Title
Patient and Radiologist Characteristics Associated With Accuracy of Two Types of Diagnostic Mammograms.

Journal
American Journal of Roentgenology, 205(2)

ISSN
0361-803X

Authors
Jackson, Sara L
Abraham, Linn
Miglioretti, Diana L
et al.

Publication Date
2015-08-01

DOI
10.2214/ajr.14.13672

Peer reviewed

HHS Public Access
Author manuscript
AJR Am J Roentgenol. Author manuscript; available in PMC 2016 March 28.

Published in final edited form as:
AJR Am J Roentgenol. 2015 August ; 205(2): 456–463. doi:10.2214/AJR.14.13672.

Patient and Radiologist Characteristics Associated With Accuracy of Two Types of Diagnostic Mammograms

Sara L. Jackson[1], **Linn Abraham**[2], **Diana L. Miglioretti**[2,3], **Diana S. M. Buist**[2], **Karla Kerlikowske**[4], **Tracy Onega**[5], **Patricia A. Carney**[6], **Edward A. Sickles**[4], and **Joann G. Elmore**[1]

[1]Division of General Internal Medicine, University of Washington, 325 Ninth Ave, Seattle, WA 98104

[2]Group Health Research Institute, Group Health Cooperative, Seattle, WA

[3]Division of Biostatistics, University of California Davis School of Medicine, Davis, CA

[4]Departments of Medicine and Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA

[5]Geisel School of Medicine at Dartmouth, Dartmouth College, Hanover, NH

[6]Departments of Family Medicine and Public Health & Preventive Medicine, Oregon Health and Science University, Portland, OR

Abstract

Objective—Earlier studies of diagnostic mammography found wide unexplained variability in accuracy among radiologists. We assessed patient and radiologist characteristics associated with the interpretive performance of two types of diagnostic mammography.

Materials and Methods—Radiologists interpreting mammograms in seven regions of the United States were invited to participate in a survey that collected information on their demographics, practice setting, breast imaging experience, and self-reported interpretive volume. Survey data from 244 radiologists were linked to data on 274,401 diagnostic mammograms performed for additional evaluation of a recent abnormal screening mammogram or to evaluate a breast problem, between 1998 and 2008. These data were also linked to patients' risk factors and follow-up data on breast cancer. We measured interpretive performance by false-positive rate, sensitivity, and AUC. Using logistic regression, we evaluated patient and radiologist characteristics associated with false-positive rate and sensitivity for each diagnostic mammogram type.

Results—Mammograms performed for additional evaluation of a recent mammogram had an overall false-positive rate of 11.9%, sensitivity of 90.2%, and AUC of 0.894; examinations done to evaluate a breast problem had an overall false-positive rate of 7.6%, sensitivity of 83.9%, and AUC of 0.871. Multiple patient characteristics were associated with measures of interpretive

Address correspondence to S. L. Jackson (sljack@uw.edu).
Supplemental Data: Available online at www.ajronline.org.

performance, and radiologist academic affiliation was associated with higher sensitivity for both indications for diagnostic mammograms.

Conclusion—These results indicate the potential for improved radiologist training, using evaluation of their own performance relative to best practices, and for improved clinical outcomes with health care system changes to maximize access to diagnostic mammography interpretation in academic settings.

Keywords

accuracy; characteristics; diagnostic mammography; patient; radiologist

Women who undergo a diagnostic mammogram have 10 times the likelihood of receiving a diagnosis of breast cancer compared with women who undergo screening mammograms [1]. Optimizing the accuracy of interpretation of these examinations with a high pretest probability of breast cancer is important. Another necessary component of accuracy is minimizing the anxiety, morbidity, and expense of additional evaluation of false-positive results. Patient and radiologist characteristics associated with the accuracy of screening mammography have been extensively evaluated [2–9], but data informing their relationship to interpretation of diagnostic mammography are more limited [10 –12].

Diagnostic mammograms are performed either for additional evaluation of a recent abnormal screening mammogram or to evaluate a breast problem. Only one prior study by the Breast Cancer Surveillance Consortium [10] assessed the interpretive performance of both types of diagnostic mammograms. That analysis focused on 107 radiologists to assess the role of interpretive volume and accuracy, and it did not report patient or radiologist characteristics, besides interpretive volume, that might be associated with accuracy [10]. Two other prior studies within the Breast Cancer Surveillance Consortium evaluated both patient and radiologist characteristics associated with the accuracy of diagnostic mammography to evaluate a breast problem [11, 12] and found that breast density, previous mammography, self-reported breast lump, and radiologists with academic affiliation were associated with interpretive performance. The larger and more recent of these studies included 123 radiologists and 35,885 diagnostic mammograms all performed before 2004 [11]. All of the studies within the Breast Cancer Surveillance Consortium incorporate some overlap of mammography data, including the cumulative larger dataset presented here.

Because of the Mammography Quality Standards Act, radiologists have specific educational and interpretive volume requirements [13, 14], and technology has changed in the 5–10 years since the previous Breast Cancer Surveillance Consortium reports, with the advent of digital mammography and computer-aided detection [15]. In this study, we expand on prior Breast Cancer Surveillance Consortium data by more than doubling the number of radiologists studied to 244, and we include more recent and a much larger number of diagnostic mammography interpretations (104,115 for additional evaluations of a recent mammogram and 170,286 to evaluate a breast problem). We assess both patient and radiologist characteristics associated with two types of diagnostic mammograms: those for additional evaluation of a recent screening mammogram and those to evaluate a breast problem.

Materials and Methods

Data Sources, Radiologist Survey, and Mammography Data

All radiologists interpreting screening mammograms in seven Breast Cancer Surveillance Consortium sites in 2005–2006 were invited to participate in a mailed self-administered survey. The Breast Cancer Surveillance Consortium registries (San Francisco Bay Area, Colorado, North Carolina, New Mexico, New Hampshire, Vermont, and Washington) include data on patient characteristics and mammography examinations performed on women 18 years old and older, linked to Surveillance Epidemiology and End Results tumor registries, state tumor registries, or pathology databases to identify breast cancer diagnoses, details of which have been previously reported [16, 17].

The radiologist survey included demographic information (age and sex), affiliation with an academic medical center, breast imaging experience (fellowship training, years of mammography interpretation, percentage of time spent in breast imaging, and malpractice experience), and self-reported annual interpretive volume (screening and diagnostic). Surveys were mailed between January 1, 2006, and September 30, 2007, and of the 364 eligible radiologists contacted, 257 (71%) consented to participation and completed the survey. One radiologist filled out a survey, but linkage information to mammography data were not available, and 12 radiologists interpreted zero diagnostic mammograms, resulting in 244 radiologists included in the analysis. Additional survey details have been published elsewhere [7].

Mammography data included breast cancer diagnoses and outcomes, as well as a woman's age, time since last mammographic evaluation, BI-RADS mammographic breast density [18], first-degree family history of breast cancer, and self-reported breast symptoms. Only diagnostic mammograms with 1 year of complete follow-up data in which the presence of a breast cancer diagnosis could be ascertained were included. Women with a history of breast cancer were included, and multiple diagnostic mammograms could be linked to the same cancer diagnosis if the diagnosis occurred within 12 months of both. We excluded mammograms performed on women with breast augmentation.

Institutional Review Board Approval

Each registry and the Statistical Coordinating Center received institutional review board 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 HIPAA compliant, and all registries and the Statistical Coordinating Center have received a Federal Certificate of Confidentiality and other protection for the identities of women, physicians, and facilities who are subjects of this research.

Breast Cancer Cases and Interpretive Performance Outcome Measures

Mammograms were classified as positive or negative using standard Breast Cancer Surveillance Consortium definitions based on the final BI-RADS assessment and recommendations assigned by the radiologist at the end of imaging workup [17]. Final assessments of BI-RADS category 4 or 5 were defined as positive. Assessments of BI-

RADS category 0 or 3 with a recommendation of biopsy, fine-needle aspiration, or surgical consultation were also defined as positive. An assessment of BI-RADS category 1 or 2 was defined as negative. An assessment of BI-RADS category 3 without a recommendation for biopsy was also considered negative, such as those with a short-interval follow-up recommendation. A BI-RADS assessment of category 0 with a recommendation for additional imaging, nonspecified work-up, or missing recommendation were considered to be missing the final result. All remaining assessments of BI-RADS category 0 were negative. Women were considered to have breast cancer if a diagnosis of invasive carcinoma or ductal carcinoma in situ occurred within 1 year of the diagnostic mammogram.

The false-positive rate was defined as the percentage of diagnostic examinations with findings interpreted as positive among all women who did not receive a diagnosis of breast cancer within the 1-year follow-up period. Sensitivity was defined as the percentage of diagnostic examinations interpreted as positive among all women who received a diagnosis of breast cancer within the follow-up period.

Statistical Analysis

All statistical procedures were performed separately by type of diagnostic mammogram: additional evaluations of a recent mammogram or to evaluate a breast problem. Univariate and multi-variate associations between outcome (false-positive rate or sensitivity) and patient and radiologist characteristics were determined using logistic regression models estimated via generalized estimating equations with an exchangeable correlation structure to adjust for the correlation within radiologist characteristics. We also estimated ROCs for unadjusted radiologist-specific sensitivity versus false-positive rate within the observed range of false-positive rates for interpretation of diagnostic mammograms.

Patient or radiologist characteristics univariately associated with a performance measure at a $p \leq 0.10$ level were included in multivariable models for that outcome and type of diagnostic mammogram. In a posthoc analysis, models were first adjusted for radiologist characteristics only (significant at the $p \leq 0.10$ level), then for both patient and radiologist characteristics, and finally for patient and radiologist characteristics except for academic affiliation to assess the effect of potential multicolinearity between this variable and other covariates. All models except univariate were adjusted for Breast Cancer Surveillance Consortium registry. All analyses were performed using SAS software (version 9.3, SAS Institute).

Results

Between January 1, 1998, and December 31, 2008, 244 radiologists interpreted 274,401 diagnostic mammograms. Of these, 104,115 were performed for additional evaluation of a recent mammogram (4663 with cancer) and 170,286 were performed for evaluation of a breast problem (7007 with cancer). For radiologists with at least one diagnostic mammogram with and one without cancer, the mean number of diagnostic mammograms performed for additional evaluation of a recent mammogram with cancer was 23 (median, 11 mammograms; range, 1–231 mammograms), and the mean number of diagnostic mammograms performed to evaluate a breast problem was 32.4 (median, 14.5 mammograms; range, 1–631 mammograms).

A total of 28.7% (70 of 244) radiologists were female, and 42% of diagnostic mammograms were interpreted by women (see Table S1, which can be viewed in the *AJR* electronic supplement to this article, available at www.ajronline.org). A total of 19.1% of radiologists had an adjunct or primary academic affiliation and they interpreted 32.4% of the additional evaluations of a recent mammogram and 36.9% of the diagnostic mammograms for evaluations of a breast problem.

Diagnostic mammograms performed for evaluation of a recent abnormal screening mammogram had a false-positive rate of 11.9% and sensitivity of 90.2%; for examinations to evaluate a breast problem, the false-positive rate was 7.6% and sensitivity was 83.9% (see Table S2, which can be viewed in the *AJR* electronic supplement to this article, available at www.ajronline.org). In univariate models for additional evaluations of a recent mammogram, many patient characteristics were associated with a false-positive rate, whereas higher sensitivity was associated with older age and breasts that are not heterogeneously dense (Table S2). For radiologist characteristics, a higher false-positive rate was univariately associated only with fellowship training, whereas higher sensitivity was associated with female sex, academic affiliation, fellowship training, 10–19 years of mammography interpretation, greater percentage of time spent in breast imaging, and greater screening and diagnostic interpretive volume.

For examinations done to evaluate a breast problem in univariate analyses, multiple patient characteristics were associated with both false-positive rate and sensitivity (Table S2). When the indication for the diagnostic mammogram was a breast lump, compared with nipple discharge or pain, interpretations were more likely to be false-positives and have higher sensitivity (Table S2). Radiologist characteristics associated with higher false-positive rates included younger radiologist age, female sex, academic affiliation, and fewer years interpreting mammograms; higher sensitivity was associated with female sex, academic affiliation, fellowship training, 10–19 years of mammography interpretation, and having never been named in a malpractice suit (Table S2).

ROC curves showing the unadjusted radiologist-specific sensitivity versus false-positive rate for additional evaluations of a recent mammogram yielded an AUC of 0.894, and for evaluation of a breast problem the AUC was 0.871 (Figs. 1A and 1B). The figures show a wide range of variability among radiologists, such that for a given false-positive rate, sensitivity varied widely.

Patient and radiology characteristics associated with false-positive rate and sensitivity for diagnostic mammograms performed for additional evaluation of a recent abnormal mammogram in multiply adjusted logistic regression models are shown in Figure 2, and those for diagnostic mammograms performed to evaluate a breast problem are shown in Figure 3. The only radiologist characteristic associated with sensitivity for both types of mammograms was primary academic affiliation.

The odds ratios describing the relationship between radiologist characteristics and interpretive performance remained unchanged when patient characteristics were also adjusted for, with the exception of academic affiliation, which resulted in an increased odds of greater than 20% for sensitivity of both types of diagnostic mammograms. When we estimated full models with all patient and radiologist characteristics (see Fig. 3 for list of all variables), except for academic affiliation, we did not see a meaningful change in the odds ratios of other radiologist characteristics for any outcome (data not shown).

Discussion

To our knowledge, this study represents the most comprehensive evaluation to date of the relationships among patient and radiologist characteristics and interpretive accuracy of two indications for diagnostic mammograms. We found that radiologists who are female and in academic practice interpret a greater relative proportion of diagnostic mammograms compared with male and nonacademic radiologists. The accuracy of interpretation of diagnostic mammograms is higher than that in an earlier study of diagnostic mammography [11] and higher than that of screening mammography [7]. Multiple patient characteristics were associated with false-positive rates and sensitivity for both indications for diagnostic mammograms, but we did not identify any radiologist characteristics associated with false-positive rates. Academic affiliation was the only radiologist factor associated with increased sensitivity for both types of diagnostic mammograms, whereas radiologists practicing for 10–19 years had higher sensitivity when interpreting additional evaluations of a recent mammogram. Most factors inherent to women are not readily modifiable, and academic radiologists cannot realistically interpret all diagnostic mammograms, but these results may inform future training and policy in diagnostic mammography interpretation.

Radiologists who affiliate with an academic center are more likely to have received fellowship training in breast imaging and to receive referrals, including more complex cases, from outside institutions [11]. Because of this increased exposure and specialization in breast imaging and specifically diagnostic mammography, it is perhaps not surprising that their cancer detection ability is greater than that of nonacademic radiologists. The purpose of analyses such as these is to identify modifiable opportunities to improve the accuracy of diagnostic tests, and because academic radiologists interpret only 6.5% of all diagnostic mammograms [11], a dramatic increase in their volume is not the answer. Short of that, these results suggest that a curricular review of training in academic settings for core content might be standardized and made available to all radiologists. This could be attained by applying core content to residency and fellowship curricula and to the 15 hours of continuing medical education in breast imaging per 36 months that all radiologists who interpret mammograms are required by the Mammography Quality Standards Act to take [14]. Recent adoption of Maintenance of Certification programs by the American Board of Medical Specialties offers a venue for self-evaluation of radiologists' own practices relative to best practice and their peers. Maintenance of Certification modules for diagnostic mammography interpretation could incorporate core content, as determined and standardized by leading academic radiologists, and could compare it with best practices in academic centers. This iterative self-reflecting environment may be able to more closely approximate the learning process in academic settings. Academic radiologists' support in improving systems and training for all radiologists who interpret breast imaging could improve the already high accuracy of diagnostic mammography in the United States.

Another tactic to improve accuracy based on this result could be to develop systems for maximizing access to academic interpretation, particularly because medical systems are coordinating services as Accountable Care Organizations. Geography, in a large country like the United States, has been a barrier compared with European countries, where this approach is already used [19], but digital and teleradiology could facilitate this process by referring diagnostic mammograms for a second opinion to breast specialists in academic or high-volume centers. Policy implications of potential models in which diagnostic mammograms were routinely interpreted by high-volume or academic centers would need to address issues regarding reimbursement when cases are passed between medical centers. In addition, if academic institutions dramatically increase their volume of diagnostic mammogram interpretation, there is potential for them to become centers specifically for diagnostic mammography interpretation and thus become less diversified. We expect that with the evolution of Accountable Care Organizations, medical delivery systems will be looking for opportunities to improve quality of care while keeping costs down. High-volume academic radiologists who are more accessible to geographically distant patients, by digital examinations and teleradiology, may be desired by these systems for improving the care of women.

By adjusting for patient characteristics, we were able to assess how the interpretive accuracy of radiologist characteristics changed once factors inherent to women were accounted for statistically. Despite many radiologist attributes being univariately associated with interpretive performance, once patient characteristics were added to statistical models, no radiologist factors were associated with false-positive rate, and only academic affiliation and 10–19 years of interpretive experience remained associated with sensitivity. Adjusting for case mix only served to strengthen the association between academic affiliation and sensitivity, and when academic affiliation was removed from the models, none of the other radiologist characteristics became statistically significantly associated with accuracy measures. Accounting for patients' characteristics when assessing variability is important, and for diagnostic mammography it appears to account for a large part of variability in interpretive accuracy. Measuring health care quality across institutions and among providers, as is occurring across many medical disciplines, should attempt to adequately account for patient characteristics, as challenging as this may prove to be.

Our results for the overall accuracy of diagnostic mammograms are higher than those reported in prior studies; an earlier report on a cohort of Breast Cancer Surveillance Consortium radiologists from three of the seven registries estimated a lower AUC of 0.80 [11]. It is possible that overall accuracy of diagnostic mammography interpretation has improved over time or that the additional registries included in our study have higher performance. Arguing for improvement over time, we did find that radiologists with 10–19 years of experience interpreting mammography had higher sensitivity for additional evaluations of a recent mammogram than those who had been interpreting for 20 years or longer. Ten to 19 years represents the period since 1992 when the Mammography Quality Standards Act was passed, and it is possible that with increased standardization of breast imaging training, accuracy has improved. Other changes over time include the adoption of digital mammography and computer-aided detection; however, studies suggest that increased false-positive results hinder the potential benefit of computer-aided detection technology on overall accuracy [20].

For women with a history of breast biopsy, screening mammograms are more likely to have lower specificity [21], a finding that we also found for both types of diagnostic mammograms. Given that cancer yield for biopsies is about 25% in the United States compared with 50–75% in Europe [22], this study reinforces the need to reduce the frequency of unnecessary benign breast biopsies, because women in the United States are at higher risk of a false-positive result compared with women in Europe.

Prior work suggests that the use of hormone therapy increases the false-positive rate and decreases sensitivity of screening mammograms [23–25], an association potentially mediated by an effect on breast density [3]. We found that, after adjustment for breast density, women using hormone therapy had lower false-positive rates than did nonusers for examinations performed to evaluate a breast problem. These results are reassuring that after adjustment for breast density, hormone therapy use is not associated with higher false-positive results for diagnostic mammograms.

One strength of this study is the opportunity to adjust for many patient characteristics that have been associated with the accuracy of screening examinations and to assess them in relation to radiologist factors. Compared with prior results, relationships among breast density, prior mammography, and the presence of a lump were associated with sensitivity and specificity in consistent patterns in this current analysis [11, 12]. This study benefits from the large volume and comprehensive outcome data available from the Breast Cancer Surveillance Consortium database. Community sampling of the largest number to date of radiologists from multiple geographic areas across the United States increases the generalizability of the findings. The breadth of patient and radiologist attributes available enhanced our ability to account for a wide variety of variables known to influence the accuracy of screening mammography but not yet applied to diagnostic examinations. We also looked at two commonly performed indications for diagnostic mammograms, because examinations for an additional evaluation of a recent examination had not been previously studied in relation to patient and radiologist characteristics.

Study limitations include the self-report of multiple patient and radiologist characteristics, though similar measures have been used in prior studies. These analyses did not validate radiologists' self-reported interpretive volume by confirming the numbers of studies performed at facilities outside the Breast Cancer Surveillance Consortium, as was done in the detailed study by Haneuse et al. [10]. In addition, we do not have detailed information about specific breast symptoms for diagnostic mammograms, and many women reported “none.” This is consistent with prior data within the Breast Cancer Surveillance Consortium [1] and likely includes women who had a history of breast cancer, biopsy, or abnormal mammogram. The mammograms included in this study were all performed before 2009; however, the technology and process of diagnostic mammography interpretation has been relatively stable, with the exception of the addition of digital imaging.

In conclusion, the overall accuracy of diagnostic mammography is high, and multiple patient characteristics are associated with interpretive performance. We did not identify any radiologist characteristic associated with false-positive rates, but academically affiliated radiologists had greater sensitivity. Potential implications for training, based on curricula for those who go in to academia, and approaches that maximize access to diagnostic mammography interpretation in academic settings might be considered on the basis of these results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the participating women, the Breast Cancer Surveillance Consortium investigators, participating mammography facilities, and radiologists for the data they have provided for this study. A list of the Breast Cancer Surveillance Consortium investigators and procedures for requesting Breast Cancer Surveillance Consortium data for research purposes are provided at breast-screening.cancer.gov.

Data analysis was supported by the National Cancer Institute–funded Breast Cancer Surveillance Consortium (grant HHSN261201100031C). Data collection was supported by the National Cancer Institute and Agency for Healthcare Research and Quality (grant R01 CA107623), the National Cancer Institute (grant K05 CA104699), the Breast Cancer Stamp Fund, and the American Cancer Society, made possible by a generous donation from the Longaberger Company's Horizon of Hope Campaign (grants SIRGS-07-271-01, SIRGS-07-272-01, SIRGS-07-273-01, SIRGS-07-274-01, SIRGS-07-275-01, and SIRGS-06-281-01). The collection of cancer and vital status 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 www.breastscreening.cancer.gov/work/acknowledgement. html. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

References

1. Sickles EA, Miglioretti DL, Ballard-Barbash R, et al. Performance benchmarks for diagnostic mammography. Radiology. 2005; 235:775–790. [PubMed: 15914475]
2. Barlow WE, Chi C, Carney PA, et al. Accuracy of screening mammography interpretation by characteristics of radiologists. J Natl Cancer Inst. 2004; 96:1840–1850. [PubMed: 15601640]
3. 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; 138:168–175. [PubMed: 12558355]
4. Smith-Bindman R, Chu P, Miglioretti DL, et al. Physician predictors of mammographic accuracy. J Natl Cancer Inst. 2005; 97:358–367. [PubMed: 15741572]
5. Elmore JG, Carney PA, Abraham LA, et al. The association between obesity and screening mammography accuracy. Arch Intern Med. 2004; 164:1140–1147. [PubMed: 15159273]
6. Paliwal P, Gelfand AE, Abraham L, Barlow W, Elmore JG. Examining accuracy of screening mammography using an event order model. Stat Med. 2006; 25:267–283. [PubMed: 16381074]
7. Elmore JG, Jackson SL, Abraham L, et al. Variability in interpretive performance at screening mammography and radiologists' characteristics associated with accuracy. Radiology. 2009; 253:641–651. [PubMed: 19864507]
8. Buist DS, Anderson ML, Haneuse SJ, et al. Influence of annual interpretive volume on screening mammography performance in the United States. Radiology. 2011; 259:72–84. [PubMed: 21343539]
9. Beam CA, Conant EF, Sickles EA. Association of volume and volume-independent factors with accuracy in screening mammogram interpretation. J Natl Cancer Inst. 2003; 95:282–290. [PubMed: 12591984]

10. Haneuse S, Buist DS, Miglioretti DL, et al. Mammographic interpretive volume and diagnostic mammogram interpretation performance in community practice. Radiology. 2012; 262:69–79. [PubMed: 22106351]

11. Miglioretti DL, Smith-Bindman R, Abraham L, et al. Radiologist characteristics associated with interpretive performance of diagnostic mammography. J Natl Cancer Inst. 2007; 99:1854–1863. [PubMed: 18073379]

12. Barlow WE, Lehman CD, Zheng Y, et al. Performance of diagnostic mammography for women with signs or symptoms of breast cancer. J Natl Cancer Inst. 2002; 94:1151–1159. [PubMed: 12165640]

13. Bassett LW, Monsees BS, Smith RA, et al. Survey of radiology residents: breast imaging training and attitudes. Radiology. 2003; 227:862–869. [PubMed: 12728182]

14.

Mammography Quality Standards Act of 1992. 42 USC §263b

15. Haygood TM, Whitman GJ, Atkinson EN, Nikolova RG, Sandoval SY, Dempsey PJ. Results of a survey on digital screening mammography: prevalence, efficiency, and use of ancillary diagnostic aids. J Am Coll Radiol. 2008; 5:585–592. [PubMed: 18359447]

16. Ballard-Barbash R, Taplin SH, Yankaskas BC, et al. Breast Cancer Surveillance Consortium. A national mammography screening and outcomes database. AJR. 1997; 169:1001–1008. [PubMed: 9308451]

17. Breast Cancer Surveillance Consortium. [Accessed May 8, 2014] National Cancer Institute. website.breastscreening.cancer.gov

18. D'Orsi, CJ.; Mendelson, EB.; Ikeda, DM., et al. Breast Imaging Reporting and Data System: ACR BI-RADS—breast imaging atlas. Reston, VA: American College of Radiology; 2003.

19. Jensen A, Geller BM, Gard CC, et al. Performance of diagnostic mammography differs in the United States and Denmark. Int J Cancer. 2010; 127:1905–1912. [PubMed: 20104518]

20. Fenton JJ, Abraham L, Taplin SH, et al. Effectiveness of computer-aided detection in community mammography practice. J Natl Cancer Inst. 2011; 103:1152–1161. [PubMed: 21795668]

21. Taplin SH, Abraham L, Geller BM, et al. Effect of previous benign breast biopsy on the interpretive performance of subsequent screening mammography. J Natl Cancer Inst. 2010; 102:1040–1051. [PubMed: 20601590]

22. Smith-Bindman R, Ballard-Barbash R, Miglioretti DL, Patnick J, Kerlikowske K. Comparing the performance of mammography screening in the USA and the UK. J Med Screen. 2005; 12:50–54. [PubMed: 15814020]

23. Kavanagh AM, Mitchell H, Giles GG. Hormone replacement therapy and accuracy of mammographic screening. Lancet. 2000; 355:270–274. [PubMed: 10675074]

24. Njor SH, Hallas J, Schwartz W, Lynge E, Pedersen AT. Type of hormone therapy and risk of misclassification at mammography screening. Menopause. 2011; 18:171–177. [PubMed: 20729766]

25. 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; 63:441–451. [PubMed: 19744825]

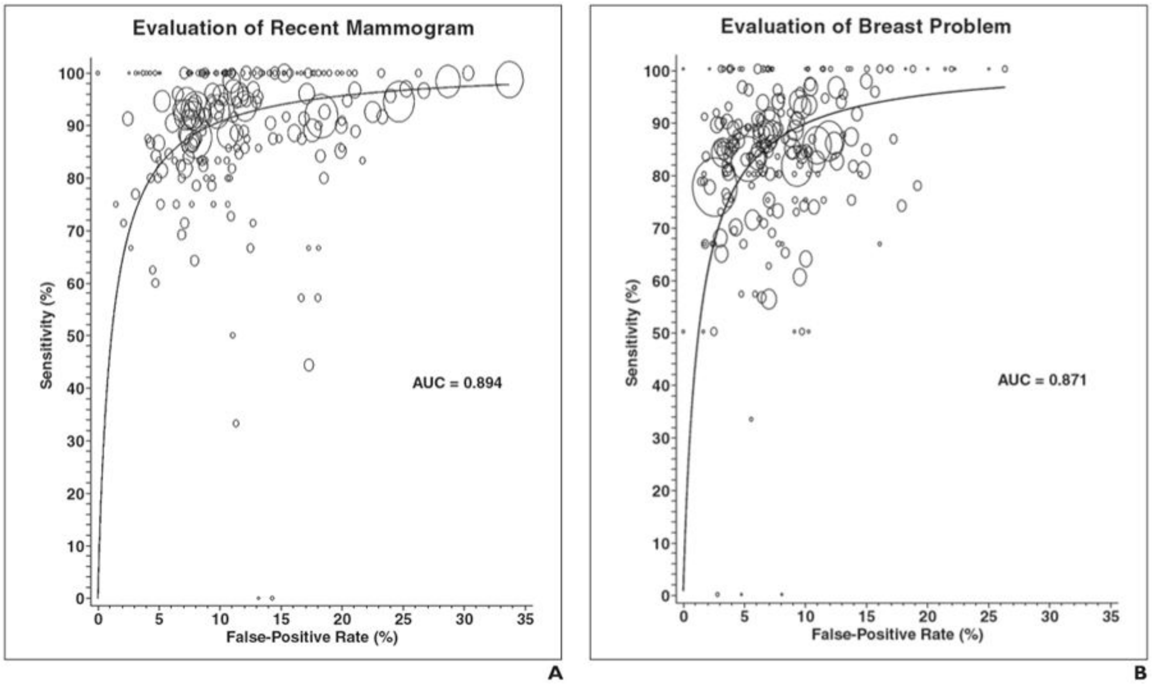

Fig. 1.

Observed (unadjusted) radiologist-specific sensitivity versus false-positive rate and corresponding ROC within observed range of false-positive rates for interpretation of diagnostic mammograms.

A and **B,** Graphs show data for additional evaluation of recent mammogram (**A**) and evaluation of breast problem (**B**). Area of each circle is proportional to number of mammograms from patients with diagnosis of breast cancer that were interpreted by that radiologist. One radiologist with false-positive rate of 60% (based on six diagnostic mammograms, five without cancer during follow-up) was excluded from these figures, but data were included in all other analyses.

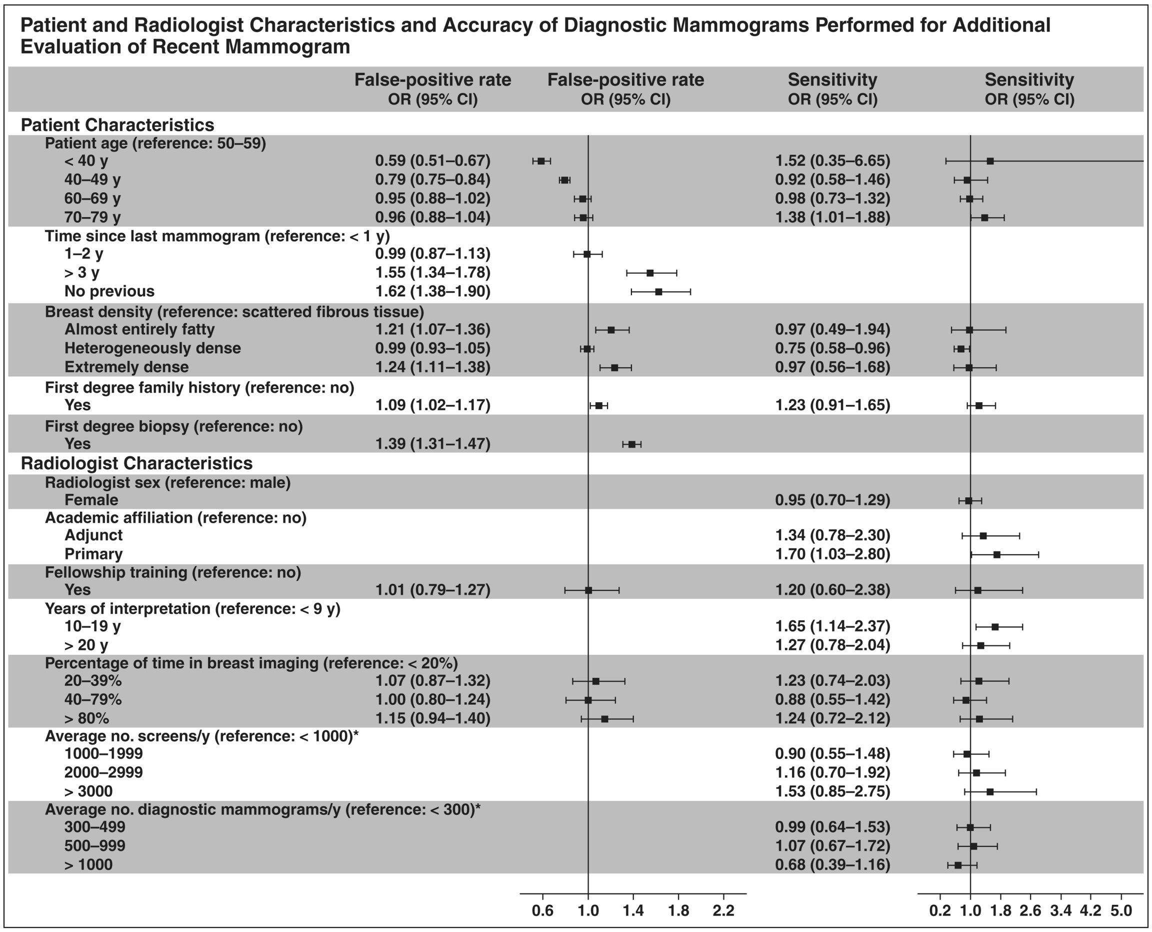

Fig. 2.

Association between patient and radiologist characteristics and false-positive and true-positive (sensitivity) diagnostic mammogram performed for additional evaluation of recent mammogram. Asterisk denotes radiologist self-reported average number of screening or diagnostic mammograms per year over past 5 years. Each model includes all variables listed (within specific column) and also adjusts for mammography registry and correlation within radiologist. OR = odds ratio.

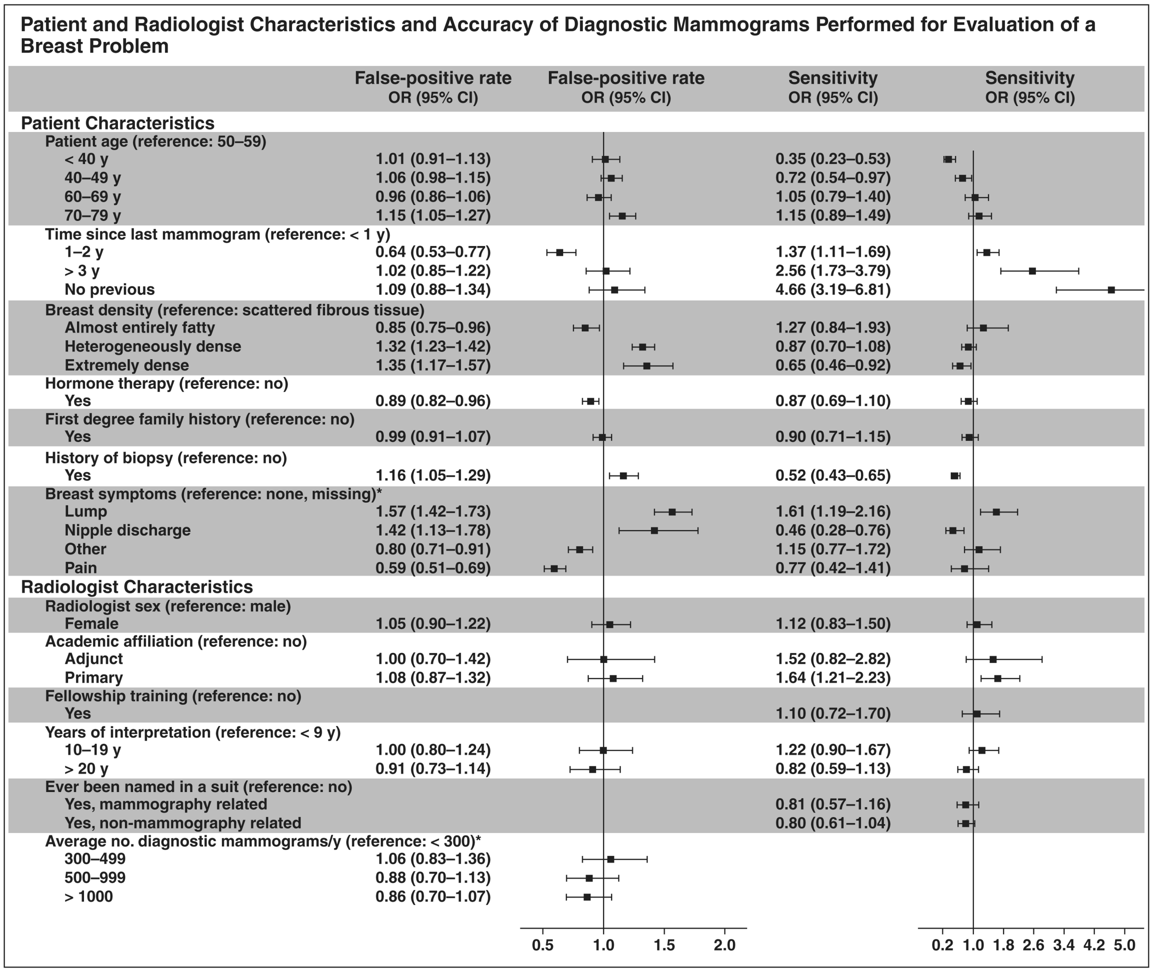

Fig. 3.

Association between patient and radiologist characteristics and false-positive and true-positive (sensitivity) diagnostic mammogram performed for evaluation of breast problem. Asterisk denotes patient self-reported breast symptoms and radiologist self-reported average number of screening or diagnostic mammograms per year over past 5 years. Each model includes all variables listed (within specific column) and also adjusts for mammography registry and correlation within radiologist. OR = odds ratio.