• left ethmoid air cells are expanded and destroyed	
Anterior extension of mass to:	[anterior extension]
• nose	• CT nose
• soft tissue of the face	• CT soft tissues of the face
• medial wall of the left orbit also destroyed	• CT medial wall of the left orbit
• fat surrounding the globe appears intact	• CT (-) fat surrounding the globe
• nasal septum is destroyed anteriorly	• CT anterior nasal septum
• anterior margin of the frontal sinus inferiorly is destroyed	
• inner table of the skull surrounding the opacified frontal sinuses is intact	• CT (-) inner table of the skull surrounding the opacified frontal sinuses,
• except inner table of skull the most inferior sections of the skull surrounding the frontal sinuses where the greater wing of the sphenoid on the left is irregular and somewhat eroded in shape in the region of the roof of the orbit	
• middle cranial fossa is intact	• CT (-) middle cranial fossa
Left extension of mass to:	[left extension]
• left medial rectus muscle (mass abuts this)	• CT left medial rectus muscle
• medial wall of the left maxillary sinus is destroyed	• CT the medial wall of the left maxillary sinus

- left lateral margin of the nasal bone destroyed
- opacification of the left sphenoid sinus
- anterior bony margin of the left sphenoid sinus (irregular)

Superior extension of mass to:

- superior aspect of the ethmoid air cells is almost completely destroyed.

Impression- A large mass also involving:

- the left ethmoid air cells
- the nasal cavity
- the left orbit
- opacification of the frontal sinuses
- some bony destruction of the greater wing of the sphenoid in the region of the roof of the orbit
- no intracrainal extension

- CT the left lateral margin of the nasal bone
- CT left sphenoid sinus w/ opacification

CT IMPRESSION

- CT the left orbit,
- CT bony destruction of the greater wing of the sphenoid in the region of the roof of the orbit
- CT (-) intracranial extension

Case B: MRI

Radiologist's Findings

Extensive soft tissue mass involving:
- opacification of the left ethmoid sinuses
- opacification of the sphenoid sinuses
- opacification of the maxillary sinuses
- opacification of the frontal sinuses bilaterally
- mass remains extra-conal
- proptosis on the left

Left extension of mass through:
- the left cribriform plate
- crosses the medial orbital wall with bony destruction

Anterior extension of mass to:
- the left cranial fossa involving the left frontal lobe medially

Posterior extension of mass to:
- no intracranial extension
- the ventricle and cisternal spaces are normal in size and configuration and midline
- no involvement of white matter throughout the cerebral hemisphere

Impression- Primary neoplasm involving:
- the paranasal sinuses
- frontal sinuses are destroyed bilaterally
- primarily the left ethmoid sinuses
- opacification of the left sphenoid sinus
- opacification of the left maxillary sinuses

MRI FINDINGS

- MR remains extra-conal ?
- MR proptosis on the left

[left extension]
- MR the left cribriform plate
- MR crosses the medial orbital wall w/ bony destruction

[anterior extension]
- MR the left cranial fossa involving the left frontal lobe medially

[posterior extension]
- MR (-) intracranially, the ventricle and cisternal spaces are normal in size and configuration and midline
- MR (-) white matter throughout the cerebral hemisphere

MRI IMPRESSION
- MR primary neoplasm involving the paranasal sinuses
- MR bilateral frontal sinuses
- MR primarily left ethmoid sinus
- MR left sphenoid sinus w/ opacification
- MR left maxillary sinus w/ opacification

- medial wall of the orbit is destroyed
- no intra-conal extension
- there is proptosis on the left as well as spreading to the anterior cranial fossa involving the left frontal lobe

- enhancement of mass with gadolinium
- no enhancement of cerebral hemisphere except frontal lobe on left
- white matter changes consistent with cerebrovascular ischemic injury in patient this age

Case 7464: CT

Radiologist's Findings

Epicenter
- the sphenoid sinus (completely filled with soft tissue mass)

Anterior extension of mass to:
- posterior nasal cavity

Posterior extension of mass to:
- posterior aspect of the ethmoid sinuses with bone destruction and infiltration
- sella turcica is completely eroded
- the brain

Right extension of mass to:
- right cavernous sinus (filled with tumor & distended)
- right optic nerve with encasement
- right vidian canal and foramen rotundum

Bilateral extension of mass to:
- both petrous apices with bone destruction

Negative for mass:
- bilateral foramen ovale & foramen spinosum
- bilateral pterygopalatine fossae

Impression- A large destructive mass
- centers in the sphenoid sinus
- mass extends to posterior aspect of the nasal cavity
- mass involves posterior right clinoid process
- mass involves right cavernous sinus
- mass involves both petrous apices with bone destruction

CT FINDINGS

- CT sphenoid sinus (epicenter)

mark who found it perhaps 4 columns w/ +

[anterior extension]
- CT posterior nasal cavity

[posterior extension]
- CT posterior aspect of the ethmoid sinuses w/ bone destruction and infiltration
- CT eroded into the sella turcica and up into the brain

[right extension]

show borders

- CT right cavernous sinus (filled with tumor & distended)
- CT right optic nerve w/ encasement
- CT right vidian canal and foramen rotundum

[bilateral extension]
- CT both petrous apices w/ bone destruction

- CT (-) bilateral foramen ovale & foramen spinosum
- CT (-) bilateral pterygopalatine fossae

CT IMPRESSION

- CT posterior right clinoid process

Case 7464: MRI

Radiologist's Findings	MRI FINDINGS
Epicenter	
• complete opacification of the sphenoid sinus	• MR sphenoid sinus
• the normal marrow within the clivus replaced with mass	• MR the clivus
Right extension of mass:	[right extension]
• through the right clinoid process to the right optic nerve	• MR through the right clinoid process to the right optic nerve
• may be encasement of optic nerves within the optic canal as well	
• opacification of the posterior ethmoid air cells	• MR posterior ethmoid air cells w/ opacification
Bilateral extension of mass to:	[bilateral extension]
• cavernous sinuses bilaterally	• MR cavernous sinuses
• encircles carotid artery within the cavernous sinus	
Negative for mass:	
• no extension of the mass to the true orbit	• MR (-) no extension to the true orbit
• within brain no mass	• MR (-) within brain
• ventricles and cisternal spaces are normal in size and configuration	
Impression- A soft tissue mass:	
• involving the sphenoid sinus	
• invading the clivus	
• extending through right anterior clinoid	
• extending to encircle right optic nerve within the optic canal	
• involving the cavernous sinuses bilaterally	

Table 1
Window Sizes in Each Study (width x height in inch)

Windows	Case B	Case 7464
CT Axial	2.78 x 2.68	2.78 x 2.75
CT Recon	2.78 x 2.58	2.78 x 2.58
MR Axial	2.78 x 2.68	2.78 x 2.75
MR Sagital	2.78 x 2.58	2.78 x 2.58
MR Coronal	2.78 x 2.58	2.78 x 2.58
MR Recon	2.78 x 2.58	2.78 x 2.58
CT-MR Axial	2.78 x 2.75	2.78 x 2.75
CT MR Recon	2.78 x 2.58	2.78 x 2.58

Table 2
Image sizes and numbers

		B lightbox	B computer	7464 lightbox	7464 computer
CT					
axial	bony window	3.38 x 3.58 (1..5)	2.78 x 2.68 (up to half)	3.78 x 3.16 (1..6)	2.78 x 2.75 (end at 3)
	soft tissue window	3.15 x 3.17 (5..10)	N/A	(7..12)	N/A
	fine cuts	3.13 x 3.9 (11?)	N/A	N/A	N/A
coronal	bony window recon	N/A	N/A	(13..14)	N/A
MR					
axial	T1 enhanced	3.13 x 3.13 (1..2)		N/A	N/A
	T1 not enhanced	3.13 x 3.15 (3..4)		3 x 3 (1..5)	
	T2	3.13 x 3.13 (5..6)		N/A	
	T1 (unknown setting)	N/A	N/A	3.13 x 3.13 (9..10)	N/A
	T2 (unknown setting)	N/A	N/A	3.13 x 3.13 (11..12)	N/A
sagital	T1	3.1 x 3.1 (7) (L->R)	(more)	3.16 x 3.16 (8)	
coronal	T1 enhanced	N/A	(back to front)	4 x 3.13 (6..7)	

Table 3

Gross timing results

Case B		CT sort	CT exam	CT cross	MR sort	MR exam	MR cross	EX sort	EX exam	EX cross	CT total	MR total	**CT&MR**	EX total	**Total**
Participant 3	Computer	8*	79	20	44	236	36	50	390	0	107	316	**423**	440	**863**
	Lightbox	105	175	0	30	246	0	0	0	0	280	276	**556**	0	**556**
Participant 4	Computer **	0	0	0	0	0	0	0	0	0	0	0	**0**	0	**0**
	Lightbox	143	287	0	21	253	7	0	0	0	430	281	**711**	0	**711**
Participant 5	Computer	0	128	17	63	491	0	90	236	0	145	554	**699**	326	**1025**
	Lightbox	118	301	0	21	176	0	0	0	0	419	197	**616**	0	**616**
Participant 6	Computer	0	15	0	0	49	0	0	140	82	15	49	**64**	222	**286**
	Lightbox	54	130	0	65	146	0	0	0	0	184	211	**395**	0	**395**

Case 7464		CT sort	CT exam	CT cross	MR sort	MR exam	MR cross	EX sort	EX exam	EX cross	CT total	MR total	**CT&MR**	EX total	**Total**
Participant 3	Computer	18	179	0	52	630	0	72	705	19	197	682	**879**	796	**1675**
	Lightbox	156	204	0	106	220	0	0	0	0	360	326	**686**	0	**686**
Participant 4	Computer	20	262	0	44	67	20	111	483	26	282	131	**413**	620	**1033**
	Lightbox	202	240	0	186	591	162	0	0	0	442	939	**1381**	0	**1381**
Participant 5	Computer	0	100	17	0	217	0	0	389	0	117	217	**334**	389	**723**
	Lightbox	122	447	0	88	351	0	0	0	0	569	439	**1008**	0	**1008**
Participant 6	Computer	18	13	0	0	68	0	0	79	0	31	68	**99**	79	**178**
	Lightbox	290	307	0	324	183	0	0	0	0	597	507	**1104**	0	**1104**

* Times are measured in seconds.
** No data available because of technical errors.

Table 4
Diagnosis accuracy summary table

		True Positive & True Negative (same)	True Positive & True Negative (different)	True Positive & True Negative (EX)	False Positive & False Negative	Total / Total Statement
Case B						
Participant 3	Computer	6	5	3	1	15/30
	Lightbox	12	4	0	1	17/30
Participant 4	Computer [a]	.	.	.	.	.
	Lightbox	8	3	0	2	13/30
Participant 5	Computer	14	1	0	1	16/30
	Lightbox	15	1	0	0	16/30
Participant 6	Computer	3	2	1	0	6/30
	Lightbox [b]	2	2	0	0	4/30
Case 7464						
Participant 3	Computer	7	4	0	0	11/18
	Lightbox	8	3	0	0	11/18
Participant 4	Computer	4	4	0	1	9/18
	Lightbox	8	2	0	0	10/18
Participant 5	Computer	5	2	3	0	10/18
	Lightbox	7	2	0	1	10/18
Participant 6	Computer	2	2	0	0	4/18
	Lightbox	6	3	0	0	9/18

[a] Technical error in videotaping resulted in no data for this situation.
[b] Wrong MR films on the lightbox for this subject resulted in these low agreement values.

Figure 1
Information sources arrangement in the computer prototype

(To be completed by Ed)

Figure 2
Single window layout

(To be completed by Ed)

Figure 3
Gross timing based on CT and MRI

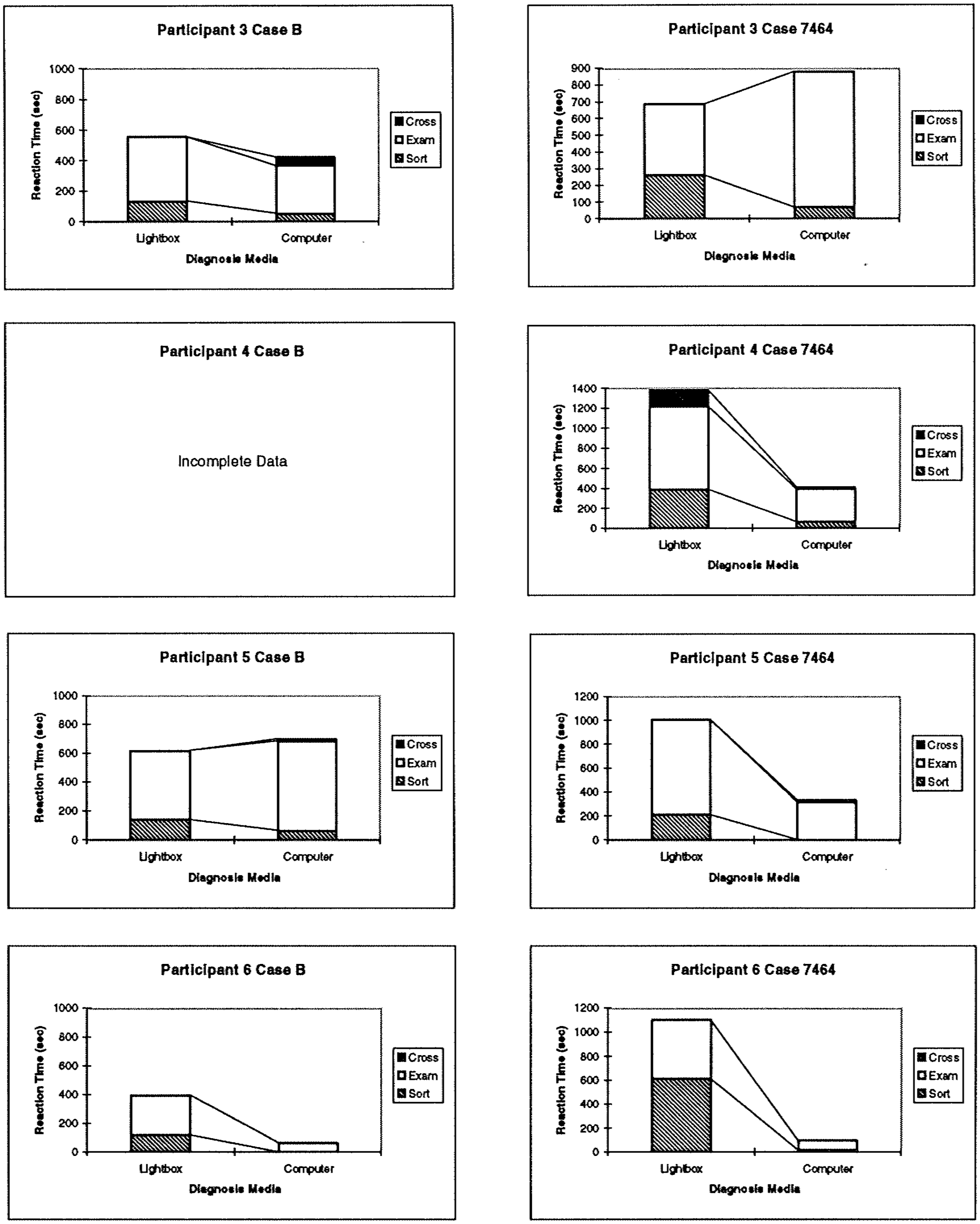

Figure 4
Time allocation among subtasks