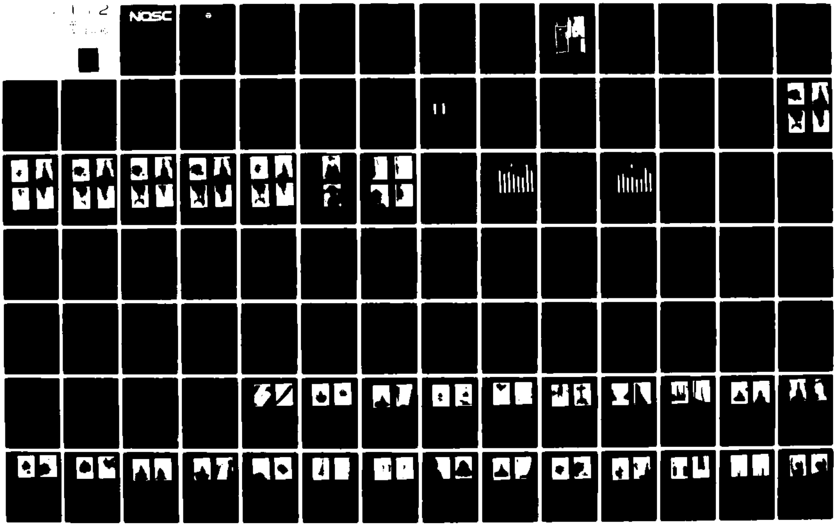
AD-A112 090   NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA   F/G 6/5

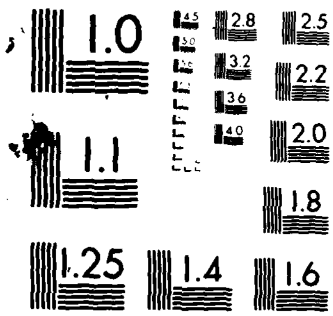
REMOTE MEDICAL DIAGNOSIS SYSTEM (RMDS) ADVANCED DEVELOPMENT MOD--ETC(U)

DEC 81   W T RASMUSSEN, P D HAYES, F H GERBER

UNCLASSIFIED   NOSC/TR-683   NL

1 of 2





MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963 A

NOSC

NOSC TR 683

Technical Report 683

REMOTE MEDICAL DIAGNOSIS SYSTEM (RMDS) ADVANCED DEVELOPMENT MODEL (ADM) RADIOLOGY PERFORMANCE TEST RESULTS

WT Rasmussen, PD Hayes, I Stevens
(NOSC)

FH Gerber
(Naval Regional Medical Center, San Diego)

JA Kuhlman, FW Hutzelman
(WESTEC Services, Inc.)

December 1981

Prepared for
Naval Medical Research and Development Command
Code 45
Bethesda, MD

Approved for public release; distribution unlimited

NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152

NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

SL GUILLE, CAPT, USN
Commander

HL BLOOD
Technical Director

ADMINISTRATIVE INFORMATION

This Technical Report is one in a series concerning work on the Remote Medical Diagnosis System (RMDS), performed under Program Element 64771N, Project M0933-PN (NOSC 512-CM38), sponsored by the Naval Medical Research and Development Command, Code 45. This report contains the test results of an experimental evaluation of image fidelity requirements for radiograph transmissions over the RMDS Advanced Development Model (ADM) terminals. It was prepared by the NOSC Bioengineering Branch (Code 5123) and WESTEC Services, Inc. (Contract N66001-78-C-0274). The evaluation testing was conducted during the period October 1978 to April 1979. Principal investigators were I Stevens, PD Hayes (NOSC), FH Gerber (NRMC San Diego), JA Kuhlman, and FW Hutzelman (WESTEC Services, Inc.), under the direction of WT Rasmussen, Head, Bioengineering Branch (NOSC Code 5123). The test plan and guidelines for this evaluation are contained in NOSC TD 396.

ACKNOWLEDGEMENTS

The authors thank the following radiologists from NRMC, San Diego, for their participation in this study.

CAPT RW Smith, MC, USN
CDR ES Amis, MC, USN
CDR SC Mills, MC, USN
LCDR LC Collins, MC, USNR
LCDR S Hilton, MC, USNR
LCDR VG Marshall, MC, USNR
LCDR LF Mazzara, MC, USNR
LCDR JR Wilcox, MC, USNR
LT VA Gueco, MC, USNR
LCDR DC Matthews, MC, USNR
LT AM Mitchell, MC, USNR
LT JJ Mocnik, MC, USNR

Released by
JM Stallard, Head
Bioacoustics and Bionics Division

Under authority of
HO Porter, Head
Biosciences Department

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER NOSC Technical Report 683 (TR 683)	2. GOVT ACCESSION NO. AD-A112090
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) REMOTE MEDICAL DIAGNOSIS SYSTEM (RMDS) ADVANCED DEVELOPMENT MODEL (ADM) RADIOLOGY PERFORMANCE TEST RESULTS	5. TYPE OF REPORT & PERIOD COVERED Test report October 1978 - April 1979
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) WT Rasmussen, PD Hayes, I Stevens (NOSC) FH Gerber (Naval Regional Medical Center, San Diego) JA Kuhlman, FW Hutzelman (WESTEC Services, Inc.)	8. CONTRACT OR GRANT NUMBER(s) N66001-78-C-0274
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Ocean Systems Center San Diego CA 92152	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE 64771N Proj M0933-PN (NOSC 512-CM38)
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Medical Research and Development Command Code 45 Bethesda MD 20014	12. REPORT DATE December 1981
	13. NUMBER OF PAGES 122
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Remote Medical Diagnosis System (RMDS)
Radiology
Telecommunication
Video-transmitted radiographs
Image processing
Medical and diagnostic [illegible] transfer

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Report of tests conducted to obtain qualitative and quantitative (statistical) data on radiology performance of the RMDS Advanced Development Models (ADMs). Data were obtained on findings confidence levels, diagnostic confidence levels, overall clinical reading accuracy, radiologist system controls, and RMDS capabilities. It was concluded that the RMDS ADMs would provide satisfactory radiographic images for radiology consultations in emergency cases with gross pathological disorders but that a higher resolution would be required in the next evolution of RMDS for cases involving subtle findings.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

OBJECTIVE

Conduct tests to obtain qualitative and quantitative (statistical) data on radiology performance of the Remote Medical Diagnosis System (RMDS) Advanced Development Models (ADMs). During testing with professional radiologists, obtain data on findings confidence levels, diagnostic confidence levels, overall clinical reading accuracy, radiologist system controls, and RMDS capabilities. Perform qualitative and quantitative analyses on the data. Analyze the clinical utility of radiographic images transferred through six possible RMDS transmission modes. View these radiographs under closed-circuit TV (CCTV) and lightbox conditions to provide a basis for comparison. Use both qualitative and statistical interpretations to evaluate the relationships within and between the viewing conditions, ie CCTV, lightbox, and RMDS transmission modes. Compile additional comments provided by the subjects on their data sheets as responses to the follow-up radiologist questionnaire.

CONCLUSIONS

1. The RMDS ADMs would provide satisfactory radiographic images for radiology consultations in emergency cases with gross pathological disorders.

2. A higher system resolution will be required in the next evolution of RMDS for cases involving subtle findings.

3. A quantization level of 6 bits per picture element should be satisfactory for RMDS consultation on emergency cases. More subtle cases may require 8 bits per picture element.

4. No single RMDS transmission mode was associated with a statistically significant increase in diagnostic accuracy. Raw data showed the best mode to be nearly equivalent to CCTV, on the average.

RECOMMENDATIONS

1. In the next phase of RMDS, the Engineering Development Models (EDMs), provide two video resolution levels: a high resolution of 512 x 512 x 8 bits, and a lower resolution of 512 x 512 x 6 bits.

2. In the EDM systems, provide the radiologist with contrast and brightness controls, zoom capabilities, and a video storage system that can store two images and display one at a time or both simultaneously.

3. Perform further studies to establish the nature and extent of the pathological cases projected for RMDS application.

4. In any future testing, incorporate additional nonpathological radiographs for control, radiographs with single or limited findings, more rigid grading procedures, and a continuous range of radiographic difficulties.

CONTENTS

CONTENTS (Continued)

ILLUSTRATIONS

TABLES

SECTION 1

INTRODUCTION

1.1 PURPOSE

This report contains the test results of the experimental evaluation of video transmissions of radiographs over the Remote Medical Diagnosis System (RMDS) Advanced Development Model (ADM) terminals, performed during the period October 1978 to April 1979. The purpose of this evaluation was to determine which, if any, of the various operational modes of the RMDS ADM terminal would satisfy the image fidelity requirements for clinical diagnosis of video-transmitted radiographs. The objective of this evaluation was to establish quantitative and qualitative values or relationships delineating the image fidelity requirements necessary for professional radiologists to make correct and confident diagnosis from video-transmitted radiographs. The test plan and guidelines for this evaluation are contained in NOSC TD 396 (ref 1).

1.2 BACKGROUND

The mission of the RMDS is to improve medical diagnosis at remote sites. This is accomplished by transmitting medical data and diagnostic information between remote ship or shore sites and full-capability medical centers. The RMDS will enable the medical personnel at a remote site to contact a physician at a diagnostic center (ashore or shipboard) and transmit a visual and auditory presentation of the medical data needed for diagnosis, such as patient history, laboratory tests, electrocardiograph (ECG) tracings, X-ray images, images of a patient injury, heart-lung sounds, and verbal descriptions. By return link, the physician will be able to send diagnosis and treatment information. The communication requirements are satisfied by any two-way, voice-grade, narrowband communication channel such as telephone line, hf or uhf radio, or a satellite link.

The system as a whole consists essentially of the RMDS terminals, the existing voice-grade communication links used to interconnect the terminals, and user personnel. Contained in the terminals is all the hardware that is unique to the system: TV camera, TV monitor, X-ray lightbox, electronic stethoscope, ECG monitor, audio tape recorder, audio handsets, and the electronics package, consisting of signal modulator, demodulator, and modems. Figure 1 illustrates the RMDS Advanced Development Model (ADM). The function and operation of the ADM terminal are described in NOSC TD 397 (ref 2).

---

1. NOSC TD 396, Remote Medical Diagnosis System (RMDS) Advanced Development Model (ADM) Test Plan for Evaluation of Image Fidelity Requirements for Radiograph Transmission, WT Rasmussen, I Stevens, PD Hayes (NOSC), and KM Newman (WESTEC Services, Inc.), July 1981.

2. NOSC TD 397, Remote Medical Diagnosis System (RMDS) Advanced Development Model (ADM) Operator's Manual, WT Rasmussen, PD Hayes, I Stevens, EW Davenport (NOSC), and JA Kuhlman (WESTEC Services, Inc.), July 1981.

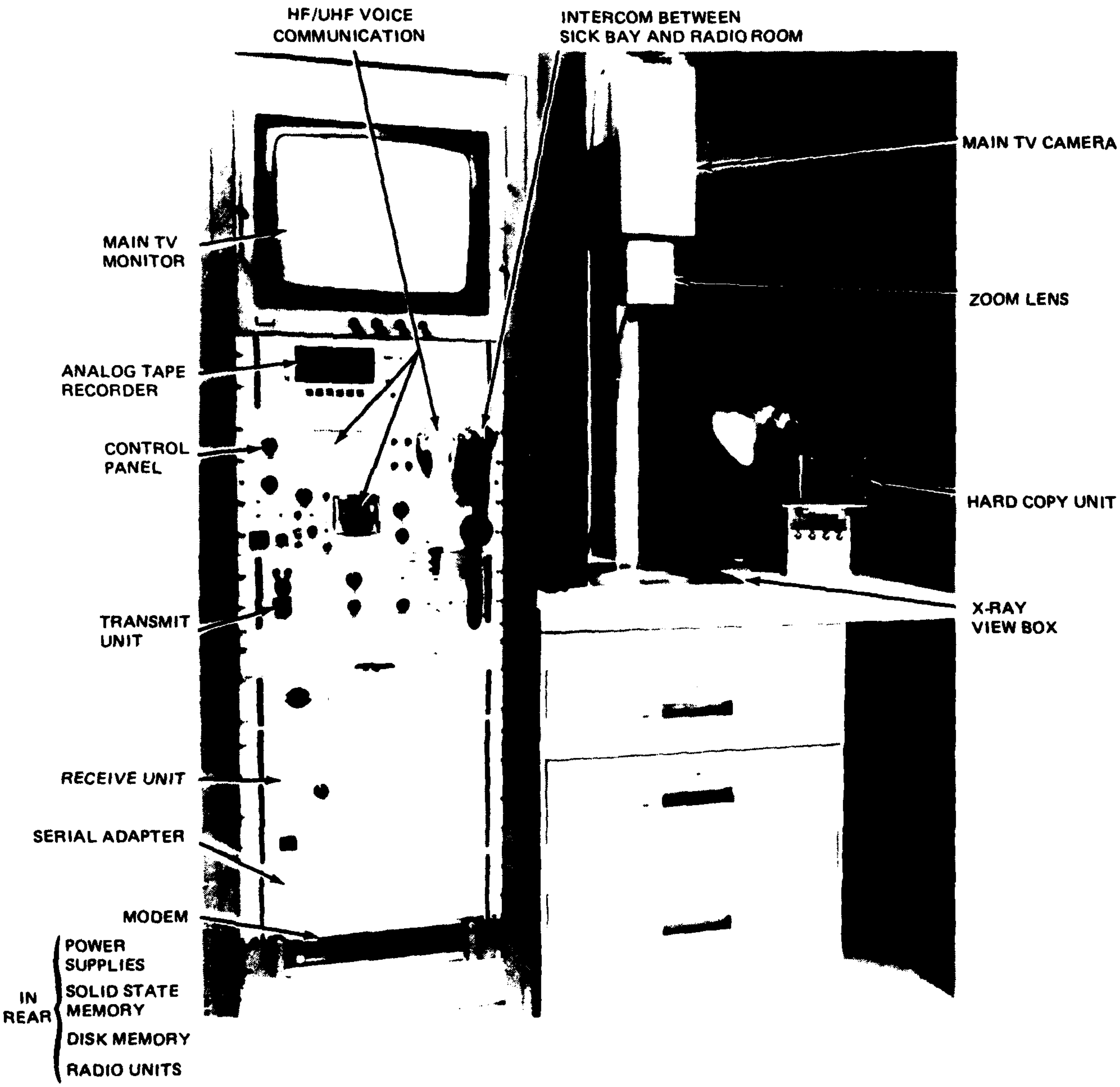


Figure 1. Advanced Development Model (ADM) RMDS terminal and related components.

Shipboard feasibility tests of an early RMDS prototype were completed during FY 75/76. This testing showed that the concept was feasible and that equipment could be developed to meet the requirements by using available technology (ref 3). Because of various constraints (narrow bandwidth, short transmission times, etc), the resolution and gray scale to be achieved in transmitting and displaying radiographic data should be kept to a minimum to meet essential requirements. The gray scale requirements for display have been satisfactorily established in existing literature to be at least 6 bits per picture element, with 8 bits preferred. In order to derive minimum resolution data, a radiology study with eight participating radiologists was conducted in FY 77. A special digital closed-circuit television system was used in this testing to simulate the RMDS equipment (ref 4).

As a result of those feasibility and radiology tests, NOSC undertook a development project to produce two Advanced Development Model RMDS terminals; the ADMs were procured in September 1977. The RMDS ADM terminals were tested for technical performance at NOSC from September 1977 to September 1979. During this period the ADMs were tested at sea, with one terminal aboard the USS ENTERPRISE (CVN 65) from 28 February to 5 March 1978; at-sea test results are documented in NOSC TR 690 (ref 5). Laboratory evaluation testing of data transmission via the ADMs was performed during the period April 1978 to May 1979; those test results are documented in NOSC TR 691 (ref 6).

1.3 EXPERIMENTAL CONSTRAINTS AND CONSIDERATIONS

a. The tests had to be structured for minimum impact on the radiologists' professional duties, without detrimental impact on good experimental design and procedural practices.

b. To arrive at valid conclusions, the tests had to provide sufficient data for statistically valid analyses.

c. The tests had to be performed with available and existing equipments.

d. The validated test results obtained from this effort were to be utilized as some of the principal inputs in developing an RMDS procurement specification.

---

3. NOSC TR 659, Feasibility Tests of the Remote Medical Diagnosis System, WT Rasmussen, I Stevens (NOSC), and JA Kuhlman (WESTEC Services, Inc.), January 1981.
4. NOSC TR 150, Resolution Requirements for Slow-Scan Television Transmission of X-rays, WT Rasmussen (NOSC), RL Crepeau (WESTEC Services, Inc.), and FH Gerber (NRMC San Diego), 19 September 1977.
5. NOSC TR 690, Remote Medical Diagnosis System (RMDS) Advanced Development Model (ADM) At-Sea Test Report, WT Rasmussen, I Stevens, PD Hayes (NOSC), and J West (WESTEC Services, Inc.), in preparation.
6. NOSC TR 691, Remote Medical Diagnosis System (RMDS) Advanced Development Model (ADM) Laboratory Test Results, WT Rasmussen, I Stevens, PD Hayes (NOSC), J West, and FW Hutzelman (WESTEC Services, Inc.), in preparation.

1.4 PARAMETRIC CONSIDERATIONS

The successful (ie, accurate and confident) diagnosis of a radiographic image by a professional radiologist depends almost exclusively on the fidelity of the displayed image. The fidelity of radiographs transmitted electronically from one location to another is influenced considerably by various parameters such as transmission mode, equipment characteristics, etc. Under test were the parameters that affect image fidelity, as follows:

a. Gray-level quantization, which can make natural gray-level changes appear as artificial edge structures or which can mask subtle gray-level changes (ie, analog vs digital transmission mode).

b. Additive noise in the transmission signal, which gives a random, textured pattern to the image, thereby possibly masking natural texture or detail (ie, high vs low signal-to-noise ratio).

c. Spatial resolution of the image presented, which affects the level of detail that can be detected under low-noise conditions (ie, fine vs coarse resolution).

The tests performed were directed toward resolving both the nature and the impact of these three parameters on radiologists' diagnostic performance in evaluating transmitted radiographs. In addition to the three principal parameters, the fidelity of the transmitted images may be affected by the settings of the contrast and brightness controls at the TV monitor. Therefore, any changes in these settings made by the radiologists were monitored and recorded during the tests.

SECTION 2

APPROACH AND DESIGN

2.1 EXPERIMENTAL APPROACH AND PARAMETERS

The experimental approach taken involved the use of two groups of test subjects (radiologists): an RMDS Radiology Test Group and a Control Radiology Test Group. The test data from each group were analyzed and compared individually as well as by groups. In the following two subsections, the experimental variables considered for each group are listed.

2.1.1 RMDS Radiology Tests Experimental Variables

Each of the three variables below had two levels:

a. Transmission mode
- Digital (64 shades of gray) transmission
- Analog transmission

b. Signal-to-noise ratio (SNR)*
- High SNR (analog 36.9 dB, digital 39.4 dB)
- Low SNR (analog 32.8 dB)

c. Resolution
- Fine (256 x 512 raster) resolution
- Coarse (256 x 256 raster) resolution

Combining the variables and levels in a true factorial design** would result in 2^3 experimental conditions. But since there is no progressive degradation of SNR in digital transmission, the theoretical combinations of digital transmission and low SNR (digital x low SNR x fine resolution; digital x low SNR x coarse resolution) are not valid. Therefore, the RMDS radiology tests consisted of the six experimental conditions shown in table 1.

Transmission Mode	Abbreviation	Label
Digital, high SNR, fine resolution	DHF	I
Digital, high SNR, coarse resolution	DHC	II
Analog, high SNR, fine resolution	AHF	III
Analog, high SNR, coarse resolution	AHC	IV
Analog, low SNR, fine resolution	ALF	V
Analog, low SNR, coarse resolution	ALC	VI

Table 1. Six RMDS radiology experimental test conditions.

---

*SNR is defined as peak-to-peak optimum signal level divided by root-mean-square (rms) broadband noise level.

**In a factorial design the effects and interactions of two or more experimental variables are observed simultaneously.

2.1.2 Control Radiology Tests Experimental Variables

Control tests were considered a necessary adjunct to the RMDS radiology tests. They compared the differences in findings confidence levels, diagnostic confidence levels, and accuracy of radiographs presented both as "pure" analog signals (not possible with the RMDS*) via closed-circuit television (CCTV) and in the "direct" manner, with a lightbox. Resolution of the CCTV image was made equivalent to that of the fine resolution mode of the RMDS tests by band-limiting the video signal. The SNR was also fixed for the above two variables. Thus, the control tests were performed under two experimental conditions:

- Lightbox -- I
- CCTV -- II

2.2 HYPOTHESES

Two principal hypotheses were tested by statistical analyses of the data:

a. With the combinations of fine resolution and high SNR, for both analog and digital transmission modes, diagnoses can be made which are not significantly different statistically in terms of confidence level and accuracy from those made using CCTV images of equivalent spatial resolution. In other words, the quantization level of 6 bits per picture element, as used in the RMDS for digital transmission, and the frozen noise of the received images for both digital and analog transmission modes degrade neither the confidence level nor the accuracy of radiographic diagnoses.

b. Fine resolution, in both the digital and analog RMDS modes, leads to statistically more confident and accurate diagnoses than coarse resolution and significantly reduces diagnostic inaccuracies.

These major hypotheses were to be accepted or rejected through statistical tests where the level of significance is 0.05 in terms of the experimental variables tested. In addition, RMDS radiographic testing was designed to provide quantitative and qualitative data on the following:

<u>Test Group</u>

- Accuracy, or Overall Clinical Reading (OCR) results, by test mode in terms of significant statistical differences.
- Accuracy (OCR) by radiologist in terms of significant statistical differences.
- Correlation between transmission modes, findings confidence levels, and diagnostic confidence levels.

---

*Since the RMDS utilizes a memory, those radiographs transmitted in the analog mode are digitized to a certain degree.

- Correlation between confidence levels and accuracy (OCR) for each transmission mode.
- The relationships between the pathology types, confidence levels, and transmission modes.
- The effects of zoom lens utilization on the OCR.

Control Group

- The accuracy and confidence differences between the results using the lightbox and CCTV monitor viewing.
- The differences between the lightbox and CCTV performance accuracy and confidence with respect to the difficulty of each radiograph.
- The confidence and accuracy results by radiologist.

The means through which the above evaluations were accomplished are discussed in section 3, and their results are presented in section 4.

2.3 SUBJECTS

The test subjects consisted of 12 staff radiologists from the Radiology Department, NRMC San Diego. The radiologists were divided into two groups of six radiologists each: the RMDS Radiology Test Group (Test Group), and the Control Radiology Group (Control Group). Each group was balanced with respect to training and experience. Each radiologist was tested with the radiographs individually; the radiologists did not know to which group they belonged.

2.4 RADIOGRAPHS

Six sets of 6 radiographs were used, for a total of 36 radiographs. These 36 radiographs were selected from over 60 case files at NRMC San Diego. The six radiographs of each set consisted of six different disorders, with all sets balanced as equivalently as possible with respect to pathology, contrast, density, and the difficulty of visual and diagnostic interpretation. The balance was achieved by assigning a Difficulty Ranking Factor (DRF) of 1 (low) to 6 (high) to each radiograph prior to its inclusion in the testing. Each radiograph was first judged independently by four individuals, then a consensus was reached on the DRF of each.

The six sets (labeled A, B, C, D, E, and F) were arranged so that each set had radiographs with DRFs of 1 through 6, then each set was balanced as far as possible with respect to pathologies and type (or zone -- ie, appendage, abdomen, chest, or skull). The sequential order of presentation of the radiographs within each set was randomized with respect to the DRF and type, and the established order was retained for all subjects throughout the testing. Table 2 shows the radiographs used and, for each radiograph, its order of presentation within the set, its assigned DRF, its zonal type, and a primary diagnosis. More details and a reduced print of each radiograph are given in appendix A.

Set	Order of Presentation	DRF*	Type (Zone)	Diagnosis
A	1	5	Appendage	Soft tissue hemangioma
	2	6	Skull	Fractured mandible
	3	3	Chest	(R) LL pneumonia
	4	2	Skull	Double floor of sella
	5	4	Abdomen	Prostatic calculi
	6	1	Abdomen	Bilateral adrenal calcification
B	1	1	Skull	Broken nose
	2	3	Appendage	Osteoid osteoma
	3	6	Chest	ASD with 4:1 shunt
	4	2	Chest	Alveolar cell calcification
	5	4	Skull	Intracranial air & fracture
	6	5	Appendage	Fx neck of femur on (R)
C	1	4	Chest	Calcified mitral annulus
	2	2	Chest	Cocci
	3	5	Skull	Parietal skull fracture
	4	1	Abdomen	Abdominal aortic aneurysm
	5	3	Appendage	Avascular necrosis of lunate
	6	6	Skull	Nasal spine Fx
D	1	4	Chest	Pancoast tumor
	2	5	Skull	Multiple myeloma
	3	2	Abdomen	Air under (R) diaphragm
	4	6	Appendage	Chondrocalcinosis
	5	1	Appendage	Fibrous cortical defect
	6	3	Chest	Pericardial calcification
E	1	3	Skull	Calvarial hemangioma
	2	6	Abdomen	Splenomegaly
	3	4	Appendage	Tibial stress Fx
	4	1	Abdomen	Osteitis condensans ilii
	5	2	Skull	Mucocoele (L) frontal sinus
	6	5	Chest	Histiocytosis-x
F	1	3	Abdomen	Abdom calcification (post traumatic splenic cyst)
	2	2	Chest	Infectious spondylitis TB
	3	4	Appendage	Cocci osteomyelitis
	4	5	Chest	Calcified myocardial infarct
	5	6	Chest	Pneumothorax on (R)
	6	1	Skull	Enlarged sella

*DRF = Difficulty Ranking Factor: 1 (low), 2 (low/med low), 3 (med low), 4 (med high), 5 (med high/high), 6 (high)

Table 2. Six sets of test radiographs.

2.5 EXPERIMENTAL DESIGN

2.5.1 RMDS Radiology Tests

The six experimental conditions (section 2.1.1, table 1), radiographs, and subjects were combined into a modified factorial treatments-by-subjects design* that was optimized for isolating the effects of the variables to be tested (table 3). The pertinent features of this design are as follows:

a. It was balanced for subjects, conditions, and radiograph sets, so that each subject was tested once under each condition with a different set of radiographs.

b. No set of radiographs was paired with any one experimental condition more than once.

c. The sequence of conditions/sets presentations was different for each of the six subjects, with respect to both experimental conditions and radiograph sets.

This type of design minimized any possible undesirable interactions between radiographs, subjects, and sequence of experimental conditions presentation.

The Test Group radiologists each completed this testing in three separate sessions, averaging about 2 hours for each session for a total of approximately 6 hours. This included the time required to set up each radiograph and time for the radiologist to record the findings and comments.

2.5.2 Control Radiology Tests

The Control Group subjects provided their findings and diagnoses based upon direct lightbox (Condition I) and CCTV (Condition II) viewing. In the direct lightbox mode, the Control Group viewed all six sets of radiographs. In the CCTV mode, the Control subjects viewed three of the six sets, three different sets for each subject. It was theorized that the results of viewing these three sets on CCTV would provide a data base of sufficient size for statistical analysis and comparison with the results of the analysis of the 36 radiographs viewed on the lightbox. As seen in table 4, each set was seen under CCTV conditions on three separate occasions. Testing was designed in such a way that when a given radiologist was scheduled to view a set under both lightbox and CCTV conditions, CCTV testing would precede lightbox viewing. In this way, CCTV viewing would not be influenced by knowledge gained from the best-case viewing conditions provided by the lightbox. A direct comparison between CCTV and the lightbox results can thus be made. Other details of the balanced treatments-by-subjects testing sequence for the Control Group are presented in table 4.

---

*In a treatments-by-subjects design, all experimental conditions are successively administered to the same subjects.

	1				2				3				4				5				6			
SUBJECT	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓
1	I(DHF)	A	1-1		II(DHC)	B	1-2		III(AHF)	C	1-3		IV(AHC)	E	1-4		V(ALF)	D	1-5		VI(ALC)	F	1-6	
2	VI(ALC)	C	2-1		V(ALF)	E	2-2		IV(AHC)	F	2-3		III(AHF)	B	2-4		II(DHC)	A	2-5		I(DHF)	D	2-6	
3	II(DHC)	F	3-1		IV(AHC)	A	3-2		VI(ALC)	E	3-3		I(DHF)	B	3-4		III(AHF)	D	3-5		V(ALF)	C	3-6	
4	III(AHF)	A	4-1		I(DHF)	C	4-2		V(ALF)	F	4-3		VI(ALC)	D	4-4		IV(AHC)	B	4-5		II(DHC)	E	4-6	
5	IV(AHC)	C	5-1		VI(ALC)	B	5-2		II(DHC)	D	5-3		V(ALF)	A	5-4		I(DHF)	F	5-5		III(AHF)	E	5-6	
6	V(ALF)	B	6-1		III(AHF)	F	6-2		I(DHF)	E	6-3		II(DHC)	C	6-4		VI(ALC)	A	6-5		IV(AHC)	D	6-6	

EXPERIMENTAL CONDITIONS:

I = DHF = Digital, High SNR, Fine Resolution
II = DHC = Digital, High SNR, Coarse Resolution
III = AHF = Analog, High SNR, Fine Resolution
IV = AHC = Analog, High SNR, Coarse Resolution
V = ALF = Analog, Low SNR, Fine Resolution
VI = ALC = Analog, Low SNR, Coarse Resolution

Table 3. RMDS Test Group experimental design.

	1			2			3			4			5			6		
SUBJECT	EXPTL. COND.	SET	BOOKLET	EXPTL. COND.	SET	BOOKLET	EXPTL. COND.	SET	BOOKLET	EXPTL. COND.	SET	BOOKLET	EXPTL. COND.	SET	BOOKLET	EXPTL. COND.	SET	BOOKLET
7	I	A	7-1	I	B	7-2	II I	C	7-3 7-3a	I	D	7-4	II I	E	7-5 7-5a	II I	F	7-6 7-6a
8	II I	B	8-1 8-1a	II I	D	8-2 8-2a	I	F	8-3	II I	A	8-4 8-4a	I	C	8-5	I	E	8-6
9	I	C	9-1	II I	A	9-2 9-2a	II I	E	9-3 9-3a	I	F	9-4	II I	B	9-5 9-5a	I	D	9-6
10	II I	D	10-1 10-1a	I	F	10-2	I	B	10-3	II I	E	10-4 10-4a	I	A	10-5	II I	C	10-6 10-6a
11	I	E	11-1	II I	C	11-2 11-2a	I	D	11-3	II I	B	11-4 11-4a	II I	F	11-5 11-5a	I	A	11-6
12	II I	F	12-1 12-1a	I	E	12-2	II I	A	12-3 12-3a	I	C	12-4	II I	D	12-5 12-5a	I	B	12-6

EXPERIMENTAL CONDITIONS:

I = DIRECT [LIGHTBOX]
II = CCTV

Table 4. RMDS Control Group experimental design.

The Control Group radiologists each completed this testing in two separate sessions, averaging about 2 hours for each session for a total of approximately 4 hours. This included the time required to set up each radiograph shown over the CCTV and time for the radiologist to record the findings and comments.

2.5.3 Findings and Diagnoses

Both the Test Group and the Control Group participants reported their findings and diagnoses for each radiographic image presented. Each, in turn, judged the radiographic images on a scale of 1 to 5 in terms of Findings Confidence Level (FCL) and Diagnostic Confidence Level (DCL). The method for determining FCL and DCL ratings were consistent for both the Test and Control Groups; a more detailed description of the method is given in section 3.1.

The Control Group radiologists' findings and diagnoses obtained by direct, lightbox viewing of each radiograph were reviewed by Dr FH Gerber, Head of Department of Radiology, NRMC San Diego, and another senior radiologist; neither was included in either testing group. Findings and diagnoses which were agreed upon by a consensus of the Control Group were used as a "standard" against which the findings and diagnoses of the RMDS Radiology Test Group and the CCTV diagnoses of the Control Group were compared. The reviewers compared each radiologist's findings and diagnoses to the "standard" and established an Overall Clinical Reading (OCR) evaluation of acceptable, marginal, or unacceptable for each radiograph in each mode.

SECTION 3

METHODS AND PROCEDURES

3.1 TEST PROCEDURES

Both RMDS radiology and control tests were carried out in the Bioengineering facilities, NOSC Code 5123. Two RMDS terminals located in the area were used back-to-back, set up in separate rooms, with one system serving as the remote terminal (transmitter) and the other one as the diagnostic center terminal (receiver). The voice communication feature of the RMDS was used in testing. The remote terminal provided zoom and selective intensity capabilities. In the control tests, the RMDS was bypassed and the radiographic image as seen by the TV camera was displayed directly on the subject's TV monitor via a bandpass filter. Figure 2 is a block diagram of the RMDS equipment experimental configuration. A cassette-type tape deck was used to record any comments made by the subjects throughout the duration of each test session.

To insure that the receive TV monitor display of each radiograph would be repeated to each radiologist at the same brightness and contrast, the camera lens aperture (K), video level (VL), and black level (BL) on the transmit RMDS terminal were recorded in advance in terms of the amount of deflection measured on a strip-chart recorder. Settings were determined that made the image displayed on the receive TV monitor as close as possible to the actual radiograph image, and these predetermined settings were used for each radiograph throughout the tests. Also, the contrast and brightness levels of the receive TV monitor were monitored and recorded to indicate how the radiologist adjusted them to make a diagnosis. Prior to each testing session, all equipments were thoroughly checked for satisfactory performance of all necessary functions and were adjusted, if required, to baseline operating parameters.

Prior to testing, each of the participating radiologists was briefed on the purpose of the RMDS; the reasons for and objectives of the test; the capabilities, features, and limitations of the hardware relevant to the experiment (such as the zoom and selective intensity capabilities); the test facility and procedures; and the test materials such as the test data sheets used for both Test and Control Groups (appendix B). The participants were told that their diagnostic comments would be tape-recorded, and they were given an opportunity to use three nontest radiographs to familiarize themselves with the equipment, the facilities, and the task to be performed. The radiologists did not know to which group they belonged.

To minimize any possible bias on the part of the subjects, the specific experimental condition under test at any one time was not revealed. The fact that different video transmission modes for radiographic transmissions were being investigated was included in the introductory briefing. During this briefing, it was emphasized that the objectives of the experiment were not to evaluate individual professional performance and capabilities, but rather to determine and assess the physical parameters necessary for the effective reading and evaluation of transmitted radiographs. The participating radiologists were instructed to read each radiograph as rapidly as possible consistent with professional responsibilities, but told that time was not a parameter under investigation. The subjects were instructed to record on their test data sheets a numerical value for their level of confidence in the findings and diagnoses arrived at for each radiograph viewed. These values could range from 1 for a low confidence level to 5 for a high confidence level, as follows.

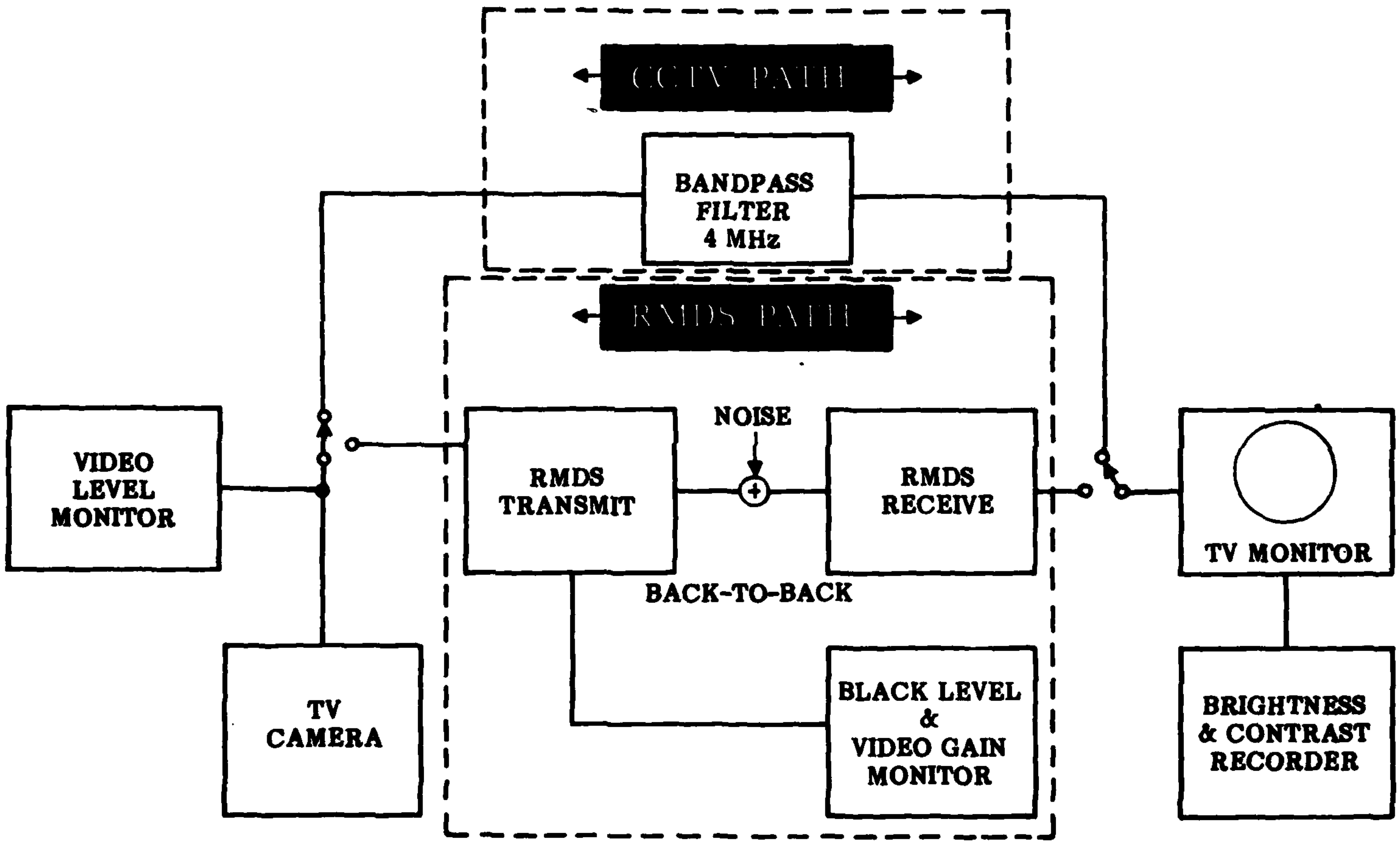


Figure 2. Block diagram of RMDS equipment configuration used in testing.

Numerical Value	Confidence Level
1	0-20%; low
2	21-40%; fairly low
3	41-60%; medium
4	61-80%; fairly high
5	81-100%; high

The subjects were also encouraged to record their comments in the space provided on the data sheet. The test sequence to be followed depended on whether the subject was part of the Test or Control Group.* The radiologists were given a break after completion of each set.

No radiologist was allowed to consult with others during the course of testing. The radiologists were also requested not to discuss observations, impressions, or diagnoses with the other participating radiologists until completion of the entire experiment. After completion of all tests, the radiologists filled out a questionnaire** in which likes, dislikes, and suggested areas of improvement could be expressed.

Following completion of the testing, two senior radiologists (including the Head of Radiology, NRMC San Diego) evaluated the findings, diagnoses, and confidence levels recorded on the test data sheets and defined an Overall Clinical Reading (OCR) score for the test data. These OCR scores (3 = acceptable, 2 = marginal, 1 = unacceptable) were differentiated as OCR_x (for Dr X's evaluation) and OCR_y (for Dr Y's evaluation) for the ensuing statistical analysis effort.

3.2 DATA ANALYSIS

3.2.1 Quantitative Analysis

The data base for the mathematical treatment of the test results consisted of the numerical confidence level scores, the OCR evaluations, and the DRFs. The first are subjective scores, arrived at by the subjects themselves, while the OCR scores are an objective measure of performance. The DRFs assess the difficulty of visual and diagnostic interpretation. Listed below are the numerical systems utilized to form the basis of analyses:

Data Type	Range
Findings Confidence Level (FCL)	1 (low) - 5 (high)
Diagnostic Confidence Level (DCL)	1 (low) - 5 (high)
Overall Clinical Reading (OCR)	1 (unacceptable) - 2 (marginal) - 3 (acceptable)
Difficulty Ranking Factor (DRF)	1 (easy) - 6 (hard)

*See appendix B of the Test Plan (NOSC TD 396) for details of the RMDS test sequences.
**Appendix B contains the two questionnaires, one for each group.

The *experimental design* (modified factorial and treatments-by-subjects) as well as the performance criteria (confidence levels, percentage errors, and difficulty ranking) allowed application of mathematical analysis. Confidence level measurements (FCL and DCL) represent interval scale data, which provide information concerning the distance between levels in addition to the relationship characteristics found in ordinal scales. This form of data allows for more powerful statistical testing. The statistical analyses performed on confidence level data included: one-way analysis of variance (ANOVA), Duncan's multiple-range test, Pearson's correlation, and Wilcoxon's matched-pairs signed-ranks test (ref 7).

The OCR data represent ordinal scale measurements, wherein the relationship between one category and the next can be classified in terms of higher or more preferred results -- ie, "unacceptable" is less desirable than "marginal," which is less desirable than "acceptable." Although the categories can be ranked from high to low, no true numerical value can be placed on the differences between them. The Wilcoxon matched-pairs signed-ranks test (ref 7) was used to analyze OCR data. Because of the small sample size and the use of only three ranks for the OCR data, a Chi-square analysis was performed, when necessary, by combining groups -- ie, 3's = 1 and 2's + 1's = 0, or 3's + 2's = 1 and 1's = 0. An analysis of variance per transmission mode and per Test Group radiologist could not be performed, since the sample size was limited. Differences between the Test Group transmission modes and radiologists were tested by combining the available data into larger groups with similar characteristics -- ie, results of the six radiologists in mode 1 versus mode 2 versus mode 3, etc.

DRF measurements fall between the ordinal and interval scales of data. They provide relationship information in conjunction with a rough estimate of the distance between levels. The DRF measurements were utilized in two basic ways: (1) as subheadings for data groupings, wherein all the data at a particular DRF level were analyzed together; and (2) for correlation purposes via Pearson's correlation.

Quantitative analysis was also performed on the raw data measurements of the time spent in diagnosing a set of six X-rays. Because these measurements included set-up and transmission times for zoom and whole-image viewings, averages for these (per mode) were determined and subtracted from the raw data to obtain a more realistic estimate of the actual time spent in diagnosis. Table 5 summarizes the statistical analysis performed on the various types of data obtained during RMDS testing. The acceptance or rejection of the proposed hypotheses was based on statistically critical values at the 0.05 significance level.

3.2.2 Qualitative Analysis

The qualitative data, represented by (1) photographic representations of the received radiographic images, (2) graphs of the collected data, (3) the brief written comments on the individual test data sheets, and (4) the questionnaire filled out by each radiologist, were used as an important corollary to the quantitative data in arriving at selected performance criteria for future Remote Medical Diagnosis Systems. The importance of a qualitative assessment of the documented radiographic evaluation

---

7. Nonparametric Statistics for the Behavioral Sciences, S Siegel, McGraw-Hill, New York, 1956.

Test	Data Type	Statistics
OCR_x vs OCR_y	OCR	Wilcoxon T
Radiographic sets	OCR	Chi-square
Lightbox vs CCTV	FCL	Wilcoxon T
	DCL	Wilcoxon T
	DRF	Wilcoxon T
	OCR	Wilcoxon T
Lightbox by radiologist	FCL	ANOVA*
	DCL	ANOVA*
	OCR	Chi-square
Lightbox correlation	FCL	Pearson's correlation
	DCL	Pearson's correlation
	OCR	Pearson's correlation
	DRF	Pearson's correlation
Test Group by mode	FCL	ANOVA*
	DCL	ANOVA*
	OCR	Chi-square
Test Group by radiologist	OCR	Chi-square
Zoom vs no zoom	OCR	Chi-square

*Duncan's multiple range test was performed if ANOVA was rejected at 0.05 level.

Table 5. Summary of statistical testing.

process should not be underestimated, since the diagnosing radiologist at the diagnostic center terminal will be an end user of the RMDS. Therefore, the radiologists' reactions to RMDS functional capabilities and performance were of prime consequence. Analysis of the qualitative data included the following concerns:

a. Determine the quality of received images by visual examination of photographs, by transmission mode.

- Full scale images
- Zoom images

b. Summarize confidence and OCR scores using mean and standard deviation, by transmission mode.

c. Evaluate the use of zoom controls by mode, testing sequence, radiologist, and OCR.

d. Establish the mean, maximum, and minimum times required to diagnose one set of six radiographs, by transmission mode.

e. Characterize test radiographs by DRF, zone, and pathology.

f. Analyze the radiologists' comments on the test data sheets.

g. Analyze the radiologists' comments on the questionnaire.

SECTION 4

RESULTS

The results of the qualitative and quantitative analyses are summarized and presented in tables, figures, and narrative form. A discussion of these results follows in section 5, and conclusions and recommendations are given in sections 6 and 7. The qualitative and quantitative findings and interpretations within this report address the following objectives:

a. Accept or reject the hypotheses (given in section 2.2).

b. Provide quantitative measures of the effects of each experimental variable on diagnostic performance, independent of the other variables. (For example, does the SNR, by itself, have a statistically significant influence on diagnostic performance? Which variable(s), if any, has (have) the most profound effect on diagnostic performance?) Quantize the level of influence exerted by the variable(s).

c. Determine statistically significant differences in diagnostic performance between each of the various transmission modes and the standard method.

d. Point out any pathologies (of those provided) which do not lend themselves to confident and accurate diagnoses from transmitted radiographs.

e. Provide guidelines to the performance parameters of digitization, resolution, and SNR for an RMDS procurement specification.

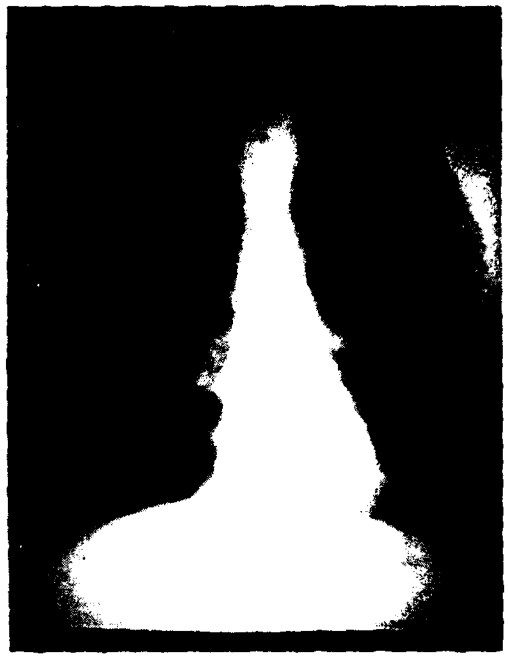
4.1 QUALITATIVE RESULTS

Photographs of radiographic images received through the actual RMDS transmission modes, taken during the RMDS testing, provide a base of data for qualitative analysis. Figures 3 through 8 illustrate the six test modes. Each shows four selected radiographs from the testing sets, representing a skull, a chest, an abdomen, and an appendage. The image quality of these received radiographs can be further compared with the photographs of the original radiographs, as shown in appendix A. The sample skull radiograph appears as D-2, page 88; the chest X-ray, B-4, page 78; the abdomen, E-2, page 94; and the foot (or appendage), F-3, page 101.

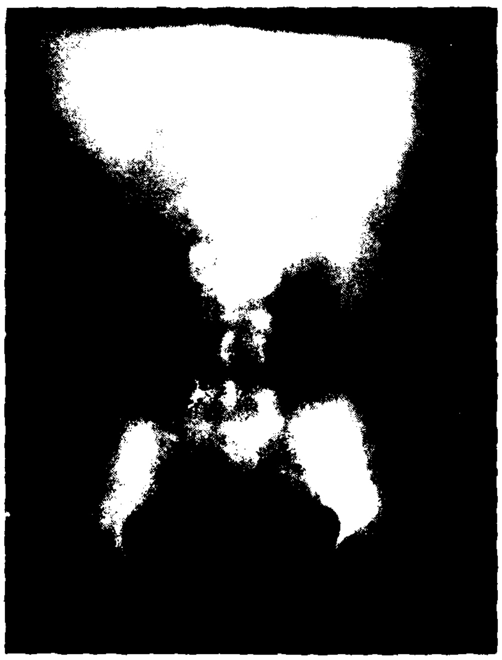
A visual examination of the image quality of figures 3 through 8 qualitatively confirms the anticipated superiority of digital vs analog, high SNR vs low SNR, and fine resolution vs coarse. (The qualitative analyses reported in the remainder of this section support these conclusions.) Unfortunately, photographic and reproductive processes degrade image quality to some degree, masking some of the more subtle variations among modes.

The image quality of radiographs under zoom or close-up conditions is also subject to analysis across test transmission modes. The photographs shown in figure 9 illustrate the quality of close-up images in the digital mode, while figure 10 presents photographs of close-up images in the analog transmission modes. These close-up

Skull



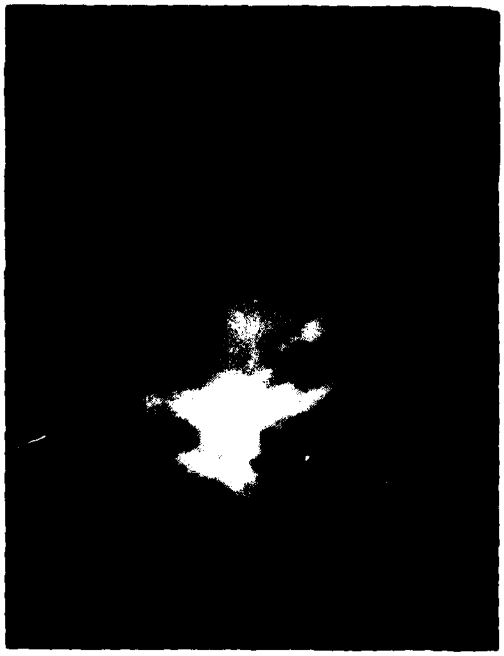
Chest



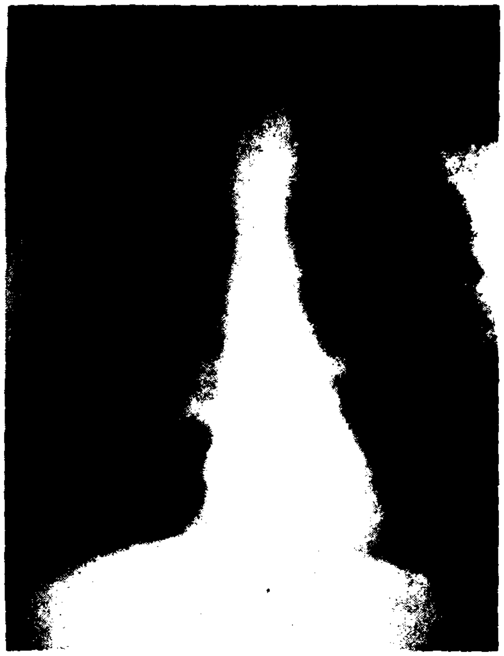
Abdomen

Appendage

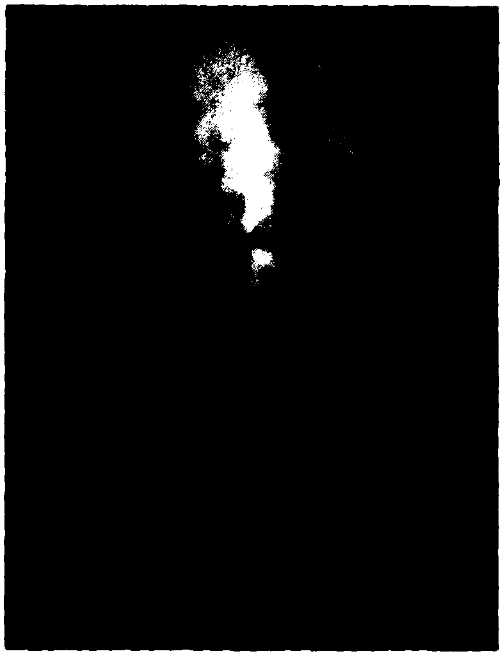
Figure 3. Received images, DHF mode.



Skull



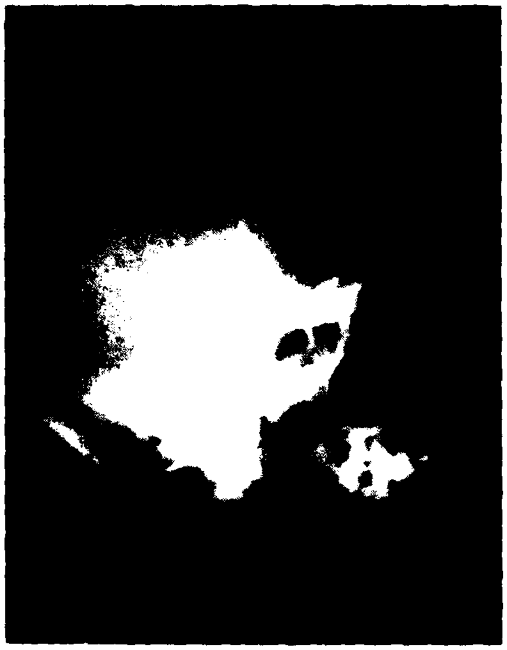
Chest



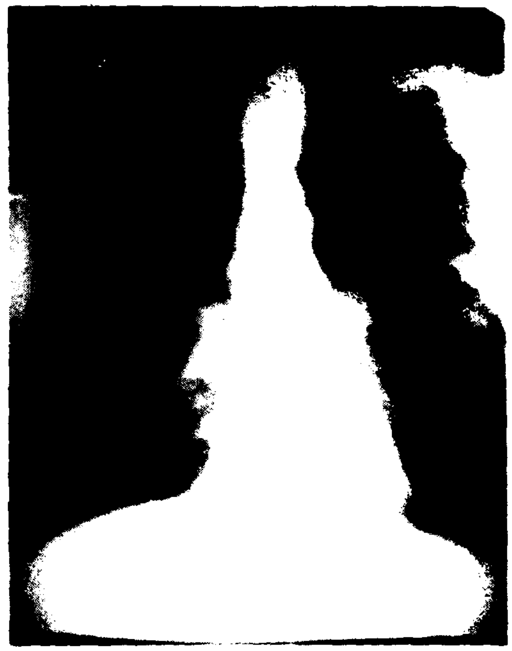
Abdomen

Appendage

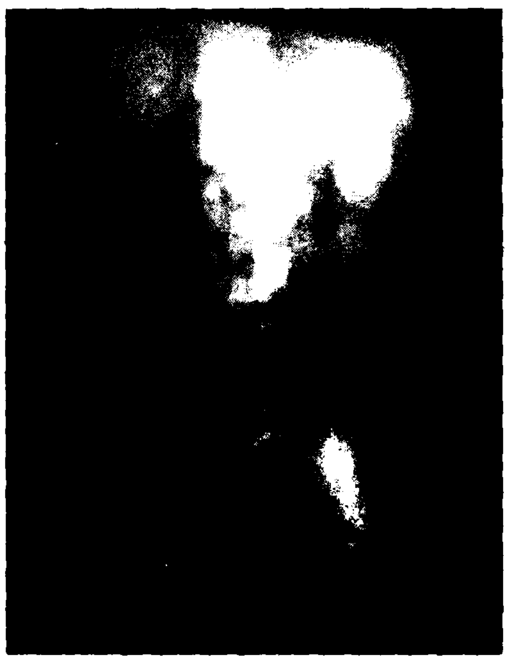
Figure 4. Received images, DHC mode.



Skull



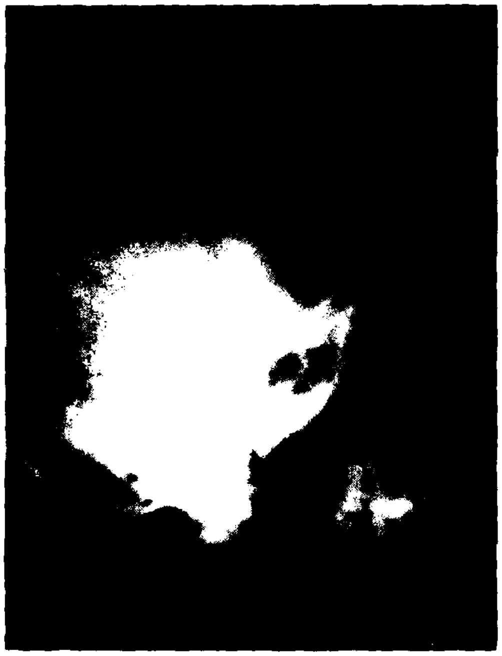
Chest



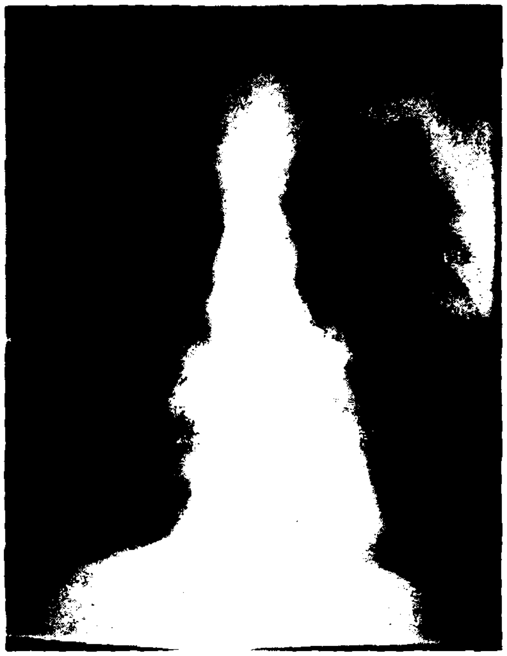
Abdomen

Appendage

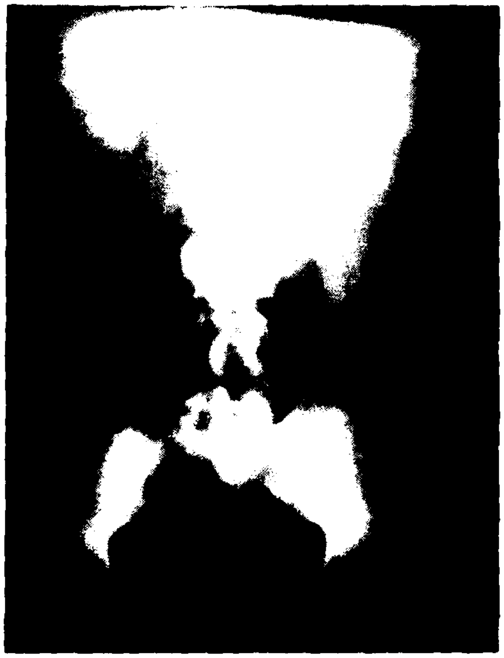
Figure 5. Received images, AHF mode.



Skull



Chest



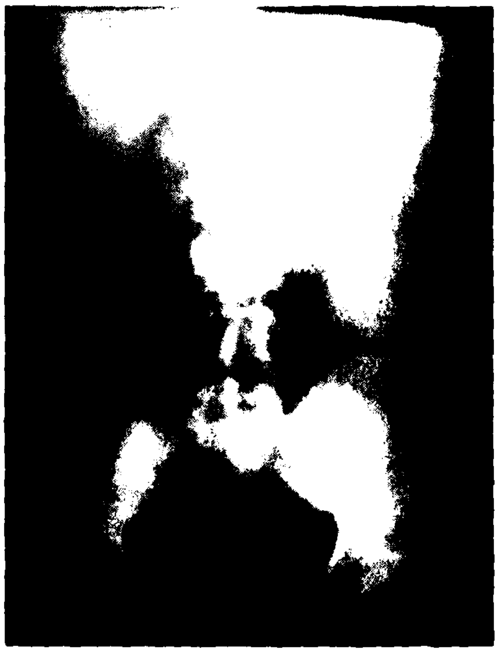
Abdomen

Appendage

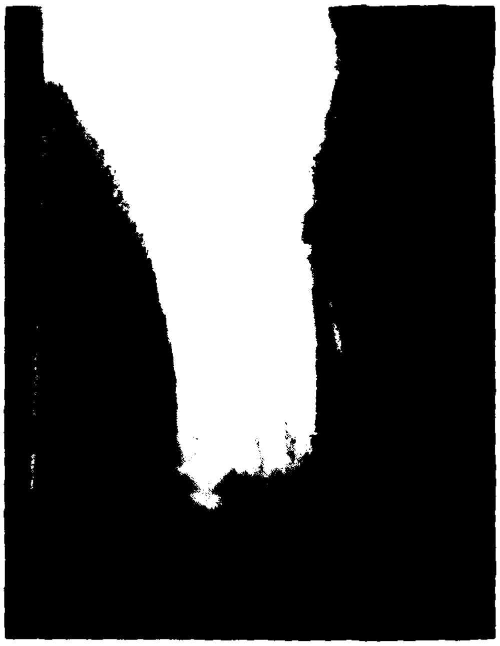
Figure 6. Received images, AHC mode.

Skull

Chest

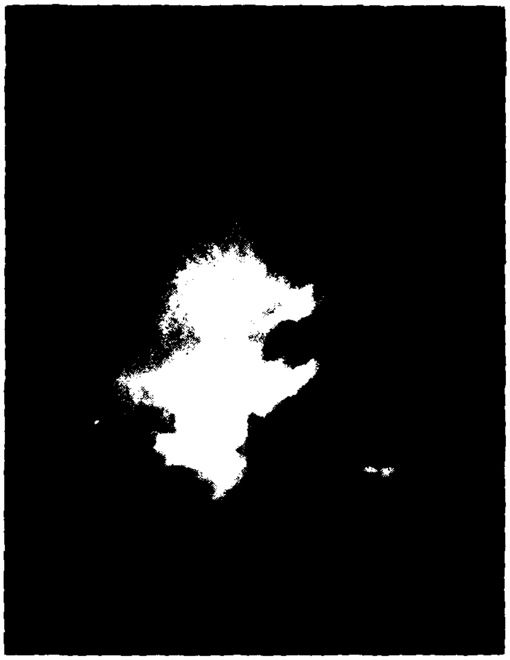


Abdomen

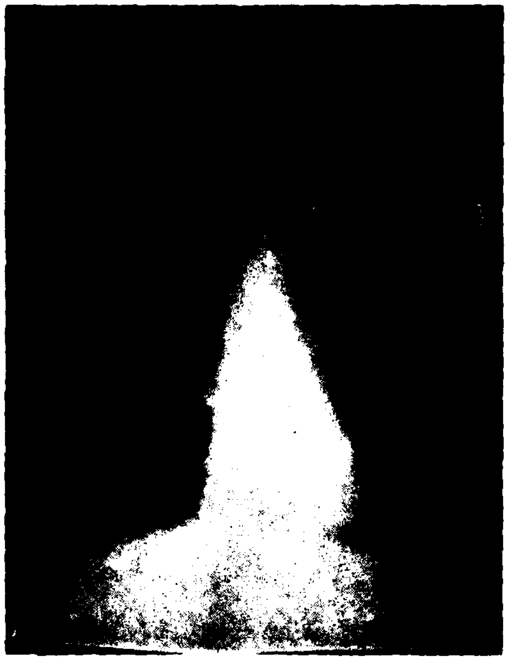


Appendage

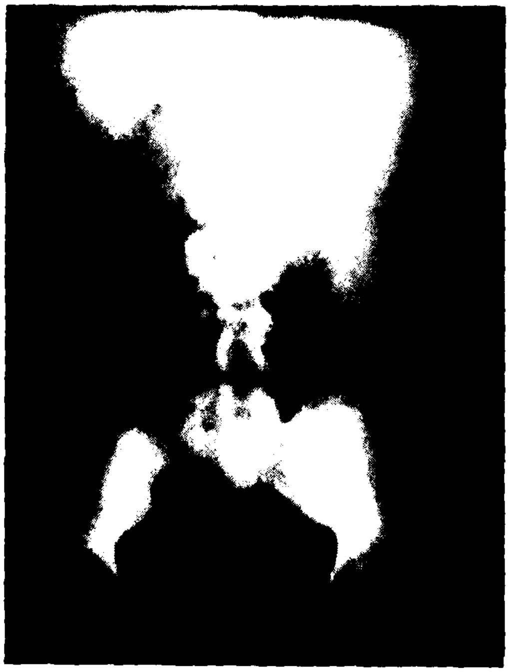
Figure 7. Received images, ALF mode.



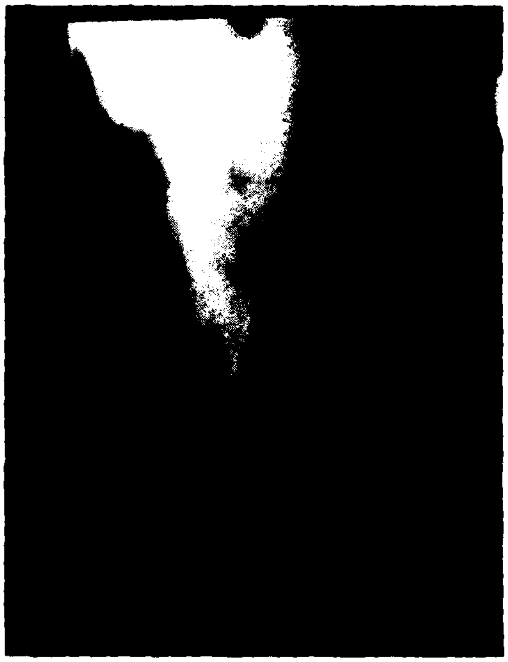
Skull



Chest



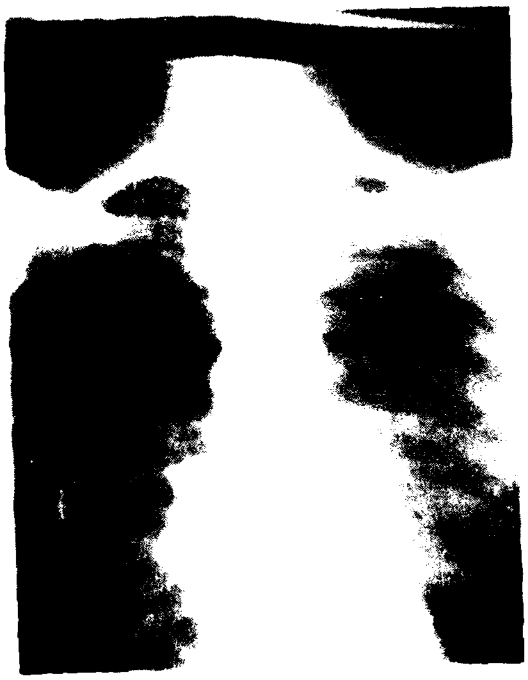
Abdomen



Appendage

Figure 8. Received images, ALC mode.

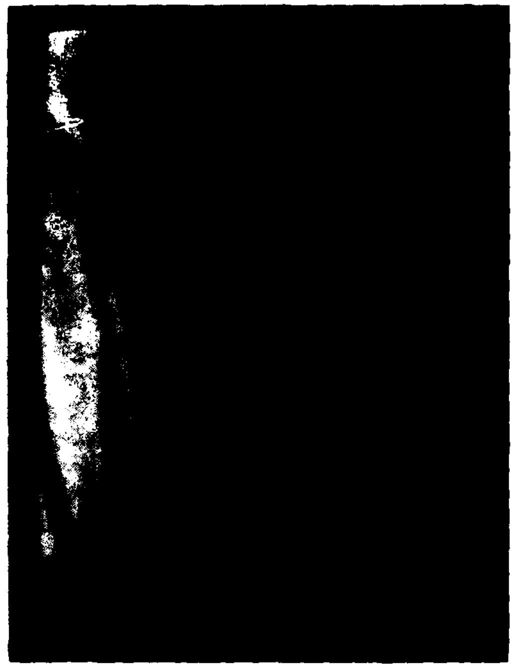
DHF



DHC

Figure 9. Close-up images, digital modes.

AHF



AHC

ALF

ALC

Figure 10. Close-up images, analog modes.

images correspond to the full-view radiographs presented in figures 3 through 8, described above. Although the comparison of a single radiographic zoom image across transmission modes would have facilitated the qualitative analysis of zoom image transmissions, this was not possible because no one radiograph was selected for zoom transmission under all six test conditions. However, each zoom image presented here may be judged against its full-scale image counterpart in figures 3 through 8.

4.1.1 Data

This subsection explores the data used in evaluating the ability of radiologists to interpret images transmitted by the RMDS equipment. Included in the data are measures of Overall Clinical Readings (OCRs), Findings Confidence Levels (FCLs), and Diagnostic Confidence Levels (DCLs); the use of the zoom or close-up capability is also discussed. Data are also provided concerning the amount of time required for diagnosis under the different transmission modes. The actual test radiographs are categorized by Difficulty Ranking Factors (DRFs), zones, and pathology types.

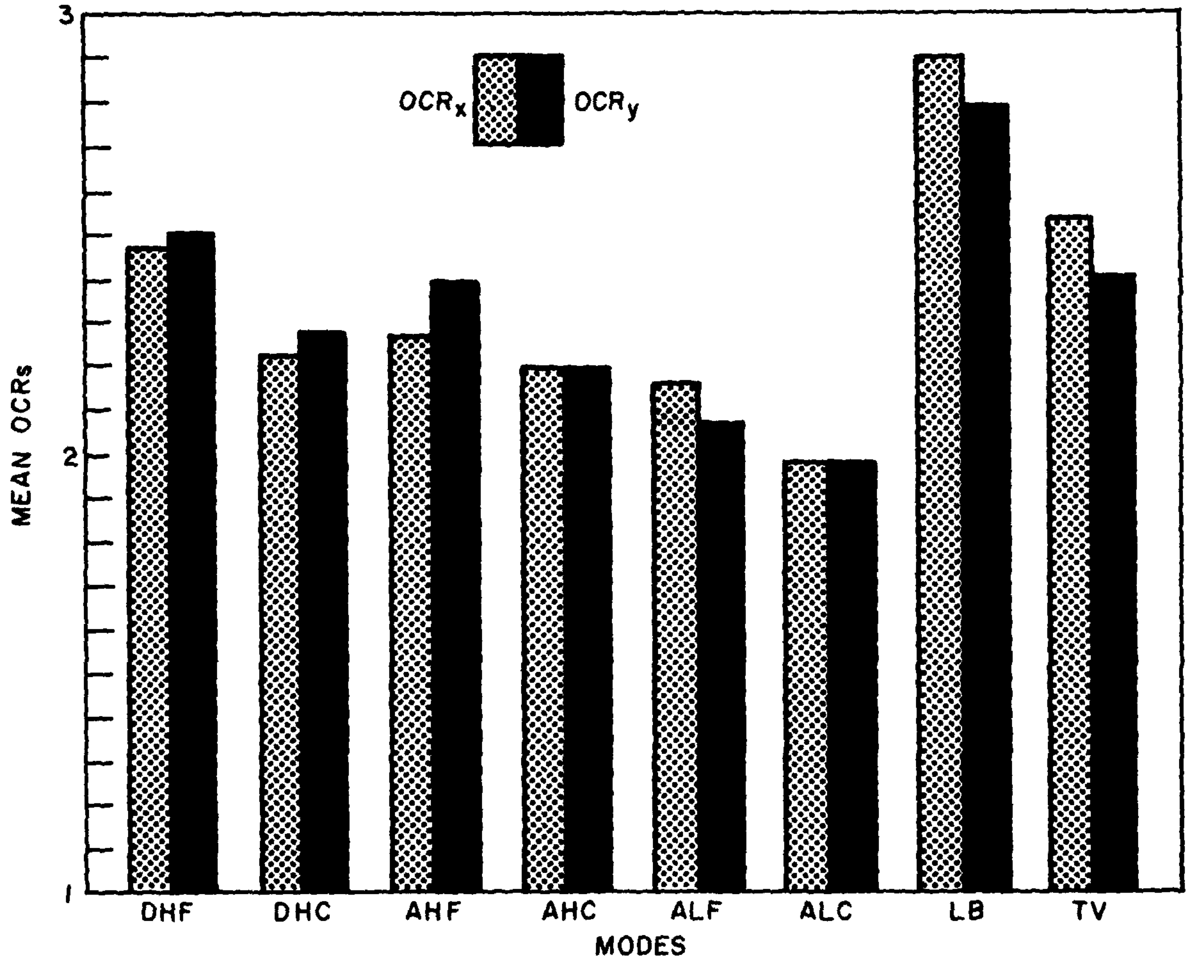
The OCR scores awarded by the two senior radiologists are not an absolute measure of accuracy but are rather a measure of the agreement between the subjects' and the experts' opinions. For ease of discussion, however, the terms "accuracy" and "OCR" will be used interchangeably. Figure 11 shows the mean accuracy scores per transmission mode, as evaluated by Dr X (OCR_x) and Dr Y (OCR_y). These are further refined in figure 12, which breaks out the relative percentages of 1's (unacceptable), 2's (marginal), and 3's (acceptable) given by Dr X and Dr Y. Average FCL and DCL measurements are graphically presented by mode in figure 13. Table 6 summarizes the OCR_x, OCR_y, FCL, and DCL data provided in the above three figures and shows the standard deviation (SD) by mode.

Utilization of the zoom controls is presented by mode, presentation sequence, radiologist, and OCR results in the figures and tables discussed below. Figure 14 shows no relationship between zoom usage and mode but supports a direct relationship between zoom usage and presentation sequence of the radiographic sets. Table 7 illustrates differences in zoom utilization by radiologist, and figure 15 relates zoom usage to the OCR scores resulting from its implementation. Since zoom usage was not a control factor and radiologists were allowed free use of zoom as desired, sample size (n) cannot be established for these open-ended zoom data.

The mean, maximum (max), and minimum (min) amounts of time required to diagnose one set of six radiographs per mode are presented in table 8. These data are organized by mode, and have been adjusted to account for variations in radiographic setup times and for both zoom and image transmission times. The unique characteristics of each of the 36 radiographs employed in RMDS testing were categorized by DRF, zone, and pathology. Table 9 provides a summary of these radiographic qualities.

4.1.2 Data Sheet Comments

Each radiologist participating in the testing had the opportunity to comment on the various radiographs at the bottom of each test data sheet. By far the most common comment offered by the radiologists of both the Test and Control Groups concerned poor radiographic image quality. Table 10 summarizes these "poor quality" comments by transmission mode. Additional comments from the participating



1 = UNACCEPTABLE
2 = MARGINAL
3 = ACCEPTABLE

n = NUMBER OF RADIOGRAPHIC IMAGES FOR EACH MODE
n = 36 (DHF, DHC, AHF, AHC, ALF, ALC)
n = 216 (LIGHTBOX)
n = 108 (CCTV)

Figure 11. Mean Overall Clinical Reading (OCR) evaluations, by mode.

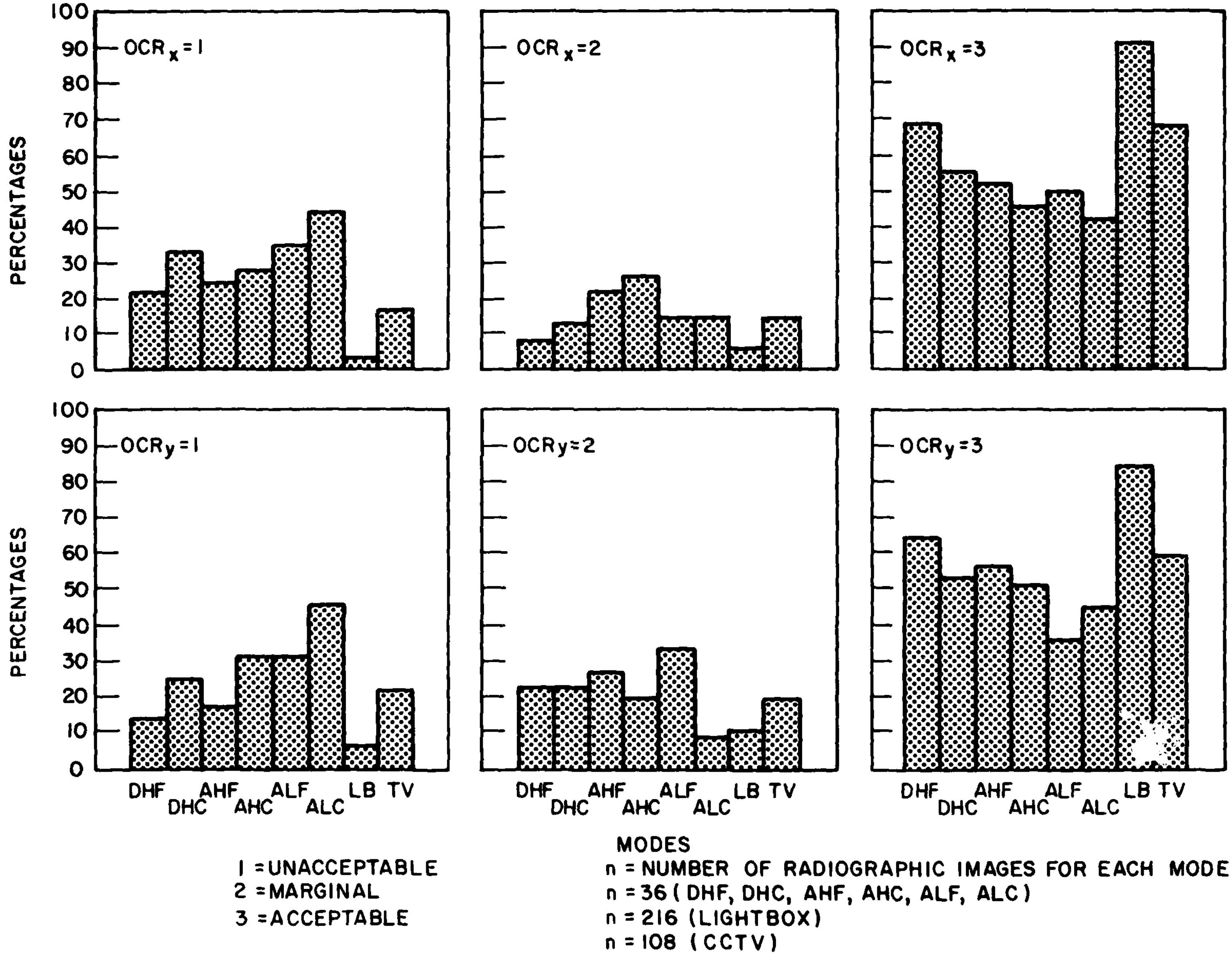
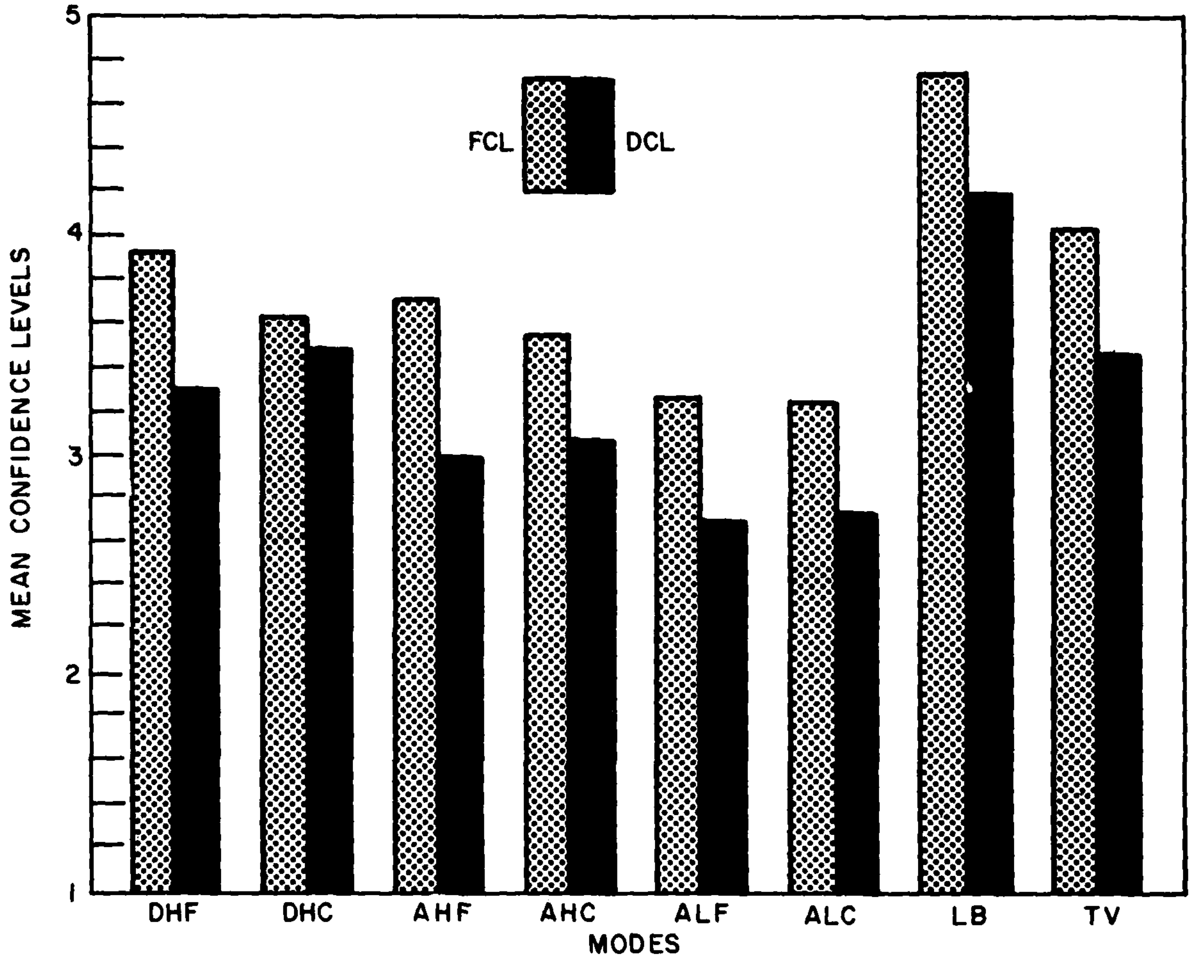


Figure 12. Overall Clinical Reading (OCR) evaluations as a percent of total, by mode.



1 = LOW CONFIDENCE
2 = MED/LOW CONFIDENCE
3 = MED CONFIDENCE
4 = MED/HIGH CONFIDENCE
5 = HIGH CONFIDENCE

n = NUMBER OF RADIOGRAPHIC IMAGES FOR EACH MODE
n = 36 (DHF, DHC, AHF, AHC, ALF, ALC)
n = 216 (LIGHTBOX)
n = 108 (CCTV)

Figure 13. Mean Findings Confidence Levels (FCL) and Diagnostic Confidence Levels (DCL), by mode.

MODE	SD OCR_x	MEAN OCR_x	SD OCR_y	MEAN OCR_y	SD FCL	MEAN FCL	SD DCL	MEAN DCL	n
DHF	2.47	0.84	2.50	0.74	3.92	1.17	3.29	1.20	36
DHC	2.22	0.93	2.28	0.85	3.61	1.10	3.49	1.06	36
AHF	2.28	0.85	2.39	0.77	3.72	0.97	3.00	1.06	36
AHC	2.19	0.86	2.19	0.89	3.57	1.08	3.08	0.92	36
ALF	2.14	0.93	2.06	0.83	3.25	1.16	2.69	1.25	36
ALC	1.97	0.94	1.97	0.97	3.22	1.17	2.72	1.17	36
LB	2.89	0.39	2.78	0.54	4.74	0.49	4.18	0.78	216
CCTV	2.53	0.77	2.40	0.81	4.03	0.98	3.46	0.96	108

Table 6. Summary of raw data.

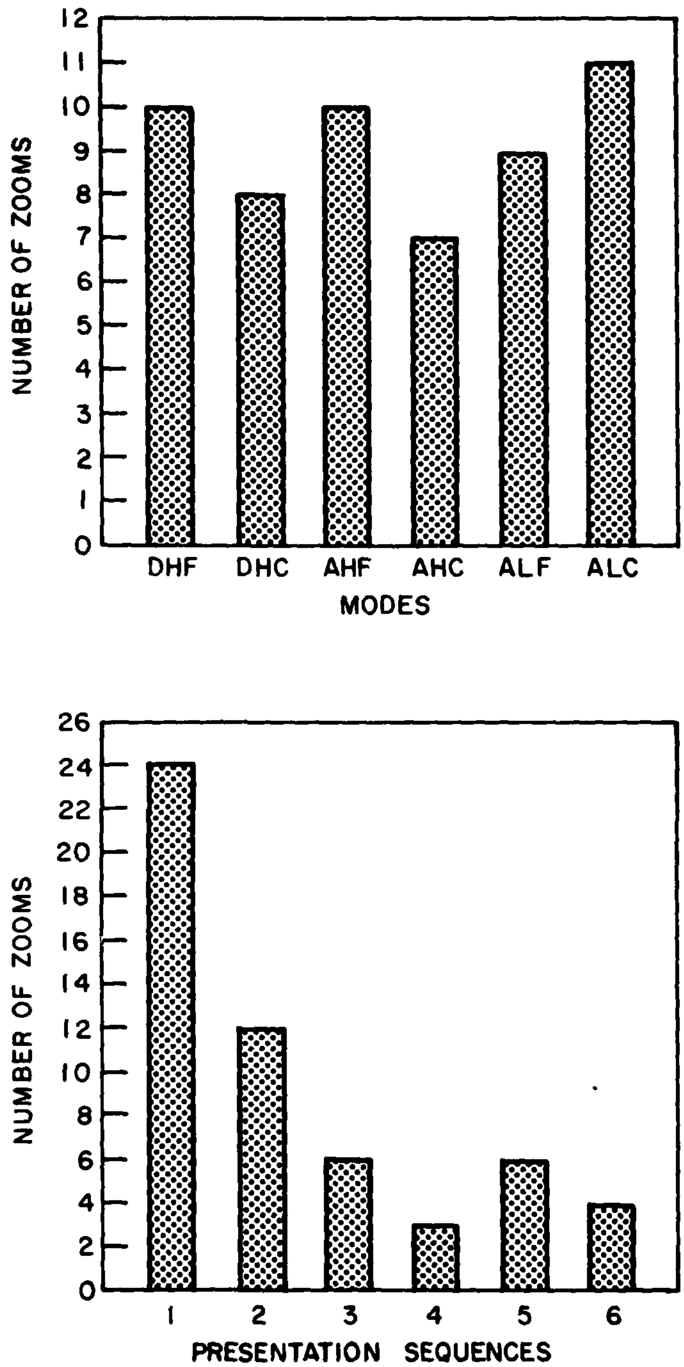


Figure 14. Zoom usage by mode and test presentation sequence.

Test Group Radiologists	Number of Zooms Employed
1	11
2	12
3	1
4	8
5	1
6	22
Control Group Radiologists	**Number of Zooms Employed**
1	22
2	15
3	10
4	7
5	14
6	13

Note: The number of times that zoom capabilities can be employed per radiograph is not limited, and therefore no sample size (n) can be reported.

Table 7. Zoom utilization per radiologist.

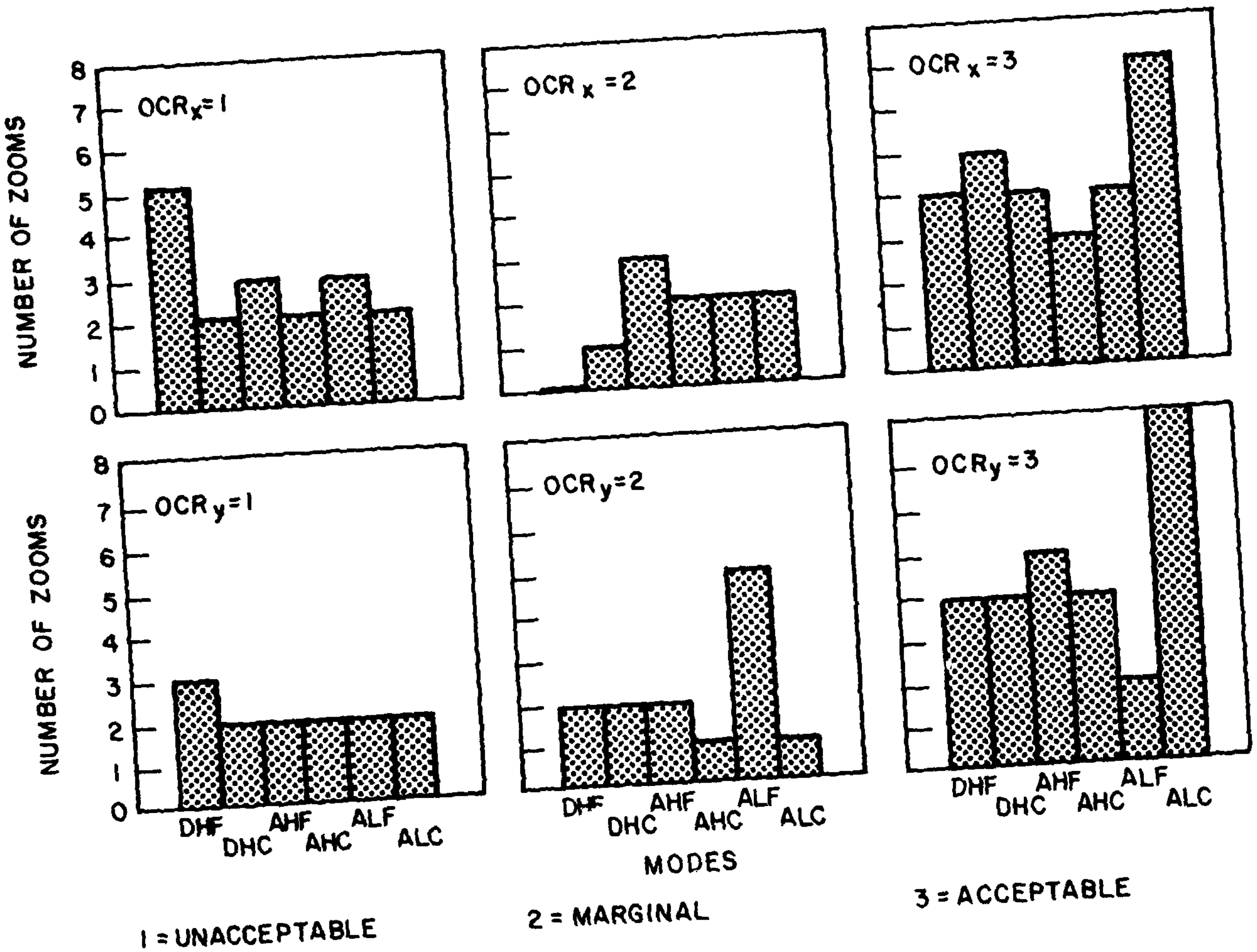


Figure 15. Overall Clinical Reading (OCR) evaluations as a function of zoom usage, by mode.

Mode	Time To Diagnose One Set, minutes			n
	Mean time to diagnose one set	Max time to diagnose one set	Min time to diagnose one set	
DHF	24.9	44.2	9.2	6 sets
DHC	19.7	36.8	7.2	6 sets
AHF	25.4	49.0	10.6	6 sets
AHC	16.3	32.0	7.2	6 sets
ALF	22.6	68.2	10.6	6 sets
ALC	20.0	28.2	9.2	6 sets
Lightbox	15.8	40.0	7.5	36 sets
CCTV	20.3	37.3	8.7	18 sets

Note: Average radiograph setup and transmission times, as well as the times required for zoom presentation, were subtracted from the raw data available per mode to produce the values listed above. These adjusted measurements reflect a more realistic image of the time actually spent in diagnosis and provide for ease in comparison across modes.

Table 8. Times required for diagnosis.

Difficulty Ranking Factor (DRF)	Zone				
	Appendage	Abdomen	Chest	Skull	Total
1. Low	1	3	0	2	6
2. Low/med low	0	1	3	2	6
3. Med low	2	1	2	1	6
4. Med high	2	1	2	1	6
5. Med high/high	2	0	2	2	6
6. High	1	1	2	2	6
Total:	8	7	11	10	36

Note: The pathologies of these 36 radiographs were of the following types:

17 - tissue
15 - skeletal
4 - tissue and skeletal

Table 9. Number of radiographs by zone and Difficulty Ranking Factor.

Mode	Number of Responses	Percent of Total	n
DHF	2	5.6	36
DHC	6	16.7	36
AHF	2	5.6	36
AHC	4	11.0	36
ALF	8	22.2	36
ALC	12	33.3	36
All test modes, total	34	15.7	216
Lightbox	4	1.9	216
CCTV	15	13.9	108

Table 10. Radiologists' data sheet comments concerning poor radiographic image quality.

radiologists were as follows (the transmission mode pertaining to each comment being shown in parentheses):

- "Changing contrast adjustment can really accentuate bony density" (ALC)
- "Bony detail is much better than soft tissue detail" (AHC)
- "Image on storage screen is better" (DHC)
- "Image shakes with increasing contrast" (DHC)
- "Mag (zoom) was crucial" (AHF)
- "Contrast range is too narrow for gray zones of a chest radiograph" (AHC)
- "Even with enlargement . . . I am not confident that these are real findings" (DHF)
- "Unable to adjust contrast and brightness to adequately evaluate soft tissue structures in the abdomen" (CCTV)

4.1.3 Radiologists' Questionnaires

At the completion of testing for both the Test and Control Groups, the participating radiologists were asked to complete a questionnaire regarding their interpretation of the overall clinical utility of the RMDS concept. Sample questionnaire sheets are included in appendix B. The following pages, broken out by individual question, address the subjects' comments and perceptions. The responses to this questionnaire are divided into Test and Control Groups. Note that the participants were not aware of the group to which they belonged.

<u>Question:</u> Do you feel that satisfactory radiology consultations for emergency cases can be made via the RMDS?

Listed below are the responses to this question:

	Test Group	Control Group
Yes	4	6
No	1	0
Other	1*	

* Did not address question

All the "yes" responses for the Test and Control Groups were accompanied by a qualifying statement. In almost every case, it was felt that the RMDS would be useful for gross pathologies with fairly specific and significant findings. Doubt was expressed as to its usefulness for various other radiological pathologies. (This issue is addressed in the next question.) The comment provided with the single "no" response (Test Group) indicated that the RMDS was not adequate to spot subtle changes and that pathologies gross enough to be transmitted by the system should be obvious to the patient's attendants.

Question: Do you feel that there are particular types of pathologies that may or may not be readily diagnosed using the RMDS?

Eleven of the twelve radiologists answered "yes," and one radiologist in the Test Group did not address the question. Listed below are the different pathologies cited by the radiologists. The number following the pathology type indicates the total number of radiologists providing a response of this type.

Not Readily Diagnosed		Readily Diagnosed	
• Bone		• Bone	
- skull lesions	(5)	- bone disease	(1)
- bone disease and tumors	(2)	- fractures	(2)
- subtle fractures & irregularities	(4)		
• Calcifications		• Calcifications	
- general	(1)	- general	(2)
- small	(1)		
- soft tissue	(1)		
• Lungs		• Lungs	
- general	(3)	- general	(2)
- pneumothorax	(2)		
- subtle findings	(2)		
- interstitial patterns	(1)		
• Abdomen	(4)	• Abdomen	(1)
• Soft Tissue		• Trauma	(1)
- subtle changes	(1)	• Sclerotic Changes	(1)
- small infiltrates	(1)	• Pneumonia	(1)

The cited pathologies varied by radiologist within the groups (Test and Control) and between groups, and no identifiable response pattern could be identified.

Question: Was the "zoom" capability (enlargement of a portion of the image) useful to you in making a diagnosis?

Listed below are the responses to this question:

	Test Group	Control Group
Yes	2	4
Sometimes	3	1
No	1	1

Several of the radiologists supplied an additional qualitative statement supporting their responses. These qualitative statements concerning zoom usefulness are listed below.

Zoom Useful

- "Improved regional resolution"
- "Allowed better resolution of abnormal area"
- "Often used in looking for calcification and periostal reaction"

Zoom Somewhat Useful

- "Zoom reconfirmed initial finding; if no pathology identified on first image, zoom did not aid in evaluation"
- "Helped make picture large, but not clearer"

Zoom Not Useful

- "Did not significantly improve resolution"

The radiologists in the Test Group were provided with the capability of storing an image for later use, which eliminated the necessity for retransmission of a radiograph. The Control Group did not have this feature, since an image could be provided without extensive transmission time. The radiologists in the Test Group were asked if they had used the "video storage" system when utilizing the zoom. All six radiologists stated that they had utilized the video storage system to allow reviewing of the original radiograph.

Question: Would one image or two simultaneous images be sufficient for this type of system?

Listed below are the responses to this question:

	Test Group	Control Group
One image	4	1
Two images	1	4
Other	1*	1**

---

*Response: Yes.
**Response: No; additional views for some subtle defects are necessary.

Of the five radiologists who stated that one image would be sufficient, four qualified their responses by indicating that some form of image retrieval system would be helpful or needed for review of the original or of additional radiographs.

Question: Do you see a need for archiving some images on disc memory for later consultation?

Listed below are the responses received for this question:

	Test Group	Control Group
Yes	—	5
Helpful/possibly	4	—
No	2	1

None of the radiologists in the Test Group responded to this question with a definite "yes." The four radiologists in the "helpful/possibly" category qualified their responses as follows:

- "If image quality was equal to diagnostic film quality"
- "Helpful but not necessary"
- "Dependent on the reliability of RMDS"
- "Useful in follow-up, but not necessary in emergency"

Question: Did you use the "reverse polarity" feature (positive or negative image)? Would such a feature be of any use to you as a radiologist?

The radiologists in the Test Group had the option of utilizing a "reverse polarity" feature on the RMDS, which reverses the black and white aspects of an image. At the beginning of each radiologist's first set of radiographs, each was instructed in its usage. Only one of the six radiologists utilized this reverse polarity feature, however, and this subject used it only once. All six of the radiologists felt that this feature was of little use.

Question: Please make any additional comments you wish.

Test Group

- "At present, the images are adequate for gross abnormalities, which should be recognized by trained individuals. Losses in image quality obscures fine details necessary for radiographic interpretation. Again, difficulty in identifying calcium in lesions was a major limitation."
- "The system is good for transmitting obvious abnormalities."

Control Group

- "I believe from reviewing the transmitted images against the actual film that 20-25 percent of diagnoses will be missed, especially abdomen and bone."

- "The pictures projected are not as clear as the films. Additional views are necessary in some lesions or defects."

- "I see two problems: (1) Resolution not high enough. (2) Ideal contrast difficult to obtain because X-rays of different parts need to be seen at different contrast levels."

- "I was biased on this exam by knowing there must be a pathology on the X-rays somewhere."

4.2 QUANTITATIVE RESULTS

Section 2.2 provided a list of the basic issues to be resolved by the RMDS radiographic testing. This section presents the results of the statistical analysis in an attempt to resolve those issues. Each issue is questioned separately and presented with the statistical results to support the acceptance/rejection of the stated hypotheses. The questions listed below are not addressed hierarchically, but reflect organization by group.

The two principal hypotheses of the RMDS radiographic test, as stated in section 2.2, will be addressed under the heading, "Principal Hypotheses Evaluation." Two tests not called for in section 2.2 were also performed on the data; these results are provided under the heading, "General Results." One of these tests was performed to validate the assumption that all six sets of radiographs were of equal difficulty. The other was performed to determine whether or not there was justification for the combination of OCR_x and OCR_y results for further statistical analysis. A summary of the test results is provided at the end of this section to clarify and consolidate the information contained herein.

4.2.1 General Results

- **Question:** Were all the sets of radiographs of equal difficulty?

  Null Hypothesis: Set A = B = C = D = E = F, with respect to difficulty.

  Results: A Chi-square (χ^2) analysis of the total number of acceptable OCR measurements obtained per set for the Control Group indicated that the null hypothesis was accepted. The probability associated with the calculated values of χ^2 lay between 0.95 and 0.90 for OCR_x, and 0.975 and 0.95 for OCR_y.

- **Question:** Were OCR_x and OCR_y scores equal per radiologist per X-ray?

  Null Hypothesis: $OCR_x = OCR_y$.

  Results: The Wilcoxon's T test was utilized to test for differences between OCR_x and OCR_y. The T test indicated that OCR_x scores were not statistically equal to OCR_y scores at the 0.05 significance level for the Lightbox and CCTV Groups, whereas OCR_x scores were equal to OCR_y scores for the Test Group. Because of the variations in OCR_x and OCR_y, they will be treated separately throughout this report.

4.2.2 Test Group Results

- **Question:** Were there any differences between OCR scores for the transmission modes?

  Null Hypothesis: DHF = DHC = AHF = AHC = ALF = ALC, with respect to OCR.

  Results: A χ^2 analysis of the total number of acceptable OCR measurements per mode and of the acceptable plus marginal OCR measurements per mode indicated that there were no statistical differences between transmission modes. Listed below are the probabilities associated with the calculated values of χ^2 for the various conditions tested.

	Acceptable only	Acceptable + Marginal
OCR_x	$0.75 > \chi^2 > 0.50$	$0.95 > \chi^2 > 0.75$
OCR_y	$0.75 > \chi^2 > 0.50$	$0.75 > \chi^2 > 0.50$

- **Question:** Were there any differences between the Test Group radiologists with respect to OCRs?

  Null Hypothesis: Radiologist 1 = 2 = 3 = 4 = 5 = 6, with respect to OCR.

  Results: A χ^2 analysis of the total number of acceptable OCR measurements per radiologist indicated that there were no statistical differences between radiologists of the Test Group. Listed below are the probabilities associated with the calculated values of χ^2 for the various conditions tested.

	Acceptable only	Acceptable + Marginal
OCR_x	$0.75 > \chi^2 > 0.50$	$0.75 > \chi^2 > 0.75$
OCR_y	$0.50 > \chi^2 > 0.25$	$0.90 > \chi^2 > 0.75$

- **Question:** Were there any differences between Diagnostic Confidence Levels (DCLs) and Findings Confidence Levels (FCLs) for the different transmission modes?

Null Hypothesis: DHF = DHC = AHF = AHC = ALF = ALC, with respect to FCL.

Results: A one-way analysis of variance (ANOVA) indicated that there were no differences between FCLs for the six transmission modes at the 0.05 significance level. The F probability value obtained from the ANOVA was 0.062, which falls outside the desired critical value of 0.05.

Null Hypothesis: DHF = DHC = AHF = AHC = ALF = ALC, with respect to DCL.

Results: A one-way ANOVA performed on DCLs indicated that a difference does exist between transmission modes (F probability = 0.017). A Duncan multiple-range test at the 0.05 critical value level established two homogeneous subsets (wherein the means of the first and last group differed by less than the critical value level for a subset of that size). Listed below are the two homogeneous subsets and the mean for each mode.

Subset A

Mode	ALF	ALC	AHF	AHC
Mean	2.69	2.72	3.00	3.08

Subset B

Mode	AHF	AHC	DHF	DHC
Mean	3.00	3.08	3.29	3.47

- Question: Was there a correlation between confidence levels and OCRs for each mode?

Null Hypotheses:

Correlation between OCR_x and FCL = 0 for DHF, DHC, AHF, AHC, ALF, ALC.

Correlation between OCR_y and FCL = 0 for DHF, DHC, AHF, AHC, ALF, ALC.

Correlation between OCR_x and DCL = 0 for DHF, DHC, AHF, AHC, ALF, ALC.

Correlation between OCR_y and DCL = 0 for DHF, DHC, AHF, AHC, ALF, ALC.

Results: A Pearson's correlation showed that there was a significant and positive correlation between OCRs and FCLs within each mode. A Pearson's correlation provided similar results for DCLs, with the exception of the AHC mode. Table 11 shows correlation coefficients

Test Group	DHF		DHC		AHF		AHC		ALF		ALC	
	Coef	Sig	Coef	Sig	Coef	Sig	Coef	Sig	Coef	Sig	Coef	Sig
OCR_x and FCL	0.7360	0.000	0.5922	0.000	0.6487	0.000	0.4810	0.001	0.7862	0.000	0.6040	0.000
OCR_y and FCL	0.6474	0.000	0.4418	0.003	0.4932	0.001	0.3886	0.010	0.6697	0.000	0.5224	0.001
OCR_x and DCL	0.6650	0.000	0.5695	0.000	0.6058	0.000	0.1961	0.126	0.5782	0.000	0.6559	0.000
OCR_y and DCL	0.6396	0.000	0.4494	0.003	0.5121	0.001	0.2238	0.095	0.6535	0.000	0.5853	0.000

Table 11. Pearson's correlation data for Test Group, by transmission mode.

and their significance for the Pearson's correlation of OCR, FCL, and DCL per mode.

- **Question:** Was there a relationship between the pathology types, confidence levels, and transmission modes?

Null Hypotheses:

DHF = DHC = AHF = AHC = ALF = ALC, with respect to FCL for appendage radiographs.

DHF = DHC = AHF = AHC = ALF = ALC, with respect to DCL for appendage radiographs.

DHF = DHC = AHF = AHC = ALF = ALC, with respect to FCL for abdomen radiographs.

DHF = DHC = AHF = AHC = ALF = ALC, with respect to DCL for abdomen radiographs.

DHF = DHC = AHF = AHC = ALF = ALC, with respect to FCL for chest radiographs.

DHF = DHC = AHF = AHC = ALF = ALC, with respect to DCL for chest radiographs.

DHF = DHC = AHF = AHC = ALF = ALC, with respect to FCL for skull radiographs.

DHF = DHC = AHF = AHC = ALF = ALC, with respect to DCL for skull radiographs.

Results: A one-way analysis of variance was performed for each of the above hypotheses. The FCL or DCL measurements for the particular pathology types were grouped per mode, and differences were tested between modes. Listed below are the F ratios and F probabilities associated with that testing.

Test	F Ratio	F Probability
FCL appendage	2.389	0.054
DCL appendage	3.021	0.020
FCL abdomen	1.417	0.242
DCL abdomen	0.864	0.515
FCL chest	0.690	0.633
DCL chest	0.669	0.649
FCL skull	0.746	0.593
DCL skull	1.664	0.159

Only the DCL of the appendage group showed significant differences at the 0.05 level between the six transmission modes. A Duncan multiple-range test was performed on this group at the 0.05 significance level, which divided the six transmission modes into the two homogeneous subsets listed below.

Subset A

Mode	ALF	ALC
Mean	1.81	2.75

Subset B

Mode	ALC	AHC	DHF	DHC	AHF
Mean	2.75	3.13	3.19	3.38	3.44

- **Question:** What effect did the utilization of the zoom have on OCR?

  Null Hypothesis: Zoom OCRs = No Zoom OCRs.

  Results: A χ^2 contingency table was constructed with Test Group OCR data (OCRs with zoom used vs OCRs with zoom not used). An analysis per transmission mode was impossible because of the small sample size, so the analysis was performed by combining the common OCR measurements for each mode. This analysis indicated that there were no statistical differences between the OCR measurements of the zoom groups and nonzoom groups. The probabilities associated with the calculated values of χ^2 for this analysis were as follows:

  OCR_x $\quad 0.75 > \chi^2 > 0.50$

  OCR_y $\quad 0.90 > \chi^2 > 0.75$

  Two additional statistical analyses were performed for the zoom data found in figures 14 and 15. A χ^2 analysis was performed to determine if more zooms were used in any one mode. The results of this testing indicated that the numbers of zooms used per mode were statistically equal ($0.95 > \chi^2 > 0.90$). The second analysis (χ^2) was performed to determine if the numbers of zooms used per test presentation sequence were statistically equal. The results of this testing indicated that the numbers of zooms per test presentation sequence were not equal for the probabilities associated with the calculated value of $\chi^2 < 0.001$. This analysis was repeated after removing the data for the first test presentation sequence; in this attempt, sequences two through six were statistically equal for the probabilities associated with the calculated value of χ^2 between 0.25 and 0.10.

4.2.3 Control Group Results

- **Question:** Were there any differences between FCL, DCL, and OCR for lightbox and CCTV monitor viewing?

Null Hypotheses: FCL of lightbox = CCTV.
DCL of lightbox = CCTV.
OCR of lightbox = CCTV.

Results: Since the radiologists in the Control Group had utilized both the lightbox and CCTV to view various radiographs, the Wilcoxon matched-pairs signed-ranks test was used to analyze the data. The results of the Wilcoxon T test indicated that there was a significant difference (> 0.001) between the lightbox and the CCTV in terms of FCL, DCL, and OCR. Figures 12 and 13 indicate that FCL, DCL, and OCR are higher (improved) for lightbox versus CCTV. This result is logical.

- **Question:** Were there any differences between the lightbox and CCTV monitor viewing in terms of FCL, DCL, and OCR per radiograph Difficulty Ranking Factor (DRF)?

Null Hypotheses: FCL of lightbox = CCTV per DRF (1 to 6).
DCL of lightbox = CCTV per DRF (1 to 6).
OCR of lightbox = CCTV per DRF (1 to 6).

Results: The data for the radiologists of the Control Group were divided into subgroups by the DRF of the radiographs. The Wilcoxon T test was then utilized to determine if differences existed between the lightbox and the CCTV in terms of DCL, FCL, and OCR. Since the number of OCR_x and OCR_y measurements per DRF was too small to allow for a meaningful analysis, a combination of the measurements (OCR_{x+y}) was used for statistical testing. Table 12 shows the results of the T testing. As indicated, the vast majority of FCL, DCL, and OCR scores showed significant differences between the lightbox and CCTV per DRF. The four cases where no differences existed between the lightbox and CCTV are not considered statistically significant in light of the propensity for rejection of the null hypotheses of the tests.

- **Question:** Were there any differences between the confidence levels (FCL and DCL)?

Null Hypotheses: Radiologist 1 = 2 = 3 = 4 = 5 = 6, with respect to FCL.
Radiologist 1 = 2 = 3 = 4 = 5 = 6, with respect to DCL.

Results: A one-way ANOVA indicated that there was a difference between the FCL and DCL of the Control Group radiologists. The F probability values obtained from the ANOVA were 0.018 for FCL and 0.004 for DCL, which fall below the critical 0.05 significance level. A Duncan multiple-range test was performed for both FCL and DCL at the 0.05 significance level. The homogeneous subsets formed by this testing are listed below.

Null Hypothesis: CCTV = Lightbox

	FCL		DCL		OCR_{x+y}	
DRF	Accept/Reject	Sig	Accept/Reject	Sig	Accept/Reject	Sig
1	Reject	T < 0.01	Reject	T < 0.01	Accept	T > 0.05
2	Accept	T > 0.05	Accept	T > 0.05	NA	(N=4)
3	Reject	T < 0.01	Reject	T < 0.01	Reject	T < 0.01
4	Reject	T < 0.01	Reject	T < 0.01	Reject	T < 0.01
5	Reject	T < 0.01	Accept	T > 0.05	Reject	T < 0.01
6	Reject	T < 0.01	Reject	T < 0.01	Reject	T < 0.01

Table 12. Wilcoxon T test data for CCTV vs lightbox, by Difficulty Ranking Factor (DRF).

FCL

Subset A

Radiologist	1	5
Mean	4.53	4.64

Subset B

Radiologist	5	3	4	2	6
Mean	4.64	4.76	4.82	4.83	4.88

DCL

Subset A

Radiologist	1	4	6	3
Mean	3.88	4.04	4.13	4.15

Subset B

Radiologist	4	6	3	5
Mean	4.04	4.13	4.15	4.29

Subset C

Radiologist	5	2
Mean	4.29	4.58

- **Question:** Were there any differences among the accuracy (OCR) results of the Control Group radiologists?

  Null Hypothesis: Radiologist 1 = 2 = 3 = 4 = 5 = 6, with respect to OCR.

  Results: A χ^2 analysis of the total number of acceptable OCR measurements per radiologist and of the acceptable plus marginal OCR measurements per radiologist indicated that there was no statistical difference between the performances of the radiologists of the Control Group. Listed below are the probabilities associated with the calculated values of χ^2 for the various conditions tested.

	Acceptable only	Acceptable + Marginal
OCR_x	$0.995 > \chi^2 > 0.99$	$\chi^2 > 0.999$
OCR_y	$0.90 > \chi^2 > 0.75$	$0.99 > \chi^2 > 0.995$

4.2.4 **Principal Hypotheses Evaluation**

- **Hypothesis:** The fine-resolution, high-SNR, analog and digital transmission modes (DHF, AHF) are statistically equal to the CCTV in terms of confidence levels (DCL and FCL) and accuracy (OCR).

  Results: A χ^2 analysis was performed on the total number of acceptable OCR measurements and on the acceptable plus marginal OCR measurements for the DHF and AHF transmission modes. An adjusted n-value was used for the CCTV as the expected frequency. This test indicated that there was no statistical difference in radiologist performance when using DHF/AHF and CCTV transmission modes. Listed below are the probabilities associated with the calculated values of χ^2 for the various conditions tested.

	Acceptable only	Acceptable + Marginal
OCR_x	$0.257 > \chi^2 > 0.10$	$0.75\ \chi^2 > 0.50$
OCR_y	$0.75 > \chi^2 > 0.50$	$0.75 > \chi^2 > 0.50$

  A one-way ANOVA was performed on FCL and DCL data for DHF, AHF, and CCTV. The results of these tests indicated that a difference did exist between the radiologists' confidence levels between the CCTV and the two RMDS transmission modes (DHF and AHF). The F probability value obtained from the ANOVA was 0.000 for both the FCL and DCL. A Duncan multiple-range test at the 0.05 critical value level established the homogeneous subsets listed below.

FCL

Subset A

Mode	AHF	DHF
Mean	3.00	3.19

Subset B

Mode	CCTV
Mean	4.03

DCL

Subset A

Mode	AHF	DHF
Mean	2.69	2.90

Subset B

Mode	CCTV
Mean	3.46

- **Hypothesis:** The analog and digital fine-resolution transmission modes are statistically better than the analog and digital coarse-resolution modes in terms of confidence (FCL and DCL) and accuracy (OCR), and they significantly reduce diagnostic inaccuracies.

  Results: As previously illustrated, there were no statistical differences found between the transmission modes with respect to FCL and OCR. Therefore, this aspect of the hypothesis must be rejected. Although differences were found between transmission modes with respect to DCL, the divisioning by homogeneous subset at the 0.05 significance level did not indicate a significant difference between coarse and *fine resolution for* either the analog or the digital transmission mode. Listed below are the two homogeneous subsets for DCL at the 0.05 level.

Subset A

Mode	ALF	ALC	AHF	AHC
Mean	2.69	2.72	3.00	3.08

Subset B

Mode	AHF	AHC	DHF	DHC
Mean	3.00	3.08	3.29	3.47

Table 13 provides a summary of the statistical analyses presented in this section.

Hypothesis	Variable	Test	Rejected/ Accepted	Significance
• **General**				
- Radiographic Set A = B = C = D = E = F	OCR_x*	χ^2	Accepted	$0.95 > \chi^2 > 0.90$
	OCR_y*	χ^2	Accepted	$0.975 > \chi^2 > 0.95$
- $OCR_x = OCR_y$	Lightbox	Wilcoxon T	Rejected	$T > 0.001$
	CCTV	Wilcoxon T	Rejected	$0.05 > T > 0.02$
	Test Group	Wilcoxon T	Accepted	$T > 0.50$
• **Test Group**				
- DHF = DHC = AHF = AHC = ALF = ALC	OCR_x*	χ^2	Accepted	$0.75 > \chi^2 > 0.50$
	OCR_x**	χ^2	Accepted	$0.95 > \chi^2 > 0.75$
	OCR_y*	χ^2	Accepted	$0.75 > \chi^2 > 0.50$
	OCR_y**	χ^2	Accepted	$0.75 > \chi^2 > 0.50$
	FCL	ANOVA	Accepted	F = 0.062
	DCL	ANOVA	Rejected	F = 0.017
	FCL appendage	ANOVA	Accepted	F = 0.054
	DCL appendage	ANOVA	Rejected	F = 0.020
	FCL abdomen	ANOVA	Accepted	F = 0.242
	DCL abdomen	ANOVA	Accepted	F = 0.515
	FCL chest	ANOVA	Accepted	F = 0.633
	DCL chest	ANOVA	Accepted	F = 0.649
	FCL skull	ANOVA	Accepted	F = 0.593
	DCL skull	ANOVA	Accepted	F = 0.159
	No. of Zooms	χ^2	Accepted	$0.95 > \chi^2 > 0.90$
- Confidence/Accuracy Correlation	OCR_x x FCL	Pearson	—	(See table 11)
	OCR_y x FCL	Pearson	—	(See table 11)
	OCR_x x DCL	Pearson	—	(See table 11)
	OCR_y x DCL	Pearson	—	(See table 11)

* Acceptable only

** Acceptable plus marginal

Table 13. Summary of statistical test results.

Hypothesis	Variable	Test	Rejected/ Accepted	Significance
- Radiologist 1 = 2 = 3 = 4 = 5 = 6	OCR_x*	χ^2	Accepted	$0.75 > \chi^2 > 0.50$
	OCR_x**	χ^2	Accepted	$0.75 > \chi^2 > 0.50$
	OCR_y*	χ^2	Accepted	$0.50 > \chi^2 > 0.25$
	OCR_y**	χ^2	Accepted	$0.90 > \chi^2 > 0.75$
- Zoom = No Zoom	OCR_x	χ^2	Accepted	$0.75 > \chi^2 > 0.50$
	OCR_y	χ^2	Accepted	$0.90 > \chi^2 > 0.75$
- Presentation Sequence 1 = 2 = 3 =4 = 5 = 6	No. of Zooms	χ^2	Rejected	$\chi^2 < 0.001$
- Presentation Sequence 2 = 3 = 4 =5 = 6	No. of Zooms	χ^2	Accepted	$0.25 > \chi^2 > 0.10$
• Control Group				
- Lightbox = CCTV	FCL	Wilcoxon T	Rejected	$0.01 > T > 0.001$
	DCL	Wilcoxon T	Rejected	$0.01 > T > 0.001$
	OCR	Wilcoxon T	Rejected	$0.01 > T > 0.001$
	DRF	Wilcoxon T	—	(See table 12)
- Lightbox Radiologist 1 = 2 = 3 = 4 = 5 = 6	FCL	ANOVA	Rejected	$F = 0.018$
	DCL	ANOVA	Rejected	$F = 0.004$
	OCR_x*	χ^2	Accepted	$0.995 > \chi^2 > 0.99$
	OCR_x**	χ^2	Accepted	$\chi^2 > 0.999$
	OCR_y*	χ^2	Accepted	$0.90 > \chi^2 > 0.75$
	OCR_y**	χ^2	Accepted	$0.999 > \chi^2$
• Principal Hypotheses				
- DHF = AHF = CCTV	FCL	ANOVA	Rejected	$F = 0.000$
	DCL	ANOVA	Rejected	$F = 0.000$
	OCR_x*	χ^2	Accepted	$0.25 > \chi^2 > 0.10$
	OCR_x**	χ^2	Accepted	$0.75 > \chi^2 > 0.50$
	OCR_y*	χ^2	Accepted	$0.75 > \chi^2 > 0.50$
	OCR_y**	χ^2	Accepted	$0.75 > \chi^2 > 0.50$

* Acceptable only

** Acceptable plus marginal

Table 13. Summary of statistical test results (continued).

SECTION 5

DISCUSSION

The RMDS radiologist testing concerned the two major areas of radiologists' diagnostic accuracy and confidence levels (findings and diagnostic) for various image processing. Two additional areas of interest pursued during the testing were the radiologist system controls (ie, contrast and brightness settings) and RMDS systems capabilities. These four topics are discussed individually in the following subsections. A treatment of both qualitative and quantitative data is provided in each subsection to consolidate the information relevant to each topic.

The overall clinical utility of the RMDS, addressed in the questionnaires filled out by the participating radiologists, produced some valuable results. Ten of the twelve radiologists felt that the RMDS terminal capabilities employed in the testing would provide satisfactory images for radiological consultations concerning gross pathologies and implied that the RMDS terminals have limitations for particular pathologies. If the RMDS terminals are to be used for diagnostic support of the more subtle pathologies, then the RMDS parametric capabilities for video fidelity must be made more stringent, ie, the system would be modified to provide increased resolution and improved SNR of received images. The key issue in answering this question is the type of pathology for which the RMDS will be required to support diagnosis.

5.1 ACCURACY

No statistical difference was found among the RMDS transmission modes. However, three general trends in accuracy were observed in the raw data (figures 11 and 12, and table 6). These tendencies take the form of decreasing accuracy when going from the digital to the analog transmission mode, from high to low SNR, and from fine to coarse resolution. The six transmission modes can be ranked roughly from best to worst in the following order: DHF, AHF, DHC, AHC, ALF, and ALC. Although no one mode was shown to be statistically superior to the others, both the limited data and electronic and communication theory would support this ranking. The lack of significant differences between transmission modes could be attributed to the fact that (1) the radiographs used in the testing were not sufficiently distributed over the spectrum of possible difficulties, and (2) the limited amount of testing precluded achieving the sample size necessary to differentiate the modes.

The accuracy of the two Control Groups (lightbox (LB) and CCTV) relative to the RMDS transmission modes indicated a ranking from best to worst as follows: LB, CCTV, DHF, AHF, DHC, AHC, ALF, and ALC. Testing showed that the difference between CCTV and the DHF transmission mode was not statistically significant, whereas the difference between LB and CCTV was significant. These results indicated that a 6-bit-per-picture-element quantization and the frozen noise of the received image were equivalent to a true analog image under similar bandwidth constraints.

For the purposes of achieving a sound experimental design, radiographs were balanced between sets according to DRF. A statistical analysis of the accuracy data for each set indicated that no differences existed among the sets. Similarly, radiologists were balanced between groups (Test or Control) based upon experience and

PRECEDING PAGE BLANK-NOT FILMED

training. The statistical analysis of the accuracy data for each radiologist per group indicated that no differences existed within either the Test Group or Control Group.

A statistical comparison of the accuracy data with zoom capacity use versus no zoom utilization indicated that no differences existed per group of radiologists. However, this is not to say that making use of the zoom capability did not aid a particular radiologist on a particular radiograph. A more detailed discussion of zoom utilization is provided in section 5.4.

A comparison of the OCR per Difficulty Ranking Factor (DRF) was performed on the CCTV and LB groups. This testing was done to determine whether radiographic difficulty affected the accuracy of both groups equally. The results of this testing indicated that of the six DRF rankings (1 to 6), only the radiographs rated as 1 (low) were viewed with equal accuracy for the LB and CCTV.

Two problems were encountered in the statistical analysis of accuracy (OCR) data. The first involved the semiordinal (1 to 3 scale) type of accuracy measure employed. Since a large number of primary and secondary findings were possible per radiograph (see appendix A), accuracy could not be measured as merely right or wrong; this inability to evaluate accuracy on a binary scale (or on a larger interval scale) resulted in statistical manipulation problems. Should further testing be required, greater care should be taken to select radiographs with more specific and clearly delineated findings. Second, comparison of the accuracy levels assigned by the two evaluating radiologists (OCR_x and OCR_y) indicated that a statistically significant difference existed between them. Based on this comparison, duplicate statistical testing was performed for OCR_x and OCR_y data. It should be noted that although OCR_x and OCR_y were statistically different, identical accepted/rejected results were obtained at the 0.05 significance level for all statistical tests involving OCR data.

5.2 CONFIDENCE LEVELS

Two types of confidence level data -- FCLs and DCLs -- were collected during RMDS testing. The distinction between these two kinds of data stems from the type of information used to establish the subjective confidence levels. The Findings Confidence Levels were based strictly on the visual information provided by the radiographic images, whereas the Diagnostic Confidence Levels were based upon visual and patient history information as well as other factors. Three additional factors could have produced variations in the radiologists' FCL and DCL responses:

a. Variations in each individual's training and experience.

b. Biases in each individual's method of assigning confidence levels.

c. Variations in overall personal confidence and attitude towards the testing.

Of these variables, only the first can be accounted for to any reasonable degree. Although the two groups (Test Group and Control Group) were balanced in terms of experience, a balance within each group in terms of experience was not attempted. A comparison of the FCL per radiologist and the DCL per radiologist for the LB mode indicated that an intrasample difference did indeed exist for both FCL and DCL values. In keeping with these above listed factors (a, b, and c) and the comparison of FCL and DCL results for the LB radiologists, limited emphasis should be placed on the absolute numerical values obtained for confidence levels.

5.2.1 Findings Confidence Levels

As shown in figure 13, FCL values correspond to the same hierarchical trend as the OCR measurements, in which the order from high to low confidence levels is arranged as follows: LB > CCTV > DHF > AHF > DHC > AHC > ALF > ALC. Pearson's testing indicated a high degree of correlation between OCR and FCL values, ie high OCR was associated with high FCL and low OCR was associated with low FCL.

An analysis comparing the three major viewing groups (LB, CCTV, and RMDS) showed (1) a statistical difference among the FCLs of the radiographic viewing groups and (2) a distinct decrease in the confidence of the radiologists proportional to the decrease in information obtained from the radiographs. As expected, the LB provided the radiologists with the most ideal conditions, with only variations in radiographic quality and pathologic disorder affecting confidence levels. Both the CCTV and RMDS viewing modes showed a decrease in confidence levels, with the larger decrease occurring for the RMDS Test Group.

A comparison of the FCL per Difficulty Ranking Factor (DRF) was performed on the CCTV and LB groups. This testing was done to determine whether radiographic difficulty affected the confidence levels of both groups equally. The results of this testing indicated that of the six DRF rankings (1 to 6), only the radiographs rated as 2 (low/medium low) were viewed with equal confidence for the LB and CCTV.

The statistical analysis comparing FCL for the six RMDS transmission modes indicated that there was no difference in FCLs between transmission modes. Nonetheless, the same general trends observed for OCR measurements were seen for FCL, ie digital transmission mode superior to analog, fine resolution superior to coarse, and high SNR superior to low.

Additional analysis was performed on FCL values to determine whether any differences existed among the six RMDS transmission modes according to pathology type (ie, appendage, abdomen, chest, and skull). For the pathology types tested, no statistical difference was found between the six transmission modes with respect to FCL.

5.2.2 Diagnostic Confidence Levels

In general, DCL values ranked lower than FCL values. Like the FCL scores, DCL values were statistically different for the three major viewing groups: confidence levels for the LB were higher than for CCTV, and those for the CCTV were higher than for the RMDS transmission modes. A statistical analysis of the six RMDS transmission modes indicated that there was a difference between modes: the digital, high-SNR modes provided higher confidence levels than the analog, low-SNR modes. These general trends were observed for DCL (figure 13): digital ranks better than analog, high SNR ranks better than low, and coarse resolution ranks better than fine. The fact that the coarse resolution mode tended to have higher DCL ratings than the fine resolution mode conflicts with the data obtained for FCL. Findings are an integral part of a diagnosis, and FCL data showed a definite trend for fine resolution to be better than coarse. Table 10 supports this conflict and shows that a larger number of "poor quality" remarks were provided for the coarse modes when the other two variables (digital/analog and high/low SNR) were constant. DCL data do show a positive correlation with OCR, but the degree of correlation is decreased from that seen for FCL values.

A comparison of DCL per Difficulty Ranking Factor (DRF) was performed on the CCTV and LB groups. The results of this testing indicated that those radiographs rated as 2 and 5, on a scale running from 1 (low) to 6 (high), were equal in terms of DCL for the LB and CCTV groups. An analysis was performed on the DCL values to determine whether there were any statistical differences between the six RMDS transmission modes by pathology type (appendage, abdomen, chest, and skull). There were found to be differences among the transmission modes for DCLs of the appendage pathology type. Statistically lower DCL values were obtained for the analog, low-SNR transmission modes in this analysis. No differences among transmission modes were observed for abdomen, chest, or skull radiograph types.

5.3 RADIOLOGIST SYSTEM CONTROLS

The TV monitor brightness and contrast control features of the RDMS terminals were evaluated during the radiographic testing. Since data obtained for brightness and contrast variations did not readily lend themselves to a statistical analysis, only qualitative evaluations are provided herein. In the beginning stage of each radiographic transmission, e`ch radiologist would experiment and vary the brightness and contrast controls over a wide range before settling at a particular level. Occasionally, a radiologist would readjust the contrast and/or brightness controls as the transmission progressed. Both brightness and contrast levels varied from radiograph to radiograph per radiologist, and from radiologist to radiologist per radiograph, with no observable patterns formed. Variations in the final contrast level selected per radiologist appeared greater than variations in brightness, although no single brightness level was consistently used throughout the testing.

Four references to contrast were supplied by certain radiologists in the remarks section of the radiographic evaluation sheets. Three of the radiologists expressed difficulties with the contrast control, while one remarked that contrast control adjustments aided in bone density viewing. In addition, two references to contrast were made in the radiologist questionnaires, both related to difficulties (lack of contrast) with the control. Only one of these questionnaire comments duplicated a remark expressed on a radiograph evaluation sheet; the other was expressed by a radiologist not represented by remarks on the radiographic comment sheet.

No remarks were provided on either the radiographic evaluation sheet or the questionnaire relating to brightness controls. It appears that radiologist brightness and contrast controls are desirable features for an RMDS terminal, although no results can be stated on the extent to which contrast and brightness variations affect clinical accuracy and confidence.

5.4 RMDS SYSTEM CAPABILITIES

Data were collected to aid in evaluating the desirability of five prospective RMDS system capabilities:

a. Zoom

b. Video storage

c. Reverse polarity

d. Disc memory

e. Simultaneous image viewing

Both groups of radiologists (Test and Control) had the opportunity to use zoom controls, whereas only the Test Group could employ the video storage and reverse polarity features. Neither group actually had the use of a disc memory system or simultaneous image viewing, but both groups were asked if they felt these systems would be useful. Qualitative data were collected for each of these features, and quantitative data were acquired and analyzed for the zoom feature. The sections that follow discuss the results of these analyses.

5.4.1 Zoom Capability

As seen in table 7 and figure 14, the use of zoom depended on two factors: the radiologist, and the sequence of radiographic set presentation. Figure 14 shows the effects of sequence on zoom use. The relationship between the number of zooms used and the presentation sequence is probably attributable to a trial or learning process for the radiologists. The fact that zoom usage did not stop entirely indicated its clinical usefulness for some radiographic cases. Ten of the twelve radiologists indicated in their questionnaires that the zoom capability was indeed useful, at least for selected cases. Figure 14 indicates that there was no relationship between zoom usage and RMDS transmission mode. A statistical analysis of the data in figure 14 supports the hypothesis of no difference between modes.

Figure 15 shows the number of zoom usages per RMDS transmission mode per OCR. Further statistical analysis was performed to determine whether the OCRs were affected by zoom usage. The results of this analysis indicated that there was no statistical relationship between OCR and zoom usage. This result must be used with caution, since zoom usage was not a controlled factor. Thus, the radiologists' use of zoom may have been only to confirm or increase their own confidence in their finding and diagnosing, a possibility attested to by their own comments (see sections 4.1.2 and 4.1.3). Although no differences were found between the OCRs with or without the use of zoom, the radiologists' questionnaire responses would tend to support the incorporation of zoom capability.

5.4.2 Video Storage Capability

Radiologists in the Test Group had the capability to store an image in "video storage" while they were viewing a second (zoom) image. In the questionnaire, the radiologists were asked (1) whether they had used this capability and (2) whether they thought it was useful. All six of the Test Group radiologists indicated that they had used this feature and stated that it was indeed helpful to have the original radiograph for reference. These responses indicated that a video image storage capability should be incorporated in future RMDS terminals used by the radiologist.

5.4.3 Reverse Polarity Capability

Radiologists in the Test Group had the opportunity to use the RMDS reverse polarity feature, which reverses the black and white aspects of an image. All six radiologists felt that reverse polarity was not useful. The responses of the Test Group radiologists indicated that the reverse polarity feature of the RMDS need not be incorporated into future models.

5.4.4 Disc Memory Capability

All twelve radiologists (Control and Test Groups) were asked whether they felt a disc memory system would be useful for later consultation of some images. Seven radiologists felt that it would be useful. Two felt that it might be useful depending on the quality of the stored image. Three radiologists offered negative responses. The feasibility of incorporating a disc memory system appears to be dependent on (1) the quality of the stored image and (2) the number of cases in which later consultation or review might be required. If the RMDS terminal is to be used primarily in emergency cases, then later consultation may not be a frequent occurrence. For those cases where later consultation is required, retransmission of the radiograph may be sufficient.

5.4.5 Simultaneous Image Viewing

All twelve radiologists were asked whether they felt that one image or two simultaneous images would be sufficient for an RMDS terminal. Five radiologists felt that two images would be preferred. Five felt that one would suffice. Four of the five radiologists who felt that one image would be sufficient qualified their responses by indicating that some form of image retrieval would be helpful or needed. One radiologist responded to the question with "yes," and one radiologist's response was "no." The responses to this question indicated that if simultaneous image viewing is not provided, the RMDS should have the capacity of storing at least one image while another image is being viewed.

SECTION 6

CONCLUSIONS

The major conclusions associated with the results of the RMDS radiology testing are as follows:

1. An RMDS with the parametric capabilities of the equipment used in this test would provide for satisfactory radiological consultations for emergency cases with gross pathological disorders. This conclusion is based primarily upon two factors. First, the average accuracy (OCR) of the radiologists' results using the DHF, AHF, and CCTV modes was statistically the same; these were evaluated as midway between marginal and acceptable. Second, the opinion of most of the radiologists was that the RMDS was satisfactory for radiological consultations for emergency cases.

2. For subtle findings, such as skull lesions, bone disease and tumors, subtle (fine) fractures, small or soft tissue calcifications, and low contrast changes in soft tissue areas (eg, pneumothorax), difficulties may arise in using an RMDS having the parametric capabilities of the equipment used in this test. Each of these subtle findings would be masked by either a low signal-to-noise ratio in the system or lower system resolution than the spatial frequency of the pathology generated. The key issue is the extent to which the RMDS would be employed to support diagnosis of these types of pathologies. This issue must be resolved by a separate analysis effort. If the RMDS is to be required to support diagnosis of the more subtle pathologies, then the system must be redesigned to have increased resolution and higher SNR than that of the equipment used in this testing (ref 6).

3. The RMDS quantization level of 6 bits per picture element (digital) and the frozen noise of the received images (digital and analog) lead to diagnostic accuracies equivalent to CCTV "pure" analog images of equal resolution. Thus, a quantization level of 6 bits per picture element should be satisfactory for RMDS consultation on emergency cases. More subtle cases may require 8 bits per picture element.

4. In the testing conducted, no single RMDS transmission mode was associated with a statistically significant increase in diagnostic accuracy. The general trend of the raw data resulted in ranking the modes, from best to worst, as: DHF, AHF, DHC, AHC, ALF, and ALC. These raw data showed the DHF mode to be nearly equivalent to the CCTV mode, on the average.

5. Other results determined from the statistical analysis are as follows:

   a. The accuracies of the diagnostic results were equivalent regardless of whether or not the zoom capability was used.

   b. Under the most ideal conditions (lightbox), significant differences were observed in the Findings Confidence Levels and Diagnostic Confidence Levels among the radiologists.

   c. A higher correlation was observed between accuracy measurements and Findings Confidence Levels than between accuracy measurements and Diagnostic Confidence Levels.

d. Findings Confidence Levels and Diagnostic Confidence Levels were significantly affected by the radiographic viewing mode (lightbox > CCTV > RMDS transmissions).

e. Findings Confidence Levels were not significantly different between RMDS transmission modes or among the pathological disorders, and they were only slightly affected by radiographic Difficulty Ranking Factors.

f. Diagnostic Confidence Levels were only slightly affected by radiographic Difficulty Ranking Factors, whereas significant differences were observed between RMDS transmission modes (digital, high-SNR > analog, low-SNR) and among the pathological disorders.

SECTION 7

RECOMMENDATIONS

The following recommendations stem from the RMDS radiology testing reported herein:

1. The next evolution of RMDS, the Engineering Development Models (EDMs), should provide two video resolution levels: a high resolution (512 x 512 x 8 bits) for difficult cases involving more subtle findings, and a lower resolution level (512 x 512 x 6 bits) which will be sufficient for most emergency cases.

2. In future RMDS EDMs, provide the radiologist with:

   a. Contrast controls

   b. Brightness controls

   c. Zoom capabilities

   d. Video storage system with the capability of storing two images and displaying either one image or two simultaneous images

3. To correlate operational requirements to system specifications, perform an analysis of the projected types of pathological cases for which the RMDS would be expected to be used.

4. To minimize the types of difficulties experienced in analyzing these testing results and to expand the quality and quantity of test information obtainable, incorporate the following improvements in future testing of a similar nature:

   a. Addition of nonpathological radiographs for control

   b. Radiographs with single or limited findings

   c. More rigid grading procedures

   d. Continuous range of radiographic difficulties

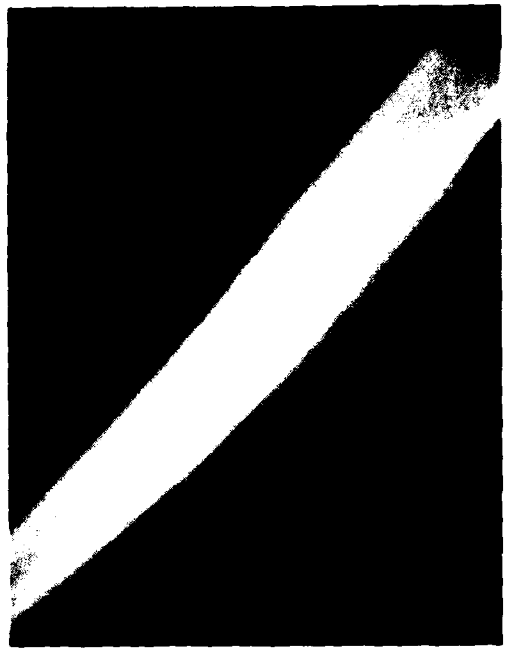
APPENDIX A: TEST RADIOGRAPHS

This appendix contains photographs of all 36 radiographic images used in the testing. Each radiograph is identified by its set letter (A, B, C, etc) and its order within the set (1, 2, 3, etc). There are two photographs for each radiograph. The photograph on the left shows the radiograph as it was normally viewed, while the photograph on the right shows a close-up image of the radiograph. Accompanying each photograph of a radiographic image are the history and the primary (P) and secondary (S) findings and diagnoses. Table A1 (repeated from table 2, section 2.4) shows the Difficulty Ranking Factor (DRF) and type (zone) of each radiograph.

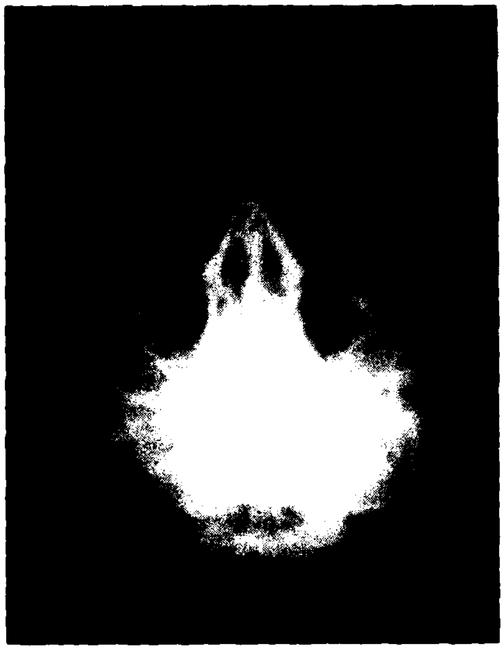
Set	Order of Presentation	DRF*	Type (Zone)	Diagnosis
A	1	5	Appendage	Soft tissue hemangioma
	2	6	Skull	Fractured mandible
	3	3	Chest	(R) LL pneumonia
	4	2	Skull	Double floor of sella
	5	4	Abdomen	Prostatic calculi
	6	1	Abdomen	Bilateral adrenal calcification
B	1	1	Skull	Broken nose
	2	3	Appendage	Osteoid osteoma
	3	6	Chest	ASD with 4:1 shunt
	4	2	Chest	Alveolar cell calcification
	5	4	Skull	Intracranial air & fracture
	6	5	Appendage	Fx neck of femur on (R)
C	1	4	Chest	Calcified mitral annulus
	2	2	Chest	*Cocci*
	3	5	Skull	Parietal skull fracture
	4	1	Abdomen	Abdominal aortic aneurysm
	5	3	Appendage	Avascular necrosis of lunate
	6	6	Skull	Nasal spine Fx
D	1	4	Chest	Pancoast tumor
	2	5	Skull	Multiple myeloma
	3	2	Abdomen	Air under (R) diaphragm
	4	6	Appendage	Chondrocalcinosis
	5	1	Appendage	Fibrous cortical defect
	6	3	Chest	Pericardial calcification
E	1	3	Skull	Calvarial hemangioma
	2	6	Abdomen	Splenomegaly
	3	4	Appendage	Tibial stress Fx
	4	1	Abdomen	Osteitis condensans ilii
	5	2	Skull	Mucocoele (L) frontal sinus
	6	5	Chest	Histiocytosis-x
F	1	3	Abdomen	Abdom calcification (post traumatic splenic cyst)
	2	2	Chest	Infectious spondylitis TB
	3	4	Appendage	Cocci osteomyelitis
	4	5	Chest	Calcified myocardial infarct
	5	6	Chest	Pneumothorax on (R)
	6	1	Skull	Enlarged sella

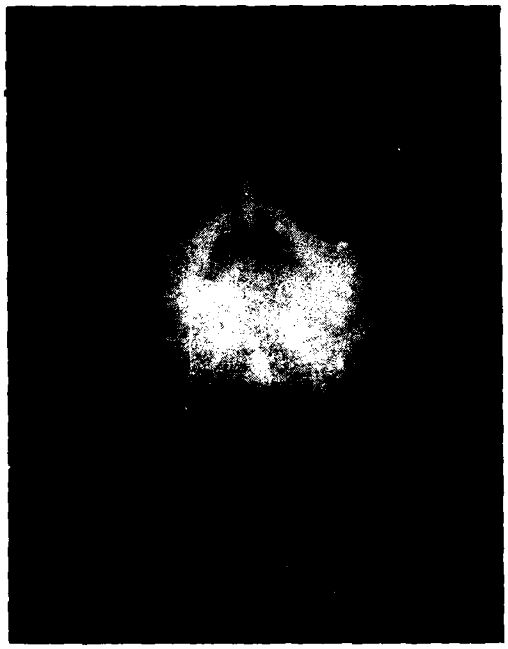
*DRF = Difficulty Ranking Factor: 1 (low), 2 (low/med low), 3 (med low), 4 (med high), 5 (med high/high), 6 (high)

Table A1. Six sets of test radiographs.

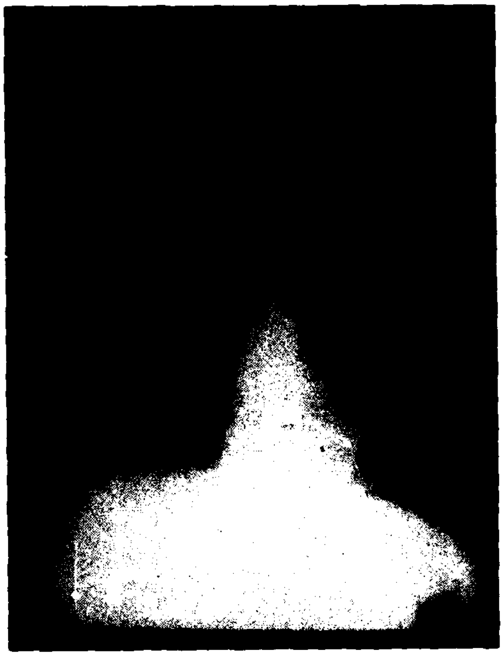


RADIOGRAPH:	A-1
HISTORY:	20-yr. old, female, with painless swelling of right arm
FINDINGS:	Phlebolith type S.T. calcifications (P)
	Soft tissue hypertrophy (P)
DIAGNOSES:	Hemangioma of soft tissue (P)
	Vascular soft tissue neoplasm (S)





RADIOGRAPH:	A-2
HISTORY:	Facial trauma, difficulty getting open-mouth view for C-spine series
FINDINGS:	Fx - angle of L mandible (P)
	Fx - mid portion body of R mandible with offset (P)
DIAGNOSES:	Fx - angle of L mandible (P)
	Fx - R mandible (P)

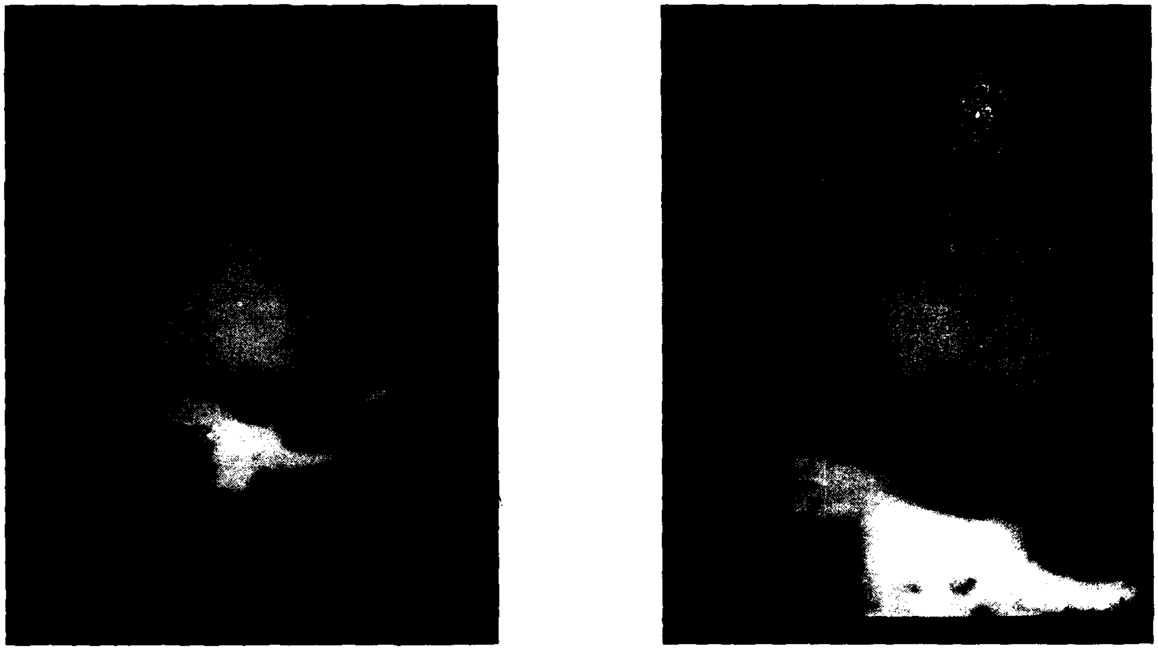


RADIOGRAPH: A-3

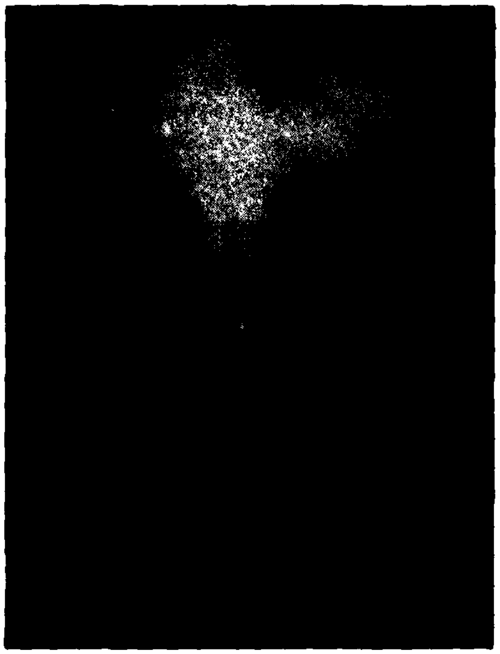
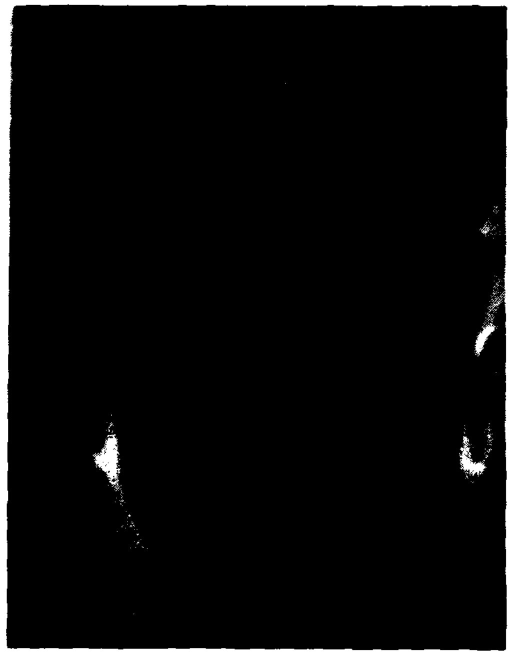
HISTORY: 3-day history of cough and fever

FINDINGS: ↑ density R lower lung silhouetting R diaphragm (P)

DIAGNOSES: R lower lobe pneumonia (P)



RADIOGRAPH:	A-4
HISTORY:	40-year old, female; galactorrhea and amenorrhea
FINDINGS:	Double floor in sella turcica (P)
	Calcifications superior and posterior to sella (S)
DIAGNOSES:	Pituitary adenoma (P)
	Pineal calcification (S)

RADIOGRAPH:	A-5
HISTORY:	Hematuria; scout film of IVP
FINDINGS:	Calcifications in pelvis (P)
	Scoliosis (S)
	Renal margin indistinct L upper pole (S)
	R renal outline indistinct (S)
DIAGNOSES:	Prostatic calculi (P)
	Scoliosis (S)
	Cannot rule out upper pole mass (S)
	Cannot evaluate R kidney (S)

RADIOGRAPH: A-6

HISTORY: 24-year old, male; 3-day history of mid-abdominal pain

FINDINGS: Suprarenal calcifications (P)

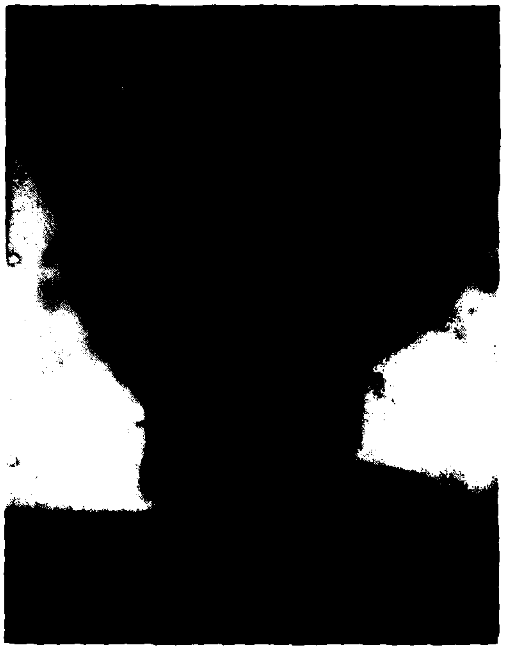
Amorphous and round calcifications L abdomen (S)

Calcific density in RUQ (S)

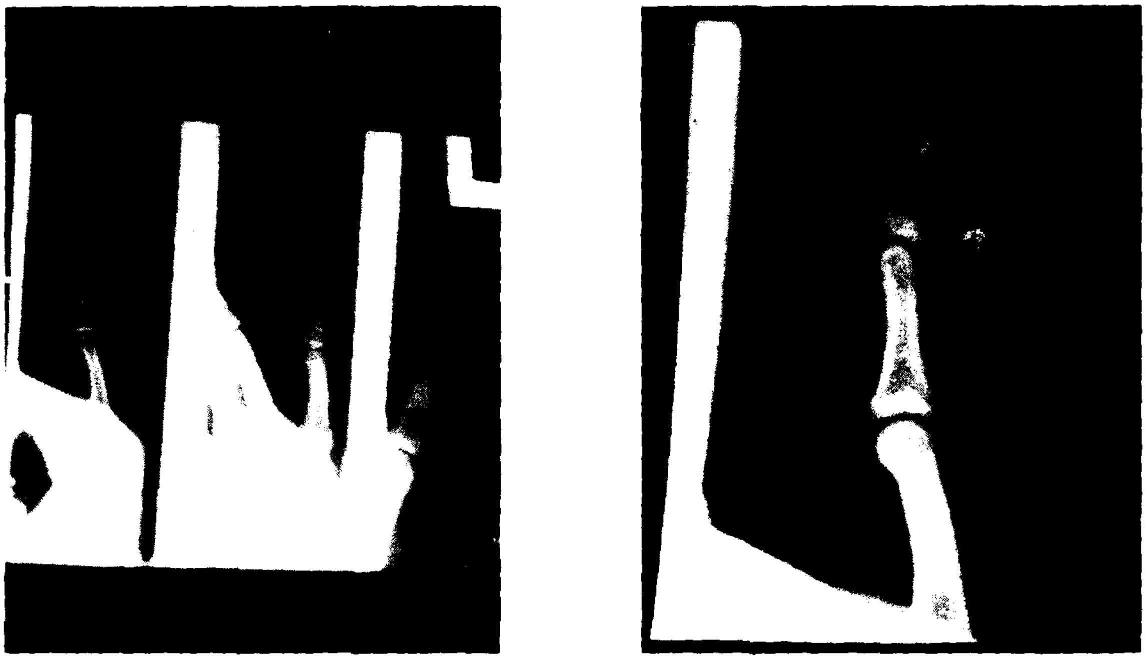
DIAGNOSES: Adrenal calcifications (P)

Etiology not determinable from film (S)

Artifact (S)



RADIOGRAPH:	**B-1**
HISTORY:	**Trauma to nose**
FINDINGS:	**FX - nose (P)**
DIAGNOSES:	**Fx - nose (nondisplaced) (P)**



RADIOGRAPH: B-2

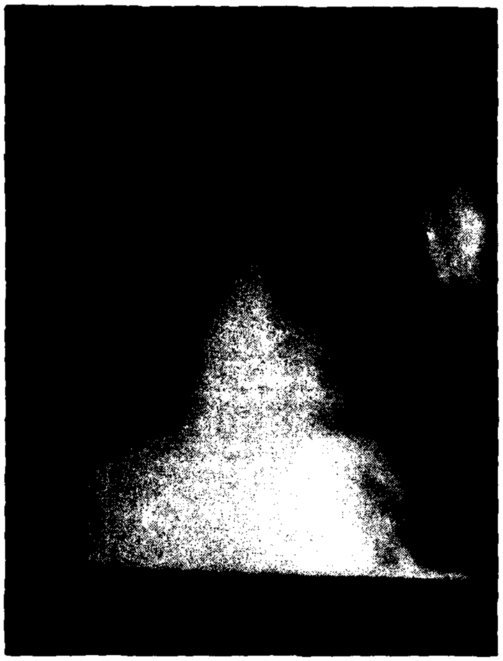
HISTORY: Pain in distal portion index finger, with no previous trauma

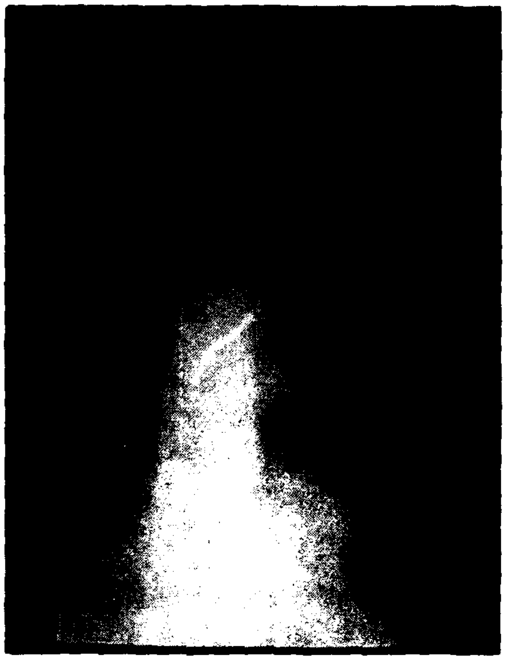
FINDINGS: Soft tissue swelling of tip of index finger (P)

Lesion of distal phalanx with dense center and lucent rim (P)

DIAGNOSES: Possibilities:

1. Palm thorn granuloma (P)
2. Epidermoid inclusion cyst (P)
3. Enchondroma (P)
4. Osteoid osteoma (P)
5. Chronic osteomyelitis (P)





RADIOGRAPH: B-3

HISTORY: 5-year old, male, with murmur and poor growth

FINDINGS: Cardiomegaly (P)

Increased pulmonary vascularity (P)

DIAGNOSES: ASD or VSD (P)

RADIOGRAPH:	B-4
HISTORY:	Weight loss, chronic cough in a 56-year old, male, with a 40 pack/year smoking history
FINDINGS:	Nodule left upper lobe (P)
	Increased lung markings and lung volume (P)
DIAGNOSES:	Bronchogenic Ca (P)
	C.O.P.D. (P)

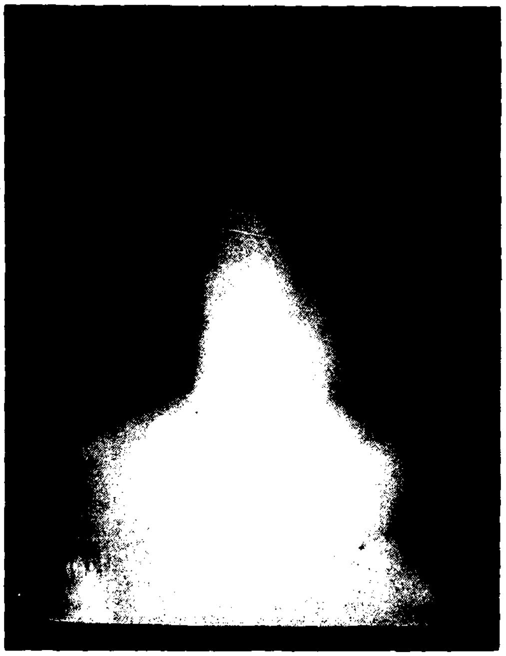
RADIOGRAPH:	B-5
HISTORY:	Motorcycle accident with head trauma
FINDINGS:	Fx - frontal area (P)
	Fx - skull base through sphenoid (P)
	Air in cranial vault (P)
DIAGNOSES:	Fx - frontal area (P)
	Basilar skull fx (P)
	Pneumocephalus (P)

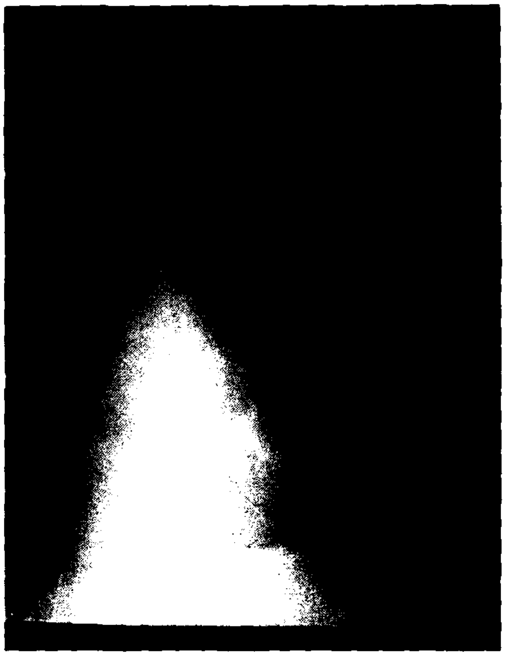
RADIOGRAPH: B-6

HISTORY: 78-year old, female; fell, now complains of right hip pain

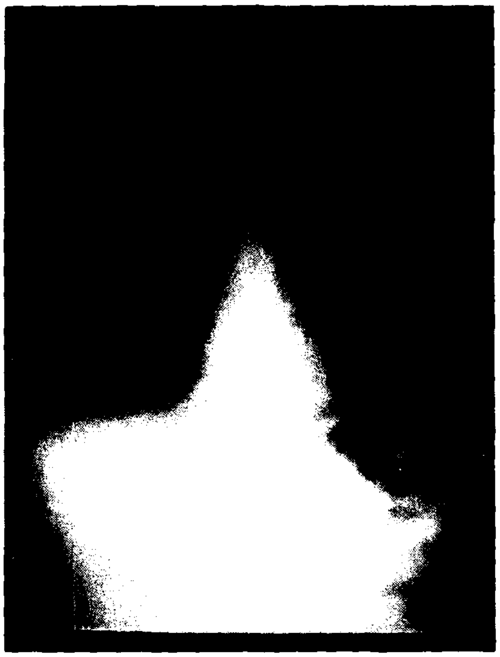
FINDINGS: Fx - femur on R (P)

DIAGNOSES: Subcapital fx femur on R (P)



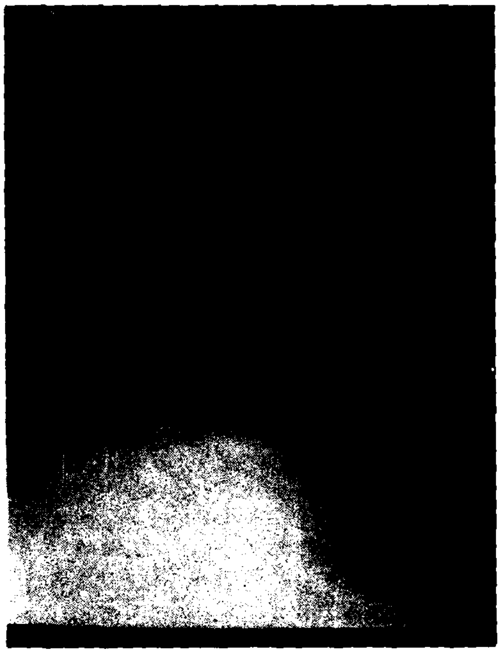


RADIOGRAPH:	C-1
HISTORY:	Routine chest film; 77-year old, female
FINDINGS:	Aortic nob calcified (P)
	Granuloma L lung, node calcifications L hilum (P)
	Cardiac valvular calcification (P)
DIAGNOSES:	Aortic nob calcification (P)
	Old granulomatous disease (P)
	Calcified valve annulus (P)

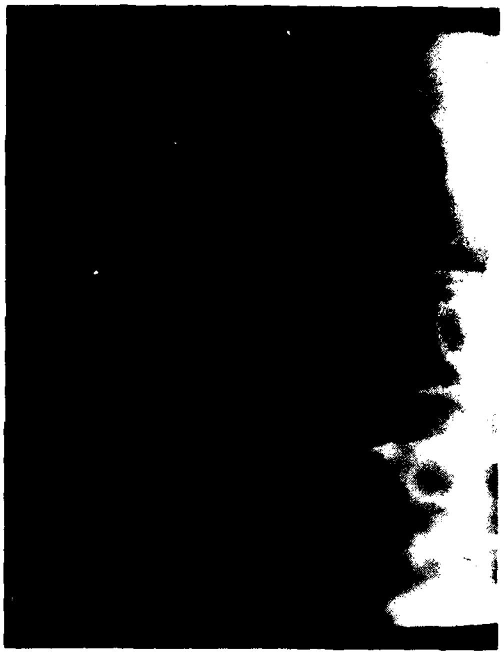
 

RADIOGRAPH:	C-2
HISTORY:	Mild upper respiratory symptoms; patient lives in San Joaquin Valley
FINDINGS:	Cavitary lesion R upper lobe with thick shaggy walls (P)
DIAGNOSES:	Possibilities:

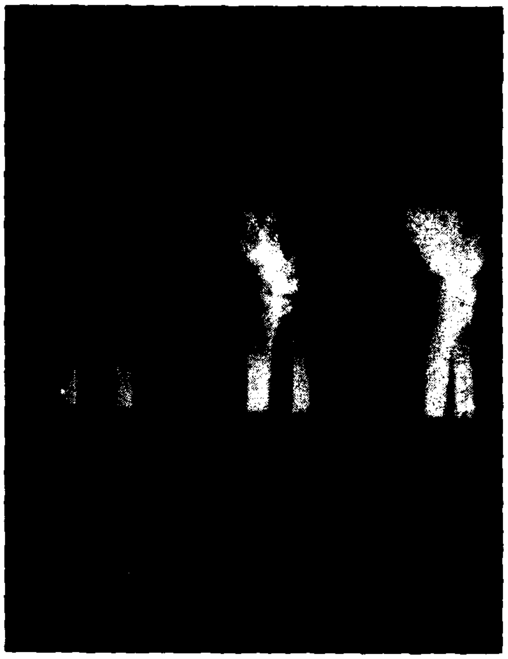
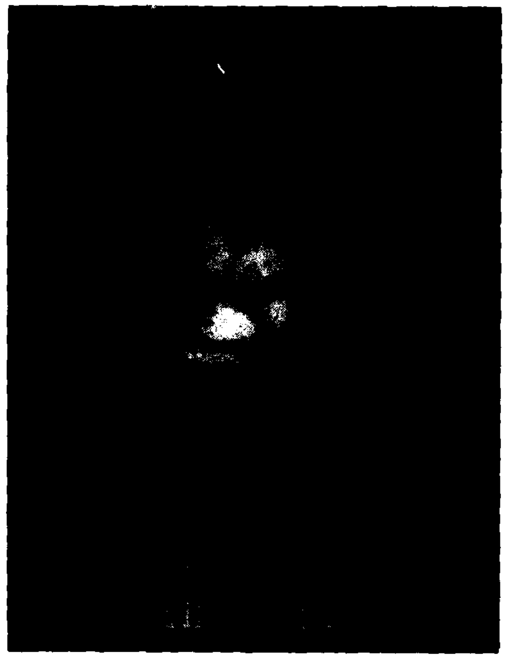
1. Coccidiodomycosis (P)
2. TB (P)
3. Other granulomatous or fungal disease (P)
4. Lung abscess (P)
5. Cavitary mass (P)



RADIOGRAPH:	C-3
HISTORY:	Patient fell from swing and hit left side of head on playground equipment
FINDINGS:	Parietal skull fx (P)
DIAGNOSES:	Parietal skull fx (P)



RADIOGRAPH:	C-4
HISTORY:	73-year old, male, with pulsatile abdominal mass and back pain
FINDINGS:	Curvilinear calcification in R abdomen (P)
	Degenerative Δs of spine (S)
	Vascular calcifications in pelvis (S)
DIAGNOSES:	Aortic aneurysm or calcification in R kidney (P)
	Degenerative spine changes (S)
	Vascular calcifications in pelvis (S)

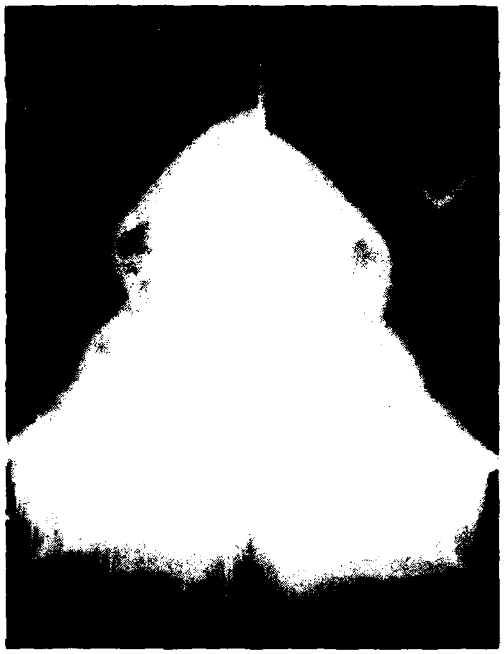
 

RADIOGRAPH: C-5

HISTORY: Trauma one year ago, persistent pain

FINDINGS: Dense lunate bone (P)

DIAGNOSES: Post-traumatic aseptic of lunate necrosis (P)

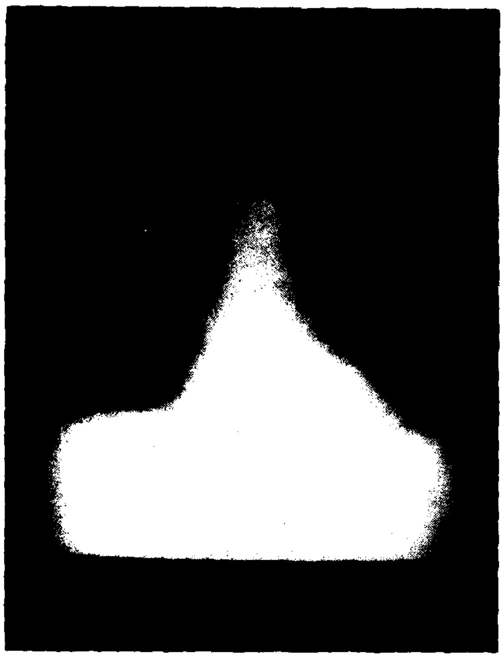
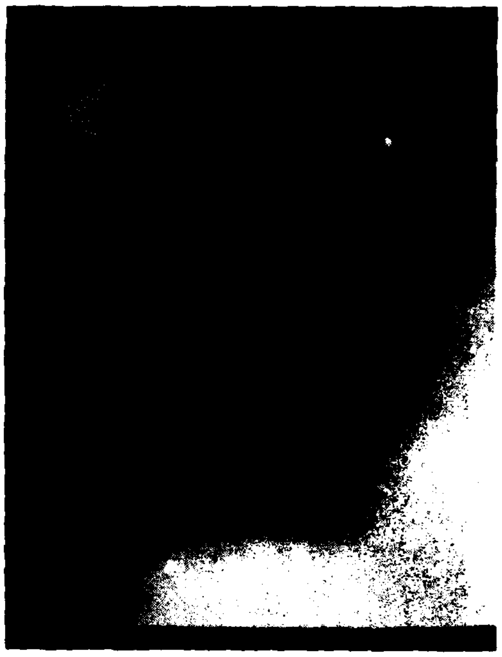


RADIOGRAPH: C-6

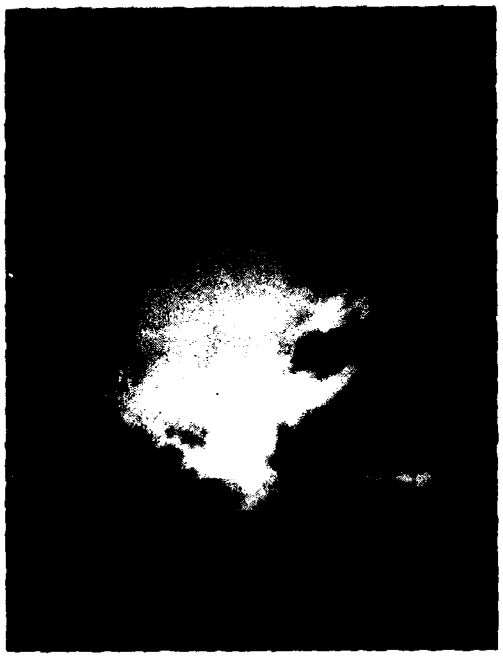
HISTORY: Trauma to nose and maxilla

FINDINGS: Fx - inferior nasal spine (P)

DIAGNOSES: Fx - inferior nasal spine (P)

RADIOGRAPH:	D-1
HISTORY:	57-year old, male, with shoulder pain (mild); seen for routine chest x-ray
FINDINGS:	Increased density R pulmonary apex (P)
DIAGNOSES:	Opacified R apex (further differential not possible w/o further studies) (P)

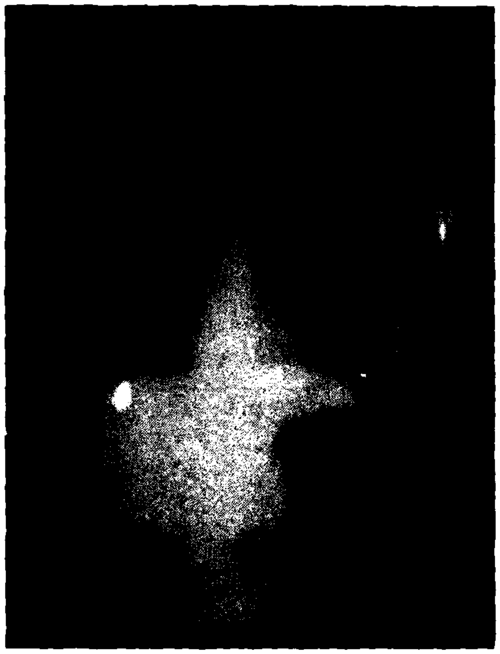


RADIOGRAPH: D-2

HISTORY: Frontal headaches for one week

FINDINGS: Multiple calvarial radiolucencies (P)

DIAGNOSES: Metastases vs multiple myeloma (P)

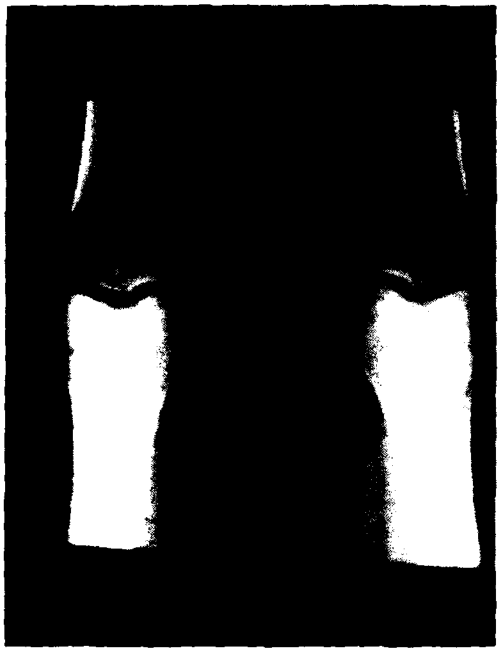
 

RADIOGRAPH: D-3

HISTORY: 23-year old, female, with acute onset of upper abdominal pain

FINDINGS: Air under diaphragm (P)

DIAGNOSES: Pneumoperitoneum (P)

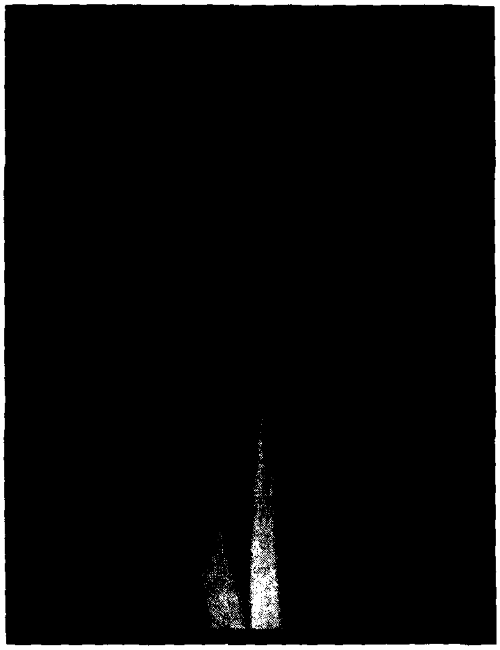
 

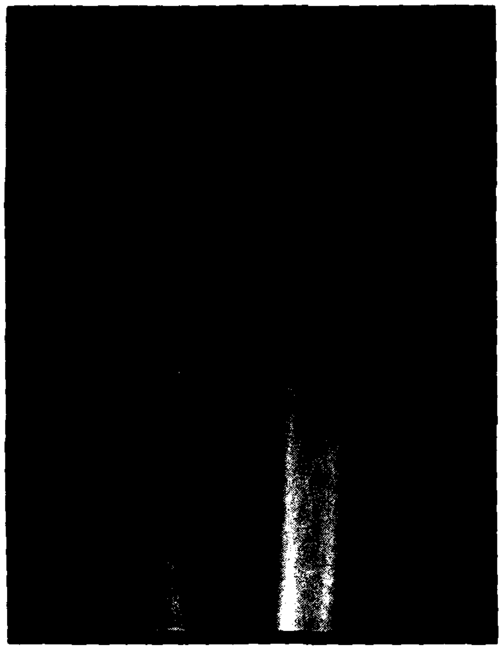
RADIOGRAPH: D-4

HISTORY: 46-year old, male, with recent onset of knee pain - bilateral

FINDINGS: Calcified articular cartilages (P)

DIAGNOSES: Chondrocalcinoisis (P)





RADIOGRAPH: D-5

HISTORY: *Minor trauma; rule out fracture*

FINDINGS: Eccentric radiolucency in distal radial diaphysis with sclerotic margins (P)

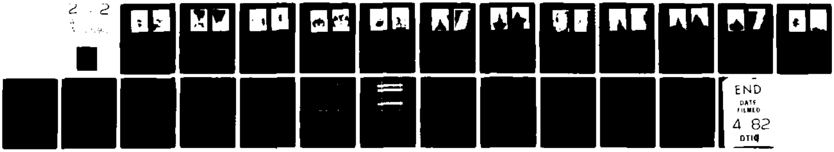
DIAGNOSES: Possibilities:

1. Fibrous cortical defect or fibroxanthoma (P)
2. Enchondroma (P)
3. Fibrous dysplasia (P)
4. Chondromyxoid fibroma (S)
5. Fibrosarcoma (S)

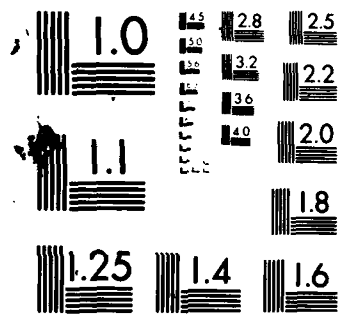
RADIOGRAPH:	D-6
HISTORY:	Asymptomatic 48-year old, female
FINDINGS:	Calcified pericardium (P)
DIAGNOSES:	TB pericarditis (P)
	Calcified aneurysm (S)
	Teratoma or dermoid cyst (S)

AD-A112 090 NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA F/G 6/5
REMOTE MEDICAL DIAGNOSIS SYSTEM (RMDS) ADVANCED DEVELOPMENT MOD--ETC(U)
DEC 81 W T RASMUSSEN, P D HAYES, F H GERBER
UNCLASSIFIED NOSC/TR-683 NL

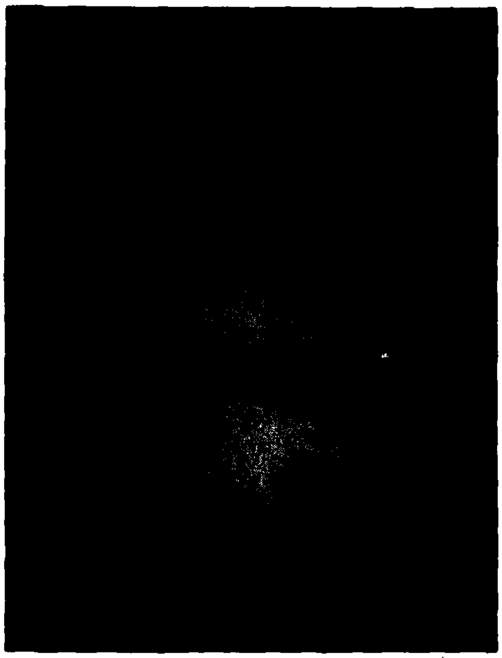
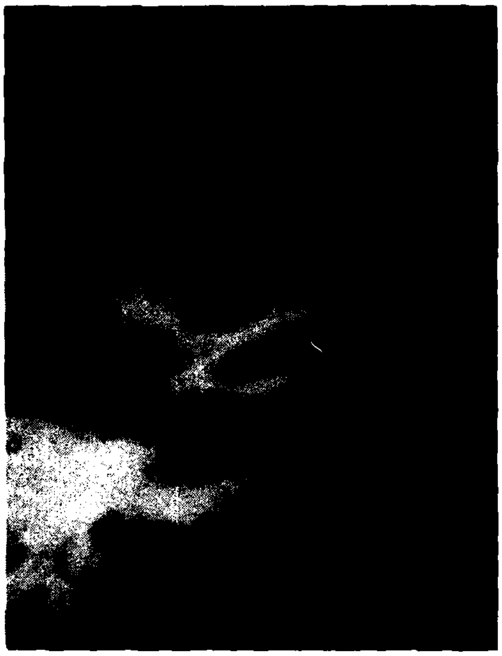
2 - 2



END
DATE
FILMED
4 82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

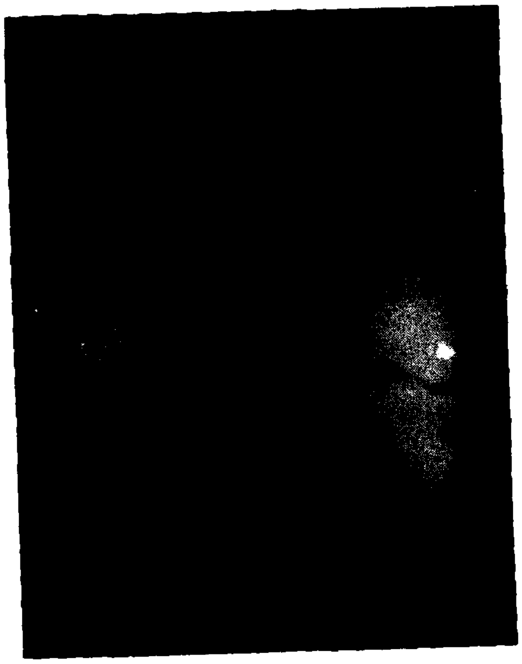
RADIOGRAPH:	E-1
HISTORY:	3-year history of occipital headaches
FINDINGS:	Lyticarea frontal area with stippled appearance (P)
	Pineal calcification (S)
	Petioclinoid ligament calcification (S)
	Cavity lower molar (S)
DIAGNOSES:	Hemangioma (P)
	Pineal calcification (S)
	Petioclinoid ligament calcification (S)
	Cavity lower molar (S)

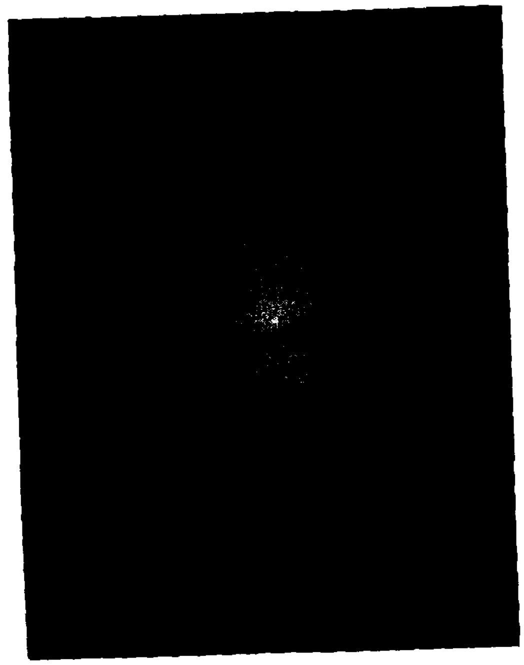
RADIOGRAPH: E-2

HISTORY: 56-year old, male, with weight loss, fatigue, intermittent left abdominal discomfort

FINDINGS: Mass LUQ (P)

DIAGNOSES: Spenomegaly (P)





RADIOGRAPH: E-3

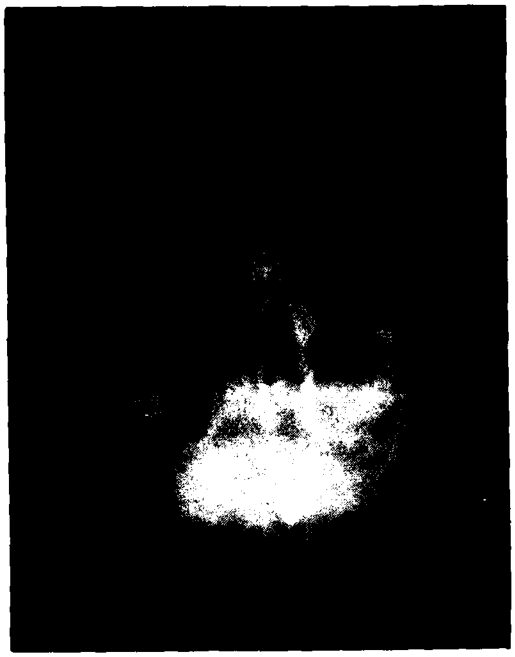
HISTORY: Rule out stress fracture

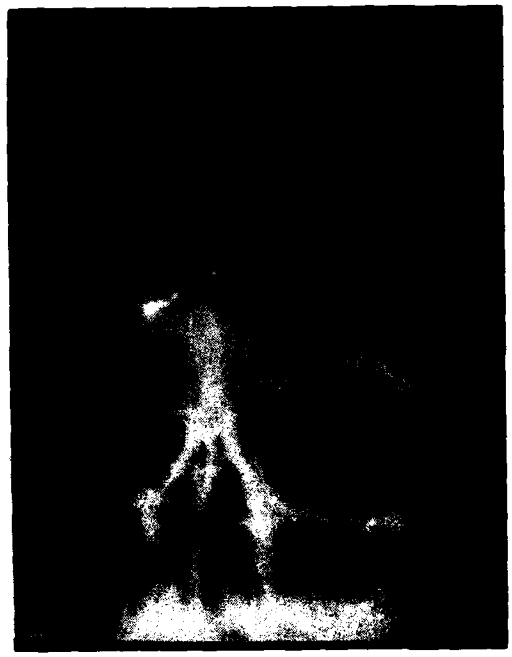
FINDINGS: Healing fx R tibia (P)

DIAGNOSES: Stress fx R tibia (P)

RADIOGRAPH:	E-4
HISTORY:	Low back pain in a 27-year old, obese, female
FINDINGS:	Sclerosis R iliac at SI joint (P)
	Sclerosis R pubis (P)
DIAGNOSES:	*Possibilities:*
	1. Osteitis condensans iliac (P)
	2. Sacroilaetis (P)
	3. Osteitis condensans pubis (P)





RADIOGRAPH: E-5

HISTORY: Left frontal headaches, allergic rhinitis

FINDINGS: Dense expanded L frontal sinus (P)

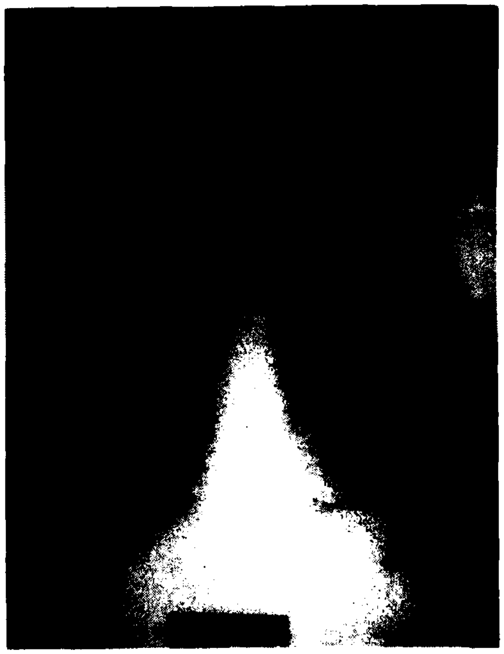
Mucosal thickening L maxillary sinus (P)

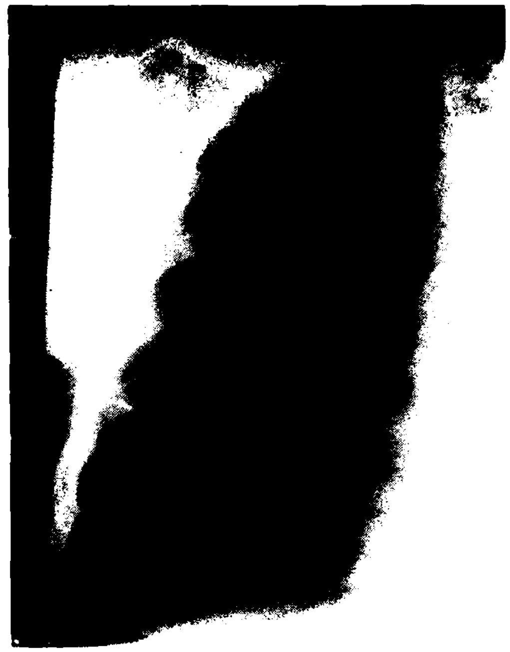
Round ↑ density R maxillary sinus (P)

DIAGNOSES: L frontal mucocoele (P)

Mucosal thickening L maxillary sinus (P)

Polyp R maxillary sinus (P)





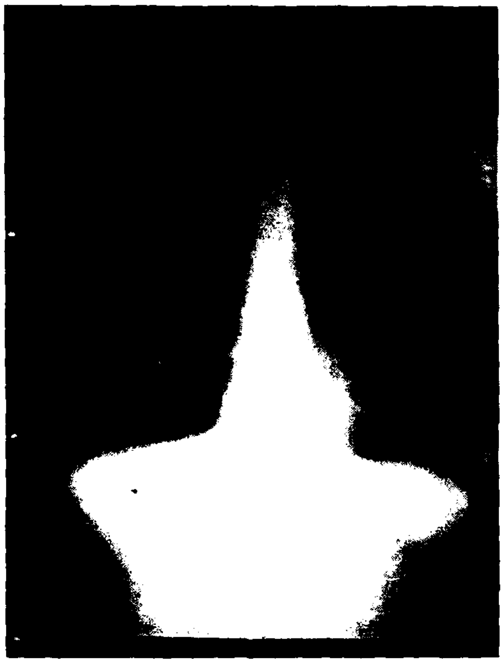
RADIOGRAPH: E-6

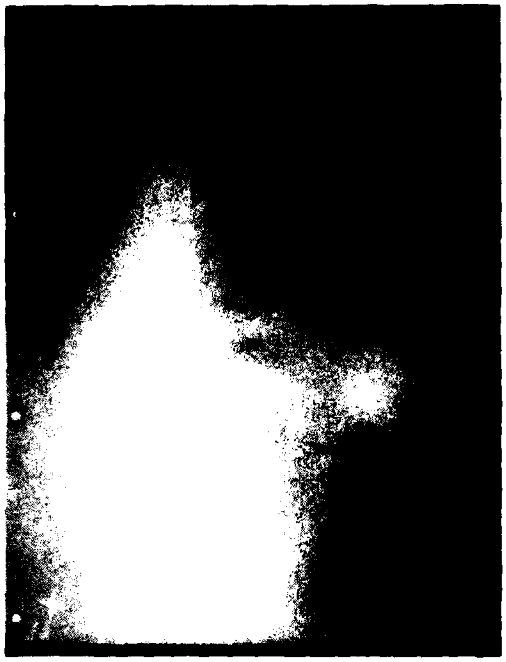
HISTORY: Asymptomatic patient in for yearly chest x-ray

FINDINGS: Bilateral interstitial lung markings (P)

DIAGNOSES: Possibilities:

1. Interstitial lung disease (P)
2. Eosinophilic granuloma (P)
3. Sarcoid (P)
4. Hamman - Rich Syndrome (P)





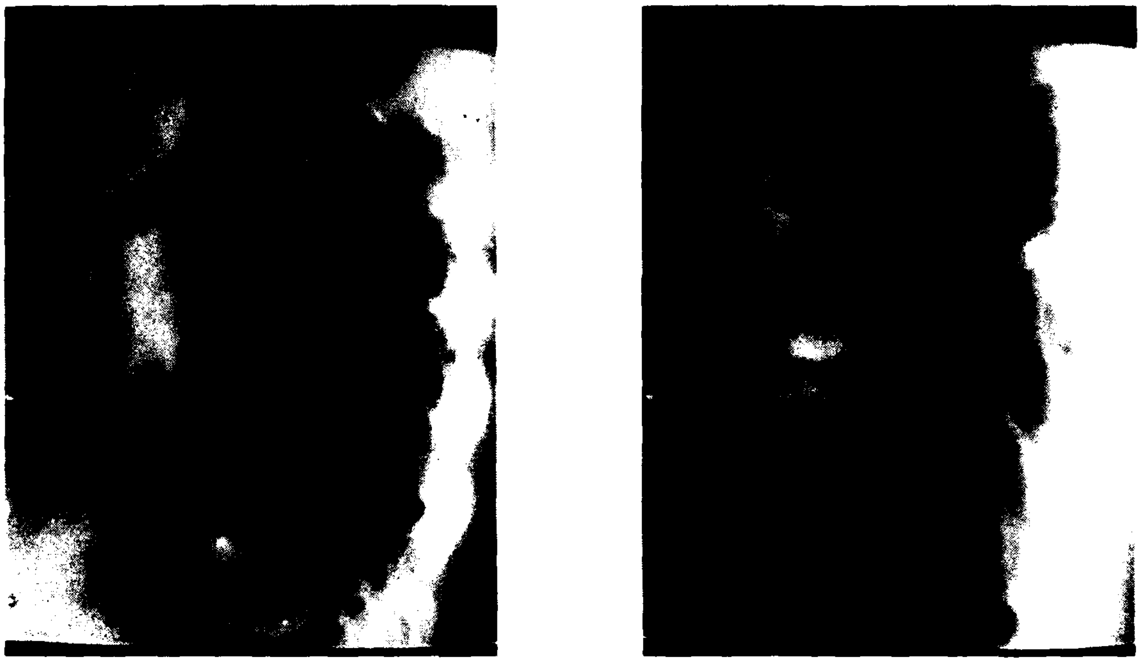
RADIOGRAPH: F-1

HISTORY: Routine chest x-ray; history of automobile wreck three years prior to date of film

FINDINGS: Abdominal curvilinear calcification (P)

DIAGNOSES: Possibilities:

1. Splenic cyst (P)
2. Echinoccocal cyst (P)
3. Abdominal curvilinear calcification (P)
4. Aneurysm (P)
5. Renal calcification (P)



RADIOGRAPH: F-2

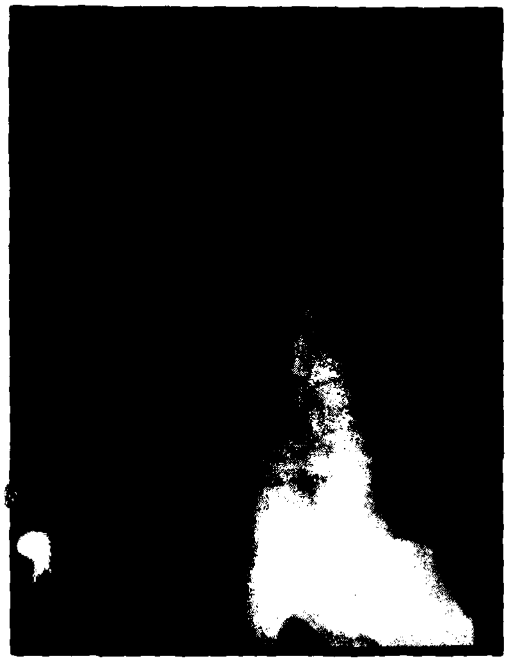
HISTORY: 34-year old, male, with back pain and positive PPD

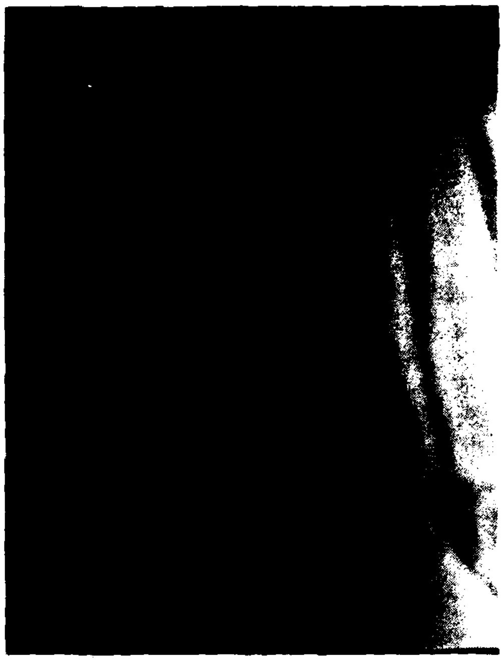
FINDINGS: Destruction of two vertebrae and intervening disc (P)

Soft tissue mass anterior to spine (P)

DIAGNOSES: Possibilities:

1. TB of spine (Pott's Disease) (P)
2. Discitis, osteomyelitis (P)





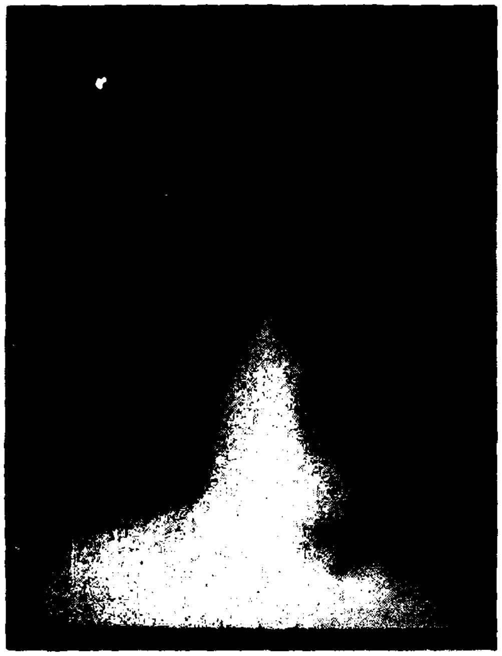
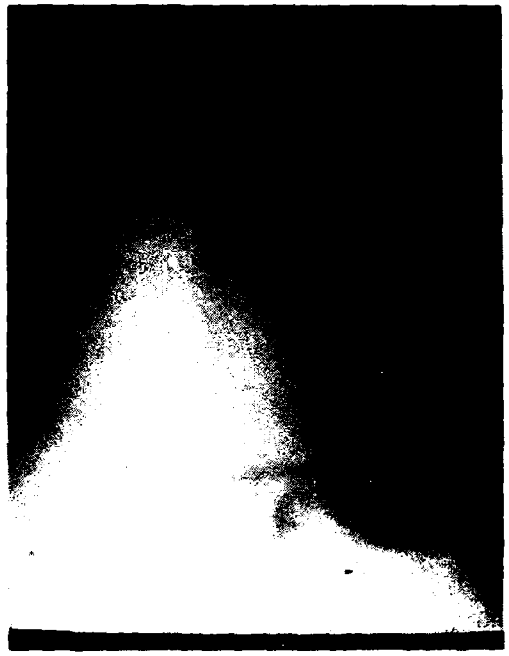
RADIOGRAPH: F-3

HISTORY: One-month history of pain in foot along lateral aspect of distal metatarsals

FINDINGS: Destructive changes in 5th metatarsal (P)

DIAGNOSES: Possibilities:

1. Osteomyelitis (P)
2. Primary or secondary bone tumor (P)
3. Osteosarcoma (P)

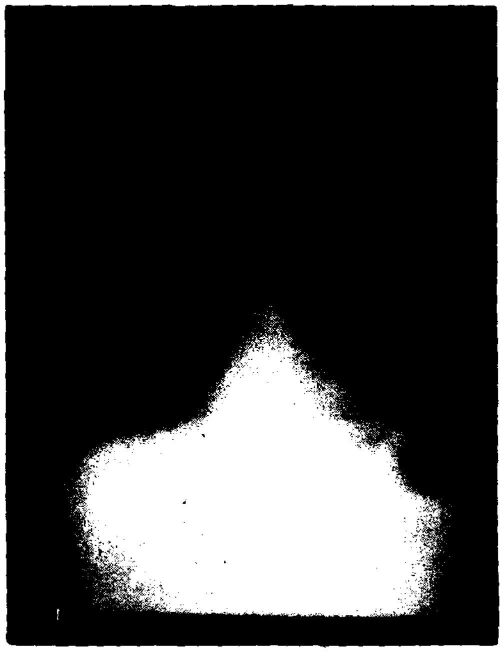
 

RADIOGRAPH: F-4

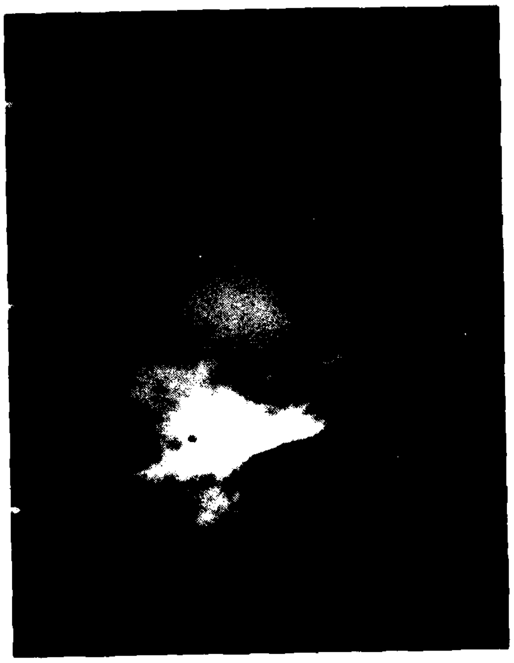
HISTORY: Yearly chest film; history of heart disease

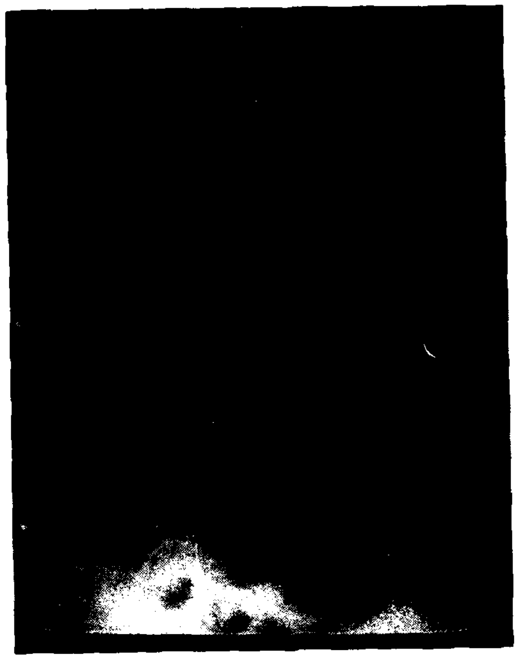
FINDINGS: Curvilinear calcification of heart border (P)

DIAGNOSES: Pericardial calcification (P)



RADIOGRAPH:	F-5
HISTORY:	Right chest pain for three hours
FINDINGS:	Pleural thickening on R (P)
	R pneumothorax (P)
DIAGNOSES:	Pleural thickening on R (P)
	R pneumothorax (P)





RADIOGRAPH: F-6

HISTORY: Increasing hat size and headaches

FINDINGS: Enlarged sella (P)

DIAGNOSES: Pituitary adenoma (P)

APPENDIX B: TEST DATA SHEETS

This appendix presents samples of all the forms used in the testing of the radiologists' ability to interpret radiographic images transmitted by RMDS equipment. A brief explanation of each form is given, in their order of presentation.

RMDS TESTS CHECK FORM; CONTROL TESTS CHECK FORM

These two forms, exhibits B1 and B2, respectively, were used during the testing procedure to insure that each radiologist subject received the radiographic image set in the proper order and under the correct experimental conditions. These forms were used by the personnel who administered the test.

RADIOLOGIST TEST DATA SHEET (TEST AND CONTROL GROUPS)

This sheet, exhibit B3, was used by each radiologist in the test to record findings, diagnoses, confidence levels, and comments for each radiograph viewed.

RADIOLOGIST TEST DATA SUMMARY SHEET (TEST GROUP); RADIOLOGIST TEST DATA SUMMARY SHEET (CONTROL GROUP)

These two summary test sheets, exhibits B4 and B5, respectively, were used by the senior radiologists to summarize, for each radiographic image, the FCL and DCL assigned by the radiologist who viewed the radiograph. Space is provided for the evaluating senior radiologist to assign an OCR for each doctor viewing the radiographic image. In the two samples shown, the actual findings and diagnoses for the particular radiograph are given.

TEST GROUP RADIOLOGY QUESTIONNAIRE; CONTROL GROUP RADIOLOGY QUESTIONNAIRE

These two questionnaires, exhibits B6 and B7, were given to each doctor of the respective group at the end of the testing. Note that the subjects were not advised as to which group they belonged to. The participating radiologists were asked to answer the questions concerning their opinion about RMDS and its features.

	1				2				3				4				5				6			
SUBJECT	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓	EXPTL. COND.	SET	BOOKLET	✓
1	I(DHF)	A	1-1		II(DHC)	B	1-2		III(AHF)	C	1-3		IV(AHC)	E	1-4		V(ALF)	D	1-5		VI(ALC)	F	1-6	
2	VI(ALC)	C	2-1		V(ALF)	E	2-2		IV(AHC)	F	2-3		III(AHF)	B	2-4		II(DHC)	A	2-5		I(DHF)	D	2-6	
3	II(DHC)	F	3-1		IV(AHC)	A	3-2		VI(ALC)	E	3-3		I(DHF)	B	3-4		III(AHF)	D	3-5		V(ALF)	C	3-6	
4	III(AHF)	A	4-1		I(DHF)	C	4-2		V(ALF)	F	4-3		VI(ALC)	D	4-4		IV(AHC)	B	4-5		II(DHC)	E	4-6	
5	IV(AHC)	C	5-1		VI(ALC)	B	5-2		II(DHC)	D	5-3		V(ALF)	A	5-4		I(DHF)	F	5-5		III(AHF)	E	5-6	
6	V(ALF)	B	6-1		III(AHF)	F	6-2		I(DHF)	E	6-3		II(DHC)	C	6-4		VI(ALC)	A	6-5		IV(AHC)	D	6-6	

<u>EXPERIMENTAL CONDITIONS:</u>

I = DHF = Digital, High SNR, Fine Resolution
II = DHC = Digital, High SNR, Coarse Resolution
III = AHF = Analog, High SNR, Fine Resolution
IV = AHC = Analog, High SNR, Coarse Resolution
V = ALF = Analog, Low SNR, Fine Resolution
VI = ALC = Analog, Low SNR, Coarse Resolution

Exhibit B1. RMDS tests check form.

SUBJECT	1 EXPTL. COND.	1 SET	1 BOOKLET		2 EXPTL. COND.	2 SET	2 BOOKLET		3 EXPTL. COND.	3 SET	3 BOOKLET		4 EXPTL. COND.	4 SET	4 BOOKLET		5 EXPTL. COND.	5 SET	5 BOOKLET		6 EXPTL. COND.	6 SET	6 BOOKLET	
7	I	A	7-1		I	B	7-2		II I	C	7-3 7-3a		I	D	7-4		II I	E	7-5 7-5a		II I	F	7-6 7-6a	
8	II I	B	8-1 8-1a		II I	D	8-2 8-2a		I	F	8-3		II I	A	8-4 8-4a		I	C	8-5		I	E	8-6	
9	I	C	9-1		II I	A	9-2 9-2a		II I	E	9-3 9-3a		I	F	9-4		II I	B	9-5 9-5a		I	D	9-6	
10	II I	D	10-1 10-1a		I	F	10-2		I	B	10-3		II I	E	10-4 10-4a		I	A	10-5		II I	C	10-6 10-6a	
11	I	E	11-1		II I	C	11-2 11-2a		I	D	11-3		II I	B	11-4 11-4a		II I	F	11-5 11-5a		I	A	11-6	
12	II I	F	12-1 12-1a		I	E	12-2		II I	A	12-3 12-3a		I	C	12-4		II I	D	12-5 12-5a		I	B	12-6	

EXPERIMENTAL CONDITIONS:

I = DIRECT [LIGHTBOX]
II = CCTV

Exhibit B2. Control tests check form.

SUBJECT NO: ______
TRIAL NO: ______
CONDITION: ______
SET: ______
RADIOGRAPH: ______
CODE: ______

HISTORY:

FINDINGS/DIAGNOSES: (Include location, extent and visual qualities of radiographic anomalies prompting the findings and diagnoses)

	CONFIDENCE LEVELS: 1=Low; 2=Fairly Low; 3=Medium; 4=Fairly High; 5=High			
	FINDINGS	CONFIDENCE LEVEL	DIAGNOSES	CONFIDENCE LEVEL
1				
2				
3				

REMARKS (Concerning the above findings/diagnoses):

Exhibit B3. Radiologist test data sheet (Test and Control Groups).

PAT	
RG	
RG#	

FINDINGS	DIAGNOSIS	CLS	TEST MODE	I	II	III	IV	V	VI
			DR#	1	2	4	3	5	6
			Z						
			NZ						
Phlebolith type S.T. calcifications	Hemangioma of soft tissue	P	FCL						
			DCL						
Soft tissue hypertrophy	Hemangioma of soft tissue	P	FCL						
			DCL						
	Vascular soft tissue neoplasm	S	FCL						
			DCL						
			FCL						
			DCL						
			FCL						
			DCL						
			FCL						
			DCL						
	OVERALL CLINICAL READING								

ABBREVIATIONS
RG = RADIOGRAPH
= NUMBER
CLS = CLASSIFICATION
FCL = FINDING CONF. LEVEL
DCL = DIAGNOSIS CONF. LEVEL
Z = ZOOM USED
NZ = NO ZOOM USED

CONFIDENCE LEVELS
1 = 0 - 20% ; LOW
2 = 21 - 40% ; FAIRLY LOW
3 = 41 - 60% ; MEDIUM
4 = 61 - 80% ; FAIRLY HIGH
5 = 81 -100% ; HIGH

CLS - CLASSIFICATION
P = PRIMARY
S = SECONDARY
T = TERTIARY

OVERALL READING
1 = ACCEPTABLE
2 = MARGINAL
3 = UNACCEPTABLE

TEST MODES					
I	II	III	IV	V	VI
DHF	DHC	AHF	AHC	ALF	ALC

Exhibit B4. Radiologist test data summary sheet (Test Group).

SET	
RG	
RG#	

FINDINGS	DIAGNOSIS	CLS	DR#	7	8		9		10	11	12	
			CONTROL MODE	I	I	II	I	II	I	I	I	II
			Z									
			NZ									
Phlebolith type S.T. calcifications	Hemangioma of soft tissue	P	FCL									
			DCL									
Soft tissue hypertrophy	Hemangioma of soft tissue	P	FCL									
			DCL									
	Vascular soft tissue neoplasm	S	FCL									
			DCL									
			FCL									
			DCL									
			FCL									
			DCL									
			FCL									
			DCL									
OVERALL CLINICAL READING												

ABBREVIATIONS
RG = RADIOGRAPH
= NUMBER
CLS = CLASSIFICATION
FCL = FINDING CONF. LEVEL
DCL = DIAGNOSIS CONF. LEVEL
Z = ZOOM USED
NZ = NO ZOOM USED

CONFIDENCE LEVELS
1 = 0 - 20% ; LOW
2 = 21 - 40% ; FAIRLY LOW
3 = 41 - 60% ; MEDIUM
4 = 61 - 80% ; FAIRLY HIGH
5 = 81 -100% ; HIGH

CLS - CLASSIFICATION
P = PRIMARY
S = SECONDARY
T = TERTIARY

OVERALL READING
1 = ACCEPTABLE
2 = MARGINAL
3 = UNACCEPTABLE

CONTROL MODES	
I	II
LIGHT BOX	LIVE TV

Exhibit B5. Radiologist test data summary sheet (Control Group).

RADIOLOGY QUESTIONNAIRE

You have been involved in a test of the Remote Medical Diagnosis System (RMDS) to determine its effectiveness in transmitting x-ray images. Different transmitting modes and resolution levels have been used, which accounts for the differences in image quality between the various sets of cases. Much of the time required between transmission of each image has been due to the data collection procedure for these tests, and should not bias your opinion of the system's usefulness. Keep in mind that the system is intended to provide an emergency diagnostic consultation capability between medical personnel at remote or isolated sites and larger medical facilities.

We would like you to take a few more minutes to answer the following questions:

1. Do you feel that satisfactory radiology consultations for emergency cases can be made via the RMDS? (Please discuss.)

2. Do you feel that there are particular types of pathologies that may or may not be readily diagnosed using the RMDS? (Please indicate such types.)

3. Was the "zoom" capability (enlargement of a portion of the image) useful to you in making a diagnosis?

Exhibit B6. Radiology questionnaire (Test Group).

4. If you used the "zoom" feature, did you save the full-sized image using the "video storage"? If so, was that useful and how?

5. Would one image be sufficient for this type of system, or are two simultaneous images required? (Please discuss.)

6. Did you use the "reverse polarity" feature (positive or negative image)? Would such a feature be of any use to you as a radiologist?

7. Do you see a need for archiving some images on disc memory for later consultation?

8. Please make any additional comments you wish.

Exhibit B6 (continued).

RADIOLOGY QUESTIONNAIRE

You have been involved in a test of the Remote Medical Diagnosis System (RMDS) to determine its effectiveness in transmitting x-ray images. Different transmitting modes and resolution levels have been used, which accounts for the differences in image quality between the various sets of cases. Much of the time required between transmission of each image has been due to the data collection procedure for these tests, and should not bias your opinion of the system's usefulness. Keep in mind that the system is intended to provide an emergency diagnostic consultation capability between medical personnel at remote or isolated sites and larger medical facilities.

We would like you to take a few more minutes to answer the following questions:

1. Do you feel that satisfactory radiology consultations for emergency cases can be made via the RMDS? (Please discuss.)

2. Do you feel that there are particular types of pathologies that may or may not be readily diagnosed using the RMDS? (Please indicate such types.)

3. Was the "zoom" capability (enlargement of a portion of the image) useful to you in making a diagnosis?

Exhibit B7. Radiology questionnaire (Control Group).

4. Would one image be sufficient for this type of system, or are two simultaneous images required? (Please discuss.)

5. Do you see a need for archiving some images on disc memory for later consultation?

6. Please make any additional comments you wish.

Exhibit B7 (continued).

APPENDIX C: LIST OF ACRONYMS AND ABBREVIATIONS

ADM	Advanced Development Model
AHC	analog, high SNR, coarse resolution
AHF	analog, high SNR, fine resolution
ALC	analog, low SNR, coarse resolution
ALF	analog, low SNR, fine resolution
ANOVA	one-way analysis of variance
BL	black level
CCTV	closed-circuit TV
DCL	Diagnostic Confidence Level
DHC	digital, high SNR, coarse resolution
DHF	digital, high SNR, fine resolution
DRF	Difficulty Ranking Factor
Dr X	senior radiologist responsible for assigning OCR_x ratings
Dr Y	senior radiologist responsible for assigning OCR_y ratings
ECG	electrocardiogram
EDM	Engineering Development Model
F	the parametric statistic in analysis of variance
FCL	Findings Confidence Level
K	camera lens aperture
LB	lightbox
n	sample size
NRMC	Naval Regional Medical Center
OCR	Overall Clinical Reading
OCR_x	Overall Clinical Reading assigned by Dr X
OCR_y	Overall Clinical Reading assigned by Dr Y
RMDS	Remote Medical Diagnosis System
rms	root-mean-square
SD	standard deviation
SNR	signal-to-noise ratio
VL	video level

PRECEDING PAGE BLANK-NOT FILMED