AD_____

Award Number DAMD17-96-1-6229

TITLE: Cancers Missed on Mammography

PRINCIPAL INVESTIGATOR: Robert M. Nishikawa, Ph.D.

CONTRACTING ORGANIZATION: University of Chicago
Chicago, Illinois 60637

REPORT DATE: October 2002

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2002	Annual (23 September 2001 - 22 September 2002)

4. TITLE AND SUBTITLE

Cancers Missed on Mammography

5. FUNDING NUNUMBER
DAMD17-96-1-6229

6. AUTHOR(S)
Robert M. Nishikawa, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Chicago
Chicago, Illinois 60637

email r-nishikawa@uchicago.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

Observer error in reading screening mammograms has been identified as a significant factor in delayed diagnosis of breast cancer. The magnitude of the problem is estimated to be about 30% of potentially detectable cancers are overlooked for one or more years before detection. Computer-aided diagnosis (CAD) programs have been developed to aid radiologists in the detection task, and pre-clinical studies have shown that CAD applied to digitized mammography films can flag about 50% of radiologists' observational oversights. Our preliminary study has also shown a wide variability in radiologist observer performance. The purpose of this investigation is to test how many additional cancers are detected by radiologists using CAD, in an observer study using an enriched mixture of cancers. A pilot study has been completed and the results were used to plan the full observer study. Using more sophisticated statistical power analysis, we estimate that we need 370 cases containing 70 cancers with 10 radiologists. We are also using ImageChecker M1000 software which has a false positive rate of 0.5 per image, compared to 3.0 for the software used in the pilot study. Observers are being recruited and reading will start in December.

14. Subject terms	15. NUMBER OF PAGES
mammography, computer-aided diagnosis, missed cancers	11
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

4. INTRODUCTION

Double reading of mammograms has been shown to significantly increase the number of cancers detected.[1-5] Computer-aided diagnosis (CAD) has been proposed as an efficient method of implementing double reading.[6] For CAD to be effective computers must find cancers that are missed by radiologists, and radiologists must react appropriately to the computer prompts. Others and we have found that computer detection schemes can find over 50% of the observational misses made by radiologists reading mammograms.[7-9] Our current study is designed to show that CAD can help detect cancers that they might otherwise be overlooked. We will collect a large database of cancers already missed by radiologists in routine clinical practice, and will test observers without and with the aid of CAD. It is expected that radiologists will detect about 10 to 15% more cancers using CAD, which would have important implications for bringing this technique into clinical practice. We will also learn much more about the reasons for and types of radiologist misses on mammography.

5. BODY OF REPORT

5.1 Tasks

There are five tasks in the Statement of Work, which are listed below.

Task 1. Preparation of review forms and finalization of eligibility characteristics for cases to be entered into the missed lesion database.

Task 2. Accumulation of database cases and copying/digitizing 100 missed malignant cases and 300 normal cases, with categorization of features and characteristics of the malignant case. Verification of missed lesion cases. Ongoing data entry.

Task 3. Computer runs producing hard copy of computer output for use in observer experiment and preparation of cases for observer experiment. Ongoing data entry of computer accuracy and truth table for missed lesion database. Final design of details of observer performance study.

Task 4. An observer experiment conducted on 15 observers at about 3 hours per session, with 6 sessions per observer spaced at 2-3 months apart. Goal is to perform 2 observation sessions and analysis minimum per week, entering observation data into a computer database. Ongoing data entry.

Task 5. Final analysis of data comparing CAD observer results with non-CAD results and observer variability, and preparation of report summarizing the results of the observer experiment and the clinical characteristics of the missed lesions.

5.1.1 Preparation of forms

A copy of the review form has been submitted previously. The eligibility criteria are as follows:

1. Patients who have had screen-film mammograms read at the participating mammography facilities.

2. For cases of missed lesions, the mammogram had to be read clinically as normal in the area where a cancer subsequently developed, and the error had to be one of observation (failure to see the lesion) rather than interpretation (seeing the lesion and categorizing it as benign). In cases where the cancer is visible on multiple examinations prior to diagnosis, the two expert mammographers reviewing the cases will collaboratively select a single representative screening exam as the index missed case.

3. Case is a minimum of 1 year old (to avoid any interference with clinical care), unless bilateral mastectomy has been performed, or unless films clinically equivalent to those entered into the study from other years are available.

4. Case is not involved in any medical-legal action.

5. No copy films will be used that include significant marks made by a previous observer prior to the copying, and no originals with such permanent marks will be used.

5.1.2 Development of database of missed lesions

The database is complete. Based on a pilot observer study, reported last year, we estimate that we need 70 cancer cases and 300 normal cases. We have collected and digitized 310 normal cases and 75 cases with a missed cancer. The extra cases will be used to create a training set, which will be augmented with 10 other cancer cases borrowed from other studies ongoing in our laboratory. Final selection of cases for the study and training set is ongoing.

Three tables in the Appendices summarize some of the characteristics of 50% of the cancers entered into our database. We are analyzing the other 50% of cancers.

5.1.3 Computer analysis of case

We determined from the pilot study that CAD schemes with fewer false positives need s to used for the study. Our current detection schemes (one for masses and one for calcifications) have about 3 false positives per image. Commercial systems average under 1 per image. Note that because of the success of commercial software, we have not developed our detection since approximately 1998. We have on loan, an R2 Technology, Inc, ImageChecker 1000 that has a sensitivity of approximately 85% with 0.5 false positives per image.[8] We will use this system in our study. All the cases have been analyzed by the R2 System. We are still computing the computer's sensitivity. The false positive rate is 0.48 per image.

5.1.4 Observer study

The formal observer study has not yet begun. Based on the pilot study, we will use 370 cases of which 70 will contain a cancer.[11-13] We will have 10 radiologists read these cases with and without the computer aid. We have identified 7 of 10 readers and are in the process of recruiting 3 more radiologists.

Each reader will be asked to answer 2 questions for each case: If you were reading this case clinically and this is all the information that is available, (i) Give your BI-RADS assessment of this case; and (ii) what is your level of confidence that the patient should be called back for further work-up or a biopsy? The later using the following confidence scale

1.0 No evidence for recalling the patient.
1.5
2.0 Some, but insufficient evidence for recalling the patient
2.5
3.0 Equivocal. [If you read this case on 10 different days, half the time you would recall.]
3.5
4.0 Sufficient evidence for recalling the patient.
4.5
5.0 Overwhelming evidence for recalling the patient.

If the radiologist gives a BI-RADS assessment is not 1, then the radiologist will be required to specify the location of the lesion and type of lesion using the computer interface.

The BI-RADS rating and lesion type and location will be used to generate sensitivity and call back rates, while the confidence scale will be used to do ROC analysis.

5.1.5 Data Analysis

Data analysis of the missed lesion/CAD study cannot begin until the observer study has been completed. The t-test will be used to compare sensitivity and call back rates. We will use the area under the ROC curve as the performance metric. We will use the Dorfman, Berbaum, Metz [14] method for testing the statistical significance of differences in the area under the ROC curve.

5.2 Discussion

We will assemble a training set of cases and then begin the observer study.

5.3 Recommendations in relation to the Statement of Work

- We do not anticipate making any further changes to the Statement of Work.

6. KEY RESEARCH ACCOMPLISHMENTS

- Pilot observer study performed
- Detailed planning of observer study complete
- Final case selection and observer recruitment are being made.

7. REPORTABLE OUTCOMES

None since last report.

8. CONCLUSIONS

Data collection is nearly complete and so we will begin to conduct our main observer study in December. Valuable data has been collected from a preliminary smaller observer study, which will influence the design of the larger scale observer study. We anticipate that we will be able to demonstrate that CAD can reduce the number of missed cancers by 50%, which has not yet been shown in a structured observer experiment. These results should provide information on which health care providers and governmental organizations can base decisions on the value of introducing this promising new technology into the clinical practice of breast cancer screening.

9. REFERENCES

1. C. E. Metz and J. H. Shen, "Gains in accuracy from replicated readings of diagnostic images: prediction and assessment in terms of ROC analysis," Medical Decision Making **12**, 60-75 (1992).

2. E. L. Thurfjell, K. A. Lernevall and A. A. S. Taube, "Benefit of independent double reading in a population-based mammography screening program," Radiology **191**, 241-244 (1994).

3. R. M. L. Warren and S. W. Duffy, "Comparison of single reading with double reading of mammograms, and change in effectiveness with experience," British Journal of Radiology **68**, 958-962 (1995).

4. C. A. Beam, D. C. Sullivan and P. M. Layde, "Effect of human variability on independent double reading in screening mammography," Academic Radiology **3**, 891-897 (1996).

5. J. Brown, S. Bryan and R. Warren, "Mammography screening: An incremental cost effectiveness analysis of double versus single reading of mammograms," British Medical Journal **312**, 809-812 (1996).

6. R. A. Schmidt and R. M. Nishikawa, "Digital screening mammography," Principles & Practice of Oncology **8**, 1-16 (1994).

7. R. A. Schmidt, R. M. Nishikawa, R. Osnis, K. L. Schreibman, M. L. Giger and K. Doi, "Computerized detection of lesions missed by mammography," in *Digital Mammography '96,* edited by K. Doi, M. L. Giger, R. M. Nishikawa and R. A. Schmidt (Elsevier Science, Amsterdam, 1996), pp. 105-110.

8. J. Roehrig, T. Doi, A. Hasegawa, B. Hunt, J. Marshall, H. Romsdahl, A. Schneider, R. Sharbaugh and W. Zhang, "Clinical results with R2 Imagechecker system," in *Digital Mammography - Nijmegen, 1998,* edited by N. Karssemeijer, M. Thijssen, J. Hendriks and L. van Erning (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998), pp. 395-400.

9. G. M. te Brake, N. Karssemeijer and J. H. C. L. Hendriks, "Automated detection of breast carcinomas not detected in a screening program," Radiology **207**, 465-471 (1998).

10. R.A. Schmidt, G.M. Newstead, M.N. Linver, G.W. Eklund, C.E. Metz, M.A. Winkler, R.M. Nishikawa, "Mammographic Screening: Sensitivity of General Radiologists," in *Digital Mammography - Nijmegen, 1998,* edited by N. Karssemeijer, M. Thijssen, J. Hendriks and L. van Erning (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998), pp. 383-388.

11. R. M. Nishikawa, M. L. Giger, C. Corbridge, R. A. Schmidt and J. Papaioannou, "Can computer-aided diagnosis help radiologists find mammographically missed screening cancers?," Proc. SPIE **4324**, 56-63, (2001).

12. S. V. Beiden, R. F. Wagner and G. Campbell, "Components-of-variance models and multiple-bootstrap experiments: an alternative method for random-effects, receiver operating characteristic analysis," Acad. Radiol. **7**, 341-349 (2000).

13. N. A. Obuchowski, "Sample size tables for receiver operating characteristic studies," Am J Roentgenol **175**, 603-608. (2000).

14. D. D. Dorfman, K. S. Berbaum and C. E. Metz, "ROC rating analysis: generalization to the population of readers and cases with the jackknife method," Invest. Radiol. **27**, 723-731 (1992).

10. APPENDICES

Table 1. Distribution of breast density in our database

Breast Density	Frequency of Occurrence
Normal	0.30
Fatty	0.21
Dense	0.37

Focal	0.09

Table 2. Distribution of subtlety on a 5-point scale, where 1 is extremely subtle.

Subtlety Rating	Frequency of Occurrence
1	0.16
2	0.39
3	0.37
4	0.05
5	0

Table 3. Distribution by lesion type*

Type of Lesion	Frequency of Occurrence
Asymmetric Density	0.29
Architectural Distortion	0.24

Developing Density	0.07
Mass	0.46
Calcifications	0.10

*numbers sum to greater than 1, because some cases have multiple lesions.

Table 4. Distribution of possible reasons for cancers being missed.*

Possible Reason	Frequency of Occurrence
Seen on only 1 view	0.48
Obscured by overlying tissue	0.40
Looks like normal tissue	0.36
"Busy" breast	0.29
Film technique	0.26
Distracting lesions	0.24
Subtle lesion	0.14
Marginal lesion	0.10
Developing density	0.10
Benign appearing lesion	0.07
Lack of prior films	0.07
Too small to prompt workup	0.05
Lucent lines	0.05
Stable lesion	0.02

*numbers sum to greater than 1, because up to three reasons were given per case.

Table 5. Summary from pilot observer study.

Reader	Unaided	With Aid	Correlation between aid and no aid
A	0.686	0.685	0.967
B	0.725	0.775	0.817
C	0.805	0.793	0.943
D	0.710	0.688	0.988
mean	0.731	0.735	0.929