AD_____

Award Number: DAMD17-97-1-7193

TITLE: Methods for Evaluating Mammography Imaging Techniques

PRINCIPAL INVESTIGATOR: Carolyn M. Rutter, Ph.D.

CONTRACTING ORGANIZATION: Group Health Cooperative
for Puget Sound
Seattle, Washington 98101-1448

REPORT DATE: December 2001

TYPE OF REPORT: Final

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
	December 2001	Final (19 May 97 - 19 Nov 01)

4. TITLE AND SUBTITLE
Methods for Evaluating Mammography Imaging Techniques

5. FUNDING NUMBERS
DAMD17-97-1-7193

6. AUTHOR(S)
Carolyn M. Rutter, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Group Health Cooperative of Puget Sound
Seattle, Washington 98101-1448

E-Mail: rutter@ghc.org

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 purpose of this award is to enable Dr. Rutter to develop biostatistical methods for evaluating the accuracy of breast cancer screening. This four year program includes advanced training in the epidemiology of breast cancer, training in clinical detection of breast cancer, development of statistical methodology, and graduate teaching. A basic knowledge of the epidemiology, disease process and detection of breast cancer guides the development of statistical methods.

During this third funding year, Dr. Rutter has continued to expand her knowledge of breast cancer epidemiology and detection. She has published two articles during the third funding year. The 1st describes bootstrap estimation of accuracy statistics when patients are assessed at multiple patient sites. The 2nd published article compares performance of mammographers in a test setting to performance in clinical practice. Dr. Rutter also has a third article that is under review by JAMA that compares changes in breast density among women who initiate, discontinue, and continue use of hormone replacement therapy.

During her fourth funding year, Dr. Rutter will teach an introductory graduate level statistics course and will focus on methods for estimating sensitivity and specificity that incorporate growth curve models.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer, CDA Award	13
	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

FOREWORD

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

___ 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.

N/A 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.

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

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

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

12-19-01

PI - Signature Date

Table of Contents

Introduction

The purpose of this Department of Defense Breast Cancer Research Program Career Development Award was to enable Dr. Rutter to develop biostatistical methods for evaluating the accuracy of breast cancer screening. This four year program included advanced training in the epidemiology of breast cancer, training in clinical detection of breast cancer, development of statistical methodology, and graduate teaching. Dr. Rutter has succeeded in gaining knowledge of the epidemiology, disease process and detection of breast cancer and this will continue to guide her work on statistical methods related to both the epidemiology of breast cancer and evaluation of screening and diagnostic tests used to detect breast cancer. Over the course of the funding period, the focus of statistical methods has changed from those originally proposed. Specifically, Dr. Rutter's research shifted away from ordinal measures based on ROC analyses and towards dichotomous outcomes. There was also a shift in emphasis, away from purely statistical research and toward epidemiological and health services research. During her fourth and final funding year, Dr. Rutter published research examining the effect of hormone replacement therapy on breast density. Dense breast tissue is a key factor related to both missed cancer and false positive mammograms. Dr. Rutter also examined estimation methods that adjust for sojourn time. Several problems with available data were revealed, and models need to be developed that account for these data issues. Dr. Rutter also taught an introductory graduate level statistics course, but was unable to teach a course specific to diagnostic testing as this was taught instead by a senior University of Washington faculty member. At the completion of this career development award, Dr. Rutter has two primary ongoing projects, one examining methods to account for sojourn time and another examining variation between radiologists and mammography centers using hierarchical models.

Achievement of Year 4 (Final) Technical Objectives

Technical Objective 1: Gain additional training in breast cancer epidemiology, detection and treatment.

Dr. Rutter has continued to expand her knowledge about breast cancer throughout the funding period, primarily through attendance at scientific seminars at the Fred Hutchinson Cancer Research Center (FHCRC) and the University of Washington (UW). Dr. Rutter also participates in a Diagnostic Methods working group that includes faculty from both FHCRC and UW and attends Breast Cancer Surveillance Consortium (BCSC) meetings [1] that have provided her with important practical information about radiologists' interpretation of mammograms, and the timing and execution of diagnostic procedures. In April 2001, Dr. Rutter presented work at the American Society for Preventive Oncology. Through additional reading and analysis of BCSC data, Dr. Rutter has gained specialized knowledge about the interrelationships among hormone replacement therapy, breast density and breast cancer.

Technical Objective 2: Develop methods for multiple patient assessments.

This objective was partly addressed during year two of this CDA award. The article describing nonparametric bootstrap ROC estimation for correlated data was published in *Academic Radiology* [2]. This article compares an iterative bootstrap estimation approach to estimating and comparing the area under the ROC curve with a non-iterative method that uses sums of squares to adjust variance estimates for correlation between observations.[3] Both methods are theoretically valid, and both perform well in a simple situation. However, the bootstrap

estimator can more easily be used in complex sampling situations that include multiple sources of correlation.

Dr. Rutter will continue work on methods for clustered mammographic screening data, now focusing on hierarchical Bayesian modeling approaches. Dr. Rutter is currently using hierarchical models to examine radiologists' use of the Breast Imaging Report and Data System (BIRADS). BIRADS was developed by the American College of Radiology as a standardized method of reporting mammographic findings [4]. Interpretations coded using BIRADS can be difficult to analyze because this coding scheme combines information from two assessment stages, resulting in an outcome that mixes binary and ordinal information. The first stage is determination of complete assessment. If the assessment is incomplete, then category 0 should be assigned to indicate the need for additional work-up before a final assessment can be made. If the radiologist determines that the assessment is complete, then a final interpretation is made. Codes for final diagnoses are on an ordinal scale ranging from 1 (negative) to 5 (highly suggestive of malignancy). Dr. Rutter is using hierarchical models to better understand radiologists' use of BIRADS category 0; to estimate the rate of radiologic follow-up among women who receive a BIRADS-0; and to determine how best to include BIRADS-0 in the ordinal interpretation scale. Specifically, this research will examine:

1. Variability in use BIRADS-0 across radiologists, mammography centers, and consortium sites;
2. Variability in the probability of follow-up imaging within one year following a BIRADS-0 interpretation, across mammography centers and consortium sites;
3. The probability of one-year cancer outcomes among women given an incomplete assessment relative to probabilities based on final BIRADS assessments.

This research will provide information about how consistently the indeterminate interpretation is used, and will help to guide the treatment of BIRADS-0 outcomes in statistical models.

Technical Objective 3: Extend exact methods for ordinal regression models.

This was abandoned as a key research objective for two reasons. First, the usefulness of ordinal regression models is limited in the mammographic screening setting. Instead, most interest focuses on the sensitivity and specificity of mammography based on particular definitions of a positive mammogram. In addition, most mammograms are assessed using BI-RADS ratings [4], and these ratings do not use a pure ordinal scale. The second reason for moving away from this objective is that Dr. Rutter is using data from the BCSC, a surveillance project that captures information from large populations of women [1]. Small sample problems are rarely an issue when evaluating screening tests using this large population-based sample.

Technical Objective 4: Develop methods to adjust for error in measurement of disease status.

Dr. Rutter will continue work on new methods that adjust for error in measurement of disease status. At the BCSC meetings in April 2001, Dr. Rutter proposed the use of hierarchical Bayesian modeling approaches to incorporate models for tumor growth rate into estimation of sensitivity and specificity. Currently, a fixed one or two year follow-up period is used to define true state, with truncation of follow-up at a subsequent screening. Defining disease state based on cancer occurring within a subsequent follow-up period ignores variability in growth rates by age and tumor type and results in misclassification of disease at the time of screening. Using this follow-up reference standard, even a perfectly accurate radiologist could 'miss' some cancers.

This simple standard can be improved by using disease characteristics collected at the time of diagnosis. For example, we could assume that cancers less than 2cm could not have been detected any earlier, that cancers between 2cm and 5cm could have been detected during the prior year, and that cancers over 5cm could have been detected within the prior 2 years. Using lesion size crudely accounts for sojourn time, the length of time when breast cancer is mammographically detectable but is not yet clinically detectable. We can further improve estimation of sensitivity and specificity by more directly estimating and accounting for sojourn time.

Mean sojourn time (MST) and test sensitivity can be jointly estimated using the number of screen detected and interval (clinically detected) cancers for a group of patients who are repeatedly tested. Estimation builds on the convolution model: $I(t)=\int_0^t g(s)f(t-s|s)ds$, where $I(t)$ is clinical incidence in the absence of screening, $g(\,)$ is pre-clinical incidence, and $f(\,)$ is the sojourn time density [5,6]. The convolution is used to estimate the number of preclinical cancers at screening, $N_{preclin}$, which is equal to the number that were missed at the last screen that have not yet become clinically detectable plus the number of newly developed preclinical cases that have not developed into clinical cancers before screening. Given $N_{preclin}$, the number of cancers detected by screening has a Binomial distribution with probability that is equal to test sensitivity. This model is used to obtain likelihood-based estimates of sensitivity and mean sojourn time [7].

Most models developed for estimation of screening sensitivity and sojourn time require several simplifying assumptions. Earlier work assumes that both incidence and sensitivity are constant over time and across individuals. It would be useful to relax these assumptions, allowing expected rates to depend on observable characteristics, such as age. Similarly, expected sojourn time may be a function of patient characteristics. For example, breast cancer in pre-menopausal women is thought to be more aggressive, that is, have a shorter sojourn time, than breast cancer occurring in older women. For example, Peer found that doubling time decreased form 157 days in women 50-70 years old down to 80 days in women under 50 [8]. Similarly, the sensitivity and specificity of mammography vary across women depending on breast density, which is related to both age and post-menopausal use of hormone replacement therapy. In addition, as breast cancer progresses, tumors become larger and more apparent, so that sensitivity increases over the preclinical screen-detectable phase. In her thesis work, Ashih (2000) jointly estimated sensitivity and specificity using a hierarchical Bayesian model that allowed random variation in growth rates across women [9,10].

Models for estimation of screening sensitivity and sojourn time might also be improved by using information from studies of breast cancer growth. These studies use mammographic tumor size to measure growth. Earlier studies of breast cancer growth used tumors that were visible on two mammograms [11] or excluded 'very small' tumors [12], possibly biasing estimated growth rates. When there are only two measurements per woman, the exponential model is used to model sojourn time[8, 11,12] .The exponential model implies a constant doubling time. When three measurements are available, it is possible to estimate models that allow deceleration of tumor growth. These models include the Gompertz model [13,14] and the generalized logistic model [15]. Although growth models use two dimensional tumor size, they generally express growth in terms of tumor volume.

Dr. Rutter used data from 67,396 women screened at Group Health Cooperative between 1996 and 1999 to examine whether tumor size could be used to improve accuracy estimates and whether estimates of tumor growth rates and limits of detectable size be used to improve cancer definitions. This cohort included 673 women were diagnosed with breast cancer. Several limitations of existing data were revealed.

A major obstacle in pursuing this work has been the quality of tumor size data. Tumor size comes from the Surveillence, Epidemiology and End Results (SEER) data [16], which reports the largest tumor diameter (mm), and uses data from pathology, operative report, physical exam, and mammography exam (in that order). However, there is no indication of how size information was collected. Missing data was another problem. Size information was missing for all women (n=115) with ductal carcinoma in situ, and 11.5% of women with invasive cancer (n=558). Data could be missing because there was no measurable tumor, or because disease was diffuse or widespread.

Reporting of the largest tumor diameter is also problematic. Tumors are generally not spherical. Further information about the distribution of tumor shapes is needed so that tumor size could be estimated from this largest diameter. Use of the SEER data would be greatly enhanced by inclusion of volumetric measurements and coding of data source.

One solution to these problems would be to get good measures of mammographic tumor size from a subset of women. This data could then be used to develop an algorithm for estimating tumor size from the maximum diameter. Ideally, this would include a large set of films, which would be split, with half used to develop the algorithm and the other half used to assess the algorithm. Although some films have been pulled for an unrelated study, funding has not been available to pay for radiologist time required for collect mammographic size information.

Another problem encountered when using growth models is that this approach requires knowledge about minimally detectable tumor size. This minimal detectable size depends on woman-level characteristics such as breast density, tumor characteristics such as presence of microcalcifications and tumor location. There have not been systematic studies of detectable tumor size to inform selection of a minimal detectable size (or sizes).

Because of these problems, this current research goal has stalled at identification of future areas of research. Dr. Rutter continues to look for funding to pay for size measurement of approximately 180 films that have already been pulled for a subsequent study. These films will be reread by a sample of 20 Group Health radiologists. Therefore, having these size measurements would allow both development of algorithms for estimation of tumor size from SEER data, and estimation of minimal clinically detectable size as a function of breast density. Unfortunately, this line of research has been discouraged by BCSC principal investigators. This group has spent a considerable amount of time coming to agreement on the use of a one year follow-up truncated at subsequent follow-up, and feels that this definition of disease state provides a reasonable standard that allows comparison of mammographic accuracy across studies.

Technical Objective 5: Develop and teach a course in methods for assessing diagnostic tests. This technical objective is no longer possible, as Dr. Margaret Pepe, a full professor at University of Washington, taught a special topics course on Medical Diagnostic Testing in Spring 2000. Instead, Dr. Rutter taught an introductory biostatistics to health services students (Biostatistics 509) in Fall 2000. While teaching a graduate level survey course was not the original intent of the training grant, this experience provided valuable training. Successfully teaching this introductory course has increased Dr. Rutter's chance of teaching special topics course in the future.

Key Research Accomplishments

During this final funding year, Dr. Rutter has continued to expand her knowledge of breast cancer epidemiology and detection; She has first authored two articles[17,18]; She coauthored a paper examining the effect of training on use of the BIRADS reporting system [19]; and she taught an introductory Biostatistics class.The papers lead by Dr. Rutter compares changes in breast density among women who initiate, discontinue, and continue use of hormone replacement therapy [17]; and describes a hierarchical model for sensitivity and specificity that is useful for meta-analysis of diagnostic and screening test data [18]. Dr. Rutter also coauthored a paper examining the effect of training on use of the BIRADS reporting system [19].

Reportable Outcomes

1. Taplin S, Rutter CM, Elmore JG, Seger D, White E, Brenner RJ. Accuracy of Screening Mammography on Single Versus Independent Double Reading, *American Journal of Roentgenology,* 2000 174:1257-62.
2. Rutter CM, Mandelson MT, Laya MB, Seger DJ, Taplin S. Changes in Breast Density associated with Initiation, Discontinuation, and Continuing use of Hormone Replacement Therapy (HRT), *JAMA*. 2001;285:171-176.
3. Lehman CD, Miller L, Rutter CM, Tsu V. The Impact of Training with the American College of Radiology Breast Imaging Data and Reporting System Lexicon on Mammography Interpretation Skills in Developing Countries, *Academic Radiology*, 2001, 8: 647-50.
4. Rutter CM, Gatsonis CA, "A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations," Rutter CM, Gatsonis, CA. *Statistics in Medicine,* 2001, 20:2865-2884.
5. Presentation: Assessing Mammographers' Accuracy: A comparison of clinical and test performance, poster, Era of Hope: Department of Defense Breast Cancer Research Program Meeting, Atlanta, GA, June 9, 2000.
6. Presentation: Changes in Breast Density Associated with Initiation, Discontinuation, and Continuing use of Hormone Replacement Therapy (HRT), American Society for Preventive Oncology, March 12, 2001, New York, New York.
7. Presentation: Tumor Size and Mammographic Screening Accuracy, Breast Cancer Surveillance Consortium Meeting, 2001.

Conclusions

- Regarding bootstrap estimation: Bootstrap estimation of the area under the receiver operating characteristic curve, sensitivity, and specificity allows simple and accurate calculation of confidence intervals for single tests and comparisons between tests.

- Regarding the use of tests data sets: Direct estimation of mammographer's clinical accuracy requires the ability to capture screening assessments and correctly identify which screened women have breast cancer. Use of screening sets offers an attractive alternative method for estimating mammographers' accuracy. Unfortunately, we found that there was little concordance between performance on a test film set and performance in clinical practice. There is the potential for bias in both types of assessment, and our research cannot distinguish which approach is best. It does, however, raises questions about construction of and use of test film sets.
- Regarding use of HRT and breast density: Initiation of hormone replacement therapy (HRT) has been shown to increase breast density [15-19]. Several lines of evidence indicate that breast density is strongly related to breast cancer risk [20-23] and that increased density decreases mammographic sensitivity [24]. Using an cohort of 5213 naturally postmenopausal women 40 to 96 years old, we used consecutive mammograms and pharmacy records to examine the relationship between initiation, cessation and continuing use of HRT on breast density. We found that women who initiated HRT were more likely than nonusers to show increases in density (OR=3.24, 95% CI (2.47,4.23)), while women who discontinued were more likely show decreases in density (0R=1.92, 95% CI (1.03,3.35)), and women who continued use of HRT were more likely to show both increases in density (OR=1.37, 95% CI (0.89,2.06)) and sustained high density (OR=1.72, 95% CI (1.50,1.98)). Continuing HRT use was more strongly associated with sustained high density among women with high BMI ($p<0.05$). These results provide strong evidence that breast density changes associated with HRT are dynamic, increasing with initiation and decreasing with discontinuation. Continued HRT use results in persistent changes, particularly among women with high BMI.
- Regarding estimation of sensitivity and specificity for screening mammography: Several lines of evidence demonstrate variability in tumor growth rates. Estimates of mammography performance can be improved through incorporation of growth rate model.

References

1. Ballard-Barbash R, Taplin SH, Yankaskas BC, et al. Breast Cancer Surveillance Consortium: A National Mammography Screening and Outcomes Database, *AJR,* 169:1001-8, 1997.
2. Rutter CM, Bootstrap Estimation of Diagnostic Accuracy using Patient-clustered Data, *Academic Radiology,* 2000 7:413-419.
3. Obuchowski NA. "Nonparametric Analysis of Clustered ROC Curve Data," *Biometrics,* 53: 567-578, 1997.
4. Breast Imaging Reporting and Data System (BI-RADS), Third Edition, American College of Radiolgy : Reston, VA, 1998.
5. Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. *Biometrics* 1984;40:1-14.
6. Pinsky P.F.. Estimation and prediction for cancer screening models using deconvolution and smoothing. *Biometrics* 2001;57:389-395.
7. Prevost TC, Launoy G, Duffy SW, Chen HH. Estimating sensitivity and sojourn time in screening for colorectal cancer: a comparison of statistical approaches. *Am J Epidemiol* 1998;148:609-19.

8. Peer PG, Verbeek AL, Straatman H, Hendriks JH, Holland R. Age-specific sensitivities of mammographic screening for breast cancer. *Breast Cancer Res Treat* 1996;38:153-60.
9. Ashih HW, Berry DA. Modeling natural history of breast cancer tumor growth, *unpublished technical report*
10. Ashih HW. Joint estimation of mammographic sensitivity and tumor growth, *unpublished Doctoral Thesis,* Duke University, 2000.
11. Fournier DV, Weber E, Hoeffken W, Bauer M, Kubli F, Barth V. Growth rate of 147 mammary carcinomas. *Cancer* 1980; 45:2198-2207.
12. Brown BW, Atkinson EN, Bartoszynski R, Thompson JR, Montague ED. Estimation of human tumor growth rate from distribution of tumor size at detection. *J Natl Cancer Inst* 1984; 72:31-8.
13. Speer JF, Petrosky VE, Retsky MW, Wardwell RH. A Stochastic Numerical Model of Breast Cancer Growth That Simulates Clinical Data. *Cancer Research*, 44: 4124-4130, 1984.
14. Norton L.. A Gompertzian model of human breast cancer growth. *Cancer Res 1* 1988; 48:7067-71.
15. Spratt JA, von Fournier D, Spratt JS, Weber EE. Decelerating growth and human breast cancer. *Cancer* 1993; 71:2013-9.
16. Reis LAG, Kosary CL, Hankey BF, Miller BA, Edwards BK. SEER Cancer Statistics Review, 1973-1995. National Cancer Institute. Bethesda, Maryland: 1998.
17. Rutter CM, Mandelson MT, Laya MB, Seger DJ, Taplin S. Changes in Breast Density associated with Initiation, Discontinuation, and Continuing use of Hormone Replacement Therapy (HRT), *JAMA*. 2001; 285:171-176.
18. Rutter CM, Gatsonis CA, "A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations," Rutter CM, Gatsonis, CA. *Statistics in Medicine,* 2001, 20:2865-2884.
19. Lehman CD, Miller L, Rutter CM, Tsu V. The Impact of Training with the American College of Radiology Breast Imaging Data and Reporting System Lexicon on Mammography Interpretation Skills in Developing Countries, *Academic Radiology*, 2001, 8: 647-50.
20. Stomper PC, Van Voorhis BJ, Ravnikar VA, Meyer JE. Mammographic changes associated with postmenopausal hormone replacement therapy: a longitudinal study. *Radiolgoy.* 1990; 174:487-90.
21. Kaufman Z, Garstin WI, Hays R, et al. The mammographic parencymal patterns of women on hormonal replacement therapy. *Clinical Radiology.* 1991; 43:389-92.
22. Laya MB, Gallagher JC, Schreiman JS, et al. Effect of postmenopausal hormone replacement therapy on mammographic density and parenchymal pattern. *Radiology.* 1995; 196: 433-7.
23. Greendale GA, Reboussin BA, Sie A, et al. Effects of Estrogen and Estrogen-Progestin on Mammographic Parenchymal Density. *Annals of Internal Medicine.* 1999; 130: 262-269.
24. Lundstrom E, Wilezek B, von Palffy Z, et al. Mammographic breast density during hormone replacement therapy: Differences according to treatment. *American Journal of Obstetrics and Gynecology.* 1999; 18:348-352.
25. Litherland JC, Stallard S, Hole D, Cordiner C. The effect of hormone replacement therapy on the sensitivity of screening mammograms. *Clin Radiol.* 1999; 54:285-8.
26. Seradour B, Esteve J, Heid P, Jacquemier J. Hormone replacement therapy and screening mammography: analysis of the results in the Bouches du Rhone programme. *J Med Screen.* 1999; 6:99-102.

27. Kavanagh AM, Mitchell H, Giles GG. Hormone replacement therapy and accuracy of mammographic screening. *Lancet.* 2000; 355:270-274.
28. Saftlas AF, Szklo M. Mammographic parenchymal patterns and breast cancer risk. *Epidemiol Review.* 1987; 9:146-74.
29. Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, White E. Breast density as a predictor of mammographic detection: Comparison of interval- and screen-detected cancers. *J Natl Cancer Inst.* 2000; 92:1081-7.

Appendices

Appendix A. Statement of Work

Technical Objective 1: Gain additional training in breast cancer epidemiology, detection and treatment.

Task 1: Months 1-4: Review of information on the epidemiology, diagnosis and treatment of breast cancer as suggested by Dr. Margaret Mandelson.

Task 2: Months 1-48: Attend seminars sponsored by the Seattle Breast Cancer Research Program.

Technical Objective 2: Statistical research, aim 1: develop methods for multiple patient assessments.

Task 3: Month 6: Review current research for generalized estimating equation and random effect approaches for nonlinear models.

Task 4: Months -11: Test bootstrap, robust covariance adjustment and generalized estimating equation methods for breast-level analyses using simulation studies.

Task 5: Months 12-21: Develop methods for woman-level analysis, possibly including software development for random effects in generalized ordinal regression models.

Technical Objective 3: Statistical research, aim 2: extend exact methods for ordinal regression models

Task 6: Month 22: Review current research in exact methods.

Task 7: Months 23-34: Extend exact methods and write computational algorithms and programs to compute distributions of sufficient statistics.

Technical Objective 4: Statistical research, aim 3: Develop methods to adjust for measurement error in disease status

Task 8: Month 36: Review current research in errors-in-measurement models.

Task 9: Months 37-48: Develop simple combined corrections for verification and follow-up bias. These methods will be extended to allow adjustments in general ordinal regression models.

Technical Objective 5: Develop and teach a course in methods for assessing diagnostic tests.

Task 10: Months 1-24: Collect relevant references and outlining lectures for the methods course. During this time, specific lectures may be presented in other University of Washington courses.

Task 11: Months 25-36: Offer methods course at University of Washington through the Department of Biostatistics.