Bias Associated With Self-Report of Prior Screening Mammography

Kathleen A. Cronin,1 Diana L. Miglioretti,3 Martin Krapcho,4 Binbing Yu,4 Berta M. Geller,5 Patricia A. Carney,6 Tracy Onega,7,8 Eric J. Feuer,1 Nancy Breen,2 and Rachel Ballard-Barbash2

1Statistical Research and Applications Branch and 2Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland; 3Group Health Cooperative, Group Health Center for Health Studies, and University of Washington, Department of Biostatistics, Seattle, Washington; 4Information Management Services, Inc., Silver Spring, Maryland; 5Office of Health Promotion Research, University of Vermont, Burlington, Vermont; 6Department of Family Medicine and Public Health and Preventive Medicine, Oregon Health & Science University, Portland, Oregon; 7Department of Community and Family Medicine, Dartmouth Medical School, Hanover and Lebanon, New Hampshire; and 8Norris Cotton Cancer Center, Dartmouth Medical School, Lebanon, New Hampshire

Abstract

Background: Self-reported screening behaviors from national surveys often overestimate screening use, and the amount of overestimation may vary by demographic characteristics. We examine self-report bias in mammography screening rates overall, by age, and by race/ethnicity.

Methods: We use mammography registry data (1999-2000) from the Breast Cancer Surveillance Consortium to estimate the validity of self-reported mammography screening collected by two national surveys. First, we compare mammography use from 1999 to 2000 for a geographically defined population (Vermont) with self-reported rates in the prior two years from the 2000 Vermont Behavioral Risk Factor Surveillance System. We then use a screening dissemination simulation model to assess estimates of mammography screening from the 2000 National Health Interview Survey.

Results: Self-report estimates of mammography use in the prior 2 years from the Vermont Behavioral Risk Factor Surveillance System are 15 to 25 percentage points higher than actual screening rates across age groups. The differences in National Health Interview Survey screening estimates from models are similar for women 40 to 49 and 50 to 59 years and greater than for those 60 to 69, or 70 to 79 (27 and 26 percentage points versus 14, and 14, respectively). Overreporting is highest among African American women (24.4 percentage points) and lowest among Hispanic women (17.9) with non-Hispanic White women in between (19.3). Values of sensitivity and specificity consistent with our results are similar to previous validation studies of mammography. Conclusion: Overestimation of self-reported mammography usage from national surveys varies by age and race/ethnicity. A more nuanced approach that accounts for demographic differences is needed when adjusting for overestimation or assessing disparities between populations. (Cancer Epidemiol Biomarkers Prev 2009;18(6):1699–705)

Introduction

National estimates of the use of cancer screening procedures are based primarily on self-reported results from the National Health Interview Survey (NHIS) and the Behavior Risk Factor Surveillance System (BRFSS). These estimates are well-known to be subject to biases such as social response bias and recall bias (1, 2).

However, quantifying the effect of overreporting on population estimates of screening behavior is difficult. Studies evaluating the validity of self-reported mammography screening have shown varying amounts of overreporting among samples or defined populations that have typically been limited in size. Validation of self-reported mammography use within the past 2 years with medical records from managed care populations have shown reasonably high agreement (approximately 70-88%) with high sensitivity (81-99%) but lower specificity (40-63%); refs. 3-8). Studies focused on minority, low-income or diverse populations have shown a wider range of agreement and variation between race and...
ethnic groups for mammography use (9-15). These results along with a recent meta-analysis (2) suggest that national survey data overestimate mammography screening usage and do not capture important racial/ethnic differences in use. An unbiased method for evaluating mammography usage at the population level is needed that will provide more accurate estimates of disparities among race/ethnic populations.

This study examines the amount of self-report bias in mammography screening rates by age, race, and ethnicity, using population-based longitudinal data on mammography usage collected by the National Cancer Institute’s Breast Cancer Surveillance Consortium (BCSC; ref. 16). More specifically, we examine the validity of survey responses for large, geographically defined populations. First, we consider the state of Vermont. The percent of women living in Vermont who received a mammogram in the years 1999 and 2000 recorded in the Vermont Breast Cancer Surveillance System or in the neighboring areas covered in the New Hampshire Mammography Network is compared with the percent of women self-reported in 2000 that they were screened in the previous 2 years from the Vermont BRFSS. Including mammograms for women who live in Vermont but received a mammogram in New Hampshire allows for a more complete accounting of mammograms received. These data allows us to directly compare self-reported screening rates to actual mammograms recorded in the BCSC data base. This analysis updates a similar comparison done by the Vermont Program for Quality in Health Care for the years 1994 to 1996, which found a 17 percentage point difference between BRFSS self-reported rates of recent mammography use and an estimate from the Vermont Mammography Registry (a component of the Vermont Breast Cancer Surveillance System) for women 52 to 64 years of age.11

Because we are interested in understanding national screening rates and Vermont is not representative of the United States as a whole, we perform a second analysis based on data from six different geographically defined locations in the BCSC representing ~5% of the U.S. population. We use a screening dissemination model to generate simulated results for the U.S. population and compare these results with national cross-sectional estimates from the year 2000 NHIS for the percent of women that received a mammogram in the previous 2 years. We update a model for mammography dissemination and usage in the U.S. population (17) based on longitudinal data from the BCSC to include additional years of data from the BCSC and to model screening use by race and ethnicity. The difference between modeled and self-reported screening rates are put into context of previously reported validation studies by comparison with reported values of sensitivity (percent of women who did have a mammogram who accurately report having a mammogram in the previous 2 years) and specificity (percent of women who did not have a mammogram who accurately report not having a mammogram in the previous 2 years).

Materials and Methods

Study Population and Data. The BCSC is a National Cancer Institute–supported research initiative that collects population-based longitudinal data on mammography usage and performance in clinical practice through mammography registries that are linked to cancer outcomes (16). Each registry and the BCSC statistical coordinating center have 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 Health Insurance Portability and Accountability Act 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. The BCSC has collected data on screening mammograms since 1994 in seven geographically defined research sites, and represents ~5% of the U.S. population. Data from these sites are transformed to a standard data format and sent to a statistical coordinating center that pools the data for analysis. Research data collection sites are Vermont Breast Cancer Surveillance System (VTBCSS), New Hampshire Mammography Network, Colorado Mammography Project, Carolina Mammography Registry, New Mexico Mammography Project, and San Francisco Mammography Registry. Group Health Center for Health Studies in Washington, one of the seven BCSC sites, was not included in this analysis because it consists of an health maintenance organization (HMO) population where the majority of women follow a 2-y screening interval with a formal reminder system.

The pooled data set includes all mammograms done at participating facilities whether or not that woman lives in the defined geographic area. It does not include records for mammograms obtained outside the geographically defined areas of the BCSC or in facilities within the defined areas that do not participate in the BCSC. Generally, research sites do not include information from all facilities where mammographic exams are given within a particular geographic area. The exception is VTBCSS, which includes all mammography facilities in the state of Vermont, allowing us to achieve complete ascertainment of mammography utilization.

Mammography exams classified as screening by the radiologist from the BCSC are used either directly or as inputs into a larger model to estimate the number of woman who received a mammogram in the years 1999 or 2000 as a percent of the population. These estimates are compared with the percent of the population who reported having a screening mammogram and the percent who had a mammogram for any reason in the past 2 y for the 2000 calendar year.

The percent of women reporting a mammogram within the past 2 y and the percent of women reporting a mammogram for any reason from BRFSS and NHIS surveys are calculated using SUDAAN to account for the complex design of these surveys.


BCSC Estimates of the Percent of Women in Vermont that Received a Screening Mammogram in 1999 or 2000. The number of screening mammography exams recorded in the BCSC for women who live in Vermont from the years 1999 and 2000 are combined with population estimates from the 2000 census to estimate age-specific screening rates. We include mammograms that were done at BCSC facilities in New Hampshire for women whose primary residence was in Vermont based on their home zip code. The probability of a mammography in the past 2 y for 10-y age groups is calculated as the number of women with at least one screening mammogram in either 1999 or 2000 recorded in the BCSC divided by the census population in that age group. These estimates are compared with the year 2000 BRFSS survey results for mammography use in Vermont.

Modeled Estimates of the Percent of Women in the United States who Received a Screening Mammogram in 1999 or 2000. With the exception of Vermont where all facilities in the state are included in the BCSC, the BCSC data set does not contain all mammography exams that occurred in a defined geographic area. Therefore, a population denominator is not available to obtain screening rates for the BCSC data set. Instead, we use a modeling approach to estimate the percent of the women in the United States that received a mammogram in a specific time period from these data. Our approach extends a previous model by allowing for different screening rates by race and ethnicity (17). The model consists of two separate components that together describe screening patterns for 5-year birth cohorts in the United States during the years 1975 to 2000. The first component describes the age at which a woman receives an initial screening exam and the second estimates the interval between successive screening exams.

Simulation is used to combine the two components of the model and generate individual screening histories representative of the U.S. population over time. This simulation output is then used to estimate the percent of women in specified age and racial/ethnic groups that had a screening mammogram in the 2-y period 1999 and 2000. Results from the model are compared with national estimates of mammography use from the 2000 NHIS. A detailed description of the modeling is available upon request.

Calculation of Sensitivity and Specificity. Because this is not a validation study, we cannot directly estimate the sensitivity and specificity of the survey respondents. However, it is possible to identify values for sensitivity and specificity that are consistent with our findings, assuming that the modeled results represent the true percent of women that had a mammogram over a 2-y period. We can then compare these values with the existing literature on the validity of self-reported mammography use. The equation below allows us to specify a value for sensitivity and calculate a corresponding value for specificity that would relate to the modeled and observed percentages.

Self – report % = sensitivity * modeled % + (1 – specificity) * (1 – modeled %).

Results

Figure 1 shows the percent of women who reported having a mammogram for any reason and specifically for screening in the past 2 years from the 2000 Vermont BRFSS and the percent of women who live in Vermont that received a screening mammogram in either 1999 or 2000 based on the BCSC data. When comparing self-report to BCSC data for the state of Vermont, differences of 15, 18, 16, and 25 percentage points are found for screening mammograms for ages 40 to 49, 50 to 59, 60 to 69, 70 to 79 years, respectively. When considering mammography for any reason, the differences are similar with percentage point differences of 16, 21, 18, and 27.

Figure 2A to D show similar results for the nationally representative 2000 NHIS survey and the percent of women receiving a screening mammogram in either 1999 or 2000 based on the mammography dissemination and usage model. Figure 2A shows all women combined, and Fig. 2B to D gives results for White, non-Hispanic White, African American, and Hispanic women. The difference in screening rates given by the model and those reported in 2000 NHIS is shown in Table 1. Modeled and self-reported estimates for the percent of women who had a mammogram in the previous 2 years are highest, and overreporting is lowest for non-Hispanic White women. This result is consistent with the majority of misclassification being attributed to women who were not screened in the defined time period, leading to an underestimation of the disparities between groups. Self-reported screening rates were similar for African Americans and Hispanics. However, a bigger difference is found between modeled and self-reported estimates among African American women, suggesting more overreporting among African American women. Overreporting is highest in younger age groups where the level of recent screening is the lowest.

Figure 3 plots the values of sensitivity and specificity consistent with the comparison between modeled results and self-reported NHIS rates by race/ethnicity.

![Figure 1](image-url)
Previously published estimates of sensitivity and specificity have been added to the graph to present our results in the context of the current literature on self-report bias related to mammography screening.

**Discussion**

Estimates of screening mammography usage rates are typically based on self-reported information because surveys are an efficient method to obtain information on a large number of individuals. Medical records, such as those obtained in the BCSC, are generally considered to provide more accurate information on mammography usage than self-reported data from state and national surveys; however, studies rarely collect both medical records and self-report to be able to actually compare the difference. This study is unusual because it examines the validity of survey response on a geographically defined population, using the Census as denominator, rather than focusing on a particular group such as members of an HMO. The approach presented provides an estimate of recent screening rates that is likely to be representative of the U.S. population.

The mammography model used is based largely on the BCSC data. Previous work has compared counties included in the BCSC to all counties in the United States to gauge representativeness of the BCSC to the United States (18). A number of county-level variables were similar between the United States and counties included in the BCSC, with BCSC counties seeming to have slightly higher income and education levels.

Our results are consistent with previously reported studies. The 15 to 25 percentage point difference between the observed rates of screening within the prior 2 years in Vermont and estimates based on self-report from the Vermont BRFSS is consistent with previous work. Values for sensitivity and specificity that would explain the differences between modeled screening rates and national survey estimates are in line with previously reported studies that sought to validate survey-based estimates for the percent of women that were screened in the previous 2 years. Estimates of screening mammography based on BCSC data are consistently lower than those reported by NHIS or BRFSS (19).

Sensitivity of self-report of health behaviors is consistently high, whereas specificity tends to be lower, resulting in an overestimation of the percent of the population actually adhering to recommended screening behavior. Therefore, misreporting occurs mainly in women who have not had a mammogram in the previous 2 years and overreporting is greatest in the groups that have the lowest screening rates. We found the largest difference between modeled and self-reported

**Figure 2.** The percent of women predicted to have a mammogram in the years 1999 and 2000 based on the mammography dissemination and usage model and the percent of women who reported have a mammogram for any reason or specifically for screening in the previous 2 y from the 2000 NHIS. Single SE bars are added to the NHIS estimates. **A.** All Women. **B.** Non-Hispanic White women. **C.** Non-Hispanic African American women. **D.** Hispanic women.
screening rates among younger women and African Americans, both of whom had the lowest screening rates.

Lower screening rates among African American women in more overreporting in African American women than in non-Hispanic White women even with similar values of sensitivities and specificities. Systematic underestimation of disparities between these two groups is a result of this pattern of misreporting. This is consistent with previous analyses that have shown a lower percent of self-reported mammogram use can be validated by medical reports for African American women compared with White women. For example, Holt et al. (11) found self-reported mammography use similar between White and African American women, but lower rates of validated mammography among African American women. McPhee et al. (13) also reported a lower validation rate for self-reported mammograms among African American women than White women.

Our results confirm other findings that Hispanic women may have different sensitivity and specificity values than White and African American women resulting in less overreporting. Hiatt et al. (6) reported lower sensitivity and higher specificity for Hispanic compared with non-Hispanic White women, and Lawrence et al. (20) showed lower sensitivity in Mexican Americans compared with Euro-Americans. Other analyses (15) have shown lower sensitivity and specificity in Puerto Rican women compared with African American or non-Hispanic White women. In contrast with our results, previous studies have shown a lower validation rate in Hispanic women compared with non-Hispanic White women (6, 13, 15, 20). Hispanic women seem to have different patterns for misreporting mammography usage, and those patterns may vary within subgroups of the Hispanic population. A large percentage of Hispanic women in the BCSC come from New Mexico.

Overreporting among women who have not received a mammogram also affects trends over time. During time-periods when screening rates have increased, the true amount of improvement will be masked because overreporting will be decreasing at the same time that screening is increasing. For the first time since mammography rates have been ascertained, there was a reported decrease in the percent of the population reporting recent mammography use between 2000 and 2005 (21). When screening rates decrease, we would expect that overreporting would increase, leading to an underestimation of the actual decrease in screening rates observed in 2005.

Women may overreport the use of mammography screening in survey situations for several reasons. The phenomenon of “telescoping” (i.e., remembering that an event occurred more recently then it actually did) can lead to systematic underreporting of the time since last mammogram and overreporting of the prevalence of women who adhere to a guideline-based screening interval, such as the past 1 or 2 years. The difference between the modeled and self-reported results may be related to the difference between being a “regular” screener and actually receiving the screening exam within the exact 2-year cutoff considered. Our model as well as previous work (22, 23) show that even regular screeners often do not achieve the recommended interval. A woman who sees herself as a regular screener may report an interval of 2 years even if the interval was slightly longer. The phenomenon of telescoping leads women to underestimate the time since their last mammography, but it is possible that self reported rates better represent women who come in regularly for exams even if it is not within the exact 2-year time frame.

Because recommendations for annual or biennial breast cancer screening are well-publicized, women may feel compelled to give a socially desirable response of having a recent screening mammogram even when untrue. Some women may lack the knowledge necessary to properly answer survey questions about prior mammography screening. Another possibility is that women...
who choose to answer the survey question have different screening behaviors than nonresponders. NHIS 2000 had an overall response rate of 72%, and the 2000 Vermont BRFSS had a response rate of 50%. Selection bias could also contribute to the differences observed.

Several data limitations were encountered when developing the mammography dissemination and usage model. The model consists of separate components for the time to first mammography examination and the time between mammography examinations. The time until a first mammography examination component is based on self-report data from surveys for whether or not a woman has ever received a mammogram, which is also subject to self-report bias. However, previous work suggests that self-report is more accurate when measuring if a woman has ever had a mammography then measuring if she had a mammogram within some specified time period (10). Although both the NHIS and BRFSS surveys obtain information on the reason for the most recent mammogram, they do not contain similar information on all mammograms ever received. Therefore, we cannot directly determine if a woman has ever had a mammogram for the purposes of screening. This may result in underestimating the age at first screening mammography, ultimately leading to an overestimation of the amount of screening in the population.

Under certain circumstances, the repeat mammography component of the model includes self-reported data. At each visit, women were asked when they had their last mammogram. If there was a discrepancy between the date of the last mammography recorded in the BCSC data and the date of self-reported last mammogram, the model included the minimum time estimate from the two sources when this discrepancy was >6 months. This inclusion of self-report data were done to allow for the possibility that a woman received a mammogram at a facility that was not covered by the BCSC. The inclusion of self-reported data may overestimate frequency of mammography and result in an underestimation of the bias between self report and registry data on mammography use.

The modeling contains uncertainty on several levels. The data used to fit model parameters are subject to the limitations described above. Given the data, the parameter estimates have an associated variance. The parameter estimates are then used to simulate outcomes representing the U.S. population. We do not include confidence intervals for the modeled screening rates because it would be difficult to quantify the true variance around these rates. Although bias in the modeled estimate may contribute to the difference reported, the results from the comparison of national estimates are very consistent with the Vermont comparison, which is not subject to the potential modeling bias.

To obtain accurate information on screening behaviors, a consistent system of electronic medical records that links patients’ records from all sources of health care and then deidentifies them for purposes of research is needed. In the absence of such a system, recommendations described in Newell et al. (1) may help maximize accuracy associated with self-report. Systematic errors in self-reported screening rates result in biased estimates of disparities. Sensitivity and specificity estimates can be used to adjust self-report data to better capture difference between groups and trends over time.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank the BCSC investigators, participating mammography facilities, and radiologists for the data they have provided for this study. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are provided at the BCSC Web Site.

References


