AD _____

CONTRACT NO: DAMD17-94-C-4082

TITLE: Increasing the Accuracy of Mammogram Interpretation

PRINCIPAL INVESTIGATOR(S): Doctor John Swets

CONTRACTING ORGANIZATION: Bolt Beranek and Newman, Inc.
Cambridge, Massachusetts 02138

REPORT DATE: October 1995

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. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1995	Annual 15 Sep 94 - 14 Sep 95

4. TITLE AND SUBTITLE
Increasing the Accuracy of Mammogram Interpretation

5. FUNDING NUMBERS
DAMD17-94-C-4082

6. AUTHOR(S)
Dr. John Swets

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Bolt Beranek and Newman, Inc.
Cambridge, Massachusetts 02138

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)*

The project's aim is to refine and evaluate practically a computer-based system that will provide decision support to the radiologist in interpreting mammograms and generate a standardized report of his/her findings for the referring clinician. The fully developed system will elicit from the radiologist via spoken prompts a spoken numerical scale value for each of the diagnostically most relevant perceptual features of the mammogram (usually on a 10-point scale). It will merge the values with optimal weights via a statistical prediction rule, to calculate a probability of malignancy as an advisory for the radiologist. From the pattern of feature values, the system will automatically construct a prose report of findings for the referring clinician; spoken recommendations and impressions will be accepted by the system to complete the report. The system is to be evaluated with two groups of radiologists and cases -- one from a diagnostic or referral setting, the other from a screening or community setting -- and corresponding groups of referring clinicians. Accomplishments in Year 1 have been to set criteria for study cases, locate them in hospital files, and process them into databases -- separately for the two clinical settings; develop an improved master list of perceptual features and associated response form; lay additional groundwork, conceptually and in software, for the automated report writer; advance the capability for data entry by speech recognition; and acquire system hardware.

14. SUBJECT TERMS
mammogram interpretation, decision-support system, checklist of perceptual features, statistical prediction rule, automated report writing, data entry by speech, study case accrual, case database, evaluation of accuracy, breast cancer

15. NUMBER OF PAGES
38

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

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.
DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA - Leave blank.
NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

*U.S.GPO:1993-0-358-779

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

__X__ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

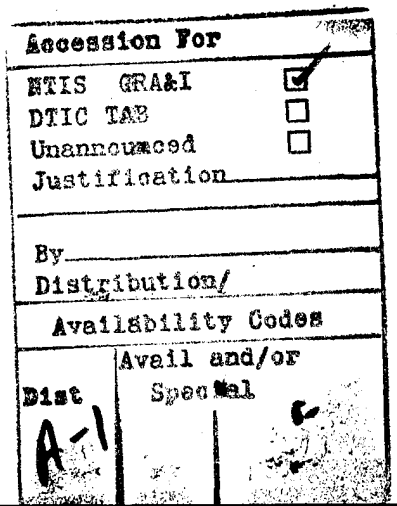
____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

__X__ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.



PI - Signature

20 Oct 95
Date

Table of Contents

INTRODUCTION

The project's aim is to refine and evaluate practically a computer-based system that will provide decision support to the radiologist in interpreting mammograms and generate a standardized report of his/her findings for the referring clinician. For each clinical case, the system will provide to the radiologist a checklist of the several perceptual features of a mammogram that have been determined to be diagnostically most relevant. The system will be developed to elicit from the radiologist via spoken prompts a spoken numerical scale value for each feature (usually on a 10-point scale). It will merge the values with optimal weights via a statistical prediction rule, to calculate a probability of malignancy as an advisory for the radiologist. Also, from the pattern of feature values, the planned system will automatically construct a prose report of findings for the referring clinician; spoken recommendations and impressions will be accepted by the system to complete the report.

Earlier laboratory experiments have shown that the parts of the system that aid mammogram reading and decision making -- that is, checklist and statistical prediction rule -- produce substantial increases in accuracy for the radiologist (Getty, Pickett, D'Orsi, and Swets, 1988; Swets, Getty, Pickett, D'Orsi, Seltzer, and McNeil, 1991). Clinically oriented system evaluations are now to be undertaken of refined versions of those components and, as well, of a refined and complete version of an automated-reporting capability, which has existed to date only in a preliminary and untested form. These evaluations will employ two groups of radiologists and cases, one representative of a community or screening setting and the other, of a referral or diagnostic setting. A significant difference between the two settings is the mix of different types of cases. Another difference is that the diagnostic center has the capability to do additional work-up (e.g., magnified views, ultrasound) while the patient is present for the mammogram, which then affect mammogram interpretation; in contrast, call-backs are required for additional work-up in the screening setting and are less frequent. It is possible that the value of the final, practical system will be viewed differently in teaching hospitals and community-oriented health maintenance organizations (HMOs).

A principal refinement of the system will be to include observed changes in perceptual features from previous to current mammograms. (These changes, as to direction and amount, will themselves be feature values for the statistical prediction rule.) Further, various algorithms for merging feature values -- i.e., various algorithms for

defining the statistical prediction rule -- will be tested. The algorithm selected, moreover, will be designed to be able to adapt over time to the mix of cases accumulating in the future in each clinical setting. Also, a speech-recognition capability will be incorporated into the system for convenient entry of case data by the radiologist. Finally, software will be developed to convert feature values into a standardized prose report of case findings; the possibility of the system's effectively recognizing concluding sentences dictated by the radiologist for the report will be explored.

Thirteen radiologists will participate as study readers, 5 in the diagnostic setting at the Brigham and Women's Hospital (BWH) and 8 in the screening setting at the Harvard Community Health Plan (HCHP). Comparison of accuracies in standard and aided conditions for both groups will be made in terms of the relative operating characteristic (ROC) (Swets, 1979; Swets and Pickett, 1982; Swets, 1988; Dorfman, Berbaum, and Metz, 1992). Focus groups of clinicians at both sites will evaluate the automatically composed reports of findings in comparison to the reports composed by the radiologist who originally read the case.

When utilized in practice, the system is expected to promote quality assurance in several respects (D'Orsi, Getty, Swets, Pickett, Seltzer, and McNeil, 1992). Its accumulating database of interpretations and outcomes will provide a detailed, quantitative, feature-by-feature basis for examining differences among multiple opinions and for evaluating an individual's ability to detect and assess perceptual features that distinguish malignant and benign conditions. It will provide a basis for evaluating an individual's ability to set appropriate decision thresholds for different levels of treatment (Swets, 1992) and for adjusting a department's average threshold, e.g., as reflected in the department's yield of biopsy recommendations. Finally, the database will allow the tailoring of individualized tutorials for continuing education (Greenes, Swets, Getty, and Pickett, 1992).

BODY OF REPORT

1. General

The 13 study radiologists (5 at BWH, 8 at HCHP) had originally agreed to flag certain mammograms among those they will be reading in their individual practices over the next few years that they would be willing to read again with the computer-based system. As of the beginning of the project, they generously agreed to interpret in both standard and enhanced modes other mammograms that we can obtain, not previously seen by them. This agreement has several desirable consequences for the project study and no apparent drawbacks. One advantage is that a set study cases at each site can now be accrued retrospectively as well as prospectively. Retrospective accrual allows stricter criteria to be set for case selection (including, for example, a longer period of follow-up to confirm normal cases and more prior mammograms) and, in general, provides better defined case sets. Therefore, better sets of cases for evaluation purposes can be achieved with smaller numbers of cases, thereby reducing somewhat the burden on our (volunteer) test radiologists. Moreover, a case set common to all of the readers in each group (BWH or HCHP) provides substantially more statistical power for a given number of verified cases (Obuchowski, 1995). Lastly, a major advantage is that the complete sets of cases used in evaluation will be available earlier in the project term and hence the test radiologists can complete their standard reading before doing any enhanced reading, thus preventing a possibly contaminating effect of enhanced-reading experience on standard reading. An incentive for the radiologists to read a common case set (common within each site) is that it permits giving each of them accuracy scores comparable across the group (anonymous except for his/her own) so that they can see where they stand relative to performances of their colleagues. We will continue to be able to assess the value of the system for individuals on cases they themselves read in a standard mode.

This refinement in study design entails some adjustments in the work schedule, which are detailed where appropriate in subsections of the report that follow.

2. Case Acquisition

Case accrual was begun at BWH early in the project and, after experience gained there, was begun a few months later at HCHP. A major part of the case-accrual process is setting the requirements for cases to be selected. A second major aspect is design and compilation of the databases of selected cases, in a manner that allows facile querying of those databases to sort groups of cases meeting certain requirements.

The data form designed to analyze and select cases at BWH is shown in the Appendix as Figure 1. The form is organized and divided into five major sections, delimited by horizontal lines. At the top, demographic information and data about the number of mammograms available are summarized. The next section lists the surgical procedures performed (for pathology proven cases) and itemizes up to 3 specific pathological diagnoses for each patient. Third from the top is a synopsis of mammogram findings from up to 6 examinations -- at the time a suspicious lesion is detected as well as for the studies before and after (for "suspicious normals"). Finally, at the bottom are two fields relating the existence of additional mammographic views or the performance of additional breast imaging studies along with their dates and findings.

The database is also kept in electronic form using the FileMaker Pro application for the Macintosh. This database software allows the user to easily search and retrieve eligible cases according to any desired criterion.

Care has been taken to assure the confidentiality of paper and electronic medical records. As can be seen from the database form, each patient has been assigned an anonymous "study number" to be used in the project. Our working database will identify patients by this study number. Only the project's research assistant will have access to the password-protected "master database" that contains patient names and medical record numbers. We will use the name and medical record numbers only to facilitate quality assurance and internal consistency within the database (e.g., ensuring that the proper pathology report is matched to the proper mammogram). The research assistant keeps all paper and electronic files containing identifiable patient information in a secure area; the rest of the project personnel and expert readers will see only the study number.

The list of minimum requirements on cases for study eligibility that was prepared to guide HCHP staff, but summarizes the inclusion criteria for the entire study, is shown in the Appendix as Figure 2. It shows the requirements for pathology-proven malignant and benign cases, and for suspicious normal cases. It indicates that for pathology-proven cases at least 2 mammograms (malignant cases) or 3 mammograms (benign cases) must be available spanning at least 1 year (malignant) or 2 years (benign). This requirement allows us to evaluate the changes in lesions over time. The more liberal criterion for malignant cases is due to the smaller number of such cases available -- a wider net needs to be cast to retrieve a sufficient number of cases for the study. For 'suspicious normal cases that do not undergo a biopsy, 2 years worth of studies are needed both before the identification of a suspicious lesion (to assess change over time) and after the lesion is discovered (to assure no change and eliminate the possibility that the lesion was actually malignant). Figure 2 also enumerates the clinical and pathology data required for all study cases.

Given our strong interest in temporal change over successive mammograms (which has a large impact on diagnosis), a requirement for all cases was at least one, and preferably two or more, prior mammograms. The suspicious normals were required additionally to have two years of imaging follow-up, to confirm normality, after first qualifying as suspicious.

At BWH, we have collected cases from the years 1992 through June, 1995, a time during which imaging technology has been constant. To date, 127 malignant cases have been identified along with 145 benign cases meeting the pathology requirement. (This number of malignant cases was available when the requirement on prior mammograms was relaxed to one; however, about 80% of the 127 cases have 2 or more prior mammograms.) Suspicious normals are cases recommended for accelerated follow-up, and at BWH 32 have been identified so far. These latter cases have met the requirements of either asymmetric density, focal lesion but seen in only one view, or focal lesion seen in two views but with residual uncertainty, e.g., about risk or from inconclusive ultrasound results. The numbers of cases identified to date at HCHP are 43 malignant, 43 pathology-proven benign, and 40 suspicious normals. We will continue to accrue cases at BWH over the final six months of 1995 and hence increase all BWH totals by about 12%. Case accrual is only partial and continues at HCHP, where we expect a greater increase over present totals.

These case-accrual results ensure adequate case samples for (1) establishing (or "training") a statistical prediction rule as a decision aid, based on the case set (and readers) from BWH, (2) standard and enhanced evaluation readings (different sets from the referral and screening sites), (3) the control condition of re-reading in the standard mode, at HCHP, and (4) pilot readings, at both sites.

3. A Note on Study Design

As just indicated, the BWH cases and readers will be used to train the statistical prediction rule. Both BWH and HCHP cases will be used to evaluate the rule. Hence, the HCHP group will provide an "experimental" cross-validation of the rule and the BWH group will give a "statistical" cross-validation (the latter requiring a jackknifing procedure to reduce optimistic bias). The "baseline" readings for the BWH radiologists will therefore be aided by a (master) checklist, but not an advisory diagnostic probability as calculated by the statistical prediction rule. Baseline readings of HCHP radiologists better justify the term: they are readings without any aid, either checklist or diagnostic probability.

4. Definition of Perceptual Features

We have made substantial progress in improving the master checklist of perceptual features. Later scaling by BWH radiologists of features on this master list for a set of proven cases, and subsequent statistical analyses, will provide the final, reduced list of features that serves as a reading aid and a specification of the features that enter the statistical prediction rule.

To date, we have placed special emphasis on features involved in "temporal change" -- on noticeable or measurable changes from one mammogram to a later one. Perhaps the greatest lack of our previous work was that it did not include temporal factors. In this connection, the project leaders have had extensive individual and joint discussions with expert mammographers and project consultants, Drs. Carl D'Orsi, Thomas Frenna, and Jack Meyer, and have reviewed typical cases with them in great detail.

Additionally, we have successfully mapped the descriptions of all perceptual features considered into the standard lexicon of the American College of Radiology. (As indicated in our proposal, our earlier work helped to define that lexicon.) We have also improved the format and logic of the questionnaire that the radiologist follows in scaling the perceptual features relevant to a case. The questionnaire now makes provision for all lesions noticed in a mammogram, not just the most worrisome one as in our preliminary work. This inclusion is required for completeness of the verbal report of findings. And we have incorporated a capability for the radiologist to indicate feature values on both current and prior mammograms, from which the system can calculate direction and amount of change from prior mammograms.

A close-to-final version of the comprehensive master checklist is shown as Figure 3 of the Appendix, titled "Response Form -- X-ray Mammography." This hard-copy version of the response form cannot adequately represent the way it will look and feel in its on-line/interactive form. The best that can be accomplished in hard copy is to show the basic components and indicate how the pieces work together. The response form consists first of a section that guides the reader through an overview of the breast images. This section is filled out first and for every case. It begins by asking for a rating of overall, technical image quality. This rating is used for study quality control and can condition the judgments on other features. Next, the form asks for a general evaluation of the entire breast parenchyma, including an assessment of overall tissue density, symmetry, and nodularity, and for reports on secondary signs, scattered calcifications, and asymmetric densities. It then prompts the reader to make a detailed inventory of all focal abnormalities and to list them in order of importance, beginning with those most worrisome with respect to cancer and continuing through to lesions, even if benign, that require at least passing comment in the report. The remaining components of the response form are sub-questionnaires aimed at each of the types of focal abnormality relevant to cancer diagnosis: mass, clustered calcifications, and architectural distortion. These sub-questionnaires are activated in the on-line version as required. The on-line version will start at the top of the inventory list and prompt the reader with the sub-questionnaire appropriate to that lesion type. In each component of the response form, we capture information on interval changes by asking that ratings be made for both the current and the immediately previous mammogram. We conclude the running of each sub-questionnaire by asking, with regard to each abnormality, for an overall, summary, diagnostic judgment of the probability of malignancy on a 100-point scale (Swets, Feehrer, Greenes, and Bynum, 1986).

This response form will be used to capture the feature ratings that will train the statistical prediction rule. It will also collect the BWH radiologist's summary diagnostic judgment for each case in a baseline condition. Only the features selected for the statistical prediction rule will compose the ultimate response form for the aided readings by HCHP radiologists.

Our work on perceptual feature analysis in radiology grew out of earlier work in our laboratories in which we studied visual representations of sounds, and some idealized visual stimuli, with support from the Office of Naval Research (e.g., Getty, Swets, and Green, 1979; Getty, Swets, and Swets, 1980, 1981; Getty and Swets, 1982).

5. Development of Automated Report Writer

Refining our prototype report writer has begun in discussions of project leaders with mammographer Carl D'Orsi and BBN Scientist Marie Meteer. The challenge here is to convert the radiologist's assignments of numerical scale values to the several perceptual features into sentences in a prose report of findings to the referring physician and surgeon. The report must flow in a natural way, must reflect interactions that arise among the particular values assigned particular features, and must be consistent with the report lexicon devised by the American College of Radiology (ACR). Dr. Meteer is expert in the generation of text from non-text data (Meteer and McDonald, 1988; Meteer, 1991; Meteer, 1991; Meteer, 1991; Meteer, 1992; Meteer, 1994). Dr. D'Orsi was co-chairman of the ACR committee that defined the lexicon (Kopans and D'Orsi, 1992).

There are two major goals in developing the automated report writer: first, to automate the process of writing the report, freeing the radiologist from that task, and second, to standardize radiology reports so that they contain the same information and that similar judgments from different radiologists are described in similar terms. Toward the first goal, we plan to use a computer-based text-generation program that can take care of the grammatical constraints, so the report is in readable English. Also, the system will be portable to new domains because the basic competence is in the generator. The *Spokesman* text generator was developed at BBN for use in automatic report writing in the military domain and has since been used in several different applications, including database query and narration of simulations. It will provide a base on which we can add a

domain-specific content selection and organization component. Toward the second goal, the project has already adopted a standard lexicon and completed a mapping of the perceptual features into that lexicon. The lexicon must be extended to include ways of expressing relative importance and the changes in perceptual features from previous to current mammograms.

In order for the automated report writer to be successful, the reports must be both readable and accepted by radiologists. Simple methods of text generation that build simple sentences from templates and produce stilted text with only a single piece of information in each sentence will not meet these goals. The generation system must combine related information and use linguistic means for foregrounding important information and backgrounding less important information. It must also select how much information is relevant to a particular case.

The first step in this process will be to work with radiologists to determine a set of "targets," that is, reports that both express the information from the questionnaire succinctly in the correct terms and lie within the linguistic competence of the Spokesman text generator. We need to determine how the language in the reports is used to indicate relative importance, such as through the order of the items described and the use of adjectives and other descriptive phrases, and how the different degrees represented by numbers in the questionnaire are expressed with words in the reports. For example, on a scale of -5 to 5, if 0 is "moderately dense" and 5 is "extremely dense," is 3 "somewhat dense," "a bit dense," or just "dense"?

The next phase will be to design and implement a content-selection component for the mammography domain. Working from the targets and the questionnaire, a domain model will be created to represent the kinds of objects and relations among them that are part of the mammography domain. This model will provide the basis for the content generation. An algorithm will be developed for scanning a questionnaire, selecting the features that should be included in the report, and determining how they should be organized. An additional mechanism that would enhance the usability of the system is a switch to set the level of detail to be included in the report. For some cases, only a brief summary of the information is necessary, as when there are no apparent abnormalities. If problems for such patients occur later, the system should be able to regenerate reports with more detail.

As mentioned, the Spokesman text generator has worked in a variety of applications and thus has a fairly wide linguistic competence. However, whenever the domain is changed there are inevitably new syntactic constructs that need to be added to the system. This work will be done in conjunction with the initial step of creating target reports, so that where possible we will use constructs already in the system.

Because the usefulness of the system will depend in a large part on the acceptance of the output by the experts, it is essential to have a final evaluation stage. We will run the generator on a selected set of data and work with focus groups of radiologists to compare the questionnaires with the generated reports and tune the output where necessary.

6. Data-Entry Modes

Concerning modes of data entry by radiologist to computer, our emphasis has been on the most desirable and challenging one, that of automatic speech recognition (in contrast to data entry by keyboard or mouse). According to our design, the radiologist in interpreting a mammogram gives vocally first the necessary case-identification information, then the various numerical scale values of the checklist's perceptual features (from both prior and current mammograms), and then treatment recommendations. The system will provide spoken prompts to guide the radiologist through the items on the checklist. A few sentences may be dictated at the end to include in the report any summary, impressions, or qualifying opinions. The prose report is to be generated immediately, on-line, for review, transmittal, and storage as desired.

7. Merging Algorithms for the Statistical Prediction Rule

Using the feature data we will obtain from our 5 BWH mammographers, we will explore the relative advantages of different merging algorithms for the statistical prediction rule. These algorithms will include both standard statistical classification techniques (e.g., linear discriminant and logistic regression classifiers) and newer neural-net techniques (e.g., back-propagation, probabilistic, and learning-vector-quantization networks). We will take advantage of experience we will gain in the next several months from our current project on prostate cancer staging, in which we will be comparing the

performance of statistical and neural-net classifiers in predicting the stage of prostate cancer using feature values generated by radiologists reading magnetic-resonance scans (Seltzer, Getty, Tempany, Pickett, Schnall, McNeil, and Swets, 1995; Swets, Seltzer, McNeil, Tempany, Getty, Pickett, and Schnall, 1996).

8. Hardware and Software

We have analyzed various possibilities for speech-recognition software, including BBN's HARK recognizer and other commercially available packages. The HARK recognizer is quite possibly the best at speaker-independent recognition of connected discourse with relatively large vocabularies. However, it works on a SUN workstation with UNIX software and has not been ported to a PC. We desire something compatible with our existing PC software for controlling a reading session and have selected the Phonetic Engine 500 package from Speech Systems Inc. for our initial purposes. The cost of this package falls within our hardware/software budget. We have also acquired the two budgeted systems based on Gateway computers. Resolving the hardware/software issues for speech recognition has required enough time to delay somewhat our production of system software and development of the automated report writer.

9. Schedule and Level of Effort

The project ran at a level of effort and expenditure in Year 1 of about 80% of budget. We anticipate being on budget and schedule after three to six months into Year 2. In general, the several aspects of the project have fallen into place nicely and we feel that it is proceeding very well.

10. Personnel

Our first research assistant, Ms. Sue Calder, left the project at the end of the year to begin medical school. We appreciate her diligent and competent efforts in case accrual. A smooth transition was made to Dr. Lisa Hermann, Pharm.D., who had been assisting Dr. Seltzer in another project at BWH. All other staff and consultants continue with the project. A broad-based team of collaborators was identified at HCHP, led by Ms. Leslie Finocchio and Ms. Betty Manong, to assist Dr. William Otto there. A second mammographer at HCHP, Dr. Jeffrey Melamed, will also work with Dr. Otto.

CONCLUSIONS

In broad summary, our major accomplishments in Year 1 have been to (1) set criteria for study cases, locate them in hospital files, and process them into databases -- separately for the BWH and HCHP sites; (2) develop an improved master list of perceptual features and associated response form; (3) lay additional groundwork, conceptually and in software, for an automated report writer; (4) advance the capability for data entry by speech recognition; and (5) acquire system hardware.

Year 2 will emphasize: (4) additional software development for data collection; (5) training the merging algorithms of the statistical prediction rule and determination of the final, reduced checklist of perceptual features for the reading/decision aid; (6) completion of the automated report writer; (7) collection of pilot reading data; (8) system integration; and (9) collection of baseline reading data from both the BWH and HCHP clinical groups. Year 3 will include (10) collection of reading data with the computer-based system; (11) collection of re-reading data without the decision aids; and (12) interviews with focus groups. The reduced budgeted effort in Year 4 will be devoted to (13) case follow-up; (14) data analysis; and (15) preparation of the final report. No unusual difficulties are anticipated in completing these several tasks.

REFERENCES AND BACKGROUND LITERATURE

Dorfman, D. D., Berbaum, K. S., and Metz, C. E. (1992) Receiver operating characteristic rating analysis: Generalization to the population of readers and patients with the jackknife method. Investigative Radiology, 27, 723-731.

D'Orsi, C. J., Getty, D. J., Swets, J. A., Pickett, R. M., Seltzer, S. E., and McNeil, B. J. (1992) Reading and decision aids for improved accuracy and standardization of mammographic diagnosis. Radiology, 184, 619-622.

Getty, D. J., Pickett, R. M., Chylack, L. T., McCarthy, D. F., and Huggins, A. W. F. (1989) An enriched set of features of nuclear cataract identified by multidimensional scaling. Current Eye Research, 8(1), 1-8.

Getty, D. J., Pickett, R. M., D'Orsi, C. J., and Swets, J. A. (1988) Enhanced interpretation of diagnostic images. Investigative Radiology, 23(4), 240-252.

Getty, D. J. and Swets, J. A. (1982) Perceptual dimensions: Similarity and identification. In Geisser, H-G., Petzold, P. (Eds.) Psychophysical judgment and the process of perception. Amsterdam: North-Holland Publishing Company; and Berlin: VEB Deutscher Verlag der Wissenschaften, 1982, pp. 239-252.

Getty, D. J., Swets, J. A., Swets, J. B., and Green, D. M. (1979) On the prediction of confusion matrices from similarity judgments. Perception and Psychophysics, 26, 1-19.

Getty, D. J., Swets, J. A., and Swets, J. B. (1980) The observer's use of perceptual dimensions in signal identification. In Nickerson R. S. (Ed.) Attention and performance, VIII, pp. 361-380. Hillsdale, NJ: Lawrence Erlbaum Associates.

Getty, D. J., Swets, J. A., and Swets, J. B. (1981) Perceptual spaces revealed by multidimensional scaling procedures. In D. J. Getty, J. H. Howard, Jr. (Eds.) Auditory and Visual Pattern Recognition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Greenes, R. A., Swets, J. A., Getty, D. J., Pickett, R. M. (1992) Computer assisted instruction in radiology. Final Report submitted to Brigham and Women's Hospital. National Institute of Health, Grant No. 5 R01 CA45574-03.

Kopans, D. and D'Orsi, C. J. (1992) ACR system enhances mammography reporting. Diagnostic Imaging, September, 125-132.

Meteer, M. and McDonald, D. D. (1988) From water to wine: Generating natural language text from today's applications programs. ACL Applied Conference, Austin Texas.

Meteer, M. (1991) Bridging the 'generation gap' between text planning and linguistic realization. Computational Intelligence, 7(4)

Meteer, M. (1991) The implications of revisions for natural language generation. In Paris, Swartuut and Mann (Eds.) Artificial Intelligence and Computational Linguistics. Klewer Academic Publishers, Boston, MA.

Meteer, M. (1991) Decision making in generation: A multi-leveled approach. Presented at the UCAI-91 Workshop on Decision Making Throughout the Generation Process, Sydney, Australia.

Meteer, M. (1992) Expressibility and the problem of efficient text planning. London: Pinter Publishers.

Meteer, M. (1994) Generating event descriptions with SAGE: A simulation and generation environment. 7th International Generation Workshop, Kennebunkport, ME.

Metz, C. E., Swets, J. A., and Al-Aish, M. (1986) Enhancing the accuracy of image-based diagnostic systems. Proposal funded under a grant from the National Cancer Institute and subcontracted from the University of Chicago. PHS Grant CA-33261.

Obuchowski, N. A. (1995) Multireader receiver operating characteristic studies: A comparison of study designs. Academic Radiology, 2, 709-716.

Seltzer, S. E., McNeil, B. J., D'Orsi, C. J., Getty, D. J., Pickett, R. M., and Swets, J. A. (1992) Combining evidence from multiple imaging modalities: A feature-analysis method. Computerized Medical Imaging and Graphics, 16(6), 373-380.

Seltzer, S. E., Getty, D. J., Tempany, C. M. C., Pickett, R, M., Schnall, M. D., McNeil, B. J., and Swets, J. A. (1995) Combined radiologist - computer system for staging prostate cancer by magnetic resonance imaging. (Submitted to New England Journal of Medicine.)

Swets, J. A. (1979) ROC analysis applied to the evaluation of medical imaging techniques. Investigative Radiology, 14(2), 109-121.

Swets, J. A. (1986a) Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. Psychological Bulletin, 99(1), 100-117.

Swets, J. A. (1986b) Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. Psychological Bulletin, 99(2), 181-198.

Swets, J. A. (1988) Measuring the accuracy of diagnostic systems. Science, 240, 1285-1293.

Swets, J. A. (1992) The science of choosing the right decision threshold in high-stakes diagnostics. American Psychologist, 47(4), 522-532.

Swets, J. A., Feehrer, C. E., Greenes, R. A., and Bynum, T. E. (1986) Use of probability estimates in medical communications and decisions. Methods of Information in Medicine, 25, 35-42.

Swets, J. A., Getty, D. J., Pickett, R. M., D'Orsi, C. J., Seltzer, S. E., and McNeil, B. J. (1991) Enhancing and evaluating diagnostic accuracy. Medical Decision Making, 11, 9-18.

Swets, J. A. and Pickett, R. M. (1982) Evaluation of diagnostic systems: Methods from signal detection theory, NY: Academic Press.

Swets, J. A., Pickett, R. M., Whitehead, S. F., Getty, D. J., Schnur, J. A., Swets, J. B., Freeman, B. A. (1979) Assessment of diagnostic technologies. Science, 205, 753-759.

Swets, J. A., Seltzer, S. E., McNeil, B. J., Tempany, C. M. C., Getty, D. J., Pickett, R. M., Schnall, M. D. (1996) Decision aids for radiologists. Paper at American Association for the Advancement of Science, Baltimore, MD, February 11.

APPENDIX

Study Number	0	**Referring MD**	
Patient Name	Test Case	**Case Status**	
Medical Record Number		**Total Mammos Available**	6
Date of Birth		**Years Spanned**	5
		Clinical Data	

Surgical Procedure	**Final Pathology**
Date of Surgery	**Pathology Diagnosis #1**
Accession Number	**Pathology Diagnosis #2**
Affected Side	**Pathology Diagnosis #3**

Mammogram Information

	Mammo A	Mammo B	Mammo C	Mammo D	Mammo E	Mammo F
Sequence Number	6	5	4	3	2	1
Date	June 20, 1995	June 19, 1995	June 18, 1995	June 17, 1995	November 4, 1990	November 3, 1990
Type	Bilateral	Bilateral	Bilateral	Right	Right	Right
Mammogram Findings #1	Microcalcifications	Microcalcifications	Microcalcifications	Dense Breasts	Microcalcifications	Microcalcifications
Mammogram Findings #2	Asymmetry	Asymmetry	Arch. Distortions	Microcalcifications	Arch. Distortions	Arch. Distortions
Mammogram Findings #3	Dense Breasts	Dense Breasts	No Suspicious Foci	Dense Breasts	Arch. Distortions	No Suspicious Foci
Location of Abnormality	R. Lower Inner	R. Lower Inner	R. Lower Inner	R. Lower Inner	R. Lower Inner	R. Lower Inner
Single View	MLO	MLO	Unspecified	CC		

Additional Views YES

	View A	View B	View C	View D	View E	View F
Additional View Date	June 20, 1995	June 19, 1995	June 18, 1995	January 1, 1993	January 1, 1994	November 4, 1990
Additional View Type #1	Spot Compression	Spot Compression	Spot Compression	Spot Compression	Spot Compression	Spot Compression
Additional View Type #2	90 Degree Lateral	90 Degree Lateral	Exag. Axillary	Spot Compression	90 Degree Lateral	90 Degree Lateral
Add. View Findings #1	Microcalcifications	Microcalcifcations	Microcalcifcations	Microcalcifcations	Dense Breasts	Microcalcifications
Add. View Findings #2	Microcalcifications	Microcalcifcations	No Suspicious Foci	Arch. Distortions	Arch. Distortions	Nodular Density
Location of Abnormality	R. Lower Inner	R. Lower Inner	R. Lower Inner	R. Upper Outer	R. Upper Outer	R. Upper Outer

Additional Studies YES

	Study A	Study B	Study C	Study D	Study E	Study F
Additional Studies Date	January 1, 1990	February 2, 1989	March 3, 1988	April 4, 1987	May 5, 1986	June 6, 1985
Additional Studies Type	Ultrasound	Ultrasound	Ultrasound	Ultrasound	Ultrasound	Ultrasound
Add. Studies Findings	Hypoechoic het.	Arch. Distortions	Anechoic Structure	Microcalcifications	Arch. Distortions	Surg.& Rad. Changes
Location of Abnormality	R. Upper Outer	R. Upper Outer	R. Upper Outer	R. Upper Outer	R. Upper Outer	R. Upper Outer

Comments THIS IS A TEST CASE ONLY1

Figure 1. Data form for characterizing study cases.

MINIMUM REQUIREMENTS FOR MAMMOGRAPHY STUDY ELIGIBILITY

Pathology Proven Cases

1. Path-proven Malignant: 2 mammograms (including the one that led to the biopsy) with 1 year interval between earliest and most recent. Cases biopsied between 1992 and 1995.

2. Path-proven Benign: 3 mammograms (including the one that led to the biopsy) with 2 year interval between earliest and most recent. Cases biopsied between 1992 and 1995.

3. Patient with previous breast biopsy is acceptable as long as the current suspicious lesion is located in a different breast or a different quadrant of the same breast.

4. The following data must be available:
   a.) Mammography reports
   b.) Localization reports
   c.) Pathology reports-note whether done by Wire Localization or Excisional Biopsy wo/pre-op loc.
   *d.) Mammograms

5. If possible, it is helpful to have localization reports and/or pathology reports from any previous breast biopsies.

6. Also helpful to know is the following relevant clinical information:
   Personal history of breast cancer or other cancer
   Family history of breast cancer
   Hormone Replacement Therapy
   Nipple abnormality-bloody discharge, spontaneous discharge, retraction
   Palpable mass
   Prior Surgery-Wire Localization, Excisional Biopsy, U/S Guided Aspiration

Suspicious Normal Cases

1. Suspicious Normal Cases (HCHP category 2 and ? category 4) have either
   a.) Accelerated follow-up - often w/ single breast mammogram
   b.) Ultrasound or other diagnostic study
   c.) Additional Views

2. Suspicious normal mammogram requirements:
   a.) 2 years mammographic follow-up to prove no change over time
   AND
   b.) 2 years of prior mammograms
   c.) The "suspicious" mammogram, with 2 years of prior mammos and 2 of follow-up mammos on either side, should have been done in 1992. Therefore the span of years will be approximately between 1990 and 1994.

3. If possible, we'd prefer cases that do not have prior breast biopsy or surgery.

4.The following data must be available:
   a.) Mammography reports
   * b.) Mammograms

5. Relevant clinical information, as listed in #6 for path-proven cases, is also very helpful.

As discussed, we'd like to identify:
   50 Malignant
   50 Benign
   50 Suspicious Normals

* Mammograms not immediately needed.

Figure 2. Minimum requirements for case inclusion in study.

Case No._____

RESPONSE FORM--X-RAY MAMMOGRAPHY

This hard-copy version of the response form can not adequately represent the way it will be structured in its on-line/interactive form. The best that can be accomplished in hard copy is to show the basic components and to indicate how the pieces work together.

The "Overview of Breast Images" component (pages 1-6) will be filled out first and for every case. In its "Inventory" section (pages 6-8), the reader will list all focal abnormalities in order of importance, beginning with those most worrisome with respect to cancer and continuing through any benign lesions or conditions that require at least passing comment in the report.

The remaining components of this hard-copy version of the response form are sub-questionnaires aimed at each type of focal abnormality relevant to cancer diagnostics -- "Mass" (3 pages), "Clustered Calcifications" (3 pages), and "Architectural Distortion" (1 page). These sub-questionnaires will be activated in the on-line version as required. The on-line version will start at the top of the inventory list and prompt the reader with the appropriate sub-questionnaire.

OVERVIEW OF BREAST IMAGES

(A) What is the Overall Image Quality?

0	1	2	3	4	5	6	7	8	9	10	
poor										excellent	OV1

(B) General Breast Background

(1) Tissue Density

Left breast:

0	1	2	3	4	5	6	7	8	9	10	
fatty										extremely dense	OV2

Figure 3. Response form -- x-ray mammography.

Right breast:

0	1	2	3	4	5	6	7	8	9	10	
fatty										extremely dense	OV3

(2) Symmetry of overall tissue distribution

0	1	2	3	4	5	6	7	8	9	10	
very asymmetrical										very symmetrical	OV4

(3) Nodularity? (e.g. fibrocystic changes, prominent ductal pattern)

Left breast:

0	1	2	3	4	5	6	7	8	9	10	
none										extreme	OV5

Right breast:

0	1	2	3	4	5	6	7	8	9	10	
none										extreme	OV6

(4) Scattered Calcifications?

Left breast:

0	1	2	3	4	5	6	7	8	9	10	
few										diffuse	OV7

2

Figure 3 (cont.)

Right breast:

0 1 2 3 4 5 6 7 8 9 10 OV8
few diffuse

(5) Regional Calcifications?

Confidence regarding presence of worrisome regional calcifications

Left breast:

0 1 2 3 4 5 6 7 8 9 10 OV9
definitely NOT present definitely present

Right breast:

0 1 2 3 4 5 6 7 8 9 10 OV10
definitely NOT present definitely present

(6) Asymmetric density?

Confidence regarding presence of asymmetric density in at least one view

Left breast:

0 1 2 3 4 5 6 7 8 9 10 OV11
definitely NOT present definitely present

Right breast:

0 1 2 3 4 5 6 7 8 9 10 OV12
definitely NOT present definitely present

3

Figure 3 (cont.)

(C) Secondary Signs

Confidence about presence of enlarged prominent duct

Left breast:

0 1 2 3 4 5 6 7 8 9 10 OV13
definitely NOT present — definitely present

Right breast:

0 1 2 3 4 5 6 7 8 9 10 OV14
definitely NOT present — definitely present

Confidence regarding presence of skin thickening

Left breast:

0 1 2 3 4 5 6 7 8 9 10 OV15
definitely NO skin thickening — definitely some skin thickening

Right breast:

0 1 2 3 4 5 6 7 8 9 10 OV16
definitely NO skin thickening — definitely some skin thickening

Figure 3 (cont.)

Confidence regarding the presence of skin retraction

Left breast:

0 1 2 3 4 5 6 7 8 9 10 OV17

definitely NOT skin retraction — definitely some skin retraction

Right breast:

0 1 2 3 4 5 6 7 8 9 10 OV18

definitely NOT skin retraction — definitely some skin retraction

Confidence regarding presence of axillary adenopathy

Left breast:

0 1 2 3 4 5 6 7 8 9 10 OV19

definitely NO axillary adenopathy — definitely some axillary adenopathy

Right breast:

0 1 2 3 4 5 6 7 8 9 10 OV20

definitely NO axillary adenopathy — definitely some axillary adenopathy

Confidence regarding presence of trabecular thickening

Left breast:

0 1 2 3 4 5 6 7 8 9 10 OV21

definitely NO trabecular thickening — definitely some trabecular thickening

5

Figure 3 (cont.)

Right breast:

0	1	2	3	4	5	6	7	8	9	10	
definitely NO trabecular thickening										definitely some trabecular thickening	OV22

(D) FOCAL OR REGIONAL ABNORMALITY INVENTORY

Are there one or more focal abnormalities visible in two views?

Current: Yes No Previous: Yes No OV23 OV24

If No, skip to OV25

If Yes, list in order of worrisomeness: (see page 8 for legend)

	Breast (R/L)	Type[1]	Frontal Locus		Depth[3]	Visible in Previous Exam? (Y/N)	
			Zone[2]	Radial Position (O'clock position)			
1.							OV25
2.							OV26
3.							OV27
.							.
.							.
.							.

6

Figure 3 (cont.)

List any non-worrisome findings to be mentioned in the report: (see page 8 for legend)

	Breast (R/L)	Type[5]	Frontal Locus		Depth[3]	Visible in Previous Exam? (Y/N)	
			Zone[2]	Radial Position (O'clock position)			
1.							OV28
2.							OV29
3.							OV30
.							.
.							.
.							.

Figure 3 (cont.)

LEGEND FOR RESPONSE FORM--X-RAY MAMMOGRAPHY

FOCAL OR REGIONAL ABNORMALITY INVENTORY (Section D, page 6, questions OV25-OV27)

Type[1]	Zone[2]	Depth[3]
M = mass	S = subareolar	A = anterior
CC = clustered calcification	C = central	M = middle
AD = architectural distortion	AT = axillary tail	P = posterior

NON-WORRISOME FINDINGS INVENTORY (Section D, page 7, questions OV28-OV30)

Type[5]	Zone[2]	Depth[3]
TD = tubular density	S = subareolar	A = anterior
ILN = intramammary lymph node	C = central	M = middle
SC = skin calcifications	AT = axillary tail	P = posterior
VC = vascular calcifications		
PLC = "popcorn-like" or coarse calcifications		
LRC = large rod-like calcifications		
RoC = round calcifications		
SpC = spherical or lucent-center calcifications		
EC = eggshell or rim calcifications		
MCC = milk of calcium calcifications		
SC = suture calcifications		
DC = dystrophic calcifications		
PC = punctate calcifications		

Figure 3 (cont.)

MASS

1. Density of mass relative to surrounding glandular tissue

0	1	2	3	4	5	6	7	8	9	10	M1
mass density much lower										mass density much higher	

2. Confidence about the presence of fat within the mass

0	1	2	3	4	5	6	7	8	9	10	M2
definitely NO fat										definitely some fat	

3. Size of mass

	Current	Previous	
a. Largest diameter in CC view	_____ mm	_____ mm	M3
b. Smallest diameter in CC view	_____ mm	_____ mm	M4
c. Largest diameter in oblique view	_____ mm	_____ mm	M5
d. Smallest diameter in oblique view	_____ mm	_____ mm	M6

4. Shape of mass

0	1	2	3	4	5	6	7	8	9	10	M7
round / oval					lobular					irregular	

1

Figure 3 (cont.)

MASS (continued)

5. Portion of the margin that is clearly circumscribed

0	1	2	3	4	5	6	7	8	9	10	
none										totally circumscribed	M8

6. Where the margin is not circumscribed, confidence that margin is indistinct

0	1	2	3	4	5	6	7	8	9	10	
margin obscured by glandular tissue										margin indistinct	M9

7. Confidence that at least a small portion of the margin is spiculated

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT spiculated										definitely spiculated	M10

8. Degree of microlobulation

0	1	2	3	4	5	6	7	8	9	10	
none										extensive	M11

9. Confidence that the mass is an intramammary node

0	1	2	3	4	5	6	7	8	9	10	
definitely NOT an intramammary node										definitely an intramammary node	M12

2

Figure 3 (cont.)

MASS (continued)

10. Confidence that the mass is a skin lesion

0	1	2	3	4	5	6	7	8	9	10	M13
definitely NOT a skin lesion										definitely a skin lesion	

11. Presence of calcium within the mass

0	1	2	3	4	5	6	7	8	9	10	M14
NO calcifications within the mass					benign, chunky calcifications within the mass					worrisome,clustered microcalcifications within the mass	

OVERALL DIAGNOSTIC JUDGEMENT

Benign vs. malignant---Rate the likelihood (as the number of chances in 100) that the mass is indicative of malignancy:

Rating (0 to 100) _____

where: 0 = certainly benign or normal

100 = certainly malignant

3

Figure 3 (cont.)

CLUSTERED CALCIFICATIONS

1. Shape of the cluster--suggestive of conformance to a lobule versus duct

0	1	2	3	4	5	6	7	8	9	10	
round, suggestive of conformance to a lobule										tubular shape, suggestive of conformance to a duct	CC1

2. Shape of elements in cluster

	Current	Previous	
punctate (<0.5 mm)	0	0	CC2
round (> 0.5 mm)	0	0	CC3
rod-like (>0.5mm)	0	0	CC4
dystrophic (>0.5 mm)	0	0	CC5
amorphous	0	0	CC6
pleomorphic (<0.5 mm)	0	0	CC7
fine-linear / branching (<0.5 mm)	0	0	CC8

3. Variability of size of elements in cluster

0	1	2	3	4	5	6	7	8	9	10	
low variability of size										high variability of size	CC9

1

Figure 3 (cont.)

CLUSTERED CALCIFICATIONS (continued)

4. Confidence that at least some of the elements are irregular in shape

0	1	2	3	4	5	6	7	8	9	10	
all elements are round										at least some elements are irregular	CC10

5. Confidence that there is at least some indication of branching

0	1	2	3	4	5	6	7	8	9	10	
definitely NO indication of branching										definitely some indication of branching	CC11

6. Largest diameter of cluster in either view

_____ mm (current) _____ mm (previous) CC12 CC13

7. Number of elements in the cluster

	Current	Previous	
below 5	0	0	CC14
5 to 10	0	0	CC15
greater than 10	0	0	CC16

2

Figure 3 (cont.)

CLUSTERED CALCIFICATIONS (continued)

OVERALL DIAGNOSTIC JUDGEMENT

Benign vs. malignant---Rate the likelihood (as the number of chances in 100) that the clustered calcifications are indicative of malignancy:

Rating (0 to 100) _____

where: 0 = certainly benign or normal

100 = certainly malignant

3

Figure 3 (cont.)

ARCHITECTURAL DISTORTION

1. Confidence regarding presence of architectural distortion in a least one view

0	1	2	3	4	5	6	7	8	9	10	AD1
definitely NO architectural distortion										definitely some architectural distortion	

OVERALL DIAGNOSTIC JUDGEMENT

Benign vs. malignant---Rate the likelihood (as the number of chances in 100) that the architectural distortion is indicative of malignancy:

Rating (0 to 100) _____

where: 0 = certainly benign or normal

100 = certainly malignant

1

Figure 3 (cont.)