Title

Collaborative Modeling of the Benefits and Harms Associated With Different U.S. Breast Cancer Screening Strategies.

Journal

Annals of Internal Medicine, 164(4)

ISSN

1056-8751

Authors

Mandelblatt, Jeanne S
Stout, Natasha K
Schechter, Clyde B
et al.

Publication Date

2016-02-16

DOI

10.7326/m15-1536

Peer reviewed

Published in final edited form as:

Ann Intern Med. 2016 February 16; 164(4): 215–225. doi:10.7326/M15-1536.

COLLABORATIVE MODELING OF THE BENEFITS AND HARMS ASSOCIATED WITH DIFFERENT U.S. BREAST CANCER SCREENING STRATEGIES

Jeanne S. Mandelblatt,
Department of Oncology, Georgetown University Medical Center and Cancer Prevention and Control Program, Georgetown-Lombardi Comprehensive Cancer Center, Washington, DC, USA.

Natasha K. Stout,
Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA.

Clyde B. Schechter,
Departments of Family and Social Medicine and Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York, USA.

Jeroen J. van den Broek,
Department of Public Health, Erasmus Medical Center, Rotterdam, the Netherlands.

Diana Miglioretti,
Department of Public Health Sciences, UC Davis School of Medicine, Davis, California, USA and Group Health Research Institute, Seattle, WA, USA.

Martin Krapcho,
Information Management Services (IMS), Calverton, Maryland, USA.

Amy Trentham-Dietz,
Carbone Cancer Center, University of Wisconsin, Madison, Wisconsin, USA.

Diego Munoz,
Departments of Biomedical Informatics and Radiology, School of Medicine, Stanford University, Stanford, California, USA.

Sandra J. Lee,

Please address correspondence to: Jeanne Mandelblatt, MD, MPH, Lombardi Comprehensive Cancer Center, 3300 Whitehaven Street, NW, suite 4100, Washington, DC 20007, Telephone 202-687-0812/facsimile 202-687-0305, mandelbj@georgetown.edu.
From the Breast Cancer Working Group of the Cancer Intervention and Surveillance Modeling Network (CISNET)* and the Breast Cancer Surveillance Consortium (BCSC)

*This work was done by six independent modeling teams from Dana-Farber Cancer Institute (PI: Lee), Erasmus Medical Center (PI: de Koning), Georgetown University Medical Center, Lombardi Comprehensive Cancer Center/A. Einstein College of Medicine (PI: Mandelblatt/Schechter), Harvard Medical School, University of Wisconsin/ Harvard Pilgrim Health Care (PI: Trentham-Dietz/Stout/ Alagoz), MD Anderson Comprehensive Cancer Center (PI: Berry) and Stanford University (PI: Plevritis). Jeanne Mandelblatt, Kathy Cronin, Harry de Koning, Diana Miglioretti, Clyde Schechter, and Natasha Stout were the coordinating committee for the project; Jeanne Mandelblatt, Natasha Stout, Clyde Schechter, and Kathy Cronin were the writing committee; Harry de Koning and Kathleen Cronin served as duel senior authors; and Eric Feuer was responsible for overall CISNET project direction. BCSC collaborated and also provided key data inputs to the modeling teams.

Portions of earlier versions of this research were presented on the US Preventive Services Task Force web site as part of a Technical Report (AHRQ Publication No. 14-05201-EF-4).

Potential Conflicts of Interest**: None disclosed**

Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard Medical School Boston, Massachusetts, USA.

Donald A. Berry,
Department of Biostatistics, University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA.

Nicolien T. van Ravesteyn,
Department of Public Health, Erasmus Medical Center, Rotterdam, the Netherlands.

Oguzhan Alagoz,
Carbone Cancer Center, University of Wisconsin, Madison, Wisconsin, USA.

Department of Industrial and Systems Engineering, University of Wisconsin, Madison, Wisconsin, USA.

Karla Kerlikowske,
Department of Medicine and Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California, USA.

Anna N.A. Tosteson,
Norris Cotton Cancer Center and The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA.

Aimee M. Near,
Department of Oncology, Georgetown University Medical Center and Cancer Prevention and Control Program, Georgetown-Lombardi Comprehensive Cancer Center, Washington, DC, USA.

Amanda Hoeffken,
Department of Oncology, Georgetown University Medical Center and Cancer Prevention and Control Program, Georgetown-Lombardi Comprehensive Cancer Center, Washington, DC, USA.

Yaojen Chang,
Department of Oncology, Georgetown University Medical Center and Cancer Prevention and Control Program, Georgetown-Lombardi Comprehensive Cancer Center, Washington, DC, USA.

Eveline A. Heijnsdijk,
Department of Public Health, Erasmus Medical Center, Rotterdam, the Netherlands.

Gary Chisholm,
Department of Biostatistics, University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA.

Xuelin Huang,
Department of Biostatistics, University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA.

Hui Huang,
Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard Medical School Boston, Massachusetts, USA.

Mehmet Ali Ergun,

Department of Industrial and Systems Engineering, University of Wisconsin, Madison, Wisconsin, USA.

Ronald Gangnon,
Carbone Cancer Center, University of Wisconsin, Madison, Wisconsin, USA.

Department of Biostatistics and Medical Informatics and Population Health Sciences, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA.

Brian L. Sprague,
Department of Surgery, College of Medicine, University of Vermont, Burlington, Vermont, USA.

Sylvia Plevritis,
Department of Radiology, School of Medicine, Stanford University, Stanford, California, USA.

Eric Feuer,
Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland, USA.

Harry J. de Koning, and
Department of Public Health, Erasmus Medical Center, Rotterdam, the Netherlands.

Kathleen A. Cronin
Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland, USA.

Abstract

Background—Controversy persists about optimal mammography screening strategies.

Objective—To evaluate mammography strategies considering screening and treatment advances.

Design—Collaboration of six simulation models.

Data Sources—National data on incidence, risk, breast density, digital mammography performance, treatment effects, and other-cause mortality.

Target Population—An average-risk cohort.

Time Horizon—Lifetime.

Perspective—Societal.

Interventions—Mammograms from age 40, 45 or 50 to 74 at annual or biennial intervals, or annually from 40 or 45 to 49 then biennially to 74, assuming 100% screening and treatment adherence.

Outcome Measures—Screening benefits (vs. no screening) include percent breast cancer mortality reduction, deaths averted, and life-years gained. Harms include number of mammograms, false-positives, benign biopsies, and overdiagnosis.

Results for Average-Risk Women—Biennial strategies maintain 79.8%-81.3% (range across strategies and models: 68.3–98.9%) of annual screening benefits with almost half the false-positives and fewer overdiagnoses. Screening biennially from ages 50–74 achieves a median

25.8% (range: 24.1%-31.8%) breast cancer mortality reduction; annual screening from ages 40–74 years reduces mortality an additional 12.0% (range: 5.7%-17.2%) vs. no screening, but yields 1988 more false-positives and 7 more overdiagnoses per 1000 women screened. Annual screening from ages 50–74 had similar benefits as other strategies but more harms, so would not be recommended.

Sub-population Results—Annual screening starting at age 40 for women who have a two- to four-fold increase in risk has a similar balance of harms and benefits as biennial screening of average-risk women from 50–74.

Limitations—We do not consider other imaging technologies, polygenic risk, or non-adherence.

Conclusion—These results suggest that biennial screening is efficient for average-risk groups, but decisions on strategies depend on the weight given to the balance of harms and benefits.

Primary Funding Source—National Institutes of Health

Introduction

Despite decades of mammography screening for early breast cancer detection, there is no consensus on optimal strategies, target populations, or the magnitude of benefits and harms.(1–11) Based on data available at the time, the 2009 US Preventive Services Task Force (USPSTF) recommended biennial film mammography from ages 50–74, with suggestions for shared decision-making about whether to start screening in the 40's.(12) Since then, there are some new data regarding screening benefits,(2,6,8,9,11,13) digital mammography has replaced plain film,(14) and increasingly effective breast cancer systemic treatment regimens targeting molecular sub-types are in widespread use.(15)

These advances have the potential to affect conclusions about optimal breast cancer screening programs.(16) There is also growing interest in personalizing screening based on breast density, risk factors, life expectancy, and patient preferences.(16–21) Modeling has the advantage of considering these factors and providing a quantitative summary of the net balance of harms and benefits that incorporate preferences (utilities), while holding selected conditions (e.g., treatment effects) constant, facilitating strategy comparisons.(22,23) Collaboration of several models provides a range of plausible effects and illustrates the impact of differences in model assumptions.(1,7,24)

We use six well-established simulation models to synthesize new data to examine outcomes of biennial or annual digital mammography screening starting at ages 40, 45, or 50 through age 74 among average-risk women. In secondary analyses we also examine how breast density, and risk- or comorbidity-level affects results, and whether utilities for health states related to screening and its downstream consequences affect conclusions. The results are intended to contribute to current practice and policy debates.

Methods

The models were developed independently within the Cancer Intervention and Surveillance Modeling Network (CISNET) (25–31) and were institutional review board approved. The models included model D (Dana-Farber Cancer Institute, Boston, Massachusetts), model E

(Erasmus Medical Center, Rotterdam, the Netherlands), model GE (Georgetown University Medical Center, Washington, DC and Albert Einstein College of Medicine, Bronx, New York), model M (MD Anderson Cancer Center, Houston, Texas), model S (Stanford University, Stanford, California), and model W (University of Wisconsin, Madison, Wisconsin and Harvard Medical School, Boston, Massachusetts).

Since our earlier analysis,(1) the models have undergone substantial revision to reflect advances in breast cancer control, including: portrayal of four distinct molecular subtypes based on estrogen receptor (ER) and human epidermal growth factor-2 receptor (HER2) status;(24) current population incidence (32) and competing non-breast cancer mortality; digital screening; and the most current therapies.(33) All models (except Model S) include DCIS. The general modeling approach is summarized below; full details are available at: https://resources.cisnet.cancer.gov/registry and (34).

The models begin with estimates of breast cancer incidence (32) and ER/HER2-specific survival trends *without* screening or adjuvant treatment and then overlay data on screening and molecular subtype-specific adjuvant treatment to generate observed US population incidence and mortality trends.(1,7,16,24,35) Breast cancers have a distribution of preclinical screen-detectable periods (sojourn time) and clinical detection points. Digital mammography performance characteristics are based on age, first vs. subsequent screen, time since last mammogram, and breast density. ER/HER2 status is assigned at diagnosis based on stage and age. Molecular sub-type- and stage-specific treatment reduces the hazards of breast cancer death (models D, GE, M, and S) or results in a cure for some cases (models E and W). Women can die of breast cancer or other causes.

Screen detection of cancer during the preclinical screen-detectable period can result in the identification (and treatment) of earlier-stage or smaller tumors than might occur via clinical detection, with a corresponding reduction in breast cancer mortality.

We used a cohort of women born in 1970 with average population risk and breast density and follow them from age 25 (since breast cancer is rare before this age [0.08% of cases]) until death.

Model Input Parameters

The models begin with a common set of age-specific variables for breast cancer incidence, digital mammography performance characteristics, ER/HER2-specific treatment effects, and average and comorbidity-level specific-non-breast cancer competing causes of death. (34) In addition, each group includes model-specific inputs (or intermediate outputs) to represent preclinical detectable times, lead-time, and age- and ER/HER2-specific stage distribution in screen- vs. non-screen-detected women on the basis of their specific model structure. (1,7,24–31) These model-specific parameters are based on assumptions about combinations of values that reproduce US trends in incidence and mortality, including proportions of DCIS that are nonprogressive and would not be detected without screening. Models M and W also assume some small nonprogressive invasive cancers. The models adopt an age-period-cohort modeling approach to project breast cancer incidence rates in the absence of

screening;(32,36) Model M uses 1975–79 SEER rates. The models assume 100% adherence to screening and the most effective treatment to isolate the effect of screening strategies.

Four models use age-specific digital mammography sensitivity values observed in the Breast Cancer Surveillance Consortium (BCSC) for detection of invasive and DCIS cancers combined (model S only uses data for invasive cancers). Separate values are used for initial and subsequent mammography performed at annual or biennial intervals. An annual interval is defined as 9–18 months between examinations and biennial as 19–30 months.(37–39) Model D uses these data as input variables (29) and models GE, S, and W use the data for calibration.(25,26,28) Models E and M fit estimates from the BCSC and other data.(27,30)

All women with ER-positive tumors receive five years of hormonal therapy (tamoxifen if diagnosed <50 years and aromatase inhibitors if ≥50 years) and an anthracycline-based regimen accompanied by a taxane. Women with ER-negative invasive tumors receive anthracycline-based regimens with a taxane. Those with HER2-positive tumors receive trastuzumab. Women with ER–positive DCIS receive hormonal therapy.(15) Treatment effectiveness is based on clinical trials and is modeled as a reduction in mortality risk or increase in the proportion cured vs. ER/HER2-specific survival in the absence of therapy; (33) estimates assume women receive local therapy.

Benefits

Screening benefits (vs. no screening or other screening strategies) are measured using percent breast cancer mortality reduction, breast cancer deaths averted, and life-years (LYs) and quality-adjusted life-years (QALYs) gained because of averted or delayed breast cancer death. Benefits (and harms) are accumulated from ages 40–100 years to capture the lifetime impact of screening strategies.

To quality-adjust life years, we applied a disutility for age- and gender-specific general population health.(40) These were further adjusted to account for additional disutilities related to undergoing screening (−0.006 for one week), having an evaluation of a positive screen (−0.105 for five weeks), initial treatment by stage (for the first 2 years after diagnosis), and distant disease (for the last year of life for all women who die of breast cancer).(34,41,42)

Harms

Harms included number of mammograms, false-positive mammograms, benign biopsies, and overdiagnosis. We defined the rate of false-positive mammograms as the number of mammograms read as abnormal or needing further work-up in women without cancer divided by the total number of screening mammograms. Benign biopsies were defined as a biopsy recommendation among women with false-positive screening results.(43) Overdiagnosis was defined as cases that would not have been clinically detected in the absence of screening (because of lack of progressive potential or death from competing mortality). The impact of overdiagnosis on QALYs is captured by the disutility for being in a cancer state but dying of other causes without a change in life expectancy.

Analysis

We evaluated eight strategies that varied by starting age (40, 45, 50) and interval (annual, biennial, and hybrid [annual in the 40's and biennial in the 50's]); all strategies stop screening at age 74. We included "no screening" as baseline for the percent mortality reduction associated with any given strategy.

We ranked strategies by the number of mammograms performed (or harms) for each model. We report the median benefit and range across models. We also obtained an efficiency frontier by plotting the sequence of points that represent the largest incremental percent mortality reduction (or LYs or QALYs) per mammograms performed or harm entailed. Screening strategies that fall on this frontier are the most efficient (i.e., no alternative exists that provides more benefit with fewer resources/harms). The most intensive strategy is usually the top anchor point on the frontier.

Four models (model D, E, GE, and W) also evaluated how results vary when different subpopulations are screened. First, we investigated subpopulations based on their breast density: entirely fatty ("a"), scattered density ("b"), heterogeneously dense ("c") and extremely dense ("d"). Based on observed age-specific prevalence rates, density was assigned at age 40, and could have remained the same or decreased by one level at age 50 and again at age 65.(44) Density modified mammography sensitivity and specificity based on age, density, interval, and first vs. subsequent screening.(34) Density also modified risk of developing breast cancer for age groups 40–49, 50–64, and 65+, using the average population density in each age group as the referent category (BCSC unpublished data).(45,46) Density did not affect molecular subtype or disease natural history. Density results were grouped into low (a and b) and high density (c and d) for results presentation.

Additionally, we evaluated results for subpopulations with higher than average risk: 1.3 (e.g., nullparity or age at first live birth >30), (18,47) 2.0 (e.g., family history of one first degree relative), (18) or 4.0 times higher than average-risk (e.g., 2 or more first degree relatives).(18,46) Higher risk levels, such as seen with BRCA 1/2 mutations, were not considered since such women have specific screening guidelines. Finally, we examined the impact on results of combinations of density and risk level. We made the simplifying assumption that risk affected incidence, but not other aspects of disease.

In other subpopulation analyses, two models (model E and GE) examined the impact of comorbidity-specific vs. overall population competing mortality on upper ages of screening cessation based on comorbidity-specific life expectancy.(21,48,49) We compare results for continuing to screen biennially past age 74 among women with lower than average comorbidity or stopping earlier than 74 for those with moderate or high comorbidity. These analyses included women who survived and did not develop breast cancer up until the point where screening is to be extended or stopped.

Four models considered the impact of varying the values for disutilities in sensitivity analyses to identify if there was a strategy where high disutility would eliminate screening benefits. Finally, we evaluated the ability of the models to independently predict external trends and results (Appendix 1).

Role of the Funding Source

We worked with US Preventive Services Task Force and Agency for Healthcare Research and Quality (AHRQ) to develop the research question, but they had no role in study conduct. NCI investigators (KC, EF) collaborated on the research in their role as scientific project officers.

Results

Benefits

The models produced consistent rankings of screening strategies (Table 1). Biennial screening from ages 50 to 74 yielded a median 25.8% reduction in breast cancer mortality compared to no screening (range: 24.1%-31.8). Annual screening led to slightly greater reductions in mortality than biennial strategies. Rankings were similar for LYs and QALYs. For all benefit metrics (Table 2), strategies that included initiation at age 45 yielded results intermediate between those beginning at 40 and 50.

Incremental benefits of starting screening at age 40 vs. 50 were slightly greater in terms of breast cancer deaths averted for annual and biennial screening (median 1.3 [range: 1.1–1.7] vs. 1.0 [0.8–1.7] per 1000 women screened, respectively). (Table 3) Biennial strategies maintained a median of 79.8%-81.3% of the mortality reduction of annual screening (range across strategies and models 68.3–98.9%) (Table 4).

Harms

All of the models projected more false-positive results, benign biopsies, and overdiagnosed cases under annual vs. biennial schedules and starting at age 40 vs. age 50 (Table 2). For instance, if biennial screening began at age 40 instead of age 50, for every 1000 women screened there would be a median of 1 more death averted, but 576 more false-positive results, 58 benign biopsies, and 2 added overdiagnosed cases.

Efficiency Frontiers

Biennial strategies starting at age 40, 45, and 50 would generally be considered efficient when examining the balance of screening benefits and harms (Figure 1 and Appendix Figure 2). The hybrid strategies of annual screening starting at age 40 or 45, followed by biennial screening at age 50 were close to being efficient. Annual screening from ages 50 to 74 yielded the same or fewer benefits than the next least intensive strategy in all or most models depending on the measure of benefits, but required more mammograms or had more harms (i.e., was dominated), so would not be considered efficient.

Subpopulation Analyses

The rankings of strategies did not change when screening was targeted to groups with increasing risk levels; annual screening from ages 50 to 74 remained dominated across all risk levels (not shown). However, the balance of harms and benefits differed by risk group, with women who had higher risk having lower rates of false positives and higher gains from screening than lower risk groups. Screening higher risk women also yielded a lower proportion of overdiagnosed cases per death averted than screening women of average-risk.

For women with a two- to four-fold increase in risk, annual screening starting at age 40 or 45 had a similar or more favorable harm to benefit ratio (based on false positives) as biennial screening average-risk women from 50–74. For instance, for every 1000 average-risk women screened biennially from 50–74, there would be 226.5 (range: 169.9–267.0) false positives per death averted. If women with a two-fold increase in risk begin annual screening at age 40, their corresponding ratio would be slightly more favorable at 200.7 (range: 177.5–232.2). For women with a 1.3 fold increase in risk, biennial screening starting at age 40 would have similar harm to benefit ratios as biennial screening of average-risk women from ages 50–74.

Considering breast density group (low vs. high) changed absolute benefits, but did not affect ranking of the strategies, and for all metrics, annual screening from 50–74 remained dominated for all breast density groups. Women in the low density group had a greater proportion of their cancers detected and therefore greater mortality reduction than those in the high density group. However, women in the high density group had a greater absolute number of cancers because the incidence of cancer was higher in these women, therefore more life years were saved among women in the high density group (not shown).

For women with no comorbidity, biennial screening could continue to age 78 or 80 and still have similar harm to benefit ratios as screening women of average comorbidity biennially from 50–74. However, for women with moderate to severe comorbidity, the comparable ratios were equivalent at about age 68 (not shown).

Sensitivity Analyses

Consideration of utilities for usual health, screening, diagnosis, and treatment decreases estimates of QALYs but does not affect ranking of strategies (Table 1). The largest decrement related to quality adjustment accrues because of declines in health as women age. There are also substantial decrements in QALYs attributed to the disutility of undergoing diagnostic evaluation of an abnormal screening exam and for having cancer. The disutility associated with undergoing screening itself has a minimal impact on QALYs. Overall, there are persistent QALY gains under all screening strategies, but the magnitude becomes smaller when the highest disutility estimates are used. (34)

Discussion

This study used six established models with differing approaches and assumptions to estimate the potential efficacy of various US screening strategies based on risk, breast density, and comorbidity. All six modeling groups project some benefits associated with screening average-risk women from 40 to 49. The models consistently rank strategies and conclude that biennial screening strategies are most frequently efficient. Screening initiation at age 40 for average-risk women has the greatest benefit, but also the greatest harms. While absolute benefits vary, the ranking of strategies was not affected by risk level or breast density. Annual screening of women with a two to four-fold increased risk from ages 40–74 had comparable harms to benefits ratios as biennial screening from age 50 to 74 in average-risk populations. Consideration of quality of life reduced, but did not eliminate the

magnitude of benefits from all strategies. Among women with severe or moderate levels of comorbidity, harms of screening outweighed benefits before age 74.

The results of this collaborative modeling research indicate that digital mammography screening of populations of average-risk women in their 40's modestly lowers mortality and extends the length of life. This conclusion was seen in all models, although the absolute benefit varied based on model structure and assumptions.

Similar to our 2009 analysis,(1) biennial strategies are most consistently on the efficiency frontier. Screening annually from ages 50–74 had fewer benefits for any given harm (or resource use) under virtually all circumstances. However, annually screening in the 40's followed by biennial screening at age 50, or the most intensive schedule evaluated (annual screening 40–74) were also on the efficiency frontier, and might be considered depending on preferences for thresholds of harms relative to benefits. While we only evaluated two starting ages in the 40's, we found that screening benefits and harms for younger women seem to exist on a continuum.

This analysis extends our prior work by explicitly considering overdiagnosis. Depending on screening strategy, the models estimated that from 2% to 12% of invasive and 30% to 50% of DCIS cases may represent overdiagnoses. While the models differed in absolute estimates, they agreed on how overdiagnosis affects the ranking of strategies and that the majority of overdiagnosed cases were DCIS. The model results for overdiagnosis are not directly comparable to other published estimates (8,50) since the models followed women for their entire lives and considered competing mortality; the models also made assumptions about input parameters. While there is no agreement on methods to estimate (51) or the true rate of overdiagnosis,(52,53) there is agreement that it can lead to harm. Active surveillance for low-risk DCIS is one potential future approach to reduce harms from overdiagnosis of DCIS. More information is also needed on consumer knowledge of and willingness to risk overdiagnosis.(54)

The balance of harms to benefits became more favorable as underlying risk levels increased, but our results did not suggest that screening strategies necessarily need to be tailored by breast density, at least when grouped into low/high categories. Our results may differ from past studies of screening based on breast density,(20,55) since we modeled established digital mammography. While optimized for density, digital sensitivity is still lower in women with dense than non-dense breasts.(56–58) Improving outcomes for women with dense breasts may require dual consideration of risk and density,(58) and/or use of technology that employs alternative approaches to tissue visualization, such as tomosynthesis (59,60) or breast-specific gamma imaging,(61,62) or identification of genetic risk markers.(63–65)

Consistent with other analyses of screening upper age limits,(21,66–68) and other recommendations,(12,69) our results suggested that the balance of harms and benefits of screening was affected by competing mortality, so that age of screening cessation should be tailored by comorbidity levels.

Overall, this study has several important strengths including collaboration of six independent modeling groups, consideration of digital technology, incorporation of increasingly effective treatments, and consideration of risk factors, breast density, and comorbidity levels. (70) The conclusions about the ranking of screening strategies are robust and should provide greater credibility than inferences based on one model alone. Despite this, there are several caveats that should be considered in evaluating our results. First, we assumed 100% adherence to screening, prompt evaluation of abnormal results, and full use of optimal treatment to evaluate program efficacy. Benefits will always fall short of the projected results since adherence is not perfect. If actual adherence varies systematically by age, risk, or other factors, it is also possible that the ranking of strategies could change. Second, we did not consider other imaging technologies, such as computer-aided detection,(71) tomosynthesis, or magnetic resonance imaging (MRI). Performance data for general populations are still emerging,(60) so this will be important to consider once additional data are available. Next, we assumed that risk factors influenced incidence but did not affect disease natural history. Since certain risk factors are associated with higher false-positive rates than average,(38) the use of non-risk adjusted test specificity could have under-estimated false-positive rates. Also, we acknowledge that certain risk factors, such as family history, are age-dependent in their effects.(18,72) Since we held relative risk levels constant over age, our benefit estimates could be over- or under-estimated for specific risk factors.(16) We did not consider polygenic risk; this is an important emerging area for future research.(73,74) Last, compared to our earlier research,(1) the models all estimated similar, but somewhat greater mortality reductions from screening (e.g., 22% vs. 25.8% median mortality reduction with biennial screening from 50–74 in 2009 vs. current models, respectively). The primary reasons for this modeled improvement relate to the increased sensitivity of digital vs. film mammography, advances in molecular sub-type directed-adjuvant therapy, and changes in underlying breast cancer trends.

Overall, the evaluation of screening strategies by the six models suggests that optimal program design for average-risk women would continue to be based on biennial intervals. Choices about optimal ages of initiation and cessation will ultimately depend on program goals, weight attached to the balance of harms and benefits,(75) and considerations of efficiency.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health under National Cancer Institute Grant U01 CA152958 and the National Cancer Institute-funded Breast Cancer Surveillance Consortium (BCSC) Grant P01 CA154292, contract HSN261201100031C, and Grant U54CA163303. The collection of BCSC cancer and vital status data used in this study was supported in part by several state public health departments and cancer registries throughout the US. For a full description of these sources, please see: http://breastscreening.cancer.gov/work/acknowledgement.html. Dr. Chang's work was supported in part by American Cancer Society Grant MRSG 14-027-01 CPHPS.

The investigators worked with US Preventive Services Task Force members and Agency for Healthcare Research and Quality (AHRQ) staff to develop the scope, analytic framework, and key questions for this research. The Task

Force, AHRQ, and the funding sources had no role in study conduct. AHRQ staff distributed earlier versions of the draft for peer review. The investigators are solely responsible for the content and the decision to submit the manuscript for publication.

The authors thank CISNET consultants Elizabeth Burnside, Allison Kurian, Donald Weaver, and Diana Buist for review of earlier versions of the models for structure, parameters, and assumptions for clinical face validity; Jennifer Croswell from the Agency for Healthcare Research and Quality for assistance with analysis and review of results; members of the U.S. Preventive Services Task Force and the Oregon Evidence-based Practice Center for comments on earlier versions of this research; and Jessica Garshell for data processing. The authors thank Adrienne Ryans for manuscript preparation.

The authors thank Stuart Baker, William Lawrence, Paul Pinsky, Tom Trikalinos, and John Wong for helpful suggestions on earlier versions of this paper.

Reference List

1. Mandelblatt J, Cronin K, Bailey S, Berry DA, de Koning H, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Inten Med. 2009; 151(10):738–747.
2. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev. 2013; 6:CD001877.
3. Biller-Andorno N, Juni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. N Engl J Med. 2014; 370(21):1965–1967. [PubMed: 24738641]
4. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet. 2002; 359(9310):909–919. [PubMed: 11918907]
5. Tabar L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, et al. The Swedish Two-County Trial twenty years later Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am. 2000; 38(4):625–651. [PubMed: 10943268]
6. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet. 2006; 368(9552):2053–2060. [PubMed: 17161727]
7. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. New Engl J Med. 2005; 353(17):1784–1792. [PubMed: 16251534]
8. Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ. 2014; 348:g366. [PubMed: 24519768]
9. Paci E, Broeders M, Hofvind S, Puliti D, Duffy SW. European breast cancer service screening outcomes: a first balance sheet of the benefits and harms. Cancer Epidemiol Biomarkers Prev. 2014; 23(7):1159–1163. [PubMed: 24991022]
10. Smith RA. The value of modern mammography screening in the control of breast cancer: understanding the underpinnings of the current debates. Cancer Epidemiol Biomarkers Prev. 2014; 23(7):1139–1146. [PubMed: 24991021]
11. Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, et al. Breast-cancer screening--viewpoint of the IARC Working Group. N Engl J Med. 2015; 372(24): 2353–2358. [PubMed: 26039523]
12. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2009; 151(10):716–236. [PubMed: 19920272]
13. Broeders M, Moss S, Nystrom L, Njor S, Jonsson H, Paap E, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. J Med Screen. 2012; 19(Suppl 1):14–25.
14. US Food and Drug Administration. Silver Spring, Maryland: US Department of Health and Human Services; Jun. Available from: http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/default.htm [Accessed January 2015]

15. National Comprehensive Cancer Network (NCCN). [Accessed January 2015] NCCN Clinical Practice Guidelines in Oncology - Breast Cancer. 2014. Available from: http://www.nccn.org/professionals/physician_gls/f_guidelines.asp
16. van Ravesteyn NT, Miglioretti DL, Stout NK, Lee SJ, Schechter CB, Buist DS, et al. Tipping the balance of benefits and harms to favor screening mammography starting at age 40 years: a comparative modeling study of risk. Ann Intern Med. 2012; 156(9):609–617. [PubMed: 22547470]
17. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med. 2007; 356(3):227–236. [PubMed: 17229950]
18. Nelson HD, Zakher B, Cantor A, Fu R, Griffin J, O'Meara ES, et al. Risk factors for breast cancer for women aged 40 to 49 years: a systematic review and meta-analysis. Ann Intern Med. 2012; 156(9):635–648. [PubMed: 22547473]
19. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA. 1996; 276(1):33–38. [PubMed: 8667536]
20. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. Ann Intern Med. 2011; 155(1):10–20. [PubMed: 21727289]
21. Lansdorp-Vogelaar I, Gulati R, Mariotto AB, Schechter CB, de Carvalho TM, Knudsen AB, et al. Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits. Ann Intern Med. 2014; 161(2):104–112. [PubMed: 25023249]
22. de Gelder R, Heijnsdijk EA, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 in the presence of adjuvant systemic therapy. Int J Cancer. 2015; 137(1):165–172. [PubMed: 25430053]
23. Mandelblatt JS, Fryback DG, Weinstein MC, Russell LB, Gold MR. Assessing the effectiveness of health interventions for cost-effectiveness analysis Panel on Cost-Effectiveness in Health and Medicine. J Gen Intern Med. 1997; 12(9):551–558. [PubMed: 9294789]
24. Munoz D, Near AM, van Ravesteyn NT, Lee SJ, Schechter CB, Alagoz O, et al. Effects of screening and systemic adjuvant therapy on ER-specific US breast cancer mortality. J Natl Cancer Inst. 2014; 106(11)
25. Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin Breast Cancer Epidemiology Simulation Model. J Natl Cancer Inst Monogr. 2006; (36): 37–47. [PubMed: 17032893]
26. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. J Natl Cancer Inst Monogr. 2006; 36:47–55.
27. Berry DA, Inoue L, Shen Y, Venier J, Cohen D, Bondy M, et al. Modeling the impact of treatment and screening on U.S. breast cancer mortality: a Bayesian approach. J Natl Cancer Inst Monogr. 2006; (36):30–36. [PubMed: 17032892]
28. Plevritis SK, Sigal BM, Salzman P, Rosenberg J, Glynn P. A stochastic simulation model of U.S. breast cancer mortality trends from 1975 to 2000. J Natl Cancer Inst Monogr. 2006; (36):86–95. [PubMed: 17032898]
29. Lee S, Zelen M. A stochastic model for predicting the mortality of breast cancer. J Natl Cancer Inst Monogr. 2006; (36):79–86. [PubMed: 17032897]
30. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. J Natl Cancer Inst Monogr. 2006; (36):56–65. [PubMed: 17032895]
31. Clarke LD, Plevritis SK, Boer R, Cronin KA, Feuer EJ. A comparative review of CISNET breast models used to analyze U.S. breast cancer incidence and mortality trends. J Natl Cancer Inst Monogr. 2006; (36):96–105. [PubMed: 17032899]
32. Gangnon RE, Sprague BL, Stout NK, Alagoz O, Weedon-Fekjaer H, Holford TR, et al. The Contribution of Mammography Screening to Breast Cancer Incidence Trends in the United States: An Updated Age-period-cohort Model. Can Epi Biomarkers Prev. 2015 Jun; 24(6):905–912.

33. Peto R, Davies C, Godwin J, Gray R, Pan HC, Clarke M, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. Lancet. 2012; 379(9814):432–444. [PubMed: 22152853]
34. Mandelblatt, JS.; Cronin, K.; de Koning, H.; Miglioretti, DL.; Schechter, CS.; Stout, N. [Accessed May 2015] Collaborative modeling of U.S. breast cancer screening strategies. AHRQ Publication No. 14-05201-EF-4. 2015 Apr. Available from: http://www.screeningforbreastcancer.org/read-the-draft-materials
35. Chang Y, Schechter CB, van Ravesteyn NT, Near AM, Heijnsdijk EA, Adams-Campbell L, et al. Collaborative modeling of the impact of obesity on race-specific breast cancer incidence and mortality. Breast Cancer Res Treat. 2012; 136:823–835. [PubMed: 23104221]
36. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. J Natl Cancer Inst Monogr. 2006; (36):19–25. [PubMed: 17032890]
37. Braithwaite D, Zhu W, Hubbard RA, O'Meara ES, Miglioretti DL, Geller B, et al. Screening outcomes in older US women undergoing multiple mammograms in community practice: does interval, age, or comorbidity score affect tumor characteristics or false positive rates? J Natl Cancer Inst. 2013; 105(5):334–341. [PubMed: 23385442]
38. Kerlikowske K, Zhu W, Hubbard RA, Geller B, Dittus K, Braithwaite D, et al. Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. JAMA Intern Med. 2013; 173(9):807–816. [PubMed: 23552817]
39. Dittus K, Geller B, Weaver DL, Kerlikowske K, Zhu W, Hubbard R, et al. Impact of mammography screening interval on breast cancer diagnosis by menopausal status and BMI. J Gen Intern Med. 2013; 28(11):1454–1462. [PubMed: 23760741]
40. Hanmer J, Lawrence WF, Anderson JP, Kaplan RM, Fryback DG. Report of nationally representative values for the noninstitutionalized US adult population for 7 health-related quality-of-life scores. Med Decis Making. 2006; 26(4):391–400. [PubMed: 16855127]
41. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst. 2006; 98(11):774–782. [PubMed: 16757702]
42. de Haes JC, de Koning HJ, Van Oortmarssen GJ, van Agt HM, De Bruyn AE, van der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. Int J Cancer. 1991; 49:538–544. [PubMed: 1917155]
43. Rosenberg RD, Yankaskas BC, Abraham LA, Sickles EA, Lehman CD, Geller BM, et al. Performance benchmarks for screening mammography. Radiology. 2006; 241(1):55–66. [PubMed: 16990671]
44. Sprague BL, Gangnon RE, Burt V, Trentham-Dietz A, Hampton JM, Wellman RD, et al. Prevalence of mammographically dense breasts in the United States. J Natl Cancer Inst. 2014; 106(10)
45. Tice JA, O'Meara ES, Weaver DL, Vachon C, Ballard-Barbash R, Kerlikowske K. Benign breast disease, mammographic breast density, and the risk of breast cancer. J Natl Cancer Inst. 2013; 105(14):1043–1049. [PubMed: 23744877]
46. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst. 2006; 98(17):1204–1214. [PubMed: 16954473]
47. Kerlikowske K, Walker R, Miglioretti DL, Desai A, Ballard-Barbash R, Buist DS. Obesity, mammography use and accuracy, and advanced breast cancer risk. J Natl Cancer Inst. 2008; 100(23):1724–1733. [PubMed: 19033562]
48. Mariotto AB, Wang Z, Klabunde CN, Cho H, Das B, Feuer EJ. Life tables adjusted for comorbidity more accurately estimate noncancer survival for recently diagnosed cancer patients. J Clin Epidemiol. 2013; 66(12):1376–1385. [PubMed: 24035494]
49. Cho H, Klabunde CN, Yabroff KR, Wang Z, Meekins A, Lansdorp-Vogelaar I, et al. Comorbidity-adjusted life expectancy: a new tool to inform recommendations for optimal screening strategies. Ann Intern Med. 2013; 159(10):667–676. [PubMed: 24247672]
50. Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. JAMA Intern Med. 2014; 174(3):448–454. [PubMed: 24380095]

51. Etzioni R, Gulati R, Mallinger L, Mandelblatt J. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. Ann Intern Med. 2013; 158(11): 831–838. [PubMed: 23732716]
52. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ. 2015; 350:g7773. [PubMed: 25569206]
53. Etzioni R, Xia J, Hubbard R, Weiss NS, Gulati R. A reality check for overdiagnosis estimates associated with breast cancer screening. J Natl Cancer Inst. 2014; 106(12)
54. Moynihan R, Nickel B, Hersch J, Beller E, Doust J, Compton S, et al. Public opinions about overdiagnosis: a national community survey. PLoS One. 2015; 10(5):e0125165. [PubMed: 25992887]
55. Vilaprinyo E, Forne C, Carles M, Sala M, Pla R, Castells X, et al. Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. PLoS One. 2014; 9(2):e86858. [PubMed: 24498285]
56. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. N Engl J Med. 2005; 353(17): 1773–1783. [PubMed: 16169887]
57. Kerlikowske K, Hubbard RA, Miglioretti DL, Geller BM, Yankaskas BC, Lehman CD, et al. Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. Ann Intern Med. 2011; 155(8):493–502. [PubMed: 22007043]
58. Kerlikowske K, Zhu W, Tosteson AN, Sprague BL, Tice JA, Lehman CD, et al. Identifying women with dense breasts at high risk for interval cancer: a cohort study. Ann Intern Med. 2015; 162(10): 673–681. [PubMed: 25984843]
59. Lee CI, Lehman CD. Digital breast tomosynthesis and the challenges of implementing an emerging breast cancer screening technology into clinical practice. J Am Coll Radiol. 2013; 10(12):913–917. [PubMed: 24295940]
60. Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA. 2014; 311(24):2499–2507. [PubMed: 25058084]
61. Park JY, Yi SY, Park HJ, Kim MS, Kwon HJ, Park NH, et al. Breast-specific gamma imaging: correlations with mammographic and clinicopathologic characteristics of breast cancer. AJR Am J Roentgenol. 2014; 203(1):223–228. [PubMed: 24951219]
62. Rechtman LR, Lenihan MJ, Lieberman JH, Teal CB, Torrente J, Rapelyea JA, et al. Breast-specific gamma imaging for the detection of breast cancer in dense versus nondense breasts. AJR Am J Roentgenol. 2014; 202(2):293–298. [PubMed: 24450668]
63. Matamala N, Vargas MT, Gonzalez-Campora R, Minambres R, Arias JI, Menendez P, et al. Tumor microRNA expression profiling identifies circulating microRNAs for early breast cancer detection. Clin Chem. 2015
64. Mavaddat N, Pharoah PD, Michailidou K, Tyrer J, Brook MN, Bolla MK, et al. Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst. 2015; 107(5)
65. Stone J, Thompson DJ, dos SSI, Scott C, Tamimi RM, Lindstrom S, et al. Novel associations between common breast cancer susceptibility variants and risk-predicting mammographic density measures. Cancer Res. 2015 Jun 15; 75(12):2457–2467. [PubMed: 25862352]
66. Walter LC, Covinsky KE. Cancer screening in elderly patients: a framework for individualized decision making. JAMA. 2001; 285(21):2750–2756. [PubMed: 11386931]
67. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. External validation of an index to predict up to 9-year mortality of community-dwelling adults aged 65 and older. J Am Geriatr Soc. 2011; 59(8):1444–1451. [PubMed: 21797837]
68. van Ravesteyn N, Stout NK, Schechter CB, Heijnsdijk EAM, Alagoz O, Trentham-Dietz A, et al. Benefits and harms of mammography screening after age 74 years: model estimates of overdiagnosis. J Natl Cancer Inst. 2015; 107(7)
69. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2012: A review of current American Cancer Society guidelines and current issues in cancer screening. CA Cancer J Clin. 2012; 62(2):129–142. [PubMed: 22261986]

70. Elmore JG, Harris RP. The harms and benefits of modern screening mammography. BMJ. 2014; 348:g3824. [PubMed: 24938686]

71. Fenton JJ, Xing G, Elmore JG, Bang H, Chen SL, Lindfors KK, et al. Short-term outcomes of screening mammography using computer-aided detection: a population-based study of medicare enrollees. Ann Intern Med. 2013; 158(8):580–587. [PubMed: 23588746]

72. Trentham-Dietz A, Sprague BL, Hampton JM, Miglioretti DL, Nelson HD, Titus LJ, et al. Modification of breast cancer risk according to age and menopausal status: a combined analysis of five population-based case-control studies. Breast Cancer Res Treat. 2014; 145(1):165–175. [PubMed: 24647890]

73. Garcia-Closas M, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, Brook MN, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. Nat Genet. 2013; 45(4):392. [PubMed: 23535733]

74. Stevens KN, Vachon CM, Couch FJ. Genetic susceptibility to triple-negative breast cancer. Cancer Res. 2013; 73(7):2025–2030. [PubMed: 23536562]

75. Hoffmann TC, Del MC. Patients' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. JAMA Intern Med. 2015; 175(2):274–286. [PubMed: 25531451]

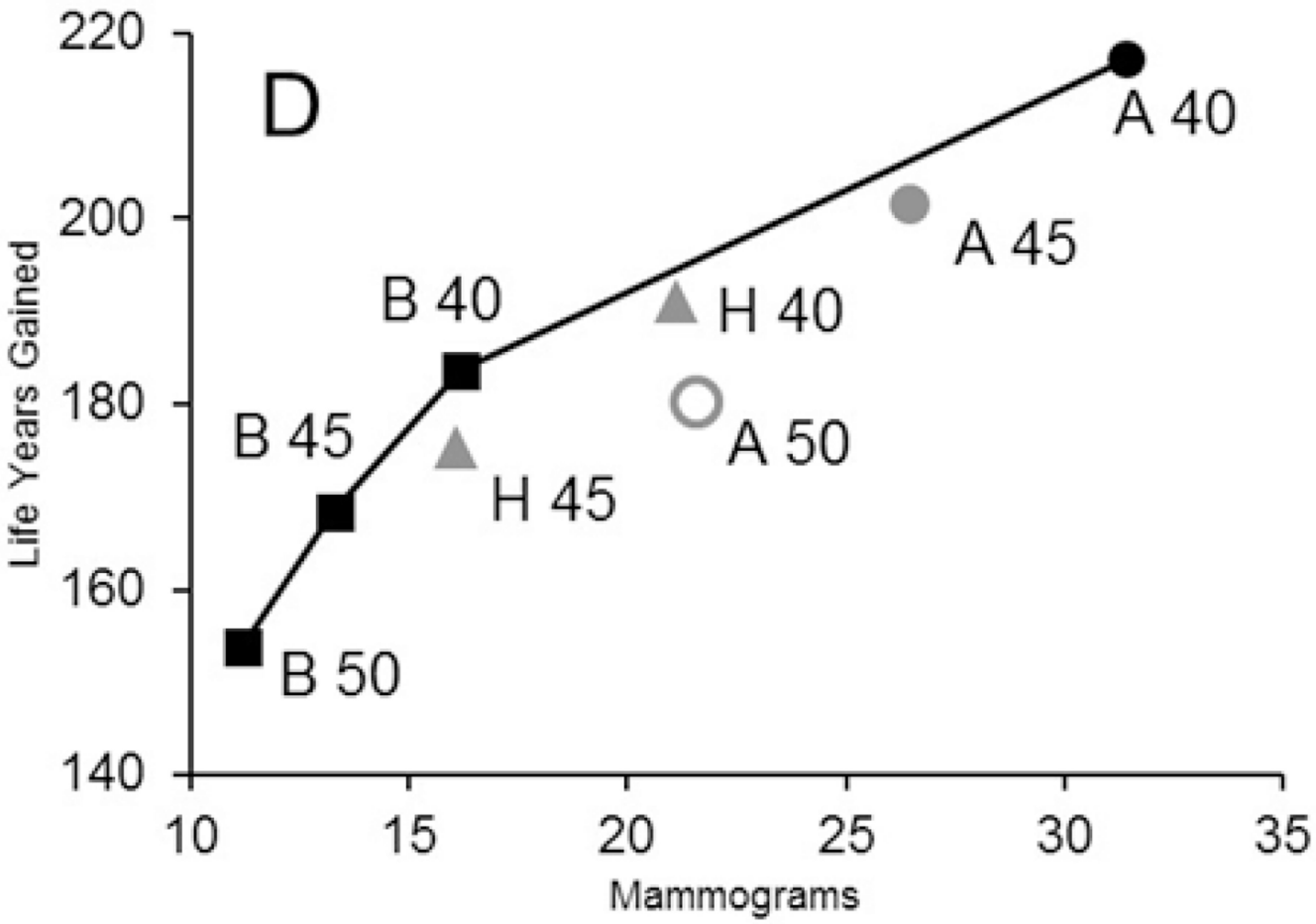

Figure 1. Efficiency Frontier for Harms (Average Number of Screening Examinations) and Benefits (Life Years Gained) for Exemplar Model by Screening Strategy

The panel shows an efficiency frontier graph for an exemplar model (model D); comparable graphs are included on Appendix Figure 2 for all 6 models.

The graph plots the average number of mammograms per 1,000 women against the life-years gained for each screening strategy (vs. no screening). We plot efficient strategies (i.e., those in which increases in mammography use results in greater life-years gained than the next least-intensive strategy). The line between strategies represents the "efficiency frontier." Strategies on this line would be considered efficient because they achieve the greatest gain per mammography used compared with the point (or strategy) immediately below it. Points that fall below the line are not as efficient as those on the line. When the slope in the efficiency frontier plot levels off, the additional life-years gained per increase in mammography are small relative to the previous strategies and could indicate a point at which additional screening might be considered as having a low return (benefit).

Black strategies are efficient; dark grey strategies are close to the efficiency frontier; and light grey strategies are dominated (inefficient). Biennial strategies are indicated with a square; hybrid strategies (annual in the 40's, followed by biennial from 50–74) are represented by a triangle, and annual strategies with a circle. Efficiency frontiers for other harm and benefit metrics can be found at: (34)

Table 1

Ranking of Benefits (Percent Breast Cancer Mortality Reduction, LYs, QALYs) by Model and Screening Strategy Per 1000 Women Screened

	Results per 1000 Women Screened							
		Percent mortality reduction (vs. no screening) by model[1]						
Strategies	# of screens*	D	E	G-E	M	S	W	Median (range across models)
B 50–74	11205	25.6%	26.0%	31.8%	26.8%	24.1%	25.4%	25.8% (24.1–31.8)
B 45–74	13301	26.6%	27.6%	33.9%	28.4%	25.9%	26.7%	27.2% (25.9–33.9)
H 45–74	16060	27.7%	29.7%	35.9%	29.2%	27.3%	30.1%	29.5% (27.3–35.9)
B 40–74	16112	28.3%	30.3%	35.9%	31.9%	28.2%	30.5%	30.4% (28.2–35.9)
H 40–74	20989	29.0%	32.3%	37.9%	31.7%	29.3%	32.8%	32.0% (29.0–37.9)
A 50–74	21447	32.1%	33.9%	37.6%	27.1%	29.1%	35.3%	33.0% (27.1–37.6)
A 45–74	26280	34.2%	37.6%	41.6%	29.4%	32.3%	39.1%	35.9% (29.4–41.6)
A 40–74	31194	35.5%	40.1%	43.6%	32.5%	34.4%	42.6%	37.8% (32.5–43.6)

	Results per 1000 Women Screened							
		Years of Life Gained (vs. no screening) by model						
Strategies	# of screens*	D	E	G-E	M	S	W	Median (range across models)
B 50–74	11205	153.8	94.0	140.5	146.5	104.2	74.6	122.4 (74.6–153.8)
B 45–74	13301	168.4	107.7	161.2	171.3	115.2	84.0	138.2 (84.0–171.3)
H 45–74	16060	175.3	117.9	170.2	171.4	125.1	95.7	147.7 (95.7–175.3)
B 40–74	16112	183.7	123.7	172.4	194.8	131.6	98.8	152.0 (98.8–194.8)
H 40–74	20989	191.1	137.6	187.2	211.5	141.0	110.9	164.1 (110.9–211.5)
A 50–74	21447	180.0	125.9	167.3	156.3	133.3	104.3	144.8 (104.3–180.0)
A 45–74	26280	201.3	149.3	196.7	177.8	154.2	123.0	166.0 (123.0–201.3)
A 40–74	31194	217.1	168.8	213.5	218.1	170.1	140.5	191.8 (140.5–218.1)

	Results per 1000 Women Screened							
		QALYs Gained (vs. no screening) by model						
Strategies	# of screens*	D	E	G-E	M	S	W	Median (range across models)
B 50–74	11205	114.5	67.3	100.1	109.6	71.9	47.1	86.0 (47.1–114.5)
B 45–74	13301	123.8	75.6	114.4	129.4	78.8	51.9	96.6 (51.9–129.4)
H 45–74	16060	126.6	80.9	118.3	128.5	84.5	58.3	101.4 (58.3–128.5)
B 40–74	16112	133.7	85.4	120.1	148.1	89.1	60.4	104.6 (60.4–148.1)
H 40–74	20989	134.2	91.0	126.1	159.4	92.5	64.8	109.3 (64.8–159.4)
A 50–74	21447	127.0	84.1	111.4	113.2	87.5	62.4	99.5 (62.4–127.0)
A 45–74	26280	138.9	97.3	129.5	129.4	99.5	71.7	114.5 (71.7–138.9)
A 40–74	31194	146.6	107.3	137.2	160.6	107.6	80.0	122.4 (80.0–160.6)

[1] Without screening, the median probability of dying of breast cancer is 2.50% (range 1.50–3.20%). Thus, if a particular screening strategy leads to a 30% reduction in breast cancer mortality, this means that the probability of breast cancer mortality was reduced from 2.50% to 1.75%. This translates into 7.5 deaths averted per 1000 women screened.

A=Annual B=Biennial H=Hybrid

[*] Strategies are ranked from the least to the most number of mammograms Average number of mammograms across models. Not all possible mammograms in the age interval are obtained since some women die from other causes before screening would occur.

[†] Model Group Abbreviations: D (Dana Farber Cancer Center), E (Erasmus Medical Center), G-E (Georgetown U. –Einstein U.), M (M.D. Anderson Cancer Center), S (Stanford U.), W (University of Wisconsin/Harvard)

[‡] Grey shaded areas in the table show strategies that are dominated (“inefficient”) within a specific model; a strategy is classified as dominated if there is another strategy that results in an equal or higher percent mortality decline/LYG/QALYs with fewer average screening exams.

[§] QALYs are adjusted for general health, diagnosis, screening and treatment.

[//] 100% of women receive adjuvant systemic therapy based on recommended stage, ER/HER2-specific adjuvant therapy for pre- and post-menopausal women.

Ann Intern Med. Author manuscript; available in PMC 2016 October 25.

Table 2

Lifetime Benefits and Harms of Screening Strategies based on Starting Ages and Screening Intervals

Strategy	Median number (range across models) per 1000 women screened (vs. no screening)*					
	Screens	Breast cancer deaths averted	False-positive screens	Benign breast biopsies	Over-diagnosed cases (invasive and DCIS) †‡	Percent of all cases over-diagnosed †‡
Biennial						
50–74	11,127	7 (4–9)	953 (830–1325)	146 (120–205)	19 (11–34)	12% (8–22)
45–74	13,212	8 (4–9)	1220 (930–1599)	168 (120–221)	19 (11–34)	12% (8–22)
40–74	16,013	8 (5–10)	1529 (1100–1976)	204 (140–264)	21 (12–38)	13% (9–24)
Hybrid						
45–74	15,966	8.0 (5–9)	1520 (1160–1968)	190 (140–250)	21 (12–40)	13% (8–25)
40–74	20,884	9 (5–10)	2106 (1480–2623)	245 (170–309)	23 (12–44)	14% (9–27)
Annual						
50–74	21,318	9 (5–10)	1798 (1706–2445)	228 (219–317)	25 (12–68)	15% (8–36)
45–74	26,136	9 (6–11)	2355 (2185–3087)	247 (230–329)	28 (12–74)	17% (9–38)
40–74	31,038	10 (6–11)	2941 (2550–3742)	303 (260–388)	30 (13–77)	18% (9–39)

*In all scenarios, 100% of women receive adjuvant systemic therapy based on recommended stage, ER/HER2-specific adjuvant therapy for pre- and post-menopausal women.

†Over-diagnosed cases are defined as cases that would not have been clinically detected in the absence of screening. The figure includes DCIS and invasive overdiagnosis. Overdiagnosis is calculated by comparing cases detected in the screening scenario to those detected in the unscreened scenario. Model S is excluded since it does not include DCIS. The percent overdiagnosis is calculated as the percent of all cases detected in the screening strategy that are overdiagnosis.

‡The upper range for all over diagnosis estimates is based on Model M results. Model M generates very high overdiagnosis based on the assumption that incidence in the absence of screening has essentially remained flat since 1975–79, with virtually all of the increases over time attributable to screening. The other models use some form of an age-period-cohort model for incidence in the absence of screening, where some of the increases in incidence are due to screening and some to changes in risk factors (e.g., use of hormone replacement therapy), generating lower rates of overdiagnosis. Other sources of variation across models are related to assumptions about the proportions of DCIS cases that never progress to invasive cancer or the number of early invasive cancers that might be nonprogressive. Generally, models that assume higher proportions of DCIS and/or invasive cancer to be nonprogressive generate higher estimates of overdiagnosis than models that assume less nonprogressive disease. Unfortunately, the underlying incidence in the absence of screening and the proportion and types of tumors that are nonprogressive are unknown and unobservable. Therefore, the different results across models based on their respective assumptions provide a range of possible overdiagnosis.

Table 3

Incremental Changes in Percentage of Reduction in Breast Cancer Mortality, by Age of Screening Initiation, Interval, and Model

	Start Age 40 (vs. 50)*				Start Age 45 (vs. 50)*			
	Difference in % breast cancer mortality reduction		Number of breast cancer deaths averted/1000 women		Difference in % breast cancer mortality reduction		Number of breast cancer deaths averted/1000	
Model	**Annual**	**Biennial**	**Annual**	**Biennial**	**Annual**	**Biennial**	**Annual**	**Biennial**
D	3.4%	2.7%	1.1	0.9	2.1%	1.0%	0.6	0.3
E	6.2%	4.3%	1.5	1.0	3.6%	1.6%	0.9	0.4
G-E	6.0%	4.1%	1.5	1.0	4.0%	2.2%	1.0	0.5
M	5.3%	5.1%	1.7	1.7	2.3%	1.6%	0.7	0.5
S	5.2%	4.1%	1.1	0.9	3.1%	1.7%	0.7	0.4
W	7.3%	5.1%	1.1	0.8	3.8%	1.3%	0.6	0.2
Median	5.7%	4.2%	1.3	1.0	3.4%	1.6%	0.7	0.4

*Incremental difference between starting at age 40 or 45 vs. 50. Annual is comparing A40–74 (or 45–74) to A50–74; biennial is comparing B40–74 (or 45–74) to B50–74. Hybrid strategies are compared to biennial 50–74, therefore for these incremental comparisons the hybrid results are the same as the annual results.

Table 4

Percent of Annual Mortality Reduction Maintained by Biennial Screening by Strategy and Model

	D	E	G-E	M[§]	S	W	Median
50–74[*]	79.8%	76.7%	84.6%	98.9%	82.8%	72.0%	81.3%
45–74[†]	77.8%	73.4%	81.5%	96.6%	80.2%	68.3%	79.0%
40–74[‡]	79.7%	75.6%	82.3%	98.2%	82.0%	71.6%	80.8%

[*] Percent of A50–74 maintained by B50–74 = percent mortality reduction B50–74/ percent mortality reduction A50–74

[†] Percent of A45–74 maintained by B45–74 = percent mortality reduction B45–74/ percent mortality reduction A45–74

[‡] Percent of A40–74 maintained by B40–74 = percent mortality reduction B40–74/ percent mortality reduction A40–74

[§] Model M does not include a natural history component. Based on a combination of assumptions regarding underlying incidence trends in the absence of screening, it essentially yields a long invasive cancer lead time; thus all cancers found with annual screening can also be detected with biennial screening.