Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography

Rebecca A. Hubbard, PhD[1,2], **Karla Kerlikowske, MD**[3,4], **Chris I. Flowers, MD**[5], **Bonnie C. Yankaskas, PhD**[6], **Weiwei Zhu, MS**[1], and **Diana L. Miglioretti, PhD**[1,2]

[1]Biostatistics Unit, Group Health Research Institute, Group Health Cooperative, Seattle, WA

[2]Department of Biostatistics, School of Public Health, University of Washington, Seattle, WA

[3]Department of Epidemiology & Biostatistics, University of California, San Francisco, CA

[4]General Internal Medicine Section, Department of Veterans Affairs, University of California, San Francisco, CA

[5]Department of Diagnostic Imaging, Moffitt Cancer Center and Research Institute, Tampa, FL

[6]Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC

Abstract

Background—False-positive mammography results are common. Biennial screening may decrease the cumulative probability of false-positive results across many years of repeat screening but could also delay cancer diagnosis.

Objective—To compare the cumulative probability of false-positive results and the stage distribution of incident breast cancer after 10 years of annual or biennial screening mammography.

Design—Prospective cohort.

Setting—Seven mammography registries in the National Cancer Institute–funded Breast Cancer Surveillance Consortium.

Participants—169,456 women who received a first screening mammogram at age 40–59 between 1994 and 2006 and 4,492 women with an incident invasive breast cancer diagnosed between 1996 and 2006.

Measurements—False-positive recalls and biopsy recommendations; stage distribution of incident breast cancer.

Results—False-positive recall probability was 16.3% at first and 9.6% at subsequent mammography. False-positive biopsy recommendation probability was 2.5% at first and 1.0% at

subsequent examinations. Availability of comparison films halved the odds of a false-positive recall (adjusted OR 0.50 (CI 0.45, 0.56)). When screening began at age 40, the cumulative probability of a woman receiving at least one false-positive recall after 10 years was 61.3% (95% CI, 59.4% to 63.1%) with annual and 41.6% (CI, 40.6% to 42.5%) with biennial screening. Cumulative probability of false-positive biopsy recommendation was 7.0% (CI, 6.1% to 7.8%) with annual and 4.8% (CI, 4.4% to 5.2%) with biennial screening. Estimates were comparable when screening began at age 50. We observed a non-statistically significant increase in the proportion of late-stage cancers with biennial compared to annual screening (absolute increase 3.3% (CI −1.1, 7.8) age 40–49, 2.3% (CI −1.0, 5.7) age 50–59) among a population of women with incident breast cancer.

Limitations—Few women underwent screening over the entire 10 year period. Radiologist characteristics influence recall rates and were unavailable. Most mammograms were film rather than digital exams. Incident cancers were analyzed in a small population of women who developed cancer.

Conclusions—After 10 years of annual screening, more than half of women will receive at least one false-positive recall, and 7–9% will receive a false-positive biopsy recommendation. Biennial screening appears to reduce the cumulative probability of false-positive results after 10 years but may be associated with a small absolute increase in the probability of being diagnosed with late stage cancer.

INTRODUCTION

Mammography is the only screening test shown to reduce breast cancer mortality in clinical trials (1–5). However, screening a healthy population confers both harms and benefits. False-positive (FP) recalls for additional imaging after screening mammography occur for 14% of women at first screening and for 8% at subsequent exams (2, 6), causing many women inconvenience and anxiety. Recommendations for fine needle aspiration or surgical biopsy after screening mammography are less common (2) but have more severe consequences (7, 8).

Women will undergo 12 screening mammography examinations in their lifetimes if, following updated U.S. Preventive Services Task Force guidelines, they start biennial screening at age 50 and stop at age 74 (9). They will undergo 17 examinations if they start biennial screening at age 40, 24 if they start annual screening at age 50, and 34 if they start annual screening at age 40. Estimates of the probability that a woman will experience at least one FP recall after 10 screening examinations range from 29% to 77% (10–12), and are about 8–9% for benign biopsy (12, 13). These estimates, however, are based on extrapolations, are limited by statistical methodology that assumes women participating in multiple screening rounds are representative of all women recommended for screening, and do not take into account factors shown in prior studies to be associated with wide variability in FP rates, such as radiologist recall rates (14–17) and patient age, breast density, hormone therapy use, and screening interval (6, 15, 18).

To address these limitations, we estimated the cumulative probability of FP recall and biopsy recommendation after 10 years of annual or biennial screening using data from the National Cancer Institute–funded Breast Cancer Surveillance Consortium (BCSC) (19), a nationally representative longitudinal sample of screening mammograms from community practice, and using newer statistical methods that account for duration of observation and informative censoring (20). In addition, we aimed to estimate how patient characteristics and variability in radiologist FP rates might affect cumulative FP probability, and compared the stage at diagnosis of incident breast cancers among women whose preceding screening interval was approximately biennial or annual.

METHODS

Study Population

We used data from seven BCSC mammography registries (http://breastscreening.cancer.gov) (19) (see Appendix A). Registries collected patient characteristics and clinical information at each mammogram, including radiologists' assessments and recommendations based on the American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS®) (21). Each registry is linked to a state cancer registry or regional Surveillance, Epidemiology, and End Results (SEER) program, which we used to determine cancer status following mammography. Six of seven sites also linked to pathology databases. Data were pooled at a central Statistical Coordinating Center. Registries and the Coordinating Center received Institutional Review Board approval for active or passive consenting processes or a waiver of consent to enroll participants, link data, and perform analysis. All procedures were Health Insurance Portability and Accountability Act compliant, and registries and the Coordinating Center received a Federal Certificate of Confidentiality and other protection for the identities of women, physicians, and facilities.

In analyses of FP probabilities we included women who were age 40–59 at first screening mammogram performed at a participating BCSC facility. Mammograms were considered screening exams if the radiologist indicated routine screening. To avoid misclassifying diagnostic mammograms as screening exams, we excluded mammograms when a breast-imaging exam occurred within the prior nine months. We included screening mammograms from 1994 to the most recent year with complete breast cancer capture, which varied from 2004 to 2007 across the seven registries.

A separate cohort was constructed for analyses of cancer stage, which included women age 40–59 at the time of diagnosis of an incident invasive breast cancer between 1996 and 2006, at or following a screening mammogram and who had at least one additional prior mammogram. We excluded women with cancer diagnoses between 1994 and 1996 to allow for capture of mammography up to two years prior to diagnosis (for women undergoing biennial screening). We also excluded women with cancer diagnoses at or after age 60 to focus on incident cancers in 10-year periods similar to those in our analyses of cumulative FP probability (ages 40–49 for women starting screening at age 40, 50–59 for women starting screening at age 50). All breast cancers were classified according to the American Joint Committee on Cancer (AJCC) staging system (22). Late-stage cancer was defined as stage IIB, III, or IV. We also restricted our sample to women with breast cancer diagnoses that occurred within a fixed follow-up period after each woman's most recent prior screening mammogram (the index mammogram): within 1 year for women with 9–18 months separating the index and next most recent prior mammograms (annual screeners); and within 2 years for women with 19–30 months between the index and most recent prior mammograms (biennial screeners). A flow diagram summarizing inclusion and exclusion criteria for the cancer cohort is provided in Appendix B.

Measures and Definitions

Patient characteristics, including birth date, postmenopausal hormone therapy use, and history of breast cancer in a first-degree relative, were collected by questionnaire at each examination. We considered mammograms first examinations if there were no prior mammograms in the BCSC database, no indication of comparison films, and no self-report of a prior mammogram. We defined screening intervals using women's self-report and information from the BCSC database on the date of the prior screening mammogram. Screening interval was categorized into annual (9–18 months), biennial (19–30 months), or

longer than biennial (>30 months). We censored 4,323 (2.5%) of women whose self-reported time since last mammogram differed from that in the database by more than six months, to ensure that women had not obtained a mammogram outside of the BCSC, and we used database-derived time since last mammogram when discrepancies of less than six months occurred. We also censored women at time of cancer diagnosis or the end of the study.

Recall was defined as a BI-RADS assessment for the initial screening mammogram (initial assessment) of 0 (needs additional imaging evaluation); 4 (suspicious abnormality); 5 (highly suggestive of malignancy); or 3 (probably benign finding) with a recommendation for immediate follow-up.

Biopsy recommendation was defined as a final BI-RADS assessment of 4 or 5, or 0 or 3 with a recommendation for biopsy, fine needle aspiration, or surgical consult after all imaging workup and within 90 days of the screening exam (final assessment) (23). Final BI-RADS assessments were set to missing and the exam was excluded from biopsy recommendation analyses if the screening exam was performed at a facility that does not capture follow-up exams or if the final assessment was 0 with recommendation for additional imaging, non-specified workup, or a missing recommendation (N = 25,045, 6.5%).

A recall or biopsy recommendation was considered false-positive when there was no diagnosis of invasive carcinoma or ductal carcinoma *in situ* within 1 year of the screening examination or before the next screening mammogram, whichever occurred first.

For our analyses of breast cancer stage, we defined the screening interval (annual or biennial) associated with cancer diagnosis as the interval between screening mammograms immediately preceding the diagnosis, specifically the interval between a woman's most recent screening mammogram prior to the date of cancer diagnosis (index mammogram) and the screening mammogram before that exam. This approach is described in a previous publication (24).

Statistical Analysis

The proportion of mammograms resulting in a FP recall or biopsy recommendation at a single screening was computed for first and subsequent screening rounds. Generalized linear mixed models estimated the effect of age, family history of breast cancer, breast density estimated using density categories defined by the BI-RADS Atlas, postmenopausal hormone therapy use, and year of first exam on odds of a mammogram resulting in a FP result at a single screening round, while accounting for BCSC registry and random variation among radiologists. Models for FP results at subsequent rounds also adjusted for availability of comparison films and time since previous mammography. Adjusted FP probabilities were estimated from these models using indirect standardization (25, 26).

We used data from women age 40–59 at first mammography to build a model for the probability of FP results at each screening round using generalized linear mixed models for FP results conditional on screening round number, total number of screening rounds before censoring, the covariates adjusted for in analyses of single screening rounds, and interpreting radiologist. We then used this to model to estimate the cumulative probability of FP results for a woman who begins screening at age 40 or age 50 after 10 years of screening, by aggregating probabilities at individual screening rounds to the individual woman level. We used the method of Hubbard et al. (20), which accounts for the informative censoring which may arise when the length of time that a subject is under observation is associated with the outcome. We assumed that covariate effects were the same at each screening round,

except for screening interval, which was assumed to have no effect at the first screening round and age, which was allowed to have a differing effect at first and subsequent screening rounds. We used indirect standardization to ensure a common distribution of BCSC registries across risk profiles (25, 26). We also report unadjusted estimates that do not account for covariates or between-radiologist variation. Appendix C has method details for estimating cumulative FP probabilities.

We report fitted values from our model for combinations of covariates chosen to represent a range of patient characteristics (age, year, hormone therapy use, family history of breast cancer, breast density, registry at the time of their first screening exams) for risk categories of low (no family history of breast cancer, BI-RADS 1 breast density), intermediate (no family history of breast cancer, BI-RADS 2 breast density), high (no family history of breast cancer, BI-RADS 3 breast density), and very high (family history of breast cancer, BI-RADS 3 breast density) and for quartiles of the distribution of radiologist FP rates. We defined the risk categories in terms of family history and BI-RADS breast density based on the results of previous studies of factors associated with FP results (15, 18). All risk profiles are for women who used no postmenopausal hormone therapy and had comparison films available for all subsequent exams. Quartiles for radiologist FP rates were constructed based on radiologist random effects from the FP risk models.

For invasive cancers, we estimated adjusted probabilities of each cancer stage and late-stage cancer using logistic regression models stratified by age at diagnosis (40–49 and 50–59) and including covariates for screening interval, race, family history, and BCSC registry. This set of adjustment variables was selected on the basis of prior research into the relationship between screening interval and risk of late stage cancer (24).

We defined statistical significance using a two-sided alpha-level of 0.05. Analyses were performed in R 2.10.1 (R Foundation for Statistical Computing, Vienna, Austria).

Role of the Funding Source

The National Cancer Institute supported this project through the BCSC cooperative agreements. All study authors and members of the BCSC Steering Committee approved the final version of the manuscript. The authors had full responsibility in designing the study, collecting the data, analyzing and interpreting the data, deciding to submit the manuscript for publication, and writing the manuscript.

RESULTS

We included 386,799 mammograms from 169,456 women interpreted by 997 radiologists. Nearly half the women (47.7%) had only 1 screening mammogram; 11.8% had 5 or more examinations (Table 1). The complete distribution of observed numbers of rounds of screening and patterns of screening intervals are provided in Appendix D. In our cohort, 9,331 women (5.5%) had only one year of follow-up, and 4,891 (2.9%) were observed for 10 or more years.

Most mammograms (78.9%) were for women aged 40–49 years at first mammogram. Median age at first screening was 42 for women who began screening in their forties and 53 for women who began in their fifties. Among subsequent mammograms, 55.6% occurred at an approximately annual screening interval (within 9–18 months of a prior mammogram) and 27.6% occurred approximately biennially (within 18–30 months of a prior mammogram). The remainder of mammograms occurred at longer than biennial intervals.

Most mammograms were assessed as negative (BI-RADS 1) or benign (BI-RADS 2); a BI-RADS score of 0, indicating need for additional imaging, was the third most common initial assessment for first and second mammograms and both age strata (Table 1). Of the 44,992 mammograms with initial BI-RADS scores of 0 across all observed screening rounds, most (71.5%) resolved to negative or benign readings; 12.3% remained a BI-RADS 0, 10.1% had suspicious abnormalities, and data were missing for 6.0% (Appendix E).

Probability of a Mammogram Leading To False-Positive Recall or Biopsy Recommendation

Unadjusted FP recall probability was 16.3% for first and 9.6% for subsequent mammograms. FP recall probabilities were higher for mammograms among women who started screening more recently, those with heterogeneously dense breasts, and in first exams only, older women and those with a family history of breast cancer (Table 2). Availability of comparison films halved the odds of FP recall on subsequent screening exams (odds ratio [OR], 0.50; 95% CI, 0.45 to 0.56), and biennial screening interval (last exam within 19–30 months) increased the risk of FP recall relative to an annual interval (within 9–18 months) (OR, 1.13; CI, 1.08 to 1.19).

Unadjusted FP biopsy recommendation probability was 2.5% for first and 1.0% for subsequent exams. FP biopsy recommendation probabilities were higher for older women and those with heterogeneously dense breasts and in first exams only, higher for those with a family history of breast cancer (Table 3).

Cumulative Probability of a Woman Experiencing a False-Positive Recall or Biopsy Recommendation After 10 Years of Screening

For a woman who starts screening at age 40 the unadjusted cumulative probability of a FP recall after 10 years of screening was 61.3% (CI, 59.4% to 63.1%) with annual and 41.6% (CI, 40.6% to 42.5%) with biennial screening (Table 4). For a woman who starts screening at age 50, the unadjusted probability was 61.3% (CI, 58.0% to 64.7%) under annual and 42.0% (CI, 40.4% to 43.7%) under biennial screening.

The adjusted cumulative probability of a FP recall under biennial screening was less than that under annual screening for each risk profile we modeled (Table 4). For women at intermediate risk of having FP recall (no family history of breast cancer, BI-RADS density 2) who start screening at age 40 and whose films are read by a radiologist with a median FP recall rate, for example, we estimated the cumulative probability of a FP recall after 10 years of biennial screening to be 37.8% compared to 52.4% for annual screening. Estimated reductions in cumulative FP probabilities with biennial compared to annual screening were comparable for women who began screening at age 50.

Estimates of a woman's adjusted cumulative probability of experiencing a FP recall after 10 years of screening increased across FP risk profiles, radiologists' recall rates, and annual compared to biennial screening (Table 4). Within each stratum of FP risk, radiologist risk, and screening frequency, 10-year risk of there was little difference in FP recall associated with age at first mammogram (e.g. 29.4% for low FP risk, low radiologist risk, annual screening begun at age 40, compared to 32.4% for low FP risk, low radiologist risk, annual screening begun at age 50).

Estimates of a woman's adjusted cumulative probability of experiencing a FP biopsy recommendation after 10 years of screening increased across FP risk profiles and with increasing radiologists' FP biopsy recommendation rates (Table 5). Probabilities were higher with annual compared to biennial screening and for older (starting age 50) compared to younger ages (starting age 40) at first mammogram.

At their first screening mammogram, 6.5% of BCSC women had BI-RADS 1 breast density and reported no family history of breast cancer. After 10 years of screening, these women would be expected to have cumulative FP probabilities similar to those of the low FP risk profile. In the BCSC sample, 37.1% had BI-RADS 2 breast density and reported no family history of breast cancer, like the intermediate FP risk profile; 39.2% had BI-RADS 3 breast density and reported no family history of breast cancer, like the high FP risk profile; and 3.7% had BI-RADS 3 breast and reported a family history of breast cancer, like the very high FP risk profile. The remainder of the BCSC sample had characteristics not reflected by the example risk profiles reported in Tables 4 and 5.

Incident Cancers

We identified 4,492 women with an incident invasive breast cancer diagnosis at age 40–59 following an annual or biennial screening interval (Appendix B). After adjusting for family history, race, and BCSC registry, a non-statistically significantly greater proportion of women who were screened biennially were diagnosed with a late-stage cancer (24.6% v. 21.3%, absolute difference 3.3% (−1.1, 7.8) for women age 40–49; 24.6% v. 21.9%, absolute difference 2.3; CI −1.0 to 5.7 for women age 50–59) (Table 6). Odds ratios for model covariates are in Appendix F.

DISCUSSION

In an analysis of probabilities of false-positive recall or biopsy recommendation using registry data collected in community practice, we estimated that the risk of a FP result was higher following a biennial screening interval than an annual interval. However, after 10 years of repeat screening at approximately annual or biennial intervals, the cumulative probability of receiving at least one FP recall or biopsy recommendation was lower with biennial compared to annual screening whether women started screening at age 40 or age 50. Compared to annual screening, biennial screening was associated with a non-statistically significant absolute increase of 2 and 3 percent in the proportion of women diagnosed with late-stage cancer in a cohort of those who developed cancer.

Our estimates of a woman's cumulative probability of a FP mammogram result after repeat screening are higher than previously reported (10, 11, 13, 15). This is partly explained by a higher probability of FP results at each exam for our cohort. Our estimate of the FP recall probability at a single screening round was 16.3% at first exam and 9.6% at subsequent exams, compared to estimates of 6.5% for other cohorts (10). Additionally, previous methods to estimate the cumulative FP probability assumed that censoring was non-informative, which leads to underestimation if women at higher risk of a FP are more likely to be observed for fewer screening rounds (20).

Our analysis identified covariates associated with FP mammography results that resemble previous reports. Positive associations between FP recall and previous breast biopsies, family history of breast cancer, postmenopausal hormone therapy use, more recent exam year, and time between screening exams have been reported previously, as have negative associations between FP recall and older age and comparison film availability (6, 15, 18, 27–29). We also identified statistically significant associations between FP recall and family history of breast cancer, exam year, time since previous mammogram, and availability of comparison films. Surprisingly, we found older age was associated with FP recall only at the first exam. This induced small differences in the cumulative probability of FP recall by starting age.

Few previous studies have estimated the cumulative probability of FP mammography results after repeat screening in U.S. community practice. We searched PubMed Central using the

terms "cumulative", "false positive", and "mammography" to identify all studies evaluating the cumulative probability of FP recall and biopsy recommendation after repeat screening mammography. From among these, we reviewed the titles and abstracts to identify all studies providing estimates of the cumulative FP probability based on screening mammography in the U.S. We then searched all references citing these papers using Web of Science and reviewed titles and abstracts of these manuscripts to identify any additional references we may have missed. This review found 7 studies reporting cumulative FP probabilities for repeat screening mammography in the U.S (10–13, 15, 20, 30). Elmore (1998) reported a 49.1% probability after 10 rounds of screening (10). Christiansen (2000) found a FP probability of 22% after five screening mammograms under biennial screening for an intermediate-risk woman and median radiologist (15), compared to estimates for our population of 38–40%. Studies of benign biopsy have found a probability of 8–9% after 10 screening rounds (12, 13), which are similar to our estimate. Based on our review, we believe ours is the first study to incorporate covariate effects and variation among radiologists into estimates of cumulative FP biopsy recommendation rates.

Our results on the risk of late-stage cancer following annual and biennial screening intervals are similar to those previously reported. A previous BCSC study found a statistically significantly higher proportion of late-stage cancers among women 40–49 participating in biennial compared to annual screening (28% vs. 21%), but no significant difference among women 50–59 (22% vs. 21%) (24). Although we found no statistically significant absolute difference in the overall proportion of late-stage cancers with biennial compared to annual screening, our findings could not exclude an increase in late stage cancer of as much as 7.8% among women in their 40s and as much as 5.7% among women in their 50s based on the upper confidence bound of the estimate of absolute difference. The relatively broad confidence limits around our estimates of difference are likely attributable to the small sample size available for our analysis of incident cancer, and a larger future study is required to exclude the possibility of a clinically significant increase in late stage cancer with biennial compared to annual screening, or even a smaller and less clinically significant decrease.

We have investigated two types of FP mammography results: recall for additional imaging and recommendation for biopsy. Our definitions of FP recall and biopsy recommendation are consistent with the BI-RADS Atlas, which distinguishes these two types of false-positives (21). Previous research on the effects of FP mammograms suggests that women receiving a FP recall or benign biopsy experienced elevated anxiety and distress (31). Benign biopsy poses additional risks of pain and scarring (32, 33). So FP recalls, although common, exert smaller effects than do FP biopsy recommendations. Both the relative frequency and severity of these two types of FP results should be considered when evaluating the harms of screening mammography.

Most screening mammograms had an initial assessment of negative or benign (BIRADS 1 or 2) or of BIRADS 0, needs additional imaging. Most in the latter category resolved on further evaluation to a negative or benign result; about 10% were interpreted as having suspicious abnormalities, and status continued to be unresolved (BIRADS 0) or was missing for about 19%. This could be because the woman did not return for follow-up imaging within 90 days of her screening mammogram or because she went to a facility outside the BCSC. In our analysis these observations have been defined as recalls but have been excluded from biopsy recommendation analyses. If the sub-group with missing final assessments is likely to go on to receive a biopsy recommendation, then this would tend to bias estimates of FP biopsy recommendation downward. However, these missing observations make up only 6% of the total sample, so the magnitude of this bias is expected to be small.

Our study has limitations. Although it was based on a large sample, it included 10 or more rounds of screening for a very small number of women, so our cumulative probability estimates after 10 years of annual screening depend on statistical modeling. However we were able to incorporate information from women with fewer than 10 exams using statistical methods developed for this purpose that accommodate informative censoring; previous methods for estimating cumulative FP probabilities are downwardly biased when FP recall is more common among women with fewer observed rounds of screening, as in our cohort (20).

We lacked information on radiologist characteristics associated with FP recall. Previous research identified variation in interpretive performance by radiologist characteristics such as fellowship training and years of experience as influencing FP recall (14, 16, 17, 34, 35). We attempted to capture differences in radiologist FP rates using random effects to estimate FP recall and biopsy recommendation variability in the middle 50% of radiologists. Variation is even larger when comparing radiologists with the highest and lowest FP rates.

Most mammograms in this analysis were film-screen exams. Digital screening mammography is rapidly becoming the predominant screening modality, with 76.2% of accredited facilities using full field digital machines as of May 1, 2011 (36). However, research on the performance of digital mammography has indicated similar specificity, and hence FP rates, for digital and film-screen exams (37, 38). A slight, non-statistically significant decrease in specificity has been observed for some sub-groups (38). This would result in increased FP probabilities relative to those observed in this study.

The study's cumulative FP risk estimates apply only to the first 10 years of screening. Over the course of a lifetime of screening, beginning screening 10 years earlier would result in an additional 10 screening mammograms under annual screening and 5 under biennial, and the lifetime risk of FP mammography results will thereby be increased. We could not estimate lifetime cumulative FP risks because doing so would require extrapolation beyond the length of observation in the current study. We found no statistical difference in FP recall probabilities among women age 60 and over and those aged 40–44 years, but estimated that FP biopsy recommendation probabilities were statistically significantly higher in women age 65 or older. Therefore, cumulative FP biopsy recommendation probabilities for the ten years beginning at age 60 might be higher than those we have reported for women who began screening at younger ages.

In summary, we estimate that after 10 years of annual screening, a majority of women will receive at least one FP recall, and 7–9% will receive a FP biopsy recommendation. Both probabilities are lowered with biennial screening. In a population of women diagnosed with cancer, we also identified a non-statistically significant increase in the proportion diagnosed with late stage cancer after biennial screening compared to annual. Biennial screening thus decreases risks but may also attenuate the benefits of routine screening. Women and physicians should be aware of the possibility of these harms associated with different screening intervals so they can make informed decisions about screening and be prepared for what to expect when they receive their results. They should also ensure that prior mammograms, when they exist, are available to the interpreting radiologist, as it seems clear from these data that availability of prior studies may halve the odds of a FP recall.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. Annals of Internal Medicine. 2002; 137:347–360. [PubMed: 12204020]

2. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med. 2009; 151(10):727–737. W237-42. [PubMed: 19920273]

3. Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? Radiologic Clinics of North America. 2004; 42:793–806. [PubMed: 15337416]

4. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet. 2002; 359(9310):909–919. [PubMed: 11918907]

5. Tabar L, Vitak B, Chen HH, et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am. 2000; 38(4):625–651. [PubMed: 10943268]

6. Yankaskas BC, Taplin SH, Ichikawa L, et al. Association between mammography timing and measures of screening performance in the United States. Radiology. 2005; 234(2):363–373. [PubMed: 15670994]

7. Montgomery M. Uncertainty during breast diagnostic evaluation: state of the science. Oncol Nurs Forum. 2010; 37(1):77–83. [PubMed: 20044342]

8. Taplin SH, Abraham L, Geller BM, et al. Effect of previous benign breast biopsy on the interpretive performance of subsequent screening mammography. Journal of the National Cancer Institute. 2010; 102(14):1040–1051. [PubMed: 20601590]

9. USPSTF. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2009; 151(10):716–726. W-236. [PubMed: 19920272]

10. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med. 1998; 338(16):1089–1096. [PubMed: 9545356]

11. Xu JL, Fagerstrom RM, Prorok PC, Kramer BS. Estimating the cumulative risk of a false-positive test in a repeated screening program. Biometrics. 2004; 60(3):651–660. [PubMed: 15339287]

12. Blanchard K, Colbert JA, Kopans DB, et al. Long-term risk of false-positive screening results and subsequent biopsy as a function of mammography use. Radiology. 2006; 240(2):335–342. [PubMed: 16864665]

13. Baker SG, Erwin D, Kramer BS. Estimating the cumulative risk of false positive cancer screenings. BMC Med Res Methodol. 2003; 3:11. [PubMed: 12841854]

14. Barlow WE, Chi C, Carney PA, et al. Accuracy of screening mammography interpretation by characteristics of radiologists. Journal of the National Cancer Institute. 2004; 96(24):1840–1850. [PubMed: 15601640]

15. Christiansen CL, Wang F, Barton MB, et al. Predicting the cumulative risk of false-positive mammograms. J Natl Cancer Inst. 2000; 92(20):1657–1666. [PubMed: 11036111]

16. Elmore JG, Jackson SL, Abraham L, et al. Variability in interpretive performance at screening mammography and radiologists' characteristics associated with accuracy. Radiology. 2009; 253(3): 641–651. [PubMed: 19864507]

17. Smith-Bindman R, Chu P, Miglioretti DL, et al. Physician predictors of mammographic accuracy. Journal of the National Cancer Institute. 2005; 97(5):358–367. [PubMed: 15741572]

18. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med. 2003; 138(3):168–175. [PubMed: 12558355]

19. Ballard-Barbash R, Taplin SH, Yankaskas BC, et al. Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database. American Journal of Roentgenology. 1997; 169:1001–1008. [PubMed: 9308451]

20. Hubbard RA, Miglioretti DL, Smith RA. Modeling the cumulative risk of a false-positive screening test. Statistical Methods in Medical Research. 2010; 19:429–449. [PubMed: 20356857]

21. American College of Radiology. Breast Imaging Reporting and Data System (BI-RADS) Breast Imaging Atlas. Reston, VA: American College of Radiology; 2003.

22. American Joint Committee on Cancer. Manual for Staging of Cancer. 6th ed. Philadelphia: JB Lippincott; 2002.

23. Breast Cancer Surveillance Consortium. BCSC Glossary of Terms. BCSC. 2010

24. White E, Miglioretti DL, Yankaskas BC, et al. Biennial versus annual mammography and the risk of late-stage breast cancer. J Natl Cancer Inst. 2004; 96(24):1832–1839. [PubMed: 15601639]

25. Graubard B, Korn E. Predictive margins with survey data. Biometrics. 1999; 55(2):652–659. [PubMed: 11318229]

26. Lane P, Nelder J. Analysis of covariance and standardization as instances of prediction. Biometrics. 1982; 38(3):613–621. [PubMed: 7171691]

27. Cook AJ, Elmore JG, Miglioretti DL, et al. Decreased accuracy in interpretation of community-based screening mammography for women with multiple clinical risk factors. J Clin Epidemiol. 2010; 63(4):441–451. [PubMed: 19744825]

28. Kerlikowske K, Carney PA, Geller B, et al. Performance of screening mammography among women with and without a first-degree relative with breast cancer. Ann Intern Med. 2000; 133(11):855–863. [PubMed: 11103055]

29. Ichikawa LE, Barlow WE, Anderson ML, Taplin SH, Geller BM, Brenner RJ. Time trends in radiologists' interpretive performance at screening mammography from the community-based Breast Cancer Surveillance Consortium, 1996–2004. Radiology. 2010; 256(1):74–82. [PubMed: 20505059]

30. Gelfand AE, Wang F. Modelling the cumulative risk for a false-positive under repeated screening events. Stat Med. 2000; 19(14):1865–1879. [PubMed: 10867676]

31. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. Annals of Internal Medicine. 2007; 146(7):502–510. [PubMed: 17404352]

32. Zagouri F, Sergentanis TN, Gounaris A, et al. Pain in different methods of breast biopsy: Emphasis on vacuum-assisted breast biopsy. Breast. 2008; 17(1):71–75. [PubMed: 17869106]

33. Yazici B, Sever AR, Mills P, Fish D, Jones SE, Jones PA. Scar formation after stereotactic vacuum-assisted core biopsy of benign breast lesions. Clinical Radiology. 2006; 61(7):619–624. [PubMed: 16784949]

34. Beam CA, Lavde PM, Sullivan DC. Variability in the interpretation of screening mammograms by US radiologists. Findings from a national sample. Archives of Internal Medicine. 1996; 156:209–213. [PubMed: 8546556]

35. Elmore JG, Miglioretti DL, Reisch LM, et al. Screening mammograms by community radiologists: variability in false-positive rates. J Natl Cancer Inst. 2002; 94(18):1373–1380. [PubMed: 12237283]

36. FDA. [Accessed May 20, 2011]
http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/
FacilityScorecard/ucm113858.htm

37. Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film
mammography for breast-cancer screening. N Engl J Med. 2005; 353(17):1773–1783. [PubMed:
16169887]

38. Pisano ED, Hendrick RE, Yaffe MJ, et al. Diagnostic accuracy of digital versus film
mammography: exploratory analysis of selected population subgroups in DMIST. Radiology.
2008; 246(2):376–383. [PubMed: 18227537]

Table 1

Characteristics of Women Aged 40–49 and 50–59 Years at Time of First Mammographic Screening, Stratified By First or Second Exam.[*]

	First exam		Second exam[‡]	
	40–49 y n (%)	50–59 y n (%)	40–49 y n (%)	50–59 y n (%)
Total	135,604	33,852	70,740	17,886
Family history of breast cancer				
Yes	7,089 (5.2)	2,181 (6.4)	4,601 (6.5)	1,410 (7.9)
No	108,346 (79.9)	25,877 (76.4)	56,430 (79.8)	13,447 (75.2)
Missing	20,169 (14.9)	5,794 (17.1)	9,709 (13.7)	3,029 (16.9)
BI-RADS[†] breast density				
Almost entirely fat	5,537 (4.1)	2,693 (8.0)	2,825 (4.0)	1,431 (8.0)
Scattered fibroglandular densities	35,372 (26.1)	11,823 (34.9)	19,930 (28.2)	6,768 (37.8)
Heterogeneously dense	43,698 (32.2)	8,115 (24.0)	24,407 (34.5)	4,639 (25.9)
Extremely dense	11,225 (8.3)	1,286 (3.8)	5,810 (8.2)	720 (4)
Missing	39,772 (29.3)	9,935 (29.3)	17,768 (25.1)	4,328 (24.2)
Current hormone therapy use				
Yes	7,292 (5.4)	5,309 (15.7)	5,991 (8.5)	4,071 (22.8)
No	111,777 (82.4)	21,673 (64.0)	56,902 (80.4)	10,826 (60.5)
Missing	16,535 (12.2)	6,870 (20.3)	7,847 (11.1)	2,989 (16.7)
Time since last mammogram				
9–18 months	--	--	28,198 (39.9)	8,247 (46.1)
19–30 months	--	--	22,659 (32)	5,116 (28.6)
>30 months	--	--	19,883 (28.1)	4,523 (25.3)
Comparison film available				
Yes	--	--	53,717 (75.9)	12,286 (68.7)
No	--	--	4,780 (6.8)	1,503 (8.4)
Missing	--	--	12,243 (17.3)	4,097 (22.9)
Year of exam				
<1997	16,162 (11.9)	6,893 (20.4)	1,251 (1.8)	1,009 (5.6)
1997–1999	41,269 (30.4)	11,503 (34)	13,477 (19.1)	5,281 (29.5)
2000–2001	27,001 (19.9)	5,966 (17.6)	16,487 (23.3)	4,185 (23.4)
2002–2004	33,231 (24.5)	6,484 (19.2)	25,678 (36.3)	5,158 (28.8)
>2004	17,941 (13.2)	3,006 (8.9)	13,847 (19.6)	2,253 (12.6)
Initial BI-RADS assessment				
1: Negative	90,505 (66.7)	22,065 (65.2)	48,592 (68.7)	12,443 (69.6)
2: Benign finding(s)	18,518 (13.7)	4,381 (12.9)	13,944 (19.7)	3,650 (20.4)
3: Probably benign	3,989 (2.9)	1,409 (4.2)	573 (0.8)	191 (1.1)
0: Needs additional imaging evaluation	21,205 (15.6)	5,513 (16.3)	7,224 (10.2)	1,488 (8.3)
4: Suspicious abnormality	514 (0.4)	195 (0.6)	138 (0.2)	51 (0.3)
5: Highly suggestive of malignancy	21 (0)	12 (0)	3 (0)	1 (0)

	First exam		Second exam[‡]	
	40–49 y n (%)	50–59 y n (%)	40–49 y n (%)	50–59 y n (%)
Missing	851 (0.6)	277 (0.8)	266 (0.4)	62 (0.3)
Final BI-RADS assessment				
1: Negative	2,699 (2.0)	937 (2.8)	719 (1.0)	201 (1.1)
2: Benign finding(s)	90,855 (67.0)	22,543 (66.6)	48,379 (68.4)	12,413 (69.4)
3: Probably benign	21,824 (16.1)	5,184 (15.3)	14,809 (20.9)	3,830 (21.4)
0: Needs additional imaging evaluation	9,376 (6.9)	2,876 (8.5)	1,967 (2.8)	472 (2.6)
4: Suspicious abnormality	2,861 (2.1)	911 (2.7)	766 (1.1)	173 (1.0)
5: Highly suggestive of malignancy	58 (0)	26 (0.1)	12 (0)	6 (0)
Missing	7,931 (5.8)	1,375 (4.1)	4,088 (5.8)	791 (4.4)
Number of screening rounds observed per woman				
1	64,636 (47.7)	16,194 (47.8)	--	--
2	27,845 (20.5)	6,831 (20.2)	--	--
3	16,463 (12.1)	4,129 (12.2)	--	--
4	10,696 (7.9)	2,626 (7.8)	--	--
5+	15,964 (11.8)	4,072 (12.0)	--	--

[*] Based on 169,456 women who received a first screening mammogram at age 40–59 between 1994 and 2006.

[†] Breast Imaging Reporting and Data System.

[‡] We present characteristics for women at their second screening exam here rather than for all subsequent screening rounds, as in Tables 2 and 3, to present data more representative of the entire population of women who underwent > 1 screening round; presenting data for all subsequent rounds here would skew characteristics toward women who underwent multiple screening rounds.

Table 2

Adjusted False-Positive Recall Probabilities at First and Subsequent Exam by Associated Characteristics with Odds Ratios and 95% Confidence Intervals.

	First exam[**]		Subsequent exam[††]	
	% False positive	OR[*] of a false-positive recall (95% CI[†])	% False positive	OR of a false-positive recall (95% CI)
Age at mammogram (y)				
40–44	16.4	1.0 (Ref)	8.9	1.0 (Ref)
45–49	19.9	1.27 (1.21, 1.33)	9.4	1.07 (1.01, 1.12)
50–54	21.4	1.39 (1.31, 1.47)	8.7	0.98 (0.92, 1.05)
55–59	19.7	1.25 (1.15, 1.36)	8.0	0.89 (0.82, 0.97)
60–64	--	--	8.3	0.92 (0.83, 1.03)
≥65	--	--	9.6	1.09 (0.89, 1.33)
Family history of breast cancer				
Yes	20.5	1.21 (1.13, 1.30)	8.8	0.98 (0.91, 1.05)
No	17.6	1.0 (Ref)	8.9	1.0 (Ref)
BI-RADS[‡] breast density				
Almost entirely fat	11.8	0.62 (0.56, 0.67)	3.7	0.45 (0.40, 0.50)
Scattered fibroglandular densities	17.8	1.0 (Ref)	7.9	1.0 (Ref)
Heterogeneously dense	19.3	1.11 (1.06, 1.16)	10.7	1.40 (1.33, 1.46)
Extremely dense	15.5	0.85 (0.79, 0.91)	9.0	1.16 (1.08, 1.25)
Current hormone therapy use				
Yes	17.3	0.96 (0.90, 1.03)	9.3	1.05 (0.99, 1.12)
No	17.8	1.0 (Ref)	8.9	1.0 (Ref)
Year of first exam				
<1997	13.6	1.0 (Ref)	8.6	1.0 (Ref)
1997–1999	15.7	1.18 (1.07, 1.31)	8.8	1.02 (0.96, 1.08)
2000–2001	17.5	1.35 (1.21, 1.50)	9.1	1.07 (1.00, 1.14)
2002–2004	19.3	1.52 (1.37, 1.69)	9.5	1.12 (1.03, 1.21)
>2004	20.7	1.67 (1.49, 1.86)	11.0	1.31 (1.10, 1.57)
Comparison film available				
Yes	--	--	8.7	0.50 (0.45, 0.56)
No	--	--	15.8	1.0 (Ref)
Time since last mammogram				
9–18 months	--	--	8.3	1.0 (Ref)
19–30 months	--	--	9.3	1.13 (1.08, 1.19)
>30 months	--	--	10.7	1.33 (1.26, 1.40)

[*] Odds ratios are from multivariable logistic regression models, including all variables presented above as well as registry and random radiologist intercepts.

[†] 95% confidence intervals.

[‡] Breast Imaging Reporting and Data System.

[**] Estimates for first exams are based on data for 81,305 women.

[††] Estimates for subsequent exams are based on data for 121,662 mammograms from 57,282 women.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Adjusted False-Positive Biopsy Recommendation Probabilities at First and Subsequent Exam by Associated Characteristics with Odds Ratios and 95% Confidence Intervals.

	First exam		Subsequent exam	
	% False positive	OR* of a false-positive biopsy (95% CI†)	% False positive	OR of a false-positive biopsy (95% CI)
Age at mammogram (y)				
40–44	2.0	1.0 (Ref)	0.8	1.0 (Ref)
45–49	2.8	1.40 (1.24, 1.57)	0.9	1.19 (1.02, 1.39)
50–54	3.5	1.75 (1.53, 2.00)	1.0	1.33 (1.11, 1.60)
55–59	3.0	1.48 (1.23, 1.79)	1.0	1.27 (1.01, 1.59)
60–64	--	--	0.8	1.09 (0.79, 1.50)
≥65	--	--	1.5	1.91 (1.15, 3.16)
Family history of breast cancer				
Yes	3.3	1.47 (1.25, 1.72)	0.8	0.91 (0.73, 1.12)
No	2.3	1.0 (Ref)	0.9	1.0 (Ref)
BI-RADS‡ breast density				
Almost entirely fat	1.6	0.67 (0.54, 0.84)	0.4	0.53 (0.38, 0.76)
Scattered fibroglandular densities	2.3	1.0 (Ref)	0.7	1.0 (Ref)
Heterogeneously dense	2.6	1.12 (1.01, 1.24)	1.1	1.47 (1.28, 1.68)
Extremely dense	2.3	0.98 (0.83, 1.16)	1.2	1.57 (1.28, 1.94)
Current hormone therapy use				
Yes	2.4	1.01 (0.86, 1.20)	0.9	1.01 (0.84, 1.22)
No	2.4	1.0 (Ref)	0.9	1.0 (Ref)
Year of first exam				
<1997	2.2	1.0 (Ref)	0.8	1.0 (Ref)
1997–1999	2.3	1.04 (0.82, 1.32)	1.0	1.14 (0.97, 1.35)
2000–2001	2.3	1.07 (0.84, 1.37)	0.9	1.12 (0.92, 1.37)
2002–2004	2.5	1.16 (0.91, 1.48)	0.8	0.95 (0.75, 1.21)
>2004	2.4	1.09 (0.84, 1.42)	0.9	1.12 (0.62, 2.01)
Comparison film available				
Yes	--	--	0.9	0.70 (0.52, 0.93)
No			1.3	1.0 (Ref)
Time since last mammogram				
9–18 months	--	--	0.8	1.0 (Ref)
19–30 months	--	--	1.0	1.22 (1.05, 1.41)
>30 months	--	--	1.3	1.60 (1.37, 1.86)

*Odds ratios (OR) are from multivariable logistic regression models including all variables presented above as well as registry and random radiologist intercepts.

†95% confidence intervals.

‡Breast Imaging Reporting and Data System.

**

Estimates for first exams are based on data for 79,397 women.

††Estimates for subsequent exams are based on data for 120,447 mammograms from 56,922 women.

Table 4

Cumulative Probability and 95% Confidence Intervals for False-Positive Recall after 10 Years of Screening under Four Screening Strategies (Start Age 40 vs. 50; Annual vs. Biennial Screening) by Radiologist's and Woman's Risk Level for False Positives[*]

	Age 40 at first mammogram		Age 50 at first mammogram	
	Annual screener	Biennial screener	Annual screener	Biennial screener
Overall	61.3 (59.4,63.1)	41.6 (40.6,42.5)	61.3 (58.0,64.7)	42.0 (40.4,43.7)
25th Percentile FP risk radiologist				
Low FP risk woman[*]	29.4 (26.5,32.3)	20.1 (18.3,21.9)	32.4 (27.7,37.1)	21.4 (19.0,23.8)
Intermediate FP risk woman	45.4 (42.3,48.5)	32.2 (30.0,34.4)	49.3 (43.8,54.8)	34.0 (31.3,36.7)
High FP risk woman	51.4 (48.1,54.7)	37.0 (34.6,39.4)	55.5 (49.8,61.2)	39.0 (35.9,42.1)
Very high FP risk woman	54.4 (50.9,57.9)	39.4 (36.9,41.9)	58.5 (52.6,64.4)	41.5 (38.2,44.8)
50th Percentile FP risk radiologist				
Low FP risk woman	34.9 (31.8,38.0)	24.1 (21.9,26.3)	38.3 (33.2,43.4)	25.6 (22.9,28.3)
Intermediate FP risk woman	52.4 (49.1,55.7)	37.8 (35.4,40.2)	56.5 (50.8,62.2)	39.8 (36.7,42.9)
High FP risk woman	58.7 (55.4,62.0)	43.2 (40.7,45.7)	62.8 (57.1,68.5)	45.2 (41.9,48.5)
Very high FP risk woman	61.7 (58.0,65.4)	45.8 (43.1,48.5)	65.8 (59.9,71.7)	47.9 (44.4,51.4)
75th Percentile FP risk radiologist				
Low FP risk woman	43.4 (39.7,47.1)	30.6 (28.1,33.1)	47.2 (41.3,53.1)	32.3 (29.2,35.4)
Intermediate FP risk woman	62.4 (59.1,65.7)	46.5 (44.0,49.0)	66.5 (60.8,72.2)	48.6 (45.3,51.9)
High FP risk woman	68.7 (65.4,72.0)	52.3 (49.6,55.0)	72.6 (67.1,78.1)	54.4 (50.9,57.9)
Very high FP risk woman	71.6 (68.1,75.1)	55.1 (52.2,58.0)	75.4 (69.9,80.9)	57.3 (53.6,61.0)

[*] False positive (FP) risk profiles are based on multivariable logistic regression models including age, year of first exam, hormone therapy use, family history of breast cancer, breast density, availability of comparison films, registry, and random radiologist intercepts. Risk profiles have year of first exam in 1997–1999, no HT and comparison films available at subsequent screenings. Levels were defined as:

Low = no family history of breast cancer, BI-RADS 1 breast density

Intermediate = no family history of breast cancer, BI-RADS 2 breast density

High = no family history of breast cancer, BI-RADS 3 breast density

Very high = family history of breast cancer, BI-RADS 3 breast density

Table 5

Cumulative Probability and 95% Confidence Interval for False-Positive Biopsy Recommendation after 10 Years of Screening under Four Screening Strategies (Start Age 40 vs. 50; Annual vs. Biennial Screening) by Radiologist and Woman Risk Level for False Positives.[*]

	Age 40 at first mammogram		Age 50 at first mammogram	
	Annual screener	Biennial screener	Annual screener	Biennial screener
Overall	7.0 (6.1,7.8)	4.8 (4.4,5.2)	9.4 (7.4,11.5)	6.4 (5.6,7.2)
25th Percentile FP risk radiologist				
Low FP risk woman[*]	3.2 (2.4,4.0)	2.4 (1.8,3.0)	4.8 (3.0,6.6)	3.4 (2.4,4.4)
Intermediate FP risk woman	5.0 (4.0,6.0)	3.7 (3.1,4.3)	7.3 (4.9,9.7)	5.3 (4.1,6.5)
High FP risk woman	6.1 (4.9,7.3)	4.5 (3.7,5.3)	9.0 (6.3,11.7)	6.5 (5.1,7.9)
Very high FP risk woman	7.6 (6.0,9.2)	5.6 (4.6,6.6)	11.1 (7.6,14.6)	8.0 (6.2,9.8)
50th Percentile FP risk radiologist				
Low FP risk woman	3.5 (2.5,4.5)	2.6 (2.0,3.2)	5.2 (3.2,7.2)	3.7 (2.7,4.7)
Intermediate FP risk woman	5.4 (4.4,6.4)	4.0 (3.2,4.8)	8.0 (5.5,10.5)	5.7 (4.3,7.1)
High FP risk woman	6.7 (5.3,8.1)	4.9 (4.1,5.7)	9.8 (6.7,12.9)	7.0 (5.4,8.6)
Very high FP risk woman	8.3 (6.5,10.1)	6.1 (4.9,7.3)	12.1 (8.2,16.0)	8.7 (6.7,10.7)
75th Percentile FP risk radiologist				
Low FP risk woman	4.2 (3.2,5.2)	3.0 (2.2,3.8)	6.1 (3.9,8.3)	4.4 (3.2,5.6)
Intermediate FP risk woman	6.4 (5.2,7.6)	4.7 (3.9,5.5)	9.4 (6.5,12.3)	6.8 (5.2,8.4)
High FP risk woman	7.9 (6.3,9.5)	5.8 (4.8,6.8)	11.5 (8.0,15.0)	8.3 (6.5,10.1)
Very high FP risk woman	9.8 (7.8,11.8)	7.2 (5.8,8.6)	14.1 (9.6,18.6)	10.2 (7.8,12.6)

[*] False positive (FP) risk profiles are based on multivariable logistic regression models including age, year of first exam, hormone therapy use, family history of breast cancer, breast density, availability of comparison films, registry, and random radiologist intercepts. Risk profiles have year of first exam in 1997–1999, no HT and comparison films available at subsequent screenings. Levels were defined as:

Low = no family history of breast cancer, BI-RADS 1 breast density

Intermediate = no family history of breast cancer, BI-RADS 2 breast density

High = no family history of breast cancer, BI-RADS 3 breast density

Very high = family history of breast cancer, BI-RADS 3 breast density

Table 6

Adjusted Proportion of Cancer Stage at Diagnosis and 95% Confidence Intervals for Annual and Biennial Screeners‡ Stratified by Age 40–49 or 50–59 at Time of Cancer Diagnosis.

	Age 40–49			Age 50–59		
	Annual screener (N = 934) Adjusted Proportion (95% CI)	Biennial screener (N = 625) Adjusted Proportion (95% CI)	Adjusted Difference in Proportions (95% CI)	Annual screener (N = 1,999) Adjusted Proportion (95% CI)	Biennial screener (N = 934) Adjusted Proportion (95% CI)	Adjusted Difference in Proportions (95% CI)
AJCC stage						
I	55.5 (52.2, 58.7)	48.3 (44.3, 52.3)	−7.1 (−12.4, −1.9)	58.9 (56.7, 61.1)	54.5 (51.2, 57.8)	−4.4 (−8.5, −0.3)
IIA	23.2 (20.6, 26.1)	27.1 (23.6, 30.8)	3.8 (−0.7, 8.4)	21.5 (19.7, 23.4)	23.6 (20.9, 26.5)	2.1 (−1.4, 5.5)
IIB	11.3 (9.4, 13.5)	9.8 (7.7, 12.4)	−1.5 (−4.7, 1.7)	9 (7.8, 10.3)	10.8 (8.8, 13.1)	1.8 (−0.7, 4.3)
III/IV	10.1 (8.3, 12.2)	14.9 (12.3, 18.0)	4.8 (1.3, 8.4)	10.6 (9.3, 12.0)	11.2 (9.2, 13.4)	0.6 (−2.0, 3.1)
Late stage	21.3 (18.7, 24.1)	24.6 (21.3, 28.2)	3.3 (−1.1, 7.8)	19.6 (17.9, 21.4)	21.9 (19.3, 24.8)	2.3 (−1.0, 5.7)

*
Late stage cancer defined as stage IIB, III, or IV.

†
Adjusted proportions comparing annual and biennial mammography are adjusted for family history, race, and BCSC registry.

‡
Annual and biennial screening intervals for this analysis were defined using the single screening interval preceding the cancer diagnosis.