Accuracy of Self-reported Breast Cancer Among Women Undergoing Mammography

Linn Abraham[a], **Berta M. Geller**[b], **Bonnie C. Yankaskas**[c], **Erin J. A. Bowles**[a], **Leah S. Karliner**[d], **Stephen H. Taplin**[e], and **Diana L. Miglioretti**[a,f]

[a] Group Health Center for Health Studies, Seattle, WA

[b] Health Promotion Research, Vermont Cancer Center, University of Vermont, Burlington, VT

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

[d] Department of Medicine, University of California, San Francisco, CA

[e] Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

[f] Department of Biostatistics, University of Washington School of Public Health and Community Medicine, Seattle, WA

Abstract

Purpose—This study estimated the sensitivity and specificity of self-reported breast cancer and their associations with patient factors and pathologic findings using data from the Breast Cancer Surveillance Consortium.

Methods—We included 24,631 women with and 463,804 women without a prior diagnosis of breast cancer who completed a questionnaire (including breast cancer history) at participating U.S. mammography facilities between 1996 and 2006. We determined "true" cancer status using cancer registries and pathology databases. Multivariable logistic regression models were used to examine associations with patient factors and pathologic findings.

Results—Sensitivity of self-reported breast cancer was higher for women with invasive cancer (96.9%) than for those with ductal carcinoma *in situ* (DCIS) (90.2%). Specificity was high overall (99.7%) but much lower for women with a history of lobular carcinoma *in situ (LCIS)* (65.0%). In multivariable models, women reporting older ages, a nonwhite race/ethnicity, or less education had lower sensitivities and specificities. Sensitivity was reduced when there was evidence of prior DCIS, especially when this diagnosis had been made more than 2 years before questionnaire completion. Women reporting a family history of breast cancer had higher sensitivity. Evidence of prior LCIS was associated with lower specificity.

Conclusion—The accuracy of self-reported breast cancer depends on the respondent's characteristics and prior diagnoses. Accuracy is lower among nonwhite women and women reporting less education. There appears to be uncertainty surrounding breast findings such as DCIS and LCIS. These results have important implications for research relying on self-report and for patient communication and care.

Keywords

Breast cancer; Self-report; Mammography; Sensitivity; Specificity

Correspondence: Linn Abraham, Group Health Center for Health Studies, 1730 Minor Ave., Suite 1600, Seattle, WA 98101. Tel: 206-287-2900; fax: 206-287-2871. E-mail: abraham.l@ghc.org.

Introduction

Epidemiologic studies often base their conclusions on self-reported data collected with questionnaires, given their ease in obtaining a large amount of information at relatively low cost. However, the validity of these studies' conclusions depends on the accuracy of the self-reported measures. Many factors beyond the design of the questionnaire, including characteristics of the respondent and the condition or disease being studied, may affect the accuracy of self-report [1,2].

The sensitivity of self-reported cancer history has been associated with patient characteristics such as age, education, and degree of urbanization [2,3]. Associations have also been found with cancer-related variables such as number of previous cancers, type of cancer, time since diagnosis, and treatment used [4,5]. Several studies have shown that women with a previous breast cancer are more likely to report their prior diagnosis accurately (sensitivity, 79%–96%) than women with a cancer in another location [2–6]. Although estimates of the sensitivity of self-reported breast cancer exist, information regarding specificity is scarce. In addition, associations between sensitivity and patient factors have not been studied in detail, nor have they been studied separately for women with *in situ* and invasive cancer.

Establishing whether patient factors are associated with reduced accuracy would benefit epidemiologists, health-services researchers, and clinicians. Lower accuracy may indicate that a woman does not understand her diagnosis, which could affect the timeliness and receipt of follow-up care [7,8] and her level of anxiety. Accuracy of self-report may also have strong implications for research studies (e.g., those predicting cancer risk[9]), particularly if misclassification depends on the outcome of interest or key covariates. The purpose of this study was to estimate the sensitivity and specificity of self-reported history of breast cancer and investigate their associations with patient factors, cancer characteristics, and pathologic findings using population-based data collected by the National Cancer Institute's Breast Cancer Surveillance Consortium (BCSC) [10].

Methods

Data Sources

This study included pooled data from five mammography registries participating in the BCSC [10,11]: Group Health Cooperative (Washington State), the New Hampshire Mammography Network, the New Mexico Mammography Project, the San Francisco Mammography Registry, and the Vermont Breast Cancer Surveillance System. Women completed questionnaires (described below) at the participating radiology facilities, usually at the time of a mammogram. The order and wording of questions differed across registries and changed over time [11]. We used data from questionnaires completed between 1996 and 2006. Women with multiple mammograms may have completed multiple questionnaires during this time period.

Each registry linked data on participating women to their state cancer registry or the Surveillance Epidemiology and End Results (SEER) cancer registry to determine whether they had ever had a diagnosis of breast cancer. Four of the registries also collected benign and malignant biopsy results from pathology databases.

Data were analyzed by a central statistical coordinating center. Each registry and the 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 were compliant with the Health Insurance Portability and Accountability Act, and all registries and the coordinating center received a Federal Certificate

of Confidentiality and other protection for the identities of women, physicians, and facilities who were subjects of this research.

Self-reported Breast Cancer History

With slight variations across mammography registries and over time, most questionnaires asked women "Have you been diagnosed with breast cancer?" and had response options of "No," "Yes," and "Don't know."

Patient Characteristics

Women reported on their questionnaire their demographic and clinical characteristics, including their date of birth, race/ethnicity, education level, zip code, and family history of breast cancer (affected first-degree relative versus none). A woman was considered to live in a rural community if $\geq$ 50% of the area defined by her zip code was deemed rural according to the 2000 U.S. Census data [12]. Receipt of any prior breast procedures (including but not restricted to biopsies, lumpectomies, and mastectomies) was ascertained from self-report, cancer registry data, and pathology data.

Breast Cancer Characteristics

We used the cancer registries and pathology databases to obtain the characteristics of any prior breast cancers, including type (invasive carcinoma or ductal carcinoma *in situ* [DCIS]), years since most recent diagnosis, American Joint Committee on Cancer (AJCC) tumor stage (I, II, III, or IV), and tumor size (<11 mm, 11–15 mm, 16–20 mm, or >20 mm). Tumor stage and size were computed only for invasive cancers. Cancer type, stage, and size were based on the most advanced cancer found before the questionnaire date.

Pathologic Findings

Using the cancer registries and pathology databases, we identified women with lobular carcinoma *in situ* (LCIS) and/or atypical ductal hyperplasia (ADH). These breast findings are considered nonobligate precursors of breast cancer but are not typically considered breast cancer. One mammography registry does not collect data on ADH, but all five registries collect data on LCIS.

Study Samples

We used three study samples to assess different aspects of the accuracy of self-reported breast cancer. Their selection is detailed below and shown in Figure 1.

Sensitivity Sample

We defined sensitivity as the percentage of women with a previous breast cancer documented in the cancer registries or pathology databases who self-reported a prior breast cancer. To estimate sensitivity, we selected all questionnaires from 1996 to 2006 for women who had a prior diagnosis of breast cancer (invasive carcinoma or DCIS) in either the cancer registry or pathology database. We excluded questionnaires completed <1 year after a cancer diagnosis to ensure workup of the cancer was complete at the time of the questionnaire. We also excluded questionnaires from which we could not determine a woman's self-reported history of breast cancer or that were missing any of the key patient factors: race/ethnicity, education, family history of breast cancer, or rural or urban status.

Given the many women and multiple questionnaires per woman, we selected one questionnaire per woman, using random selection to reduce potential bias. Choosing the most recent questionnaire may produce higher estimates of accuracy since the woman has had more time to understand her diagnosis, whereas choosing the questionnaire closest to her diagnosis may

produce lower estimates of accuracy since a woman may not fully understand her diagnosis at that time. Selecting a questionnaire randomly should yield the most representative estimate. After this selection process, the main analysis of sensitivity included 24,631 women.

Specificity Sample

We defined specificity as the percentage of women without a previous documented breast cancer *and* with a prior (assumed accurate) self-report of no breast cancer who, on a subsequent questionnaire, self-reported no prior breast cancer. To estimate specificity, we used the cancer registry and pathology database to identify women who did not have a prior breast cancer. However, some women may have had a cancer that neither source captured, especially those who had a cancer diagnosis before our data capture began. To overcome this concern, we restricted our specificity analysis to women who had at least two questionnaires (from 1996 to 2006) where, at the time of the first questionnaire, we found no prior breast cancer in our databases and the woman gave a response of "No" when asked about her history of breast cancer. If we again found no prior cancer diagnosis before a subsequent questionnaire, we used self-report on the subsequent questionnaire to compute specificity.

We excluded questionnaires from years in which cancer information was incomplete or from which self-reported history of breast cancer could not be determined. Additional questionnaires missing the key patient factors were also excluded. We again randomly selected one questionnaire per woman, resulting in the inclusion of 463,804 women in the analysis of specificity.

Sensitivity Sample, Subanalysis

To test whether we had accurately identified the cancer-free status of our specificity sample, we conducted a subanalysis of sensitivity based on a restricted sample similar to that used for the specificity analysis. That is, we restricted this analysis to women who we were certain initially did not have breast cancer but subsequently received a diagnosis of breast cancer. If this subanalysis produced higher estimates of sensitivity than the main analysis of sensitivity, this could indicate we were also overestimating specificity as a result of our sampling procedure. For this analysis, we randomly selected one questionnaire per woman after an initial accurate self-report of no prior breast cancer and a subsequent diagnosis of breast cancer. The subanalysis of sensitivity included 11,217 women.

Statistical Analyses

We computed raw overall sensitivities and specificities for each patient and pathologic factor. Because sensitivity has been associated with type of breast cancer [5], sensitivity was additionally analyzed separately for women with a prior invasive cancer and women with a prior DCIS.

We computed odds ratios (ORs) and 95% confidence intervals (CIs) from logistic regression models summarizing the association between each factor and the probability of a true positive report (sensitivity) and the probability of a true negative report (specificity). Factors significant at the $P < 0.05$ level in models adjusted only for mammography registry were retained in the final models. To determine the impact of changes in questionnaires over time, we also assessed the effect of adding the year of the questionnaire to these models. The logistic regression models were fit using the SAS procedure GENMOD [13].

Results

Sample Characteristics

Most women were 40 to 69 years old, were white, and had some college education or more (Tables 1 and 2). As expected, women in the analysis of sensitivity were older and more likely to have a family history of breast cancer compared with women in the analysis of specificity.

Sensitivity

The overall sensitivity of self-reported breast cancer was 95.7% (Table 1). Although sensitivity was high, 1,054 women in our sample (4.3%) reported no prior history of breast cancer despite having a breast cancer diagnosis in a cancer registry or pathology database. Of particular note, 36 women who reported having no prior breast cancer had a prior stage III or IV breast cancer.

Sensitivity was slightly lower for women who were aged ≥ 80 years (93.1%), reported a nonwhite race/ethnicity (range, 90.1%–95.0%), had some high school education or less (91.0%), and did not have a family history of breast cancer (95.4%) (Table 1). We did not find a consistent association between sensitivity and stage or tumor size.

Women with a prior invasive cancer diagnosis were more likely to report their diagnosis correctly than those with a prior DCIS diagnosis (96.9% versus 90.2%) (Table 1). The differences in sensitivity across race/ethnicity were even more pronounced in women with a prior DCIS. Women who received a DCIS diagnosis in the 2 years preceding their questionnaire were more likely to report their history of breast cancer than women who received the same diagnosis more than 2 years before their questionnaire.

Among the subset of women with a previous accurate self-report of no prior breast cancer followed by a subsequent breast cancer diagnosis, the overall sensitivity of self-reported breast cancer was 96.2% (Table 1). Although the sensitivities were generally slightly higher in the subanalysis of sensitivity as compared with the main analysis, the relationships were comparable.

Specificity

The specificity of self-reported breast cancer was 99.7% overall and was high for most groups (Table 2). Higher specificities were seen in women reporting younger ages, higher levels of education, and no family history of breast cancer, although these differences were quite small. Lower specificities were found in women with a prior breast procedure (99.0%) or a previous finding of a pathologic risk factor for breast cancer (LCIS, 65.0%; ADH, 95.5%).

Independent Determinants of Sensitivity

Table 3 presents the odds of accurately reporting a prior breast cancer diagnosis from the multivariable models. In the model that included all cancers, women reporting an age of 70 years or older or a nonwhite race/ethnicity had reduced odds of correctly reporting a history of breast cancer compared with women 40–49 and white women, respectively. Women who had a college education or more, or a family history of breast cancer had higher odds of accurately reporting a history of breast cancer than their respective counterparts.

The adjusted odds of correctly reporting a prior history of breast cancer were more than three times higher for women with a prior invasive cancer diagnosis compared with those with a prior DCIS diagnosis (adjusted OR = 3.54; 95% CI: 3.10–4.04) (Table 3). Women with a history of DCIS reporting an African American, Hispanic, Asian/Pacific Islander, or Native American, race/ethnicity were less likely to report a prior history of cancer. Although a higher level of education was associated with increased sensitivity in both models, this association was

stronger in the prior invasive cancer group compared with the prior DCIS group. Family history of breast cancer was strongly associated with sensitivity in women with prior invasive cancer but was not associated with this measure in women with prior DCIS. Women who had received a diagnosis of DCIS 3–9 years before their questionnaire had lower odds of reporting their DCIS compared with those who received this diagnosis within 2 years of their questionnaire.

Independent Determinants of Specificity

The registry that did not collect history of ADH included many women of minority races/ethnicities. As race/ethnicity was an important covariate in this analysis, the final model for specificity included all registries but excluded ADH.

Table 4 presents the odds of correctly reporting no prior breast cancer diagnosis from the multivariable models. The odds of correctly reporting no prior breast cancer diagnosis were lower among women reporting an age of ≥ 80 years compared with their counterparts reporting an age of 40–49 years; a Hispanic ethnicity or an Asian/Pacific Islander race compared with a white non-Hispanic race/ethnicity; or a family history of breast cancer compared with no family history. Women with a college education or more had higher odds of accurately reporting no history of breast cancer than those with some high school education or less. The odds were markedly lower among women who had undergone a prior breast procedure (adjusted OR = 0.15; 95% CI: 0.13–0.16) or had evidence of a prior LCIS (adjusted OR = 0.02; 95% CI: 0.01–0.02) relative to those without these traits.

Discussion

This study found the overall sensitivity and specificity of self-reported breast cancer to be quite high. However, certain groups of women were less accurate at reporting their cancer history, and these populations may benefit from more education and improved communication at the time of breast procedures, diagnosis, or both.

As reported by others [2–6], women who were older, less educated, or of nonwhite race/ethnicity had the lowest sensitivities. These differences could be due to communication issues, differential access to follow-up care, or a different interpretation of the survey question itself. Hence, sensitivity might be increased further if appropriate education and explanation were provided at the time of the diagnosis, administration of the questionnaire, or both. Screening literature suggests individually tailored interactive interventions or linguistically and culturally informed patient navigation may improve the use of screening and follow-up after abnormal screens [14,15]. Similar methods could be used at the time of a diagnosis or survey. For example, culturally appropriate mailed material, a telephone call, or an office visit clearly explaining a woman's diagnosis, in her native language, might increase her chance of understanding her results. Further work is needed to determine whether and how much these approaches might improve knowledge and the accuracy of self-report. Also, as the likelihood of a false-positive mammogram is not negligible [16], it is important to provide greater clarity to women about the meaning of their benign breast findings after such a mammogram.

As in other studies [4,5], women with prior invasive cancer were more likely to report a history of cancer compared to women with a prior DCIS. In addition, women who previously had DCIS and reported a nonwhite race/ethnicity were much more likely to inaccurately indicate that they had not had breast cancer compared with women who previously had invasive breast cancer. Receipt of a DCIS diagnosis more remotely in the past was also associated with a reduced sensitivity compared with receipt of the same diagnosis closer to the time of the questionnaire. These findings suggest that physicians', and subsequently women's, interpretations of DCIS may vary, especially over time. They may also reflect a lack of uniform agreement regarding the best way to view, treat, and discuss DCIS with patients [17–19]. A

diagnosis of DCIS may be confusing, and some affected women may not understand their future risk and the need for follow-up care.

Women without a documented DCIS or invasive cancer were more likely to report they had breast cancer if they had a prior breast procedure or a finding of LCIS. In fact, about 35% of women who had previously had LCIS indicated that they had had breast cancer, although no record of cancer was found in the registry or pathology data. Although LCIS is registered as a cancer in cancer registries, it is usually not treated. Therefore, physicians do not always tell their patients they have cancer. LCIS is thought to increase the risk of developing invasive breast cancer by 3.0- to 4.2-fold, although it is often considered a marker of an increased risk of breast cancer rather than a precursor lesion [20]. In the past, LCIS was treated with mastectomy; therefore, some older women may report this pathologic finding as a cancer. Recently, chemoprevention has been offered to women with LCIS, and some women may have misinterpreted this as cancer treatment [21]. The lower specificity in our analysis reflects the ambiguity surrounding the nature of LCIS and the understandable confusion women may have about this diagnosis.

A self-reported family history of breast cancer was associated with an increased sensitivity and decreased specificity. This suggests that women with a family history may be more likely to respond in the affirmative when asked about a history of breast cancer, or are more likely to understand when they have a breast cancer diagnosis because of familiarity with the disease.

One limitation of this study was the variability of questionnaires among registries and within registries by facility and time. However, adding the year of the questionnaire to the multivariable models did not alter the results. Another limitation is that our sensitivity sample excluded women with double mastectomies since we included only women who returned for a mammogram after diagnosis with invasive cancer or DCIS; however, only 2% of women in the BCSC with a breast cancer diagnosis have double mastectomies, so their exclusion is unlikely to have affected the findings. In addition, our samples were chosen from women receiving mammograms. However, prior literature indicates that at least 81% of women aged 50 years or older have previously had a mammogram [22]. Thus, our estimates are generalizable to many women in the United States. Another limitation is imperfect matching between mammography registry data and cancer registry data, which might potentially lead to underestimations in both sensitivity and specificity. It is also possible that diagnoses were made outside of our catchment areas. To reduce the likelihood of missing such diagnoses, we computed specificity on the subset of women who initially reported having no history of breast cancer.

This study has three major strengths. First, compared with other investigators, we included many more records for women of varying ages and races/ethnicities from diverse geographic locations, increasing the generalizability of these results. Second, our questionnaires prospectively collected extensive information, enabling us to examine associations with patient characteristics and benign breast findings that have not been previously examined. Third, our study estimated both sensitivity and specificity, whereas most similar studies have computed only sensitivity. We were able to estimate specificity by selecting a subset of women who we were fairly certain did not have a previous breast cancer. However, we may have overestimated specificity because these women previously reported they did not have a history of breast cancer, and sensitivity in this subpopulation was slightly higher than sensitivity in the overall sample (96.2% versus 95.7%).

In conclusion, our study shows that women undergoing mammography self-report a personal history of breast cancer with considerable accuracy. However, 4.3% of women with a documented breast cancer and 0.3% of women without a documented breast cancer gave a

response on a questionnaire conflicting with the evidence found in our cancer registry and pathology databases. Researchers using self-reported history should be aware that accuracy varies with patient characteristics, cancer characteristics, and previous benign breast findings. There also seems to be uncertainty surrounding breast findings such as LCIS and DCIS, particularly among nonwhite women and among women more than 3 years past their diagnosis of DCIS. Prior studies have shown that women seen by physicians who explained and documented the necessary follow-up care were more likely to receive the appropriate follow-up [7,8]. Better communication could increase the likelihood that a woman understands her diagnosis and will improve her chance for appropriate treatment and health care in the future. Our findings highlight the importance of clinician-patient communication regarding breast cancer diagnosis, particularly for potentially vulnerable ethnic minority populations.

Acknowledgments

References

1. Newell SA, Girgis A, Sanson-Fisher RW, Savolainen NJ, Hons BA. The accuracy of self-reported health behaviors and risk factors relating to cancer and cardiovascular disease in the general population: a critical review. Am J Prev Med 1999;17:211–229. [PubMed: 10987638]

2. Schrijvers CT, Stronks K, van de Mheen DH, Coebergh JW, Mackenbach JP. Validation of cancer prevalence data from a postal survey by comparison with cancer registry records. Am J Epidemiol 1994;139:408–414. [PubMed: 8109575]

3. Bergmann MM, Calle EE, Mervis CA, Miracle-McMahill HL, Thun MJ, Heath CW. Validity of self-reported cancers in a prospective cohort study in comparison with data from state cancer registries. Am J Epidemiol 1998;147:556–562. [PubMed: 9521182]

4. Desai MM, Bruce ML, Desai RA, Druss BG. Validity of self-reported cancer history: a comparison of health interview data and cancer registry records. Am J Epidemiol 2001;153:299–306. [PubMed: 11157418]

5. Parikh-Patel A, Allen M, Wright WE. California Teachers Study Steering Committee. Validation of self-reported cancers in the California Teachers Study. Am J Epidemiol 2003;157:539–545. [PubMed: 12631544]

6. Dominguez FJ, Lawrence C, Halpern EF, et al. Accuracy of self-reported personal history of cancer in an outpatient breast center. J Genet Couns 2007;16:341–345. [PubMed: 17508275]

7. Poon EG, Haas JS, Puopolo AL, et al. Communication factors in the follow-up of abnormal mammograms. J Gen Intern Med 2004;19:316–323. [PubMed: 15061740]

8. Kerner JF, Yedidia M, Padgett D, et al. Realizing the promise of breast cancer screening: clinical follow-up after abnormal screening among Black women. Prev Med 2003;37:92–101. [PubMed: 12855208]

9. Barlow WE, White E, Ballard-Barbash R, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst 2006;98:1204–1214. [PubMed: 16954473]

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

11. Breast Cancer Surveillance Consortium. National Cancer Institute. 2008 [Accessed 3 Dec 2008]. Available at: http://breastscreening.cancer.gov/

12. United States Census 2000. Geographic Products and Information. 2007 [Accessed 3 Dec 2008]. Available at: http://www.census.gov/geo/www/census2k.html

13. SAS Institute Inc. SAS/STAT® users Guide, version 9. SAS Institute Inc; Cary, NC: 2002.

14. Champion VL, Springston JK, Zollinger TW, et al. Comparison of three interventions to increase mammography screening in low income African American women. Cancer Detect Prev 2006;30:535–544. [PubMed: 17110056]

15. Battaglia TA, Roloff K, Posner MA, Freund KM. Improving follow-up to abnormal breast cancer screening in an urban population. A patient navigation intervention. Cancer 2007;109(2 suppl):359–367. [PubMed: 17123275]

16. 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;338:1089–1096. [PubMed: 9545356]

17. Wiechmann L, Kuerer HM. The molecular journey from ductal carcinoma in situ to invasive breast cancer. Cancer 2008;112:2130–2142. [PubMed: 18383519]

18. Nápoles-Springer AM, Livaudais JC, Bloom J, Hwang S, Kaplan CP. Information exchange and decision making in the treatment of Latina and white women with ductal carcinoma in situ. J Psychosoc Oncol 2007;25:19–36.

19. Prinjha S, Evans J, McPherson A. Women's information needs about ductal carcinoma in situ before mammographic screening and after diagnosis: a qualitative study. J Med Screen 2006;13:110–114. [PubMed: 17007650]

20. Li CI, Malone KE, Saltzman BS, Daling JR. Risk of invasive breast carcinoma among women diagnosed with ductal carcinoma in situ and lobular carcinoma in situ, 1988–2001. Cancer 2006;106:2104–2112. [PubMed: 16604564]

21. Cuzick J, Forbes JF, Sestak I, et al. International Breast Cancer Intervention Study I Investigators (2007) Long-term results of tamoxifen prophylaxis for breast cancer—96-month follow-up of the randomized IBIS-I trial. J Natl Cancer Inst 99:272–282. [PubMed: 17312304]

22. Cronin KA, Yu B, Krapcho M, et al. Modeling the dissemination of mammography in the United States. Cancer Causes Control 2005;16:701–712. [PubMed: 16049809]
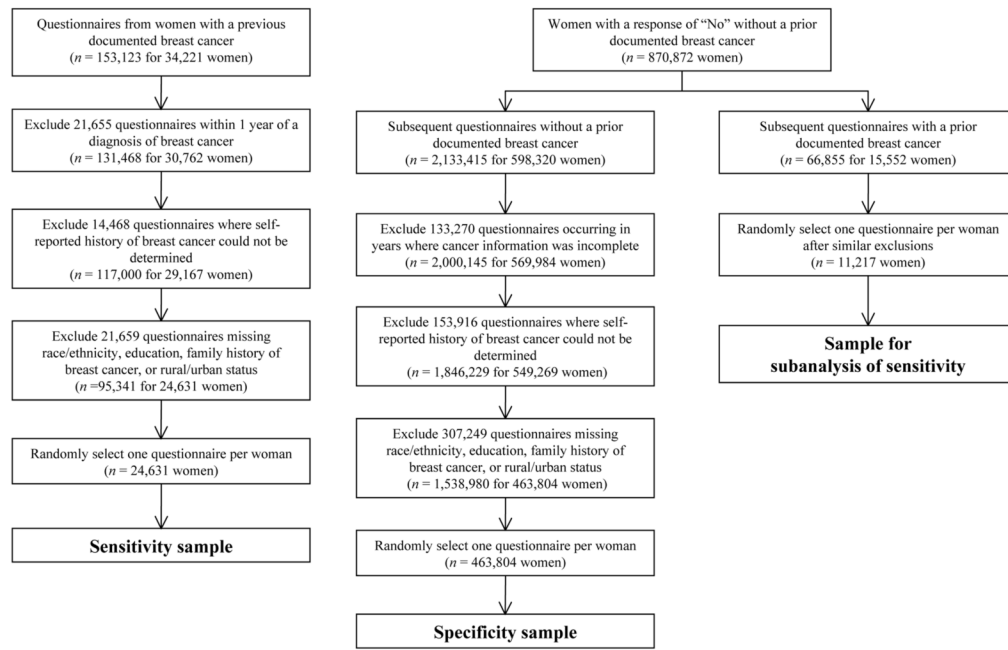
Figure 1.
Study samples (numbers of questionnaires and women remaining after each exclusion)

Table 1

Sensitivity of self-reported history of breast cancer (based on women with questionnaires from 1996–2006 in the Breast Cancer Surveillance Consortium; 24,631 women with a prior breast cancer; 20,303 with a prior invasive cancer, 4,328 with a prior DCIS)

Variable	Distribution No.	Distribution (%)	Sensitivity[a] Main Analysis All %	All 95% CI	Invasive %	Invasive 95% CI	DCIS %	DCIS 95% CI	Sensitivity[a] Subanalysis[b] %	95% CI
Overall	**24,631**	**100.0**	**95.7**	**95.5–96.0**	**96.9**	**96.6–97.1**	**90.2**	**89.3–91.1**	**96.2**	**95.9–96.6**
Residence										
Urban	20,281	82.3	95.5	95.2–95.8	96.8	96.5–97.1	89.5	88.5–90.5	96.2	95.8–96.6
Rural	4,350	17.7	96.6	96.0–97.1	97.3	96.7–97.8	93.4	91.4–95.0	96.3	95.6–97.0
Age at questionnaire, years										
18–39	492	2.0	98.2	96.6–99.2	98.9	97.4–99.6	91.5	79.6–97.6	99.0	96.3–99.9
40–49	3,205	13.0	96.1	95.4–96.8	97.3	96.6–97.9	91.7	89.3–93.7	96.7	95.8–97.5
50–59	6,304	25.6	96.8	96.3–97.2	97.8	97.4–98.2	92.4	90.8–93.8	96.8	96.2–97.4
60–69	6,000	24.4	96.1	95.6–96.6	97.2	96.7–97.7	90.9	89.0–92.6	96.9	96.2–97.5
70–79	5,514	22.4	95.1	94.5–95.7	96.4	95.9–97.0	88.3	86.0–90.3	95.4	94.4–96.2
≥80	3,116	12.7	93.1	92.2–94.0	94.6	93.7–95.4	83.5	79.6–86.9	93.4	91.9–94.8
Race/ethnicity										
White, non-Hispanic	19,425	78.9	96.7	96.4–96.9	97.5	97.3–97.8	92.5	91.6–93.4	97.1	96.7–97.4
African American, non-Hispanic	817	3.3	92.7	90.6–94.3	96.0	94.2–97.3	77.6	69.9–84.0	92.4	88.8–95.1
Hispanic	1,364	5.5	93.4	92.0–94.7	94.8	93.4–96.0	85.2	79.6–89.8	91.7	88.9–94.0
Asian/Pacific Islander	2,415	9.8	91.0	89.8–92.1	93.5	92.3–94.6	82.8	79.4–85.8	92.3	90.4–93.8
Native American	172	0.7	90.1	84.6–94.1	93.2	87.8–96.7	72.0	50.6–87.9	97.4	86.5–99.9
Mixed/other	438	1.8	95.0	92.5–96.8	96.1	93.6–97.9	89.5	80.3–95.3	93.7	89.9–96.4
Education										
Some high school or less	2,246	9.1	91.0	89.7–92.1	92.5	91.3–93.7	82.5	78.2–86.3	93.0	91.2–94.6
High school graduate	5,665	23.0	94.8	94.2–95.3	96.4	95.8–96.9	87.0	84.7–89.0	95.0	94.1–95.8
Some college	6,565	26.7	96.5	96.0–96.9	97.5	97.1–97.9	91.4	89.6–93.0	96.7	96.0–97.3
College or postgraduate	10,155	41.2	96.8	96.4–97.1	97.8	97.4–98.1	92.6	91.3–93.7	97.3	96.8–97.7
Family history of breast cancer										
No	18,818	76.4	95.4	95.1–95.7	96.6	96.3–96.8	89.7	88.6–90.8	96.1	95.7–96.5
Yes	5,813	23.6	96.8	96.4–97.3	98.0	97.5–98.4	91.7	89.8–93.3	96.5	95.7–97.2
Type of prior breast cancer (most advanced)										
Invasive	20,303	82.4	96.9	96.6–97.1					97.5	97.2–97.8
DCIS	4,328	17.6	90.2	89.3–91.1					91.2	90.0–92.4
Time since most recent breast cancer, years										
1–2	9,421	38.2	96.1	95.6–96.4	97.0	96.6–97.3	92.1	90.8–93.3	95.9	95.4–96.3
3–4	5,342	21.7	95.7	95.1–96.2	97.2	96.7–97.7	88.9	86.8–90.8	96.8	96.1–97.4
5–9	5,766	23.4	95.5	94.9–96.0	96.9	96.4–97.4	88.6	86.5–90.5	96.6	95.7–97.4
≥10	4,102	16.7	95.3	94.6–95.9	96.3	95.6–96.9	89.1	86.3–91.6	93.8	69.8–99.8
Tumor size, mm[c]										
<11	4,965	28.6			96.8	96.3–97.3				
11–15	4,212	24.2			97.6	97.1–98.1				
16–20	2,992	17.2			97.7	97.1–98.2				
>20	5,218	30.0			97.3	96.8–97.7				
Missing	2,916									
AJCC tumor stage[d]										
Stage I	9,855	56.2			97.2	96.8–97.5				
Stage II	6,572	37.5			97.5	97.1–97.9				
Stage III or IV	1,100	6.3			96.7	95.5–97.7				
Missing	2,776									

CI, confidence interval; DCIS, ductal carcinoma *in situ*; AJCC, American Joint Committee on Cancer.

[a] Sensitivity = (women with prior breast cancer in database and with a self-report of a prior breast cancer)/(women with a prior breast cancer in database).

[b] Subanalysis is based on questionnaires from women with a diagnosis of breast cancer after an initial accurate report of no prior history of breast cancer (includes 11,217 women).

[c] Based on prior invasive cancer with largest tumor size.

[d] Based on prior invasive cancer with highest AJCC tumor stage.

Table 2

Specificity of self-reported history of breast cancer (based on 463,804 women without a prior breast cancer having questionnaires from 1996–2006 in the Breast Cancer Surveillance Consortium)

Variable	Distribution		Specificity[a]	
	No.	%	%	95% CI
Overall	**463,804**	**100.0**	**99.7**	**99.7–99.7**
Residence				
Urban	334,206	72.1	99.7	99.6–99.7
Rural	129,598	27.9	99.7	99.7–99.7
Age at questionnaire, years				
18–39	14,442	3.1	99.8	99.7–99.9
40–49	148,253	32.0	99.8	99.8–99.8
50–59	145,948	31.5	99.7	99.7–99.8
60–69	81,286	17.5	99.6	99.5–99.6
70–79	54,005	11.6	99.4	99.3–99.4
≥80	19,870	4.3	99.1	99.0–99.2
Race/ethnicity				
White, non-Hispanic	368,453	79.4	99.7	99.7–99.7
African American, non- Hispanic	12,243	2.6	99.5	99.4–99.6
Hispanic	29,512	6.4	99.4	99.3–99.5
Asian/Pacific Islander	41,961	9.0	99.5	99.5–99.6
Native American	2,744	0.6	99.7	99.4–99.8
Mixed/other	8,891	1.9	99.6	99.5–99.7
Education				
Some high school or less	37,971	8.2	99.3	99.2–99.4
High school graduate	108,268	23.3	99.6	99.6–99.7
Some college	124,099	26.8	99.7	99.7–99.7
College or postgraduate	193,466	41.7	99.7	99.7–99.8
Family history of breast cancer				
No	390,425	84.2	99.7	99.7–99.7
Yes	73,379	15.8	99.5	99.4–99.5
Prior breast procedure				
No	346,814	74.8	99.9	99.9–99.9
Yes	116,990	25.2	99.0	99.0–99.1
Evidence of prior LCIS				
No	463,521	99.9	99.7	99.7–99.7
Yes	283	0.1	65.0	59.1–70.6
Evidence of prior ADH				
No	335,564	99.8	99.7	99.7–99.8
Yes	651	0.2	95.5	93.7–97.0
Not collected[b]	127,589			

CI, confidence interval; LCIS, lobular carcinoma *in situ*; ADH, atypical ductal hyperplasia.

[a] Specificity = (women with no prior breast cancer in database and a prior accurate self-report of no breast cancer, with a self-report of no prior breast cancer)/(women with no prior breast cancer in database and a prior accurate self-report of no breast cancer).

[b] One mammography registry did not collect data on benign breast pathology.

Table 3

Logistic regression models predicting an accurate self-report of prior breast cancer (based on women with questionnaires from 1996–2006 in the Breast Cancer Surveillance Consortium; 24,631 women with a prior breast cancer; 20,303 with a prior invasive cancer, 4,328 with a prior DCIS)

Variable	Sensitivity[a] (All) OR	95% CI	Sensitivity[a] (Invasive) OR	95% CI	Sensitivity[a] (DCIS) OR	95% CI
Age at questionnaire, years						
18–39	1.84	0.98–3.95	2.54	1.12–7.29	1.02	0.39–3.54
40–49	1.00		1.00		1.00	
50–59	1.11	0.88–1.40	1.17	0.86–1.59	1.11	0.77–1.57
60–69	0.91	0.72–1.15	0.93	0.69–1.26	0.94	0.65–1.35
70–79	0.70	0.55–0.88	0.71	0.52–0.95	0.76	0.52–1.10
≥80	0.41	0.32–0.52	0.40	0.29–0.55	0.46	0.30–0.69
Race/ethnicity						
White, non-Hispanic	1.00		1.00		1.00	
African American, non-Hispanic	0.51	0.39–0.69	0.70	0.47–1.08	0.37	0.24–0.58
Hispanic	0.74	0.57–0.96	0.82	0.61–1.14	0.62	0.40–0.99
Asian/Pacific Islander	0.41	0.34–0.50	0.38	0.30–0.49	0.47	0.35–0.63
Native American	0.45	0.27–0.79	0.65	0.34–1.37	0.23	0.10–0.62
Mixed/other	0.61	0.40–0.99	0.60	0.36–1.09	0.69	0.34–1.59
Education						
Some high school or less	1.00		1.00		1.00	
High school graduate	1.45	1.19–1.77	1.68	1.32–2.14	1.07	0.75–1.52
Some college	1.96	1.59–2.42	2.16	1.67–2.80	1.61	1.11–2.32
College or postgraduate	2.11	1.71–2.58	2.26	1.76–2.91	1.81	1.26–2.57
Family history of breast cancer						
No	1.00		1.00		1.00	
Yes	1.45	1.23–1.71	1.66	1.34–2.09	1.17	0.91–1.51
Type of prior breast cancer (most advanced)						
Invasive	3.54	3.10–4.04				
DCIS	1.00					
Time since most recent breast cancer, years						
1–2	1.00		1.00		1.00	
3–4	0.95	0.80–1.13	1.14	0.91–1.44	0.71	0.54–0.93
5–9	0.92	0.77–1.09	1.04	0.84–1.29	0.74	0.56–0.97
≥10	0.97	0.80–1.19	1.01	0.80–1.29	0.92	0.65–1.30

OR, odds ratio; CI, confidence interval; DCIS, ductal carcinoma *in situ*.

[a] Model adjusts for mammography registry in addition to all factors listed in the table.

Table 4

Logistic regression model predicting an accurate self-report of no prior breast cancer (based on 463,804 women without a prior breast cancer having questionnaires from 1996–2006 in the Breast Cancer Surveillance Consortium)[a]

Variable	OR	95% CI
Age at questionnaire, years		
18–39	1.10	0.75–1.68
40–49	1.00	
50–59	0.78	0.67–0.92
60–69	0.53	0.45–0.63
70–79	0.37	0.31–0.43
≥80	0.24	0.20–0.29
Race/ethnicity		
White, non-Hispanic	1.00	
African American, non-Hispanic	0.78	0.59–1.04
Hispanic	0.66	0.55–0.80
Asian/Pacific Islander	0.59	0.49–0.71
Native American	0.82	0.45–1.72
Mixed/other	0.74	0.53–1.08
Education		
Some high school or less	1.00	
High school graduate	1.49	1.27–1.76
Some college	1.58	1.33–1.87
College or postgraduate	1.89	1.60–2.24
Family history of breast cancer		
No	1.00	
Yes	0.66	0.59–0.74
Prior breast procedure		
No	1.00	
Yes	0.15	0.13–0.16
Evidence of prior LCIS		
No	1.00	
Yes	0.02	0.01–0.02

OR, odds ratio; CI, confidence interval; LCIS, lobular carcinoma *in situ*.

[a] Model adjusts for mammography registry in addition to all factors listed in the table.