Innovations in decreasing recall rates for screening mammography

CLINICAL RELEVANCE:

Performance metrics are critical for assessing quality in mammography. Innovations that facilitate decreased recall rates while maintaining quality metrics benefit patient care while decreasing costs, adding value to the screening program.

INTRODUCTION:

Screening mammography is widely used in the United States in average risk women over the age of 40 and high-risk women at younger ages. Recall rates from screening mammography have an accepted standard of 10% or less in the United States, with estimated recall costs on the order of $1238 per case. Our work reviews a method to reduce recall rates, therefore improving performance while decreasing costs.

METHODS AND MATERIALS:

We first measured baseline recall rates before implementation of an intervention for both 2D and 3D mammography. The first intervention surveyed 10 breast regarding recall rates and perceptions of performance. Following discussions of survey results, the second intervention required all breast imagers to review personal recalls, record outcomes, and identify patterns that enhanced cancer detection and false positive results. The third intervention implemented consensus double reading of all recalls. This required that two radiologists agreed whether or not a recall was necessary. Recall rates, cancer detection rate and positive predictive values (PPV) were recorded and compared over the time frame of the interventions.

RESULTS:

An average of 2.3 minutes per case were spent on consulting another radiologist for each recall. 52,637 2D and 5,597 3D mammograms were included in the pre-intervention period from 1/3/2012 to 12/31/2014. The baseline recall rate, cancer detection rate, and PPV were 12.0%, 1.9/1000, and 25.3%, respectively, for 2D mammography, and 8.3%, 3.4/1000, and 43.2%, respectively, for 3D mammography. 8,452 2D mammograms and 8,695 3D mammograms were included in the post-intervention period from 2/1/2015 to 11/8/2015. The post-intervention recall rate, cancer detection rate, and PPV changed to 9.1%, 2.2/1000, and 33.3%, respectively, for 2D mammography, and 6.7%, 5.7/1000, and 39.2%, respectively, for 3D mammography.
CONCLUSIONS:

Reduced recall rates from screening mammography are desirable if performance metrics remain favorable or improve. Our method of radiologist education and awareness improved overall performance, and is possible to implement in any breast imaging setting.

Authors
Corresponding: Babita Panigrahi (Johns)

Babita Panigrahi (Johns), Jacqueline Hollada (The Johns Hopkins Bloomberg School of Public Health), William Speier (University of California, Los Angeles), Susan Harvey (Johns Hopkins Hospital)

Screening breast cancer detection rates before and after complete conversion to digital breast tomosynthesis

Purpose:

Reader studies as well as randomized trials have demonstrated digital breast tomosynthesis (DBT) to be more sensitive than full field digital mammography (FFDM). While compelling, these studies retain potential for selection bias. Having undergone total conversion from FFDM to DBT four years ago, our experience allows comparison before and after conversion in a single setting without selection bias.

Methods:

DBT was introduced in October 2011 and conversion was completed in April 2012. Data was acquired prospectively as part of MQSA audit. After IRB approval, HIPPA-compliant re-tabulation organized the data into three 12-month periods before DBT and three after complete conversion, excluding the transition. Only patients presenting for screening mammography were included. Patients presenting with clinical signs or symptoms were excluded. The number of patients included for each 12-month period of FFDM was 22888, 21537 and 20664. The number for DBT was 20709, 21584, and 21846.

Cancer detection rates (CDR) and invasive cancer detection rates (ICDR) were calculated for each period.

Results:

During the three 12-month periods of FFDM, CDR was 0.48%, 0.44%, and 0.49% (overall 0.47% or 304 cancers/65089), while ICDR was 0.29%, 0.25% and 0.24% (overall 0.26% or 169 invasive cancers/65089). In the three years following full implementation of DBT, CDR increased to 0.62%, 0.60% and 0.51% (overall 0.58% or 370 cancers/64139 patients) while ICDR was 0.43%, 0.47% and 0.35% (overall 0.42% or 267 invasive cancers/64139). DCIS cases trended lower with 131 before DBT and 101 after DBT.

Increased detection of invasive cancers was statistically significant with improvements for the 1st, 2nd
and 3rd periods calculated at 67.64% (Fisher p value ≤ 0.0161), 80.22% (≤ 0.0001), and 33.99% (≤ 0.039) over the average for FFDM. The trend was evident across all ages.

Conclusion:

The improvement in ICDR after implementation of DBT is unequivocal. Detection improved still further during the second year. The less striking ICDR in year three is explained by the fact that 85% of diagnosed patients had had prior DBT examinations. ICDR can be predicted to return to pre DBT levels unless the incidence of interval cancers decreases.

Clinical relevance statement:

While DBT is gaining acceptance and is increasingly being offered alongside FFDM, its effectiveness when provided to all patients needs to be understood. In addition, examining its impact on a screening program over time will also be essential to determine whether DBT could or should replace FFDM.

Authors

Corresponding: Alan Semine (Newton Wellesley Hospital, Tufts University School of Medicine)

Laura Semine-Misbach (Beth Israel Deaconess Medical Center), Alan Semine (Newton Wellesley Hospital, Tufts University School of Medicine)

Trends in screening mammography interpretation over 4 years with digital breast tomosynthesis: are recall and cancer detection rates maintained?

Purpose: Digital breast tomosynthesis (DBT) has been shown to reduce recall rates and increase cancer detection in screening compared to 2D-mammography. Current studies have primarily assessed results in prevalence DBT screens. There is little data that address DBT performance in subsequent years of interpretation. The purpose of this study was to assess trends in the recall rates (RR), cancer detection rates (CDR) and positive predictive values for recalls (PPV1) in DBT screening over a 4-year period and evaluate individual radiologist’s performance over time.

Materials and Methods: A HIPAA-compliant search of the electronic database (PenRad, MN) at a single institution was performed to identify DBT screens interpreted by 7 breast-subspecialized radiologists since introduction of DBT at our facility. Data were divided into one-year periods: (1)8/1/11–7/31/12, (2)8/1/12–7/31/13, (3)8/1/13–7/31/14, (4)8/1/14–7/31/15. Overall RR, CDR and PPV1 for each period and for individual radiologists were assessed. Percentages of in-situ vs invasive cancers were evaluated.
Results: 16,587 screening DBT exams were interpreted by 7 radiologists during the 4-year period. Overall, there was a trend towards decreasing RR over time after an initial increase in RR in year two: (1)7.6%, (2)9%, (3)8.2%, (4)7.2%. A similar trend in RR was observed on an individual basis in 4 of the 7 radiologists. Two radiologists had stable RR over time, while one had slightly increased RR.

Overall CDR remained stable for the first three years and increased significantly during year 4 (7 vs 4.7-5.8 per 1000 exams). PPV1 was maintained for the duration of the study period (6.3%-7.5%).

A higher percentage of invasive cancers was detected over time (38% of screen-detected cancers in year 1 vs 75% in year 4).

Conclusion: Our results upheld current DBT prevalence data on performance measures and showed an improvement of these measures over time. DBT was associated with a trend towards decreasing RR, increasing CDR and increasing proportion of screen-detected invasive cancers over a 4-year period. A similar trend in RR was observed on an individual basis.

Clinical Relevance: As radiologists’ experience and performance with DBT continue to improve, RR and CDR (especially for invasive cancers) are expected to improve accordingly.

Authors
Corresponding: Christine Chen (Yale-New Haven Hospital)

Christine Chen (Yale-New Haven Hospital), Melissa Durand (Yale School of Medicine), Madhavi Raghu (Yale-New Haven Hospital), Liane Philpotts (Yale-New Haven Hospital)

Effect of Mammography Screening on Incidence Based Breast Cancer Mortality
Purpose: To understand the burden of breast cancer in terms of deaths arising from cancers detectable at various ages, i.e., incidence-based mortality (IBM). To estimate the influence of different mammography screening regimens on IBM.

Materials: The Wisconsin CISNET Breast Cancer Model was utilized. Modern data on sensitivity and specificity of mammography versus age and breast density, data on effectiveness of treatment versus size and stage of cancer.

Methods: The model was run for a birth cohort of 2,000,000 women to estimate the number and age distribution at which breast cancers arise, tumor size and stage at detection. The number of breast
cancer deaths and age of death from breast cancer or the age of death from other causes were calculated and deaths expressed versus the age at which the initial breast cancer became detectable. From life tables, the number of life-years (LY) lost due to breast cancer was estimated and expressed as incidence-based LY lost. Estimates were performed for an unscreened cohort and for women whose screening began at age 40 or 50, ended at age 69 or 74 and was annual, biennial or triennial or an age-dependent screening interval. The burden of breast cancer deaths and LY lost and the impact of screening were examined using the incidence-based construct.

Results: Depending on the screening regimen, mammography screening is associated with a reduction in breast cancer mortality by between 4 and 10 deaths per 1000 women, corresponding to 28% to 53% mortality reduction compared to No Screening. From the perspective of incidence-based mortality, the breast cancer deaths come largely from cancers detectable between the mid forties to the mid seventies. When expressed in terms of life-years (LY) lost due to breast cancer, the distribution of the burden in an unscreened population shifts downward by 6.5 years to between the early forties and the late sixties. The lifesaving effect of screening behaves in a similar manner; the distribution of the benefit of screening, i.e., the recovery of life-years lost due to breast cancer is centered at a lower age than that of breast cancer deaths averted. In the presentation the specific contributions associated with different screening regimens will be discussed.

Conclusion: Insight into the age-dependence of the burden of breast cancer as well as the potential of different screening regimens to reduce breast cancer deaths or life-years lost can be gained by considering IBM predictions derived from modeling.

Authors
Corresponding: Martin Yaffe (Sunnybrook Research Institute, University of Toronto)

Martin Yaffe (Sunnybrook Research Institute, University of Toronto)

The Risk of Risk-Based Breast Cancer Screening: Inability of Family History and Density to Predict Malignancy
Purpose: To evaluate screen-detected breast cancers and correlate with risk factors such as family history and breast density.

Materials and Methods: Following IRB approval, retrospective search of records at a single institution NCCN-designated cancer center identified 256 screen-detected breast cancers in 250 patients. Inclusion criteria were screen-detected in-situ and invasive breast cancers in asymptomatic females between
1/1/2011 and 12/31/2012. Patients with clinical symptoms, outside mammogram, non-primary breast cancers, or abnormal recent imaging study such as PET were excluded from the non-palpable cancer database. Medical records were reviewed for clinical history, family history, pathology, cancer size, nodal status, density and mammogram findings. Mammograms were interpreted by 1 of 14 MQSA certified breast imaging radiologists with 2-25 years of experience. Patient ages ranged from 36-87 (mean = 60) years. Positive family history was defined as a first-degree relative with breast cancer. Dense mammograms included heterogeneous or extremely dense classifications.

Results: 256 breast cancers were detected with screening mammography during the study period for an overall detection rate of 5.3/1000 mammograms (256 cancers/48205 screening mammograms). 71.9% (184/256) of breast cancers occurred in patients without family history and 27.7% (71/256) occurred in patients with family history. One patient was adopted. 40.2% (103/256) of breast cancers were detected in patients with dense breasts. 59.8% (153/256) of the breast cancers were detected in patients with non-dense breasts. 24.8% (62/250) of the patients had a personal history of breast cancer or high-risk lesion (e.g. ADH). 29.2% (73/250) of screen detected cancer patients had no family, density, or personal risk factors. 20.7% (53/256) of the breast cancers occurred in women < 50 years old. 36.7% (94/256) of the breast cancers were DCIS and 63.2% (162/256) were invasive cancers. Both the mean and median size of the invasive cancers was 11 mm. 172 patients had axillary staging; 87.2% (150/172) had N0 disease and 12.8% (22/172) had > N1 disease.

Conclusion: Most screen detected breast cancers developed in patients without dense tissue, a family history of breast cancer, or a personal history of breast cancer/high risk lesion. Exclusive use of risk-based screening would result in delayed breast cancer detection for many women.

Clinical Relevance: Risk-based screening mammography based on family history or dense tissue has been promoted, but would result in delayed diagnosis of many breast cancers. Many cancers occurred in patients without these risks. Screening mammography should be applied broadly to maximally reduce breast cancer morbidity and mortality.

Authors
Corresponding: Colleen Neal (University of Michigan)

Wahida Rahman (University of Michigan), Colleen Neal (University of Michigan), Mitra Noroozian (University of Michigan), Renee Pinsky (University of Michigan), Mark Helvie (University of Michigan)
Screening Outcomes of Digital Breast Tomosynthesis versus Digital Mammography in an Underserved Population

Purpose: The purpose of this report is to compare screening outcomes in an underserved population with two different imaging techniques: digital breast tomosynthesis (DBT) and standard digital mammography (DM). Materials and Methods: All cancer screens were performed according to American Cancer Society guidelines by technologists certified in mammography. All mammographic images were digital and were read by one of two board certified radiologists, specialists in mammography and trained to read tomosynthesis images. For this report, the screening results are reported in terms of the Breast Imaging-Reporting and Data System (BI-RADS). Of special interest is a category of BI-RADS 0, which warrants either an effort to ascertain prior imaging for comparison or to call the patient back for additional views and/or higher quality images. For the population we serve, in the remote rural areas, or the working poor in the non-rural areas, recalls are a significant extra burden on the patients. Results of screening from July 1, 2014 through June 30, 2015, with both mobile units and the fixed clinic are presented. Chi-Squared tests were used to calculate p values. For the meeting presentation the numbers will be updated to include more patients screened and demographics and area health measures will be included.

Results: For DM (fixed and mobile clinic combined) there were 2,319 patients screened during the year, with 277 (11.9%) recalls. There were 11 cancers diagnosed with DM, a screening detection rate of 4.74 cancers per 1,000 screened. With DBT, there were 1,077 patients screened and 44 (4.1%) recalls and 6 cancers detected for a screening detection rate of 5.57 per 1,000 screened. The recall rate was significantly lower for DBT, p < 0.0001. The cancer detection rate was 18% higher with DBT but was not significant. Approximately 23% of the patients seen, on the mobile units, with a BI-RADS score of 0 did not return for follow-up. For the fixed clinic the no-show rate was 5.2%, significantly lower than the mobile clinic patients, p < 0.0001. This strikingly high difference in no-show rates illustrates the magnitude of the barriers to breast cancer screening in the rural, underserved populations of Louisiana.

Conclusion: This study supports the use of tomosynthesis as a valuable asset in mobile screening.

Clinical Relevance: Our data supports digital breast tomosynthesis as a screening method in underserved populations.

Authors
Corresponding: Stacey Massey (LSU Health Shreveport)

Stacey Massey (LSU Health Shreveport), Stacy Lee (LSU Health Shreveport), Jerry McLarty (LSU Health Shreveport), Jennifer Lance (LSU Health Shreveport), Nelson Luraguis (LSU Health Shreveport)
Perspectives on Screening: An Analysis of Breast Screening in the United Kingdom

Purpose: The structure of the United States breast screening program is currently a topic of national discussion. Any evaluation of breast screening efficacy may be strengthened by exposure to and understanding of alternative models of screening. The purpose of this study is therefore to examine the parameters of the United Kingdom national population based screening program in order to initiate discussion regarding practices and assumptions of screening in the United States.

Materials and Methods: Policies and outcomes measures of the United Kingdom National Breast Screening Service are reviewed with respect to the UK breast imaging including: screening interval and roundlength, age of women to be screened, uptake rates, and target recall rates, as well as UK mammography reporting. Reading systems, including double reading, arbitration, and performance assessment are reviewed. Methods and results of data collection and audit, including statistics on cancer detection rates and interval cancer evaluation are also discussed. Finally UK quality assurance and multidisciplinary systems for audit (for example, local multidisciplinary team meetings, regional quality assurance oversight) are investigated.

Results: The United Kingdom screened 2.74 million women from ages 50-70 in 2013-2014. A randomized trial for age extension (47-73) is currently active. Screening occurs every 3 years for average risk women. Overall uptake was 72.1% in 2013-2014 [minimum standard 70%]. Minimum standard recall rate is <10%. Mammograms for recalled women are categorized on a 1-5 scale which is correlated to clinical exam and biopsy pathology results. In 2013-14, there were 17,961 cancers detected (8.6 cancers/1000 women screened) of which 7175 were small invasive cancers (<15 mm). Regional and national outcomes are periodically reviewed with quality assurance meetings and audits.

Conclusion: Appropriate guidelines for screening are currently a controversial and debated topic. Screening outcomes for distinct national populations may not be directly comparable; however, understanding and awareness of international screening programs may help inform debate regarding United States screening practices.

Clinical Relevance: United States breast screening guidelines are a topic of national discussion. Understanding international models of screening and outcomes may help inform current debates.
Authors
Corresponding: Samantha Heller (new)

Samantha Heller (new), Louise S. Wilkinson (St. George’s Hospital)