



Published in final edited form as:

Med Care. 2014 July ; 52(7): e44–e51. doi:10.1097/MLR.0b013e318269e0f5.

Distinguishing Screening from Diagnostic Mammograms using Medicare Claims Data

Joshua J. Fenton, MD, MPH¹, Weiwei Zhu, MS², Steven Balch, MA, MBA², Rebecca Smith-Bindman, MD, MPH³, Paul Fishman, PhD², and Rebecca A. Hubbard, PhD^{2,4}

¹Departments of Family and Community Medicine, Radiology, and the Center for Healthcare Research and Policy, University of California, Davis, Sacramento, CA

²Group Health Research Institute, Seattle, WA

³Departments of Radiology, Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA

Department of Biostatistics, University of Washington, Seattle, WA

Abstract

Background—Medicare claims data may be a fruitful data source for research or quality measurement in mammography. However, it is uncertain whether claims data can accurately distinguish screening from diagnostic mammograms, particularly when claims are not linked with cancer registry data.

Objectives—To validate claims-based algorithms that can identify screening mammograms with high positive predictive value (PPV) in claims data with and without cancer registry linkage.

Research Design—Development of claims-derived algorithms using classification and regression tree analyses within a random half-sample of bilateral mammogram claims with validation in the remaining half-sample.

Correspondence to: Joshua Fenton, MD, MPH, Department of Family and Community Medicine, University of California, Davis, 4860 Y Street, Suite 2300, Sacramento, CA 95817, (916) 734-3164; Fax (916) 734-5641, Joshua.fenton@ucdmc.ucdavis.edu.
Joshua J. Fenton, MD, MPH, Department of Family and Community Medicine, University of California, Davis Medical Center, 4860 Y Street, Suite 2300, Sacramento, CA 95817, (916) 734-3164, Fax (916) 734-5641, Joshua.fenton@ucdmc.ucdavis.edu
Weiwei Zhu, MS, Group Health Research Institute, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101, (916) 287-2066; fax (206) 287-4710, Zhu.w@ghc.org
Steven Balch, MA, MBA, Group Health Research Institute, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101, (916) 287-2066; fax (206) 287-4710, Zhu.w@ghc.org
Rebecca Smith-Bindman, MD, MPH, Departments of Radiology and Epidemiology and Biostatistics, University of California, San Francisco, Box 1667, San Francisco, CA 94143-1667, (415) 253-2573; fax (415) 885-7876, rebecca.smith-bindman@radiology.ucsf.edu
Paul Fishman, PhD, Group Health Research Institute, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101, (916) 287-2066; fax (206) 287-4710, Fishman.p@ghc.org
Rebecca Hubbard, PhD, Group Health Research Institute, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101, (916) 287-2066; fax (206) 287-4710, hubbard.r@ghc.org

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

Conflicts of Interest: None declared.

Subjects—Female fee-for-service Medicare enrollees age 66 years and older who underwent bilateral mammography from 1999 to 2005 within Breast Cancer Surveillance Consortium (BCSC) registries in four states (CA, NC, NH, and VT), enabling linkage of claims and BCSC mammography data (N=383,730 mammograms obtained by 146,346 women).

Measures—Sensitivity, specificity, and PPV of algorithmic designation of a “screening” purpose of the mammogram using a BCSC-derived reference standard.

Results—In claims data without cancer registry linkage, a three-step claims derived algorithm identified screening mammograms with 97.1% sensitivity, 69.4% specificity, and a PPV of 94.9%. In claims that are linked to cancer registry data, a similar three-step algorithm had higher sensitivity (99.7%), similar specificity (62.7%), and higher PPV (97.4%).

Conclusions—Simple algorithms can identify Medicare claims for screening mammography with high predictive values in Medicare claims alone and in claims linked with cancer registry data.

Keywords

Breast Cancer Screening; Mammography; Validation Studies; Medicare; Quality Assessment

INTRODUCTION

Despite advances in breast imaging, mammography will remain the cornerstone of breast cancer screening programs for the foreseeable future. American women receive forty million mammograms annually,¹ and one-third of these are received by older women enrolled in Medicare.^{2,3} In addition, more than half of incident breast cancer diagnoses occur among women older than 65 years,⁴ so the improvement of mammography services remains a public health priority both for the general population and the Medicare program.

Comparative effectiveness research and mammography quality measurement may be facilitated by data infrastructures that enable efficient comparisons within real populations, such as Medicare claims or the linked Surveillance, Epidemiology, and End Results (SEER)-Medicare data. However, uncertainty regarding the ability to distinguish screening from diagnostic mammograms has limited the utility of these data for clinical mammography research. Accurate distinction between screening and diagnostic mammography is crucial for research because the interpretive accuracy of mammography and subsequent cancer incidence varies markedly based on examination purpose.^{5,6} Distinction is also fundamental to claims-based assessment of screening utilization.⁷

One source of difficulty was the slow adoption of procedure codes for screening mammography after the extension of Medicare coverage to screening mammography in 1991. Early studies suggested that many claims for women undergoing screening mammography had procedure codes for either diagnostic or unilateral examinations.^{8,9} Investigators subsequently proposed algorithms that incorporated information on prior diagnoses or procedures to assist in identifying screening examinations even when coded as diagnostic.⁷

Among 2,593 mammograms received by women who eventually were diagnosed with breast cancer from 1991 to 1999, Smith-Bindman, et al. found that an algorithm based upon four claims-derived variables identified screening mammograms with 87% sensitivity and 89% specificity compared to a reference standard derived from mammography data from two Breast Cancer Surveillance Consortium (BCSC) registries.¹⁰ However, all women in the sample eventually developed breast cancer; it is uncertain whether similar performance would be observed in a general screening population in which the large majority of women do not develop breast cancer.¹¹ The algorithm also requires claims linked to cancer registry data, enabling identification of women with prevalent breast cancers for whom mammography is performed for diagnostic purposes. However, analysts may wish to identify screening mammography claims that are not linked with cancer registries (e.g., Medicare claims from non-SEER regions).

Capitalizing on a recently developed data infrastructure – the linked BCSC-Medicare data – we evaluated whether claims-based algorithms could accurately distinguish screening from diagnostic mammograms in a general Medicare screening population, including populations with and without linkage to cancer registries. We hypothesized that claims-derived algorithms could distinguish screening from diagnostic mammograms with high positive predictive value both with and without cancer registry linkage.

METHODS

Data

We used data from Medicare claims files (the Carrier Claims, Outpatient, and Inpatient files) and the Medicare denominator file, which provides demographic, enrollment, and vital status data. While Medicare mammography claims nearly always appear in the Carrier file, we assessed both the Carrier and Outpatient files to capture the minority of claims present only in the Outpatient file (~3%).¹⁰ We used Healthcare Common Procedure Coding System (HCPCS) procedure codes to identify bilateral mammograms and breast imaging and procedures in the year prior to mammography. Medicare claims also include International Classification of Diseases, 9th Edition, Clinical Modification (ICD-9-CM) diagnosis codes, which we used to identify breast symptoms and comorbidities.

Medicare claims from 1998 to 2006 were linked with BCSC mammography data derived from regional mammography registries in four states (North Carolina; San Francisco Bay Area, CA; New Hampshire; and Vermont) (<http://breastscreening.cancer.gov/>). BCSC facilities transmit prospectively collected patient and mammography data to regional registries, which link the data to breast cancer outcomes ascertained from cancer registries. The BCSC has established standard definitions for key variables and multiple levels of data quality control and monitoring.¹² BCSC sites have received institutional review board approval for 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 BCSC sites have received a Federal Certificate of Confidentiality to protect the identities of patients, physicians, and facilities. Among women aged 65 years and older and BCSC enrolled during the study period, over 87% were successfully matched to Medicare claims.

Subjects

We identified a matched sample of bilateral mammograms captured in both Medicare claims and the BCSC among women who were aged 66 or older on mammography dates from January 1, 1999 to December 31, 2005. We identified bilateral mammograms based upon Medicare claims with HCPCS codes 76091, 76092, G0202-G0205 (encompassing film-screen and digital screening and diagnostic mammograms) and considered mammograms to have matching BCSC records if claims and BCSC records had the same date of service. When Medicare mammogram claims were present in both the Carrier and Outpatient files, we collapsed the claims to represent a single mammogram, maintaining codes for diagnostic mammograms if present in one but not both files.

From this matched sample, we then selected mammograms for women with continuous enrollment in fee-for-service Medicare (parts A and B) for twelve months before and after mammography, enabling longitudinal assessment of outpatient claims for clinical events that might suggest a diagnostic purpose for mammograms. We excluded mammograms when BCSC data on examination purpose was either missing or unknown. We randomly divided the matched sample into two half-samples, one for training and one for validation in classification and regression tree (CART) analyses.

Because researchers using isolated Medicare claims would not have access to cancer registry data, we included in the overall sample mammograms for women with prevalent breast cancer based upon BCSC data, which are derived from cancer registries. We used this sample to develop classification algorithms for use when women with breast cancer cannot be easily excluded from the sample. However, to address the needs of researchers with linked cancer registry data (e.g., SEER-Medicare data), we performed all analyses after excluding from mammograms for women with any prior breast cancer diagnosis as identified using BCSC data.¹³ We excluded women with prior breast cancers because mammography may often serve a mixed-purpose in these women, including screening, surveillance, and diagnosis.

Reference Standard

We developed a reference standard classification of mammograms as either “screening” or “diagnostic/other” based on two steps. First, we used the standard BCSC definition of “screening” mammograms as bilateral mammograms performed on asymptomatic women that are designated as “routine screening” by the interpreting radiologist.¹² The BCSC further specifies that screening mammograms must be performed at least nine months after the most recent prior mammogram based on either patient self-report, radiologist report, or BCSC mammography data. Thus, we classified mammograms for women with a prior history of breast cancer or who report breast symptoms or signs at the time of examination are classified as “diagnostic/other.” Similarly, we classified mammograms performed within nine months of a prior mammogram are classified as “diagnostic/other.”

In the second step, we re-classified mammograms as “diagnostic/other” if there were Medicare claims for mammography within nine months even if BCSC data indicated no mammography in the prior nine months. We included the latter step because BCSC data on

prior mammography may be either incomplete (e.g., due to in-migration of women to BCSC registries) or susceptible to error (e.g., patient recall bias).

Claims-Based Algorithms for Defining Screening and Diagnostic Mammograms

We first categorized mammograms as screening and diagnostic based on a four-step algorithm (updated to include contemporaneous coding) that sequentially considers the following claims data: 1) mammography claims within the prior nine months; 2) whether the claim HCPCS code was for a “screening” rather than a “diagnostic” mammogram; 3) claims containing codes for breast symptoms or procedures within the prior ninety days; and 4) the incidence of breast cancer within six months of mammography.¹⁰

We then conducted CART analyses in attempts to identify algorithms with superior performance.¹⁴ CART is a non-parametric decision tree methodology that identifies sequential binary partitions in independent variables that optimally predict the dependent variable. In this case, the dependent variable was the reference standard definition of a mammogram as screening (vs. diagnostic/other). We included as potential independent variables the following claims-derived variables: age; a modification of the Charlson comorbidity index¹⁵; mammogram code signifying screening purpose (76092, G0202, any GG modifier, G0203/05 in 2001); mammogram code signifying diagnostic purpose (76090, 76091, G0204, G0206); days from any prior Medicare mammogram; days from any prior breast ultrasound, magnetic resonance imaging or other breast imaging; days from any prior breast biopsy or breast-directed surgery; days from any prior claim with ICD-9-CM codes for breast signs or symptoms (611.7x); days from any prior encounter with ICD-9-CM codes for breast cancer (174.x, 233.0, V103); and the total number of outpatient visits, visits with primary care physicians, and visits with obstetrician/gynecologists in the past one and twelve months. We defined physician visits based on Berenson-Eggers Type of Service codes and physician specialty based on Health Care Financing Administration specialty codes on claims.¹⁶ Because we sought an algorithm that could be used in studies of claims events ensuing after mammography (e.g., subsequent breast imaging), we only considered claims events preceding the mammogram date. Specific codes used in the CART analyses are available from the authors.

We performed CART on the training half-sample of the matched mammogram set. The CART algorithm selected splits in independent variables on the basis of the Gini index, and continued growing trees until no further splits improved the Gini index by more than 0.00001.¹⁴ To minimize overfitting, we pruned trees to optimal complexity based on cross-validation. The resultant trees were extremely complex, limiting their practicality. We therefore selected simpler trees that included the first three splits of the pruned tree.

Analyses of Classification Accuracy

Within validation sub-samples, we created cross-tabulations to compare the classification of mammograms as screening vs. diagnostic using claims-based algorithms versus the reference standard. We quantified accuracy using: sensitivity (the proportion of screening mammograms classified as screening); specificity (the proportion of diagnostic mammograms classified as diagnostic); positive predictive value (PPV, or the proportion of

mammograms classified as screening also classified as screening by the reference standard); negative predictive value (the proportion of mammograms classified as diagnostic also classified as diagnostic by the reference standard); and Cohen's kappa. To understand potential underlying causes of misclassification, we performed descriptive analyses to identify Medicare claims characteristics that led to disagreement in claims-based and reference standard classifications. 95% confidence intervals around all point estimates were negligibly small, so we report only point estimates. Study estimates and confidence intervals were similar when derived by using bootstrapped mammogram samples from the validation cohort. We performed statistical analyses using R, version 2.12.0 (R Foundation for Statistical Computing, Vienna, Austria) The study was approved by the Group Health Research Institute Institutional Review Board.

RESULTS

Mammogram Samples

We identified a sample of 383,730 mammograms with matched Medicare claims and BCSC records. The mammograms were obtained by 146,346 women who received an average of 2.62 mammograms during the study period (range: 1–9). On the date of mammography, women had a mean age of 73.9 years (SD: 5.8). The women were ethnically diverse, and mammograms in the training and validation samples had similar characteristics (Table 1). Of all mammograms, nearly 8.5% were performed on women with a prior breast cancer diagnosis based on BCSC data. The reference standard classified 328,069 mammograms (85.5%) as “screening,” and the remaining 55,661 (14.5%) mammograms as “diagnostic/other.”

Performance of Claims-Based Algorithm

We first examined the performance of the adapted four-step sequential algorithm.¹⁰ Originally developed using a test set with cancer registry linkage, the algorithm had very high sensitivity (99.9%) but low specificity (34.1%) and moderate PPV (89.9%) within the claims sample that included prevalent breast cancers (Table 2). Within the testing sample with prevalent cancers excluded, the algorithm remained highly sensitive (99.7%) but still had low specificity (52.1%), a PPV of 96.7%, and moderate agreement beyond chance (Cohen's kappa = 0.649).

Using CART analyses, we identified alternative claims-based algorithms with improved PPVs of a screening designation. Compared to the four-step sequential algorithm,¹⁰ a three-step sequential algorithm had a higher specificity (69.4% vs. 34.1%), PPV (94.9% vs. 89.9%), and agreement beyond chance (Cohen's kappa = 0.704 vs. 0.469) in Medicare claims alone (Table 3). The three sequential nodes are: 1) whether the mammography HCPCS code was for a “screening” rather than “diagnostic” examination; 2) whether the woman had received mammography in the prior nine months (≤ 270 days); and 3) any ICD-9-CM code for breast cancer in the prior year (Figure 1). The performance of this three-step algorithm was similar to the performance of a fifteen-step cross-validation pruned CART algorithm that reclassified mammograms with a diagnostic HCPCS code as “screening” mammograms based on days since prior mammography, breast symptoms in the

prior year, numbers of outpatient visits in the prior month and prior year, numbers of visits with primary care physicians and obstetrician/gynecologists in the prior month, days since prior breast imaging other than mammography, and patient age.

Within a test sample that excluded mammograms for women with prevalent cancer (based upon linked cancer registry data), CART identified a three-step algorithm with slightly higher positive predictive value of a screening designation than the adapted four-step algorithm (97.4% vs. 96.7%) (Table 3). The algorithm includes the following sequential nodes: 1) whether the woman had received mammography in the prior nine months (≤ 270 days); 2) whether the mammography HCPCS code was for a “screening” rather than “diagnostic” examination; and 3) whether the woman had any breast symptoms in the prior 349 days (Figure 2). Compared to the reference standard, the three-step algorithm had high sensitivity (99.7%), moderate specificity (62.7%), and excellent agreement beyond chance ($\kappa = 0.739$). Performance of this three-step algorithm was comparable to the performance of an eighteen-step cross-validation pruned algorithm that further classified mammograms with ICD-9-CM codes for breast symptoms in the prior 349 days using eleven additional retrospective variables (outpatient visits in the prior month and prior year; visits with primary care physicians, obstetrician/gynecologists and total visits in the prior month and prior year; days since a prior breast biopsy, prior non-mammographic breast imaging, and a prior breast cancer diagnosis; the Charlson comorbidity index; and patient age).

Reasons for Mammogram Misclassification

Among mammograms without linkage with breast cancer data, 4,806 screening mammograms were classified as diagnostic by the three-step algorithm (Table 4). In most cases, misclassification occurred because the mammogram had a diagnostic HCPCS code. In the remaining cases, mammograms were misclassified as diagnostic because of a breast cancer diagnosis code in claims in the prior year, although BCSC data indicated that women receiving these screening mammograms were breast cancer-free.

The three-step algorithm misclassified 8,542 diagnostic mammograms as screening, usually because there were no prior claims with diagnoses of breast cancer despite BCSC data indicating prior breast cancer (Table 4). Additional reasons for misclassification of diagnostic mammograms were either patient self-report or a BCSC record of a prior mammogram within nine months (despite absent prior mammography claims) and a non-screening indication for the mammogram in BCSC data. In addition, all of these misclassified claims had a “screening” HCPCS code despite BCSC data suggesting a diagnostic purpose.

Among mammograms with cancer registry linkage, few (0.3% of total) were misclassified as diagnostic when classified as screening by the reference standard, and this misclassification was always explained by the use of a diagnostic HCPCS code. Among 4,329 diagnostic mammograms that were misclassified as screening, most had a non-screening indication in the BCSC data, yet no claims evidence of prior breast symptoms or signs. The remaining diagnostic mammograms were misclassified as screening because of either self-report or BCSC records indicating prior mammography within nine months, although there were no corresponding prior mammography claims.

DISCUSSION

In samples of bilateral mammograms with corresponding Medicare claims and mammography registry data, we evaluated the ability of algorithms based on claims data to classify mammograms as screening versus diagnostic. Using CART analyses, we identified three-step, claims-based algorithms that identified screening mammograms with higher PPVs (and agreements beyond chance) than a four-step algorithm that was previously validated among women who eventually were diagnosed with breast cancer.¹⁰ Additionally, we validated algorithms for use with Medicare claims both with and without linkage with breast cancer registry data. With improved positive predictive value of a screening designation, the simpler, three-step algorithms presented here seem better suited than the earlier algorithm for claims-based studies of screening mammography.

The three-step algorithms had high sensitivities for screening mammography (>97%). Small fractions of BCSC screening mammograms were misclassified as diagnostic by the algorithms, most commonly because screening mammograms had claims with codes for diagnostic mammography. Historically, lower Medicare fees for screening as compared to diagnostic mammography may have encouraged widespread use of diagnostic codes when mammograms were actually performed for screening.^{8,9} However, our analyses of claims from 1999 to 2005 suggest that many diagnostic mammograms, including mammograms for women with prior breast cancer, are often now coded as screening examinations.

In contrast to the high sensitivity of the claims algorithms, specificities were relatively low, implying that many diagnostic mammograms are misclassified as screening. The underlying reasons for misclassification of diagnostic mammograms differed based on whether the algorithms were designed for use in claims with vs. without cancer registry linkage. For claims unlinked to cancer registries (i.e., that include mammograms for women with prevalent breast cancer), the most common reason for misclassification of BCSC diagnostic mammograms as screening was a prior breast cancer diagnosis that was not identified based on review of claims diagnoses during the prior year. Consistent with prior research,¹⁷ this finding suggests that prior claims are imperfectly sensitive for identifying prevalent breast cancers. A longer look-back for breast cancer diagnoses may have increased identification of prevalent breast cancers but at the cost of introducing differential misclassification. Because younger Medicare enrollees will usually have fewer years of prior claims than older enrollees, algorithms with longer or unlimited look-backs would likely misclassify more younger than older women as breast cancer-free (leading to lower algorithm specificity and PPV among younger women).

For claims that are linked with breast cancer registry data, the most common reason for misclassification of BCSC diagnostic mammograms as screening was the absence of claims evidence of prior breast symptoms or signs despite BCSC data indicating that the mammogram was performed to evaluate breast symptoms or signs or other diagnostic evaluation. Breast symptoms are common among women in the community,¹⁸ and many women may report relatively mild breast symptoms at the time of screening mammography that were not previously brought to medical attention. Mammography facilities may perform diagnostic examinations for these symptomatic women yet submit claims for routine

screening mammography, billing Medicare for diagnostic mammography only when the examination was scheduled specifically for diagnostic purposes.

With high sensitivities and PPVs, the pragmatic, three-step algorithms can be useful tools for researchers seeking to identify samples of screening mammogram claims within the Medicare population. However, with a PPV of >97%, up to 3% of mammograms identified with the algorithms may have been performed for diagnostic purposes, and investigators must consider the potentially confounding impact of these mammograms on study outcomes, particularly because breast cancer incidence rates are much greater after diagnostic than screening mammography.⁵ More complex, cross-validation pruned CARTs had slightly higher PPVs, but the gains would not seem to justify the implementation efforts for most applications.

Study strengths include the inclusion of large mammography claim samples from geographically diverse setting that were linked with high-quality external mammography data, yielding well-powered, rigorous validation analyses. We also developed algorithms for use both with and without claims data linkage to cancer registries. Because cancer registries such as SEER encompass only 25% of the U.S. population,¹⁹ the alternative algorithms may enable mammogram sampling for research or quality improvement across the entire Medicare program regardless of claims linkage with cancer registry data.

The BCSC classification of mammograms as screening versus diagnostic may be an imperfect reference standard. Indeed, because prior mammography within nine months is integral to BCSC definitions of “screening” and “diagnostic” mammography, it is predictable that this variable was identified as an important classification node in CART analyses. We also modified the BCSC reference standard based on the presence of a mammogram claim within the prior nine months, although BCSC records and patient self-report did not indicate prior mammography within nine months. In these instances, patients may have received mammography outside of BCSC facilities and erroneously reported the time since prior mammography. Our results also derive from mammography claims of fee-for-service Medicare enrollees within four U.S. regional mammography registries. Algorithms may not generalize to non-Medicare claims or to Medicare enrollees outside these regions. Additionally, because study algorithms exclude women with prior breast cancers, they are not suitable for assessing mammography purpose among breast cancer survivors.

Our results suggest that simple, three-step algorithms can identify Medicare claims for screening mammography with very high predictive value in claims samples both with and without linkage with cancer registry data. These algorithms should be useful to researchers designing studies of screening mammography based on Medicare claims.

Acknowledgments

The authors thank Dr. Karen Lindfors for advice on screening mammography coding practices. We also thank the participating women, mammography facilities, and radiologists for the data they have provided for this study. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are provided at: <http://breastscreening.cancer.gov/>.

Sources of support: This work was supported by the National Cancer Institute-funded Breast Cancer Surveillance Consortium (U01CA63740, U01CA86076, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, U01CA70040, HHSN261201100031C) and the National Cancer Institute grants R21CA158510 and RC2CA148577. 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: <http://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. U.S. Food and Drug Administration. [Accessed May 15, 2010] MQSA National Statistics. 2010. <http://wwwfdagov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/FacilityScorecard/ucm113858htm>
2. U.S. Census Bureau. [Accessed July 5, 2008] Annual Estimates of the Population by Sex and Five-Year Age Groups for the United States: April 1, 2000 to July 1, 2007 (NC-EST2007-01). 2000. www.census.gov/popest/national/asrh/NC-EST2007/NC-EST2007-01.xls
3. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2008: a review of current American Cancer Society guidelines and cancer screening issues. *CA Cancer J Clin.* 2008 May-Jun;58(3):161–179. [PubMed: 18443206]
4. Ries, LAG.; Melbert, D.; Krapcho, M., et al., editors. SEER Cancer Statistics Review, 1975–2004. Bethesda, MD: National Cancer Institute; 2007.
5. Sickles EA, Miglioretti DL, Ballard-Barbash R, et al. Performance benchmarks for diagnostic mammography. *Radiology.* 2005 Jun; 235(3):775–790. [PubMed: 15914475]
6. Rosenberg RD, Yankaskas BC, Abraham LA, et al. Performance benchmarks for screening mammography. *Radiology.* 2006 Oct; 241(1):55–66. [PubMed: 16990671]
7. Randolph WM, Mahnken JD, Goodwin JS, Freeman JL. Using Medicare data to estimate the prevalence of breast cancer screening in older women: comparison of different methods to identify screening mammograms. *Health Serv Res.* 2002 Dec; 37(6):1643–1657. [PubMed: 12546290]
8. Blustein J. Medicare coverage, supplemental insurance, and the use of mammography by older women. *N Engl J Med.* 1995 Apr 27; 332(17):1138–1143. [PubMed: 7700288]
9. Trontell AE, Franey EW. Use of mammography services by women aged > or = 65 years enrolled in Medicare--United States 1991–1993. *MMWR. Morbidity and mortality weekly report.* 1995 Oct 20; 44(41):777–781. [PubMed: 7565562]
10. Smith-Bindman R, Quale C, Chu PW, Rosenberg R, Kerlikowske K. Can Medicare billing claims data be used to assess mammography utilization among women ages 65 and older? *Med Care.* 2006 May; 44(5):463–470. [PubMed: 16641665]
11. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med.* 1978 Oct 26; 299(17):926–930. [PubMed: 692598]
12. 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 Oct; 169(4):1001–1008. [PubMed: 9308451]
13. Randolph WM, Goodwin JS, Mahnken JD, Freeman JL. Regular mammography use is associated with elimination of age-related disparities in size and stage of breast cancer at diagnosis. *Ann Intern Med.* 2002 Nov 19; 137(10):783–790. [PubMed: 12435214]
14. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med.* 2003 Dec; 26(3):172–181. [PubMed: 14644693]
15. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol.* 1992 Jun; 45(6):613–619. [PubMed: 1607900]
16. Bach PB, Pham HH, Schrag D, Tate RC, Hargraves JL. Primary care physicians who treat blacks and whites. *N Engl J Med.* 2004 Aug 5; 351(6):575–584. [PubMed: 15295050]
17. Nattinger AB, Laud PW, Bajorunaite R, Sparapani RA, Freeman JL. An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health Serv Res.* 2004 Dec; 39(6 Pt 1):1733–1749. [PubMed: 15533184]

18. Barton MB, Elmore JG, Fletcher SW. Breast symptoms among women enrolled in a health maintenance organization: frequency, evaluation, and outcome. *Ann Intern Med.* 1999 Apr 20; 130(8):651–657. [PubMed: 10215561]
19. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care.* 2002 Aug; 40(8 Suppl):IV-3–IV-18.

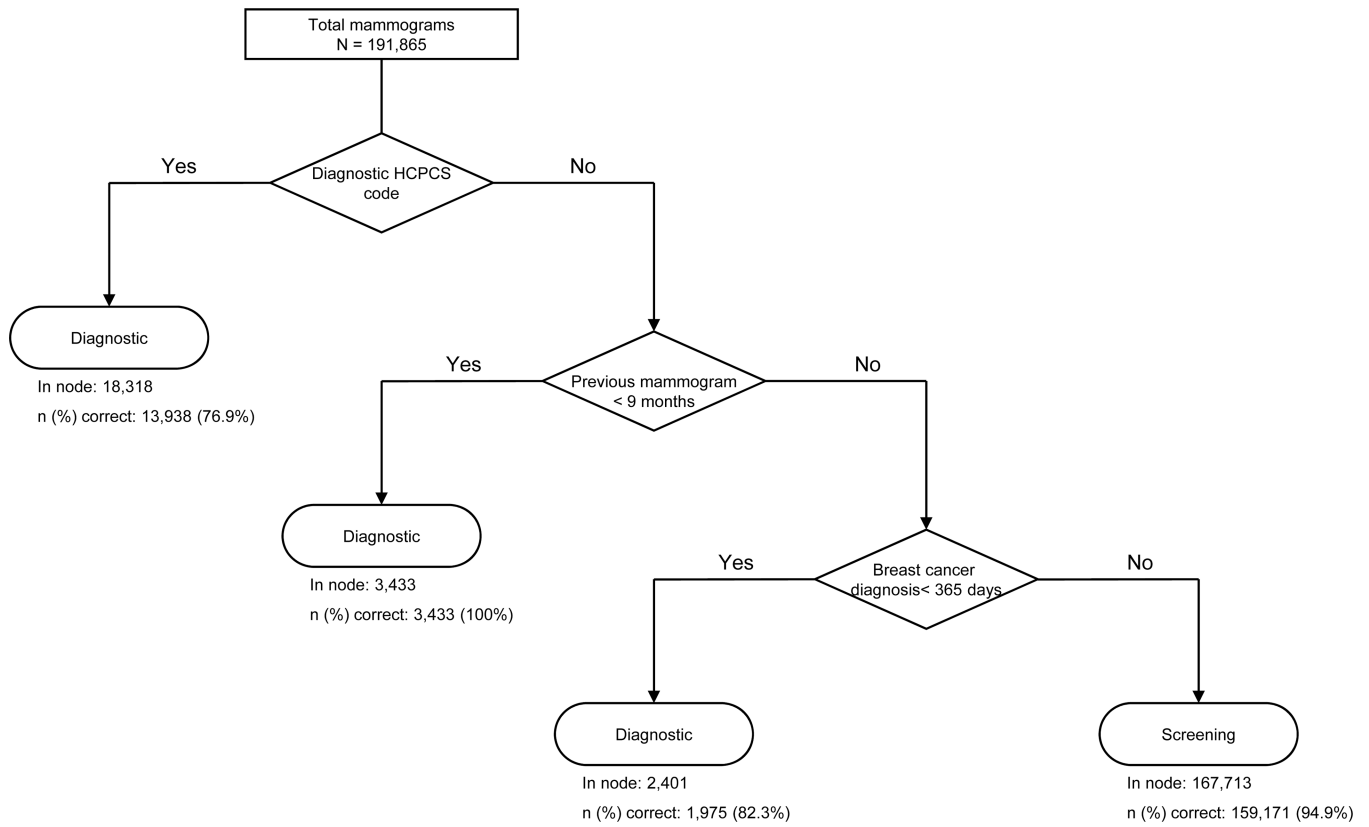


Figure 1. Allocation of mammograms by screening vs. diagnostic purpose using the three-step algorithm for claims without linkage to cancer registry data

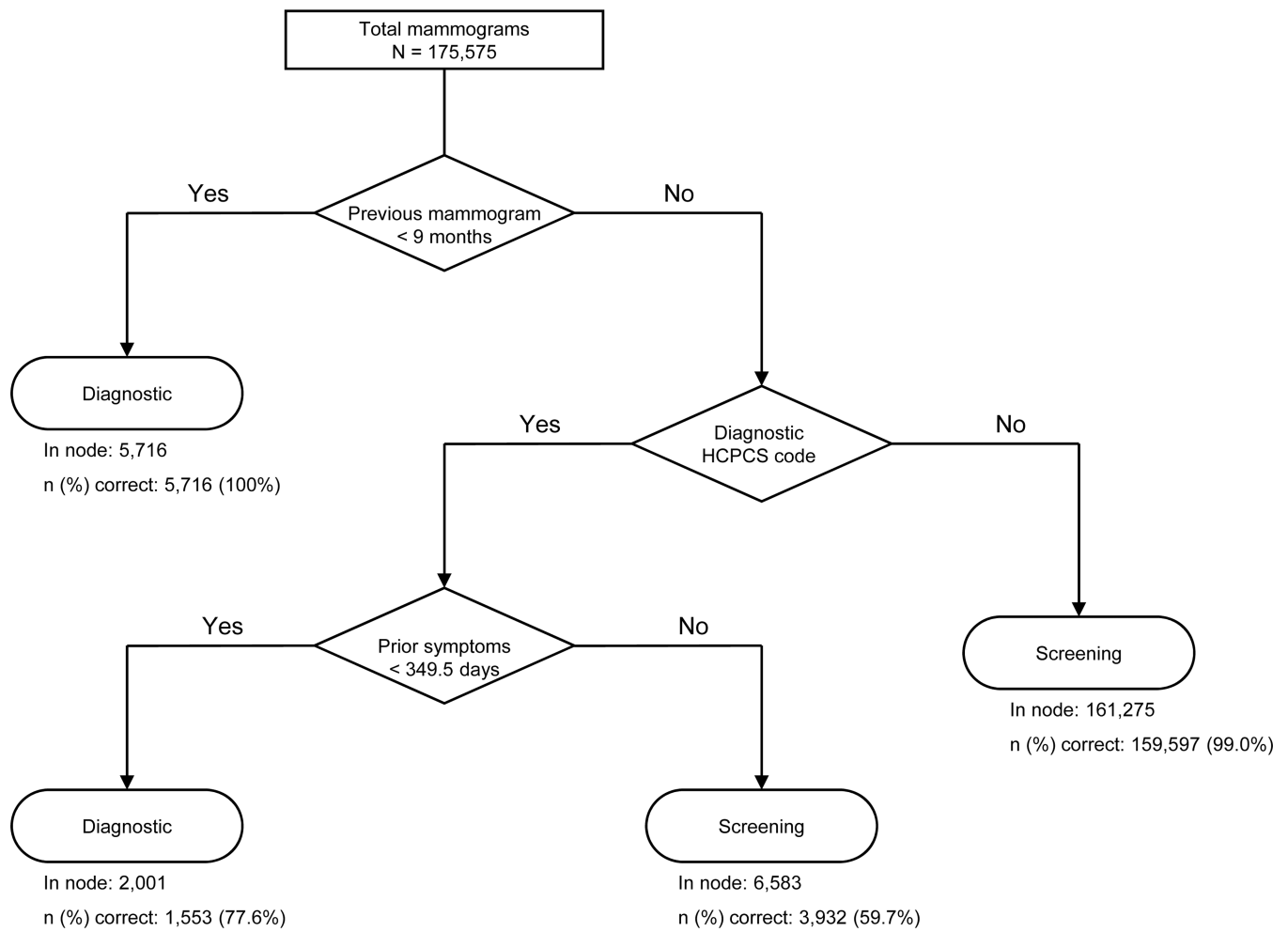


Figure 2. Allocation of mammograms by screening vs. diagnostic purpose using the three-step algorithm for mammography claims with linkage to cancer registry data

Table 1

Patient and Mammogram Characteristics in the Matched Breast Cancer Surveillance Consortium and Medicare Claims Samples *

Characteristic	All (N=383,730)	Training (N=191,865)	Testing (N=191,865)
Age	n (%)	n (%)	n (%)
66–74	223,155 (58.2)	111,654 (58.2)	111,501 (58.1)
75–84	140,364 (36.6)	70,010 (36.5)	70,354 (36.7)
85+	20,211 (5.3)	10,201 (5.3)	10,010 (5.2)
Race			
White, non-Hispanic	310,397 (80.9)	155,180 (80.9)	155,217 (80.9)
Black	30,466 (7.9)	15,243 (7.9)	15,223 (7.9)
Asian/Pacific Islander	8,905 (2.3)	4,457 (2.3)	4,448 (2.3)
American Indian/Alaskan Native	1,624 (0.4)	776 (0.4)	848 (0.4)
Hispanic	6,149 (1.6)	3,105 (1.6)	3,044 (1.6)
Other/mixed/unknown	26,189 (6.8)	13,104 (6.8)	13,085 (6.8)
Months since previous mammogram			
0–18	282,898 (73.7)	141,352 (73.7)	141,546 (73.8)
19–30	49,192 (12.8)	24,673 (12.9)	24,519 (12.8)
31–42	15,086 (3.9)	7,581 (4)	7,505 (3.9)
>42	14,456 (3.8)	7,246 (3.8)	7,210 (3.8)
Missing/unknown	22,098 (5.8)	11,013 (5.7)	11,085 (5.8)
Mammogram Type			
Film-screen	356,848 (93)	178,299 (92.9)	178,549 (93.1)
Digital	26,871 (7)	13,562 (7.1)	13,309 (6.9)
Missing/unknown	11 (0)	4 (0)	7 (0)
Year of Mammogram			
1999	45,582 (11.9)	19,735 (11.2)	22,829 (11.9)
2000	47,361 (12.3)	20,326 (11.6)	23,714 (12.4)
2001	53,436 (13.9)	24,455 (13.9)	26,683 (13.9)
2002	64,901 (16.9)	30,267 (17.2)	32,588 (17)
2003	62,430 (16.3)	31,308 (16.3)	31,122 (16.2)
2004	57,094 (14.9)	28,578 (14.9)	28,516 (14.9)
2005	52,926 (13.8)	26,513 (13.8)	26,413 (13.8)
Charlson comorbidity index**			
0	279,265 (72.8)	139,809 (72.9)	139,456 (72.7)
1	78,213 (20.4)	38,915 (20.3)	39,298 (20.5)
>=2	26,252 (6.8)	13,141 (6.8)	13,111 (6.8)
Prior History of Breast Cancer			
No	351,166 (91.5)	175,591 (91.5)	175,575 (91.5)
Yes	32,564 (8.5)	16,274 (8.5)	16,290 (8.5)

* Descriptive data derived from Medicare claims except for months since previous mammography, mammogram type, and prior history of breast cancer, which were derived from Breast Cancer Surveillance Consortium (BCSC). In Classification and Regression Tree analysis, classification variables were derived only from Medicare claims. Individual women may have received more than one mammogram.

Table 2
Performance of Four-Step Claims-based Algorithm for Distinguishing Screening from Diagnostic Mammograms

Algorithmic Designation*	Without cancer registry link (n=191,865 mammograms)		Excluding mammograms for women with prevalent breast cancer (n=175,575)	
	Screening	Diagnostic/Other	Screening	Diagnostic/Other
	163,935	18,380	163,462	5,555
	42	9,508	515	6,043
Performance Measures[†]				
Sensitivity, %		99.97		99.69
Specificity, %		34.09		52.10
PPV of "Screening" Designation, %		89.92		96.71
NPV of "Diagnostic" Designation, %		99.56		92.15
Kappa		0.469		0.649

Abbreviations: BCSC=Breast Cancer Surveillance Consortium; PPV=Positive Predictive Value; NPV=Negative Predictive Value.

* Developed with claims data linked to cancer registries,¹⁰ the algorithm includes four steps: 1) mammography claims within the prior nine months; 2) whether the claim HCPCS code was for a "screening" rather than a "diagnostic" mammogram; 3) claims containing codes for breast symptoms or procedures within the prior ninety days; and 4) the incidence of breast cancer within six months of mammography.

[†] 95% confidence intervals on all performance measures are very narrow owing to the large sample size.

Table 3
Performance of Three-Step Claims-based Algorithms for Distinguishing Screening from Diagnostic Mammograms

Algorithmic Designation	Claims Without Cancer Registry Linkage (N=191,865 mammograms)*		Excluding Prevalent Breast Cancers Using Cancer Registry Data (N=175,575 mammograms)†	
	Screening	Diagnostic/Other	Screening	Diagnostic/Other
Screening	159,171	8,542	163,529	4,329
Diagnostic	4,806	19,346	448	7,269
Performance Measures				
Sensitivity, %		97.07		99.73
Specificity, %		69.37		62.67
PPV of "Screening" Designation, %		94.91		97.42
NPV of "Diagnostic" Designation, %		80.10		94.19
Kappa		0.704		0.739

Abbreviations: BCSC=Breast Cancer Surveillance Consortium; PPV=Positive Predictive Value; NPV=Negative Predictive Value.

* Developed with Medicare claims data without linkage to breast cancer registry data (and therefore including claims for women with prevalent breast cancers prior to mammography), the algorithm includes three steps: 1) whether the claim HCPCS code was for a "diagnostic" rather than a "screening" mammogram; 2) mammography claims within the prior nine months; and 3) claims containing diagnostic codes for breast cancer within the year prior to mammography.

† Developed with Medicare claims data linked to breast cancer registry data (and therefore excluding claims for women with prevalent breast cancers prior to mammography), the algorithm includes three steps: 1) mammography claims within the prior nine months; 2) whether the claim HCPCS code was for a "diagnostic" rather than a "screening" mammogram; and 3) claims containing diagnostic codes for breast symptoms or signs within the 349 days prior to mammography.

Table 4

Reasons for Misclassification by Three-Step Algorithms

Type of Misclassification	Mammograms without Cancer Registry Linkage (N=191,865)		Excluding mammograms for women with prevalent breast cancer (N=175,575)	
	Reason for Misclassification	n (%)	Reason for Misclassification	n (%)
Misclassified as diagnostic	Diagnostic HCPCS code Breast cancer diagnoses on claims in prior year *	4,380(2.3) 426 (0.2)	Diagnostic HCPCS code	448 (0.3)
Misclassified as screening	Prior history of breast cancer not detected by claims BCSC data indicating breast imaging within 9 mo Non-screening indication †	6,802 (3.5) 947 (0.4) 793 (0.4)	Non-screening indication † BCSC data indicating breast imaging within 9 mo	3,350 (1.9) 979 (0.5)
Total misclassified		13,348 (7.0)		4,777 (2.7)

Abbreviations: HCPCS=Healthcare Common Procedure Coding System; ICD-9-CM=International Classification of Disease, 9th Edition, Clinical Modification; BCSC=Breast Cancer Surveillance Consortium; mo=months

* Breast cancer diagnosis codes can appear on claims of women without breast cancer (e.g., during evaluation to “rule-out” breast cancer).

† Clinical indications for mammogram based upon BCSC mammogram data. Non-screening indications include evaluation of breast symptoms/signs and short interval follow-up.