Published in final edited form as:

J Clin Epidemiol. 2012 February ; 65(2): 219–230. doi:10.1016/j.jclinepi.2011.06.010.

USE OF CLINICAL HISTORY AFFECTS ACCURACY OF INTERPRETIVE PERFORMANCE OF SCREENING MAMMOGRAPHY

Patricia A. Carney, PhD¹, Andrea J Cook, PhD^{2,3}, Diana L. Miglioretti, PhD^{2,3}, Stephen A. Feig, MD⁴, Erin Aiello Bowles, MPH², Berta M. Geller, EdD⁵, Karla Kerlikowske, MD⁶, Mark Kettler, MD⁷, Tracy Onega, PhD⁸, and Joann G. Elmore, MD, MPH⁹

¹ Department of Family Medicine, Oregon Health & Science University, Portland, OR

² Group Health Research Institute, Seattle, WA

³ Department of Biostatistics, University of Washington, Seattle, WA

⁴ Department of Radiological Sciences, University of California, Irvine, Irvine, CA

⁵ Office of Health Promotion Research, University of Vermont, Burlington, VT

⁶ Department of Internal Medicine, University of California, San Francisco, San Francisco, CA

⁷ Department of Radiology, Oregon Health & Science University, Portland, OR

⁸ Department of Community and Family Medicine, Dartmouth Medical School, Lebanon, NH

⁹ Department of Internal Medicine, University of Washington, Seattle, WA

Abstract

Objective—To examine how use of clinical history affects radiologist's interpretation of screening mammography.

Study Design and Setting—Using a self-administered survey and actual interpretive performance, we examined associations between use of clinical history and sensitivity, false-positive rate, recall rate, positive predictive value, after adjusting for relevant covariates using conditional logistic regression.

Results—The majority of radiologists the 216 radiologists surveyed (63.4%) reported usually or always using clinical history when interpreting screening mammography. Compared to radiologists who rarely use clinical history, radiologists who usually/always use it had a higher false-positive rate with younger women (10.7 vs. 9.7), denser breast tissue (10.1 for heterogeneously dense to 10.9 for extremely dense versus 8.9 for fatty tissue), or longer screening intervals (> prior 5 years) (12.5 vs. 10.5). Effect of current hormone therapy (HT) use on false-positive rate was weaker among radiologists who use clinical history compared to those who did not ($p=0.01$), resulting in fewer false-positive exams and a non-significant lower sensitivity (79.2 vs. 85.2) among HT users

© 2011 Elsevier Inc. All rights reserved

Correspondence to: Patricia A. Carney, PhD, Professor of Family Medicine, and Public Health and Preventive Medicine - MC-FM 3181 SW Sam Jackson Park Rd. Oregon Health & Science University Portland, OR 97239-3098 Telephone: 503-494-9049 Fax: 503-494-2746 carney@ohsu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conclusions—Interpretive performance appears to be influenced by patient age, breast density, screening interval and hormone therapy use. This influence does not always result in improved interpretive performance.

Keywords

Mammography screening; interpretive acumen; breast cancer detection; breast cancer risk

INTRODUCTION

Radiologists vary in their interpretation of screening and diagnostic mammography (1–3). Research has focused on the extent to which this variability is attributable to the characteristics of the women being screened (4–6) and the radiologists interpreting the mammograms (7–11). Little is known about the process of interpretation such as how radiologists use clinical history in their interpretation of screening mammograms and whether this use affects accuracy. Prior research on the use of clinical history in mammography has revealed conflicting results, with two studies showing improvements in the accuracy of detection of breast cancer (12,13) and one showing no improvement (14). These prior studies were limited by small samples of radiologists (n=2–10) in some cases and use of test sets, which may not represent the use of clinical history in actual practice (12–14). In addition, the studies differed in the elements of the clinical history that they examined, such as patient age or results of prior clinical breast exam.

A recent analysis using data from the Breast Cancer Surveillance Consortium (15) assessed the impact of women's breast cancer risk factors on radiologists' mammographic interpretive performance (16). This study noted that having one or more clinical risk factors was associated with higher recall rates and lower specificity on screening mammography without a corresponding improvement in sensitivity and only a small increase in positive predictive value. A weakness of this study was its inability to discern whether the changes in radiologists' interpretive performance were due to their knowledge of patient risk factors during the interpretive process.

We know of no studies that have examined the use and impact of clinical history in interpreting screening mammograms in community settings. Therefore, we used a self-administered survey to assess radiologists' use of women's clinical history (e.g., age, family history of breast cancer, current hormone therapy use, screening history and previous biopsy) when they interpret mammograms, and we linked the results to the same radiologists' actual clinical performance in community practice. We hypothesized that knowledge of the clinical history may alter a radiologist's level of suspicion without improving interpretive performance.

Methods

Study population

Seven mammography registries that are part of the National Cancer Institute-funded Breast Cancer Surveillance Consortium (BCSC; further information available at <http://breastscreening.cancer.gov>) contributed data for this study. These registries collect patient demographic and clinical information each time a woman receives a mammography examination at a participating facility. This information is linked to regional cancer registries and pathology databases to determine cancer outcomes. Data from the registries were pooled at the BCSC Statistical Coordinating Center (SCC) for analysis. Each registry and the SCC received IRB approval for either active or passive consenting processes or a waiver of consent to enroll participants, link data, and perform analytic studies. All

procedures are compliant with Health Insurance Portability and Accountability Act (HIPAA) and all registries and the SCC have received a Federal Certificate of Confidentiality and other protection for the identities of women, physicians, and facilities that are subjects of this research.

Radiologists who interpreted mammograms at a facility contributing to any of the seven registries between January 2005 and December 2006 were invited to participate in a mailed survey in 2006, using survey methods previously described (17, 18). Of the 257 radiologists who responded to the survey ($257/364 = 71\%$ response rate), we excluded radiologists with no screening mammograms in the database during the study years ($n=10$) and those who were missing information on radiologist's use of clinical history ($n = 1$). Included in the analysis were screening mammograms for women aged 40 years or older interpreted by a participating radiologist at a BCSC site between January 1, 2000 and November 1, 2007. A screening mammogram was defined as a bilateral mammogram designated as screening in women without a history of breast cancer (because these examinations are surveillance rather than screening exams) or breast augmentation. To avoid misclassifying diagnostic mammograms as screening, we excluded mammograms performed among women who had breast imaging within the prior 9 months. The initial study population included 1,454,035 screening mammograms performed by 246 radiologists from 180 facilities (Figure 1). Previously published results showed that the interpretive performance and patient characteristics of the radiologists who completed the survey compared to the entire BCSC radiologist population did not differ strongly (17).

Ten facilities from the initial 180 were excluded from these analyses because they did not collect information on a single clinical risk factor $>90\%$ of the time or because they did not collect information on any of the clinical risk factors $\geq 75\%$ of the time. The clinical risk factors of interest were current use of hormone therapy (HT), first-degree family history of breast cancer, and previous breast biopsy. An additional 28 facilities were excluded because they did not collect Breast Imaging Reporting and Data System (BI-RADS®) mammographic breast density (19) on $>75\%$ of the mammograms. These exclusions removed 29 radiologists at 38 facilities and 249,938 mammograms leaving a total of 1,204,097 mammograms interpreted by 217 radiologists at 142 facilities (Figure 1).

In addition to excluding mammograms from facilities that did not regularly collect the information of interest, we removed individual mammograms missing information on the woman's breast density ($n=58,231$, 4.8%) or time since last mammogram ($n=107,827$, 9.0%), because these were deemed important potential patient-level confounders associated with interpretive performance; this exclusion removed 4 radiologists leaving a total of 216 radiologists. We then excluded mammograms missing all three clinical history risk factors of interest ($n=10,542$, 1.0%). We removed these mammograms because we could not determine whether the data were available to the radiologist at the time of screening or if the data were just not entered into the data system. Our final analysis included the remaining 1,027,497 screening mammograms, which were interpreted by 216 radiologists at 142 facilities (Figure 1).

Measurements

Data on radiologist characteristics were obtained from the self-administered mailed survey described above, which included the question, "When you are interpreting screening mammograms, do you use the women's' clinical history (e.g., risk factors such as age, family history, estrogen use, previous biopsies)?" The three response categories to this question were: 1) never/rarely; 2) yes, but only if an abnormality is noted; 3) yes, most of the time or always when the information is available. We stratified our analyses by responses to this question. The survey also asked radiologists about their demographic characteristics,

experience, and clinical practice characteristics in the prior year. We examined several radiologist characteristics previously found to be associated with the accuracy of screening mammography (17, 20): radiologist age and gender, full- or part-time status, affiliation with an academic institution, fellowship training years of experience, percentage of time working in breast imaging, and self-reported number of mammograms interpreted in the prior year.

Patient risk factor information collected at the time of the mammogram included age, time since last mammogram, current HT use, ever having a breast biopsy and first-degree family history of breast cancer. We derived a clinical history risk factor score using three risk factors available to the radiologist at the time of the mammogram: current HT use (Yes/No), family history of breast cancer (Yes/No), and ever having a breast biopsy (Yes/No). We treated any missing clinical risk factors as being unavailable to the radiologist at the time of screening and, therefore, contributing 0 to the overall clinical risk factor score, which ranged from 0 to 3. Clinical risk factor score values of 2 and 3 were collapsed into one category because very few women had all 3 risk factors under study. The clinical risk score was used to examine the influence of different levels of patient risk in the analysis (no risk factors, one risk factor, or 2 or 3 risk factors).

We also obtained the BI-RADS® (19) assessment and recommendation and the BI-RADS® mammographic breast density category (entirely fatty, scattered fibroglandular tissue, heterogeneously dense, or extremely dense) assigned by the radiologist. Because breast density is a radiologist-defined variable that represents how the radiologist estimates the percentage of mammographically dense tissue and may vary from radiologist to radiologist (21–25), we did not include it in the patient clinical history studied.

We used standard definitions developed by the BCSC to measure interpretive performance (26). We classified as *positive* those mammograms given an initial BI-RADS® assessment of 0 (needs additional imaging), 4 (suspicious abnormality), or 5 (highly suggestive of cancer). An initial BI-RADS® assessment of 3 (probably benign) with a recommendation for immediate follow-up was also considered *positive*. We classified as *negative* those mammograms given an initial BI-RADS® assessment of 1 (negative), 2 (benign), or 3 (probably benign) without a recommendation for immediate follow-up (typically occurring within weeks). *Recall rate* was defined as the percentage of positive examinations among all screening mammograms.

We linked women to tumor registry data to determine whether they were diagnosed with invasive breast cancer or ductal carcinoma *in situ* (DCIS) within one year of the mammography examination and before the next screening mammogram, again using standard definitions developed by the BCSC (26) and considered a mammogram given a positive assessment to be a *true positive* if breast cancer was diagnosed within the follow-up period. We considered a mammogram given a negative assessment to be a *true negative* if breast cancer was not diagnosed within the one-year follow-up period. We defined the *cancer rate* as the number of cancers per 1,000 screening mammograms (regardless of interpretation) and the *cancer detection rate* as the number of true positive assessments per 1,000 screening mammograms. *Sensitivity* was defined as the percentage of true-positive examinations among women diagnosed with breast cancer. *False-Positive Rate* was defined as the percentage of false-positive examinations among women without a breast cancer diagnosis. *Positive predictive value* (PPV) was defined as the percentage of true positive examinations among women with positive examinations.

Statistical Analysis

We calculated frequencies of radiologist characteristics stratified by their use of available clinical history. We tested for relationships between radiologists' characteristics and use of

available clinical history by applying a Pearson's Chi-Square Test for difference between groups. We calculated unadjusted rates of cancer detection per 1,000 screening mammograms, recall, sensitivity, false-positive rate, and PPV by the risk factors of the women and characteristics of the radiologists. Recall rate was included as an outcome variable in this study because it can influence other interpretive performance measures, especially specificity, and could likely be improved if it was out of recommended ranges.

To assess the relationship between clinical risk factors and interpretive performance we fit separate conditional logistic regression models stratified on radiologist for each performance measure. Conditional logistic regression uses the statistical technique of conditioning to remove the effects of any heterogeneity among radiologists. Using this technique, we assessed whether an individual radiologist's mammography performance differed by the presence of the patient's clinical risk factors, controlling for the effects of any radiologist-level characteristics such as radiologist experience and registry site. Importantly, this approach adjusts for differences in case-mix across radiologists (e.g., radiologists who primarily interpret mammograms for women at high risk of breast cancer may have a different mammography performance compared to radiologists who see women with fewer risk factors). In this way, conditional logistic regression estimates associations solely due to varying covariates such as clinical risk factors within radiologists. We then fit conditional logistic regression models to assess interactions between the effect of radiologist use of available clinical history and women level risk factors available to the radiologist at the time of the mammography examination in relation to interpretive performance. We do not include our results for recall rate in the tables as they were almost identical to the results for false-positive rate.

We performed multivariable analyses to assess the relationship between performance measures and the effect of varying woman-level risk factors adjusted for other woman-level characteristics (e.g., woman's age), controlling for radiologist-level effects also using conditional logistic regression. We fit multivariable analyses only for recall rate and false-positive rate because these were the only two outcomes showing significant relationships with patient risk factors. We included main effects and interactions with radiologist's use of clinical history, which were statistically significant at the 0.10 level in the bi-variable analyses. We did not include the risk factor score in the multivariable model because individual level clinical risk factors had stronger associations than the risk factor score alone and the interpretation of odds ratios when adjusting for the risk factor score is not straightforward.

Data analyses were conducted using SAS® software, Version 9.1 (SAS institute, Cary, NC). *P*-values are two-sided, and we considered *p*-values less than 0.05 to be statistically significant.

Results

The majority of radiologists (63.4%) reported usually or always using clinical history when interpreting screening mammography (Table 1), 29.2% reported using it only when they identified an abnormality, and 7.4% reported never or rarely using it. The majority of radiologists were male, age 50 or older, working full time in radiology, not affiliated with an academic medical center, not fellowship trained and with more than 10 years' experience interpreting mammography (Table 1). Only one radiologist characteristic was associated with use of clinical history: full-time radiologists reported being more likely to use clinical history routinely (e.g., mostly or always) compared to part-time radiologists.

Table 2 outlines the characteristics of women who underwent screening mammography interpreted by study radiologists, including the number of mammograms and the number of cancers detected. Cancer detection rates increased with increasing patient age, increasing breast density, longer time since last mammogram, having a family history of breast cancer, having a prior breast biopsy, and increasing summary score for clinical history factors. Cancer detection rates were lower among radiologists who interpreted more than 20 years, but were not associated with any other characteristic of the radiologists including the use of clinical history.

Table 3 shows performance indices (sensitivity, false-positive rate, and positive predictive value) by women's characteristics with all patients combined and then according to use of the clinical history by the interpreting radiologist and Table 4 shows odds ratios from oneway interaction analyses comparing performance by women's characteristics and radiologist's use of clinical history. We found no significant associations between use of clinical history and recall ($p=0.72$), sensitivity ($p=0.37$), false-positive ($p=0.70$), or PPV ($p=0.29$). However, we found that use of clinical history changed the magnitude of the associations between performance measures and numerous women-level risk factors.

Overall, regardless of whether the radiologist used the clinical history, younger women were more likely to be recalled ($p<0.001$) with lower sensitivity ($p<0.001$), higher false positive rate ($p<0.001$), and lower PPV ($p<0.001$) compared to older women (Table 3). Among women without cancer, the decrease in false-positive rates observed with increasing patient age was stronger among radiologists who use the clinical history compared to those who never or rarely use it, but this interaction was only borderline significant ($p=0.07$). This resulted in a higher false-positive rate for women aged 40–49 (10.7 and 10.1 vs. 9.7) and a lower false-positive rate for women older than 70 (6.5 and 6.5 vs. 6.9) when mammograms were interpreted by radiologists who mostly/always or only if abnormality used clinical history compared to radiologists who did not use clinical history (Adjusted odds ratio by use of clinical history: never/rarely 0.66 (95% confidence intervals: 0.60, 0.73) compared to abnormality only 0.60 (95% CI: 0.58, 0.63); $p=0.079$ and mostly/always 0.60 (95% CI: 0.58, 0.61); $p=0.37$) (Tables 3 and 4).

As expected, women with denser breasts were recalled more often ($p<0.001$) and had a lower sensitivity ($p<0.001$), higher false-positive rate ($p<0.001$), and lower PPV ($p<0.001$). The increase in false-positive rates observed with increasing breast density was stronger for radiologists who use clinical history compared to those who never or rarely use it. This resulted in more women being recalled without cancer who had heterogeneously or extremely dense breasts by radiologists who used clinical history compared to those who never use clinical history (Adjusted OR by use of clinical history: heterogeneously dense vs. scattered fibroglandular tissue: [never/rarely use clinical history 1.32 (95% confidence intervals: 1.23, 1.42) compared to use history only if an abnormality is noted 1.48 (95% CI: 1.44, 1.53); $p=0.003$ and mostly/always use the history 1.47 (95% CI: 1.44, 1.50); $p=0.006$] and extremely dense vs. scattered fibroglandular tissue: [never/rarely 1.13 (95% confidence intervals: 1.00, 1.29) compared to abnormality only 1.32 (95% CI: 1.25, 1.39); $p=0.030$ and mostly/always 1.33 (95% CI: 1.28, 1.38); $p=0.017$]) Table 4

Overall, radiologists recalled more women with a longer time since last mammogram ($p<0.001$) with a higher sensitivity ($p<0.001$), higher false-positive rate ($p<0.001$), and a higher PPV ($p=0.002$) compared to women with shorter times since last mammogram. The increase in false-positive rate with increasing screening interval length was stronger for radiologists who used the clinical history. Radiologists who used clinical history recalled a higher proportion of women without cancer with at least 3 years since their last mammogram relative to women with ≤ 2 years since their last mammogram compared to

radiologists who never use clinical history (Adjusted OR by use of clinical history: 3–4 yrs vs. ≤ 2 yrs since last mammogram: [never/rarely 1.15 compared to abnormality only 1.22; $p=0.15$ and mostly/always 1.26; $p=0.03$] and ≥ 5 yrs vs. ≤ 2 yrs since last mammogram: [never/rarely 1.46 compared to abnormality only 1.78; $p=0.004$ and mostly/always 1.74; $p=0.010$]) (Table 4). This interaction between clinical history use and time since last mammogram resulted in higher observed false-positive rates among women who had not been screened in the prior 5 years (12.5 vs. 10.5) for radiologists who always used clinical history compared to those who did not (Table 3).

Current HT use was associated with higher recall ($p<0.001$) and false-positive rates ($p<0.001$), but not associated overall with sensitivity ($p=0.29$) or PPV ($p=0.43$). The associations were not as strong for radiologists who use clinical history compared to those who never use it. As a result, among HT users, fewer false-positive exams (9.2 vs. 9.8) and a corresponding, non-significant lower sensitivity (79.2 vs. 85.2) occurred for mammograms interpreted by radiologists who used clinical information compared to those who rarely or never used it (Table 3).

Women with a family history of breast cancer had significantly higher recall ($p<0.001$), false-positive rate ($p=0.002$), and PPV ($p<0.001$) compared to those without a family history, but there was no significant difference in sensitivity ($p=0.12$) (Table 3). Similarly, women with a previous benign breast biopsy had a significantly higher recall rate ($p<0.001$), false-positive rate ($p<0.001$) and PPV ($p<0.001$), and a significantly lower sensitivity ($p=0.028$) compared to women without a previous benign biopsy. Table 3 A radiologist's use of clinical history had no significant effect on the relationship between any of the performance measures and family history of breast cancer or a previous biopsy. Women with more clinical risk factors (current HT use, family history, and benign breast biopsy) had a higher recall rate ($p<0.001$), lower sensitivity ($p=0.005$), higher false-positive rate ($p<0.001$), and higher PPV ($p<0.001$) compared to women with no clinical risk factors. Reported use of clinical history by the radiologist did not change these associations.

All results from the multivariable analyses that adjusted for the other patient risk factors were similar to the bi-variable results except that use of clinical history no longer significantly changed the magnitude of the effect of patient age on recall or false-positive rate (data not shown).

DISCUSSION

To our knowledge, this is the largest and likely most generalizable study on radiologists' use of clinical history while interpreting screening mammograms in the clinical setting. This study included over 200 radiologists, mostly community-based, who interpreted 1,027,497 screening mammograms performed in more than 140 mammography facilities in the U.S. We examined the radiologists' use of clinical history for risk factors individually and then as a global measure.

We found no significant overall associations between use of clinical history and interpretive performance. However, what is new in our study is that we found the use of clinical history changed the magnitude of the association between performance measures and numerous women-level risk factors. First the association between age and false-positive rate was stronger for radiologists who used clinical history. Age is a strong risk factor for breast cancer and is one that is not evident on the images themselves, except possibly through its inconsistent relationship with breast density. Those who use clinical history were more likely to recall women with dense breasts, which may indicate their concerns about missing a cancer as it is well known that breast density can mask tumors.

Use of clinical history also changed the associations between screening interval and both recall rate and false-positive rate. Radiologists who use clinical history recalled a higher proportion of women who were not recently screened compared to radiologists who did not use clinical history. This suggests that knowing a woman has not been recently screened may decrease radiologists' thresholds for working up an area of concern, though not without an increase in false-positive examinations. Perhaps radiologists are concerned that infrequently screened women may not return for another screening mammogram for many years. Radiologists may be aware of a woman's screening interval in two ways. First, when prior studies are available, they are placed side-by-side with the current study. Thus, radiologists observe the dates on prior images. Second, when prior studies are not immediately available, radiologists may read the clinical history where they may learn that a prior exam was performed but the images are not yet available. Radiologists may be less likely to recommend supplementary imaging if they believe that prior studies will soon be obtained from a different facility. Knowledge that a concerning mammographic finding has been stable, particularly over longer intervals (2–3 years), substantially reduces a radiologist's level of concern and will usually obviate the need for further evaluation.

Both a family history of breast cancer and a previous benign biopsy were associated with increased recall and false-positive rates; however, surprisingly, we found that use of clinical history did not change this association. This suggests that the higher false-positive rates may be due to the presence of more benign masses and calcifications which require additional imaging and/or biopsy for further assessment (27–32), and not to radiologists' knowledge or concern that these women are at increased risk of breast cancer. Also surprisingly, we found that radiologists who use clinical history recalled a lower proportion of women without cancer who were current HT users compared to other radiologists. Although HT is known to increase cancer risk and cancer growth rates, knowledge that a patient has received HT may provide a reassuring explanation for observation of new or enlarging masses and focal asymmetries that might otherwise require additional imaging or even biopsy.

When we grouped clinical risk factors (current HT use, family history, and benign breast biopsy), we found a higher recall rate, lower sensitivity, higher false-positive rate, and higher PPV among women with these risk factors than among women without them and that use of clinical history did not change these associations.

We conducted a previous analysis (16) demonstrating that women with clinical risk factors who undergo screening mammography are more likely to be recalled for a false positive evaluation without an associated increase in sensitivity. This prior analysis was limited by our lack of knowledge regarding the extent to which radiologists were aware of the patient's risk factors when they made their interpretations. In the current study we found that most radiologists reported mostly or always using clinical history when interpreting screening mammograms. We have speculated that women with clinical risk factors may have different breast architecture and may be more likely to have benign lesions that require work-up to rule out cancer compared to women with no risk factors, which has been extensively reported in the literature (27–33).

Our findings are also similar to those found in a small study by Elmore et al (13), which assessed whether interpretations by 10 radiologists on a test set of 100 women's screening and diagnostic mammography exams were affected by the patient's clinical history. In this study, overall diagnostic accuracy was not altered by the presence of clinical history during the interpretation, but recommendations were affected for appropriate further diagnostic workup: an alerting history (e.g., breast symptoms or family history of breast cancer) was associated with an increased number of workups recommended in women without cancer

($p=.01$); and a non-alerting history was associated with fewer recommended workups in the women with cancer ($p=.02$).

Strengths of the present study include the large number of community-based radiologists and screening mammograms and the high response rate to the radiologists' survey assessing clinical history use. One weakness is that we did not directly assess reference to clinical history at the time of the individual mammography interpretations and we asked a more global question about use of clinical history in general rather than asking about use of each clinical history feature separately. Another weakness is the possibility of social response bias such that radiologists may have over-reported their use of clinical history if they thought that was the desired response. If they did over-report their use of clinical history, this bias could have affected the results in either direction because the radiologists did not have their interpretive performance data readily available to influence their responses in either direction, toward improved or lower interpretive performance measures. Future research should address more comprehensively the clinical thinking that radiologists use when they access and process patient history and how patient risk may be influencing their decision making at the actual time of mammographic interpretation, which may clarify some of the complexities identified in this study.

In conclusion, most radiologists report using clinical history when interpreting screening mammograms, but use does not necessarily lead to improved interpretive performance. Radiologists who use clinical history appear to have a lower false positive rate when evaluating older women and women with current HT use, but they have a higher false positive rate when evaluating younger women and women with longer times elapsed since their last mammogram.

Acknowledgments

This work was supported the National Cancer Institute (1R01 CA107623; 1K05 CA104699; Breast Cancer Surveillance Consortium: U01CA63740, U01CA86076, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, U01CA70040), the Breast Cancer Stamp Fund, the Agency for Health Care Research and Quality (R01 CA107623), and the American Cancer Society, made possible by a generous donation from the Longaberger Company's Horizon of Hope Campaign (SIRGS-07-271-01, SIRGS-07-272-01, SIRGS-07-273-01, SIRGS-07-274-01, SIRGS-07-275-01, SIRGS-06-281-01). The collection of cancer data used in this study was supported in part by several state public health departments and cancer registries throughout the U.S. For a full description of these sources, please see: <http://breastscreening.cancer.gov/work/acknowledgement.html>. The authors had full responsibility in the design of the study, the collection of the data, the analysis and interpretation of the data, the decision to submit the manuscript for publication, and the writing of the manuscript. We thank the participating women, mammography facilities, and radiologists for the data they have provided for this study. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are provided at: <http://breastscreening.cancer.gov/>.

REFERENCES

1. Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability of radiologists' interpretation of mammograms. *NEJM*. 1994; 331:1521–1522. [PubMed: 7969306]
2. Elmore JG, Nakano CY, Koepsell TD, Desnick LM, D'Orsi CJ, Ransohoff DF. International variation in screening mammography interpretations in community-based programs. *JNCI*. 2003; 95(18):1384–1393. [PubMed: 13130114]
3. Elmore JG, Miglioretti DL, Reisch LM, Barton MB, Kreuter W, Christiansen CL, Fletcher SW. Screening mammograms by community radiologists: variability in false-positive rates. *JNCI*. 2002; 94(18):1373–1380. [PubMed: 12237283]
4. Kerlikowske K, Carney PA, Geller B, Mandelson MT, Taplin SH, Malvin K, Ernster V, Rosenberg R, Cutter G, Ballard-Barbash R. Performance of Screening Mammography Among Women with and without a Family History of Breast Cancer. *Annals of Internal Medicine*. 2000; 133(11):855–863. [PubMed: 11103055]

5. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter C, Geller BM, Abraham LA, Dignan M, Cutter G, Ballard-Barbash R. Individual and Combined Effects of Breast Density, Age, and Hormone Replacement Therapy Use on the Accuracy of Screening Mammography. *Annals of Internal Medicine*. 2003; 138(3):168–175. [PubMed: 12558355]
6. Elmore JG, Carney PA, Abraham LA, Barlow WE, Egger JR, Fosse JS, Cutter GR, Hendrick E, D'Orsi CJ, Paliwal P, Taplin SH. The Association Between Obesity and Screening Mammography Accuracy. *Archives of Internal Medicine*. 2004; 164:1140–1147. [PubMed: 15159273]
7. Barlow WE, Chi chen, Carney PA, Taplin SH, Gutter G, Fosse J, Elmore JG. Accuracy of screening mammography interpretation by characteristics of radiologists. *JNCI*. 2004; 96(24):1840–1850. [PubMed: 15601640]
8. Fenton JJ, Egger J, Carney PA, Cutter G, D'Orsi C, Sickles EA, Fosse J, Abraham L, Talpin SH, Barlow W, Hendrick RE, Elmore JG. Reality Check: Perceived versus Actual Performance of Community Mammographers. *AJR*. July.2006 187:42–46. [PubMed: 16794153]
9. Carney PA, Yi J, Abraham LA, Miglioretti DL, Aiello EJ, Gerrity MS, Reisch L, Sickles EA, Berns EA, Elmore JG. Reactions to Uncertainty and the Accuracy of Diagnostic Mammography. *J Gen Int Med*. 2007; 22(2):234–241.
10. Miglioretti DL, Smith-Bindman R, Abraham LA, Brenner RJ, Carney PA, Aiello EJ, Buist DSM, Elmore JG. Radiologist Characteristics Associated with Interpretive Performance of Diagnostic Mammography. *JNCI*. 2007; 99:1854–1863. [PubMed: 18073379]
11. Aiello Bowles EJ, Miglioretti DL, Sickles E, Abraham L, Carney PA, Yankaskas B, Elmore JG. Accuracy of short-interval follow-up mammograms by patient and radiologist characteristics. *AJR*. 2008; 190:1200–1208. [PubMed: 18430832]
12. Lo JY, Baker JA, Kornguth PJ, Floyd CE Jr. Effect of patient history data on the prediction of breast cancer from mammographic findings with artificial neural networks. *Academic Radiology*. 1999; 6(1):10–15. [PubMed: 9891147]
13. Elmore JG, Wells CK, Howard DH, Feinstein AR. The impact of clinical history on mammographic interpretations. *JAMA*. 1997; 277(1):49–52. [PubMed: 8980210]
14. Knight JA, Libstug AR, Moravan V, Boyd NF. An assessment of the influence of clinical breast examination reports on the interpretation of mammograms in a breast screening program. *Breast Cancer Research and Treatment*. 1998; 48(1):65–71. [PubMed: 9541190]
15. Ballard-Barbash R, Taplin SH, Yankaskas BC, Ernster VL, Rosenberg RD, Carney PA, Barlow WE, Geller BM, Kerlikowske K, Edwards BK, Lynch CF, Urban N, Chrvala CA, Key CR, Poplack SP, Worden JK, Kessler LG. Breast Cancer Surveillance Consortium: A national mammography screening and outcomes database. *AJR*. 1997; 169:1001–1008. [PubMed: 9308451]
16. Cook AJ, Elmore JG, Miglioretti DL, Cutter GR, Aiello-Bowles EJ, Carney PA. Influence of patient risk and radiologists factors on the interpretation of screening mammography in community-based settings. *Journal of Clinical Epidemiology*. In Press.
17. Elmore JG, Jackson SL, Abraham L, Miglioretti DL, Carney PA, Geller BM, Yankaskas BC, Kerlikowske K, Onega T, Rosenberg RD, Sickles EA, Buist DSM. Variability in Interpretive Performance of Screening Mammography and Radiologist Characteristics Associated with Accuracy. *Radiology*. 2009; 253:641–651. [PubMed: 19864507]
18. Breast Cancer Surveillance Consortium: FAVOR Surveys. <http://breastscreening.cancer.gov/collaborations/favor.html> (accessed 6/1/11)
19. American College of Radiology (ACR). Breast Imaging Reporting and Data System Atlas (BI-RADS® Atlas). Reston; VA: 2003.
20. Miglioretti DL, Gard CC, Carney PA, Onega T, Buist DSM, Sickles EA, Kerlikowske K, Rosenberg RD, Yankaskas BC, Geller BM, Elmore JG. When radiologists perform best: The learning curve in screening mammography interpretation. *Radiology*. 2009; 253:632–640. [PubMed: 19789234]
21. Benichou J, Byrne C, Capece LA, et al. Secular stability and reliability of measurements of the percentage of dense tissue on mammograms. *Cancer Detect Prev*. 2003; 27:266–274. [PubMed: 12893074]

22. Prevral S, Shepherd JA, Smith-Bindman R, Cummings SR, Kerlikowske K. Accuracy of mammographic breast density analysis: results of formal operator training. *Cancer Epidemiol Biomarkers Prev.* 2002; 11:1389–1393. [PubMed: 12433716]
23. Lehman C, Holt S, Peacock S, White E, Urban N. Use of the American College of Radiology BI-RADS guidelines by community radiologists: concordance of assessments and recommendations assigned to screening mammograms. *AJR Am J Roentgenol.* 2002; 179:15–20. [PubMed: 12076896]
24. Kerlikowske K, Grady D, Barclay J, et al. Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. *J Natl Cancer Inst.* 1998; 90:1801–1809. [PubMed: 9839520]
25. Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B, Carozzi F, Catarzi S, Lamberini MP, Marcelli G, Pellizzoni R, Pesce B, Risso G, Russo F, Scorsolini A. Categorizing breast mammographic density: intra- and interobserver reproducibility of BI-RADS density categories. *The Breast.* 2005; 14:269–275. [PubMed: 16085233]
26. Breast Cancer Surveillance Consortium Data Elements, Questionnaires and Definitions. <http://breastscreening.cancer.gov/data/elements.html> (accessed 6/1/11)
27. Rohan TE, Negassa A, Chlebowski RT, Habel L, McTiernan A, Ginsberg M, et al. Conjugated equine estrogen and risk of benign proliferative breast disease: a randomized controlled trial. *J Natl Cancer Inst.* 2008; 100:563–71. [PubMed: 18398105]
28. Collins LC, Baer HJ, Tamimi RM, Connolly JL, Colditz GA, Schnitt SJ. The influence of family history on breast cancer risk in women with biopsy-confirmed benign breast disease: results from the Nurses' Health Study. *Cancer.* 2006; 107:1240–7. [PubMed: 16902983]
29. Webb PM, Byrne C, Schnitt SJ, Connolly JL, Jacobs T, Peiro G, et al. Family history of breast cancer, age and benign breast disease. *Int J Cancer.* 2002; 100:375–8. [PubMed: 12115556]
30. Berg JC, Visscher DW, Vierkant RA, et al. Breast cancer risk in women with radial scars in benign breast biopsies. *Breast Cancer Res Treat.* Mar; 2008 108(2):167–74. [PubMed: 18297395]
31. Worsham MJ, Raju U, Lu M, Kapke A, Cheng J, Wolman SR. Multiplicity of benign breast lesions is a risk factor for progression to breast cancer. *Clin Cancer Res.* 2007; 13:5474–9. [PubMed: 17875777]
32. Degnim AC, Visscher DW, Berman HK, et al. Stratification of breast cancer risk in women with atypia: a Mayo cohort study. *J Clin Oncol.* Jul 1; 2007 25(19):2671–7. [PubMed: 17563394]
33. Menes TS, Kerlikowske K, Jaffer S, Seger D, Miglioretti DL. Rates of atypical ductal hyperplasia have declined with less use of postmenopausal hormone treatment: Findings from the Breast Cancer Surveillance Consortium. *Cancer Epidemiology Biomarkers Prev.* Nov; 2009 18(11): 2822–8.

Figure 1.
Summary of Exclusions/Inclusions of Radiologists, Mammograms and Facilities

Table 1

Radiologists Demographics/Characteristics by How They Use Available Clinical History When Interpreting Screening Mammography Exams.

	N (%)	Use of available Clinical History		
		Never/ Rarely	Yes, abnormality	Yes, Mostly/Always
TOTAL	216 100.0	16 (7.4)	63 (29.2)	137 (63.4)
Demographics				
Physician's age (years)				
34–44	55 (25.5)	4 (7.3)	12 (21.8)	39 (70.9)
45–54	77 (35.6)	5 (6.5)	27 (35.1)	45 (58.4)
≥55	84 (38.9)	7 (8.3)	24 (28.6)	53 (63.1)
Physician's gender				
Male	153 (70.8)	13 (8.5)	47 (30.7)	93 (60.8)
Female	63 (29.2)	3 (4.3)	16 (25.4)	44 (69.8)
Practice Type				
Work full-time				
No	45 (21.1)	5 (11.1)	20 (44.4)	20 (44.4)
Yes (≥40 hours per week)	168 (78.9)	11 (6.5)	43 (25.6)	114 (67.9)
Missing	3			
Affil. with an academic medical center				
No	172 (66.7)	13 (7.6)	51 (29.7)	108 (62.8)
Yes	42 (19.6)	3 (7.1)	12 (28.6)	27 (64.3)
Missing	2			
Breast Imaging Experience				
Fellowship training in breast imaging				
No	200 (92.6)	15 (7.5)	59 (29.5)	126 (63.0)
Yes	16 (7.4)	1 (6.3)	4 (25.0)	11 (68.8)
Years of mammography interpretation				
<10	38 (17.7)	2 (5.3)	13 (34.2)	23 (60.5)
10–19	73 (34.0)	7 (9.6)	18 (24.7)	48 (65.8)
≥20	104 (48.4)	7 (6.7)	32 (30.8)	65 (62.5)
Missing	1			
Time working in breast imaging (%)				
<20%	55 (26.3)	2 (3.6)	15 (27.3)	38 (69.1)
20–39%	55 (26.3)	6 (10.9)	14 (25.5)	35 (63.6)
40–79%	35 (16.7)	2 (5.7)	11 (31.4)	22 (62.9)
80–100%	64 (30.6)	5 (7.8)	21 (32.8)	38 (59.4)
Missing	7			
Number of Mammograms interpreted in 1 year				

	N (%)	Use of available Clinical History		
		Never/ Rarely	Yes, abnormality	Yes, Mostly/Always
<1,000	23 (11.2)	2 (8.7)	5 (21.7)	16 (69.6)
1,001–2,000	71 (34.6)	4 (5.6)	22 (31.0)	45 (63.4)
>2,000	111 (54.1)	9 (8.1)	34 (30.6)	68 (61.3)
<i>Missing</i>	<i>11</i>			

All row percents except for first column is column percents

Table 2

Characteristics of Women Obtaining Screening Mammography Exams in 7 U.S. States Included in the Study

	Number of Mammograms N (%)	Number of Cancers N (%)	Cancer Rate per 1000 mamms % (95%CI)	Cancer Detection Rate per 1000 mamms % (95% CI)
TOTAL	1,027,497	4,796	4.7 (4.4, 4.9)	3.9 (3.6, 4.1)
Characteristics of Women				
Age				
40–49	273,184 (0.27)	798 (0.17)	2.9 (2.7, 3.2)	2.2 (2.0, 2.4)
50–59	335,853 (0.33)	1,344 (0.28)	4.0 (3.7, 4.3)	3.4 (3.1, 3.6)
60–69	218,410 (0.21)	1,275 (0.27)	5.8 (5.4, 6.3)	4.9 (4.4, 5.3)
70+	200,050 (0.19)	1,379 (0.29)	6.9 (6.4, 7.4)	5.9 (5.4, 6.4)
Breast Density				
Almost entirely fat	97,675 (0.10)	249 (0.05)	2.5 (2.2, 3)	2.2 (1.8, 2.7)
Scattered fibroglandular tissue	467,406 (0.45)	1,984 (0.41)	4.2 (4, 4.5)	3.7 (3.4, 4.0)
Heterogeneously dense	390,928 (0.38)	2,185 (0.46)	5.6 (5.2, 6)	4.5 (4.2, 4.9)
Extremely dense	71,488 (0.07)	378 (0.08)	5.3 (4.6, 6)	3.8 (3.2, 4.4)
Time since Last Mammogram				
≤ 2 years	812,371 (0.79)	3,533 (0.74)	4.3 (4.1, 4.6)	3.5 (3.3, 3.8)
3 – 4 years	165,614 (0.16)	878 (0.18)	5.3 (4.9, 5.8)	4.6 (4.2, 5.0)
≥ 5 years or No Previous	49,512 (0.05)	385 (0.08)	7.8 (7, 8.7)	7.1 (6.3, 8.0)
Clinical History Risk Factors				
Current HT Use				
No	752,275 (0.77)	3,328 (0.75)	4.4 (4.2, 4.7)	3.7 (3.4, 3.9)
Yes	219,183 (0.23)	1,104 (0.25)	5.0 (4.5, 5.6)	4.0 (3.6, 4.5)
Missing	56,039	364		
Family History of Breast Cancer				
None	794,342 (0.84)	3,365 (0.78)	4.2 (4, 4.5)	3.5 (3.3, 3.8)
Any	152,273 (0.16)	969 (0.22)	6.4 (5.9, 6.9)	5.1 (4.7, 5.6)
Missing	80,882	462		
Prior Benign Breast Biopsy				
No	791,981 (0.79)	3,284 (0.70)	4.1 (3.9, 4.4)	3.5 (3.2, 3.7)
Yes	211,097 (0.21)	1,394 (0.30)	6.6 (6.2, 7)	5.3 (5.0, 5.7)
Missing	24,419	118		
Summary Clinical History Risk Factor Score (# of Risk Factors)*				
HT use, Family History, and Biopsy				
0	565,239 (0.55)	2,192 (0.46)	3.9 (3.6, 4.1)	3.3 (3.0, 3.5)
1	365,404 (0.36)	1,917 (0.40)	5.2 (4.9, 5.6)	4.3 (4.0, 4.7)
2+	96,854 (0.09)	687 (0.14)	7.1 (6.5, 7.8)	5.6 (5.0, 6.2)

	Number of Mammograms N (%)	Number of Cancers N (%)	Cancer Rate per 1000 mamms % (95%CI)	Cancer Detection Rate per 1000 mamms % (95% CI)
Characteristics of Radiologists				
Years interpreting mammography				
<10	269,654 (0.26)	1,314 (0.28)	4.9 (4.5, 5.3)	4.0 (3.7, 4.4)
10–19	439,556 (0.43)	2,165 (0.45)	4.9 (4.6, 5.3)	4.1 (3.8, 4.5)
≥20	313,432 (0.31)	1,294 (0.27)	4.1 (3.6, 4.7)	3.3 (2.9, 3.9)
<i>Missing</i>	4,855	23		
Number of Mammograms interpreted in 1 year				
< 1,000	39,175 (0.04)	176 (0.04)	4.5 (3.9, 5.2)	3.8 (3.2, 4.6)
1,001–2,000	189,770 (0.20)	918 (0.20)	4.8 (4.4, 5.3)	4.1 (3.7, 4.6)
>2,000	732,335 (0.76)	3,449 (0.76)	4.7 (4.4, 5.1)	3.9 (3.5, 4.2)
<i>Missing</i>	66,217	253		
Work full-time				
No	247,973 (0.25)	1,197 (0.25)	4.8 (4.3, 5.4)	4.1 (3.6, 4.6)
Yes (≥40 hours per week)	762,596 (0.75)	3,519 (0.75)	4.6 (4.3, 4.9)	3.8 (3.5, 4.1)
<i>Missing</i>	16,928	80		
Use of Clinical History				
Never/Rarely	56,651 (0.06)	269 (0.06)	4.7 (4.2, 5.3)	4.1 (3.6, 4.6)
Yes, when an abnormality is noted;	309,707 (0.30)	1,448 (0.30)	4.7 (4.4, 5.0)	3.8 (3.5, 4.1)
Yes, most of the time or always when information is available	661,139 (0.64)	3,079 (0.64)	4.7 (4.1, 5.3)	3.9 (3.4, 4.5)

†Breast Density is a risk factor assigned by radiologist and not part of the clinical history presented to the radiologist

* Risk factor score is based on available clinical history at time of mammogram

Table 3
Interpretive Performance by Women's Characteristics (All Patients Combined) and Clinical History Use Categories

	Sensitivity				False Positive Rate				PPV			
	Clinical History Use				Clinical History Use				Clinical History Use			
	AH Pts	Never/Rarely	Yes, Abnormality	Yes, mostly/ always	All Pts	Never/Rarely	Yes, Abnormality	es, mostly/ always	All Pts	Never/Rarely	Yes, Abnormality	yes, mostly/ always
TOTAL	82.7	85.5	83.6	82.0	8.4	7.9	8.2	8.6	4.4	4.9	4.6	4.3
Woman Level Characteristics												
Age												
40–49	75.2	76.3	71.4	76.5	10.5	9.7	10.1	10.7	2.1	2.2	1.9	2.1
50–59	84.2	82.4	83.6	84.7	8.5	7.7	8.3	8.8	3.8	4.1	3.9	3.7
60–69	83.1	85.4	85.2	81.8	7.5	7.2	7.3	7.7	6.1	7.7	6.5	5.7
70+	85.2	94.1	88.3	83.1	6.5	6.9	6.5	6.5	8.4	7.3	8.6	8.4
Breast Density												
Almost entirely fat	86.3	--	88.5	84.8	4.2	3.8	4.3	4.2	5.0	7.0	4.8	5.0
Scattered fibroglandular tissue	86.7	82.3	89.4	85.9	7.2	7.2	7.0	7.4	4.9	4.7	5.1	4.8
Heterogeneously dense	80.7	87.9	80.5	80.2	10.7	8.9	10.4	10.9	4.1	5.3	4.3	3.9
Extremely dense	71.2	86.7	70.6	70.5	9.9	8.8	9.5	10.1	3.7	3.9	4.0	3.5
Time since Last Mammogram												
≤ 2 years	80.7	82.7	81.6	80.1	8.0	7.4	7.7	8.2	4.2	4.8	4.6	4.0
3–4 years	86.9	96.3	87.9	85.4	9.4	9.0	9.1	9.6	4.7	5.5	4.5	4.7
≥ 5 years or No Previous	91.7	83.3	91.2	92.5	12.3	10.5	12.2	12.5	5.5	5.0	5.2	5.7
Clinical History Risk Factors												
Current HT Use												
No	82.8	85.1	84.0	82.1	8.2	7.5	7.8	8.4	4.3	5.2	4.4	4.2
Yes	79.9	85.2	80.4	79.2	9.1	9.8	8.9	9.2	4.2	4.5	4.5	4.1
Family History of Breast Cancer												
None	82.6	82.4	83.8	82.0	8.4	7.8	8.0	8.7	4.0	4.3	4.3	3.9
Any	80.7	91.8	80.8	79.5	8.8	8.2	8.3	9.1	5.5	7.4	5.6	5.4
Prior Benign Breast Biopsy												

	Sensitivity				False Positive Rate				PPV			
	Clinical History Use				Clinical History Use				Clinical History Use			
	AH Pts	Never/ Rarely	Yes, Abnormality	Yes, mostly/ always	All Pts	Never/ Rarely	Yes, Abnormality	es, mostly/ always	All Pts	Never/ Rarely	Yes, Abnormality	yes, mostly/ always
No	83.2	84.9	84.9	82.3	7.8	7.2	7.5	8.0	4.2	4.9	4.5	4.1
Yes	80.8	87.7	79.8	80.8	10.6	10.1	10.2	10.8	4.8	5.0	4.8	4.3
Risk factor Score (# of Risk Factors) *												
HT use, Family History, and Biopsy												
0	82.8	81.7	87.0	83.2	7.8	7.0	7.6	8.0	4.0	4.4	4.3	3.9
1	81.4	89.6	81.4	82.4	8.8	8.4	8.4	9.1	4.7	5.5	4.7	4.6
2+	76.3	86.5	79.8	77.3	10.7	11.1	10.4	10.8	5.0	5.3	5.2	4.8

--Indicates that either no observations or no cancers in the group

Bold indicates statistical significance ($P < 0.05$) from a conditional logistic regression model only done within all patients column.

Unadjusted rates are presented throughout the table

Study Population: 216 radiologists, 1,027,497 screening mammograms, 90,307 total positive assessments, 4,796 cancers which 3,220 had positive assessments, and 1,022,701 non-cancers with 936,361 negative assessments.

* Risk factor score is based on available clinical history at time of mammogram

Table 4
Odds Ratios and 95% Confidence Intervals from Conditional Logistic Regression Models on Use of Clinical History, Accuracy Indices and Characteristics of Women.

	Sensitivity			False Positive Rate			PPV		
	Never/ Rarely	Yes, Abnormality	Yes, mostly/ always	Never/ Rarely	Yes, Abnormality	Yes, mostly/ always	Never/ Rarely	Yes, Abnormality	Yes, mostly/ always
Woman Level Characteristics									
Age									
40–49	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
50–59	1.7(0.6, 4.8)	2.1(1.4, 3.2)	1.7(1.3, 2.3)	0.8(0.7, 0.8)	0.8(0.8, 0.8)	0.8(0.8, 0.8)	1.9(1.2, 3.0)	2.1(1.7, 2.5)	1.8(1.6, 2.1)
60–69	2.0(0.7, 5.6)	2.3(1.5, 3.5)	1.4(1.0, 1.8)	0.7(0.7, 0.8)	0.7(0.7, 0.7)	0.7(0.7, 0.7)	3.8(2.4, 5.8)	3.6(2.9, 4.3)	2.8(2.5, 3.2)
70+	6.0(1.6, 22.1)	3.1(2.0, 4.8)	1.6(1.2, 2.2)	0.7(0.6, 0.7)	0.6(0.6, 0.6)	0.6(0.6, 0.6)	3.7(2.3, 5.8)	4.7(3.9, 5.8)	4.3(3.8, 4.9)
Breast Density									
Almost entirely fat	--	0.9(0.4, 2.0)	1.0(0.6, 1.6)	0.6(0.5, 0.7)	0.5(0.5, 0.6)	0.5(0.5, 0.6)	1.7(0.7, 4.0)	0.9(0.7, 1.2)	1.2(1.0, 1.4)
Scattered fibroglandular tissue	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Heterogeneously dense	1.8(0.8, 4.0)	0.4(0.3, 0.6)	0.6(0.4, 0.7)	1.3(1.2, 1.4)	1.5(1.4, 1.5)	1.5(1.4, 1.5)	1.3(0.9, 1.7)	0.8(0.7, 0.9)	0.8(0.7, 0.9)
Extremely dense	1.1(0.2, 5.9)	0.2(0.1, 0.4)	0.3(0.2, 0.4)	1.1(1.0, 1.3)	1.3(1.3, 1.4)	1.3(1.3, 1.4)	0.9(0.5, 1.6)	0.7(0.6, 0.9)	0.7(0.6, 0.9)
Time since Last Mammogram									
≤ 2 years	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
3 – 4 years	4.7(1.1, 20.6)	1.6(1.1, 2.4)	1.5(1.1, 2.0)	1.1(1.1, 1.2)	1.2(1.2, 1.3)	1.3(1.2, 1.3)	1.2(0.8, 1.6)	0.9(0.8, 1.1)	1.1(1.0, 1.2)
≥ 5 years or No Previous	1.1(0.3, 4.1)	2.4(1.2, 4.8)	2.9(1.3, 4.7)	1.5(1.3, 1.7)	1.8(1.7, 1.9)	1.7(1.7, 1.8)	1.0(0.6, 1.7)	1.1(0.9, 1.3)	1.3(1.1, 1.5)
Clinical History Risk Factors									
Current HT Use									
No	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Yes	1.1(0.4, 2.9)	0.8(0.6, 1.1)	0.9(0.7, 1.2)	1.2(1.1, 1.3)	1.2(1.2, 1.3)	1.1(1.1, 1.1)	0.9(0.6, 1.3)	1.0(0.9, 1.2)	1.0(0.9, 1.2)
Family History of Breast Cancer									
None	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Any	2.1(0.8, 5.6)	0.8(0.6, 1.2)	0.8(0.6, 1.0)	1.0(0.9, 1.1)	1.0(1.0, 1.1)	1.0(1.0, 1.1)	1.8(1.3, 2.5)	1.3(1.1, 1.5)	1.4(1.3, 1.6)
Prior Benign Breast Biopsy									
No	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Yes	1.1(0.5, 2.5)	0.8(0.6, 1.0)	0.8(0.7, 1.0)	1.4(1.3, 1.5)	1.4(1.3, 1.4)	1.3(1.3, 1.4)	1.0(0.8, 1.4)	1.1(0.9, 1.2)	1.2(1.1, 1.3)

Sensitivity		False Positive Rate				PPV			
		Clinical History Use		Clinical History Use		Clinical History Use		Clinical History Use	
Never/ Rarely	Yes, Abnormality	Yes, Abnormality	Yes, mostly/ always	Never/ Rarely	Yes, Abnormality	Yes, Abnormality	Yes, mostly/ always	Never/ Rarely	Yes, Abnormality
Risk Factor Score (# of Risk Factors) *									
HT use, Family History, and Biopsy									
0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1	1.8(0.8, 4.0)	0.7(0.5, 1.0)	1.0(0.8, 1.2)	1.2(1.1, 1.2)	1.1(1.1, 1.2)	1.1(1.1, 1.1)	1.3(1.0, 1.8)	1.1(1.0, 1.3)	1.3(1.2, 1.4)
2+	1.3(0.4, 4.0)	0.6(0.4, 0.9)	0.7(0.5, 0.9)	1.5(1.4, 1.7)	1.4(1.4, 1.5)	1.3(1.3, 1.4)	1.3(0.8, 1.9)	1.2(1.0, 1.5)	1.3(1.2, 1.5)

-- Indicates that either no observations or no cancers in the group

Bold Indicates statistical significance ($P < 0.05$) from a conditional logistic regression model of an interaction between clinical history and a women characteristic. A single bold item within a row signifies that the magnitude of the Odds Ratios within that use of clinical history is statistically different from both of the other clinical history uses. Two bold values within a row signify that the magnitude of those two Odd ratios are statistically significant.

Study Population: 216 radiologists, 1,027,497 screening mammograms, 90,307 total positive assessments, 4,796 cancers which 3,220 had positive assessments, and 1,022,701 non-cancers with 936,361 negative assessments.

* Risk factor score is based on available clinical history at time of mammogram