Annals of Internal Medicine ORIGINAL RESEARCH

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

Jeanne S. Mandelblatt, MD, MPH; Natasha K. Stout, PhD; Clyde B. Schechter, MA, MD; Jeroen J. van den Broek, MS; Diana L. Miglioretti, PhD; Martin Krapcho, BS; Amy Trentham-Dietz, PhD, MS; Diego Munoz, PhD, MS; Sandra J. Lee, ScD; Donald A. Berry, PhD; Nicolien T. van Ravesteyn, PhD; Oguzhan Alagoz, PhD; Karla Kerlikowske, MD; Anna N.A. Tosteson, ScD; Aimee M. Near, MPH; Amanda Hoeffken, MPH; Yaojen Chang, DrPH, MS, MPH; Eveline A. Heijnsdijk, PhD; Gary Chisholm, MS; Xuelin Huang, PhD; Hui Huang, MS; Mehmet Ali Ergun, MSc; Ronald Gangnon, PhD; Brian L. Sprague, PhD; Sylvia Plevritis, PhD; Eric Feuer, PhD; Harry J. de Koning, MD, PhD; and Kathleen A. Cronin, PhD, MPH

Background: Controversy persists about optimal mammography screening strategies.

Objective: To evaluate screening outcomes, taking into account advances in mammography and treatment of breast cancer.

Design: Collaboration of 6 simulation models using national data on incidence, digital mammography performance, treatment effects, and other-cause mortality.

Setting: United States.

Patients: Average-risk U.S. female population and subgroups with varying risk, breast density, or comorbidity.

Intervention: Eight strategies differing by age at which screening starts (40, 45, or 50 years) and screening interval (annual, biennial, and hybrid [annual for women in their 40s and biennial thereafter]). All strategies assumed 100% adherence and stopped at age 74 years.

Measurements: Benefits (breast cancer–specific mortality reduction, breast cancer deaths averted, life-years, and quality-adjusted life-years); number of mammograms used; harms (false-positive results, benign biopsies, and overdiagnosis); and ratios of harms (or use) and benefits (efficiency) per 1000 screens.

Results: Biennial strategies were consistently the most efficient for average-risk women. Biennial screening from age 50 to 74 years avoided a median of 7 breast cancer deaths versus no screening; annual screening from age 40 to 74 years avoided an additional 3 deaths, but yielded 1988 more false-positive results and 11 more overdiagnoses per 1000 women screened. Annual screening from age 50 to 74 years was inefficient (similar benefits, but more harms than other strategies). For groups with a 2- to 4-fold increased risk, annual screening from age 40 years had similar harms and benefits as screening average-risk women biennially from 50 to 74 years. For groups with moderate or severe comorbidity, screening could stop at age 66 to 68 years.

Limitation: Other imaging technologies, polygenic risk, and nonadherence were not considered.

Conclusion: Biennial screening for breast cancer is efficient for average-risk populations. Decisions about starting ages and intervals will depend on population characteristics and the decision makers' weight given to the harms and benefits of screening.

Primary Funding Source: National Institutes of Health.

Ann Intern Med. 2016;164:215-225. doi:10.7326/M15-1536 **www.annals.org**
For author affiliations, see end of text.
This article was published at www.annals.org on 12 January 2016.

Despite decades of mammography screening for early detection of breast cancer, there is no consensus on optimal strategies, target populations, or the magnitude of harms and benefits (1–11). The 2009 US Preventive Services Task Force (USPSTF) recommended biennial film mammography from age 50 to 74 years and suggested shared decision making about screening for women in their 40s (12). Since that recommendation was formulated, new data on the benefits of screening have emerged (2, 6, 8, 9, 11, 13, 14), digital mammography has essentially replaced plain film (15), and increasingly effective systemic treatment regimens for breast cancer have become standard (16). There has also been growing interest in consumer preferences and personalized screening approaches (17–20). These factors could each affect the outcomes of breast cancer screening programs or alter policy decisions about population screening strategies (17).

Modeling can inform screening policy decisions because it uses the best available evidence to evaluate a wide range of strategies while holding selected conditions (such as treatment effects) constant, facilitating strategy comparisons (21, 22). Modeling also provides a quantitative summary of outcomes in different groups and assesses how preferences affect results. Collaboration of several models provides a range of plausible effects and illustrates the effects of differences in model assumptions on results (1, 7, 23).

We used 6 well-established simulation models to synthesize current data and examine the outcomes of digital mammography screening at various starting ages and intervals among average-risk women. We also examined how breast density, risk, or comorbidity levels affect results and whether preferences for health

See also:

Related articles . . . 205, 226, 236, 244, 256, 268, 279
Editorial comment 303

Web-Only
CME quiz

EDITORS' NOTES

Context

Multiple alternative mammography screening strategies exist.

Contribution

This modeling study estimated outcomes of 8 strategies that differed by starting age and interval. Biennial screening from age 50 to 74 years avoided a median of 7 breast cancer deaths; in contrast, annual screening from age 40 to 74 years avoided an additional 3 deaths but yielded 1988 more false-positive results and 11 more overdiagnoses per 1000 women screened. Annual screening from age 40 years for high-risk women had similar outcomes as screening average-risk women biennially from 50 to 74 years of age.

Caution

Imaging technologies other than mammography and nonadherence were not modeled.

Implication

Biennial mammography screening for breast cancer is efficient for average-risk women.

states related to screening and its downstream consequences affected conclusions.

METHODS

Strategies

We evaluated 8 strategies that varied by starting age (40, 45, or 50 years) and interval (annual, biennial, and hybrid [annual for women in their 40s and biennial thereafter]); all strategies stop screening at age 74 years. We included "no screening" as a baseline.

Model Descriptions

The models used to evaluate the screening strategies were developed within the Cancer Intervention and Surveillance Modeling Network (CISNET) (24–30), and the research was institutional review board–approved. They were named model D (Dana-Farber Cancer Institute, Boston, Massachusetts), model E (Erasmus MC, University Medical Center Rotterdam, 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). The **Appendix** (available at www.annals.org) provides information on model validation.

Since earlier analyses (1), the models have undergone substantial revision to reflect advances in breast cancer control, including portrayal of molecular subtypes based on estrogen receptor (ER) and human epidermal growth factor-2 receptor (HER2) status (23); current population incidence (31) and competing non–breast cancer mortality; digital screening; and the most current therapies (32). All models except model S include ductal carcinoma in situ (DCIS).

The general modeling approach is summarized in this article; full details, including approach, construction, data sources, assumptions, and implementation, are available at https://resources.cisnet.cancer.gov/registry and reference 33. Additional information is available on request, and the models are available for use via collaboration.

The models begin with estimates of breast cancer incidence (31) 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 incidence and breast cancer–specific mortality trends in the U.S. population (1, 7, 17, 23, 33, 34). Breast cancer has a distribution of preclinical screen-detectable periods (sojourn time) and clinical detection points. Performance characteristics of digital mammography depend on age, first versus subsequent screen, time since last mammogram, and breast density. ER/HER2 status is assigned at diagnosis on the basis of stage and age. Molecular subtype– and stage-specific treatment reduces the hazard 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-risk and average breast density and followed them from age 25 years (because breast cancer is rare before this age [0.08% of cases]) until death or age 100 years.

Model Input Parameters

The models used a common set of age-specific variables for breast cancer incidence, performance of digital mammography, treatment effects, and average and comorbidity-specific non–breast cancer causes of death (20, 35). The parameter values are available at www.uspreventiveservicestaskforce.org/Page/Document/modeling-report-collaborative-modeling-of-us-breast-cancer-1/breast-cancer-screening1 (33). In addition, each group included model-specific inputs (or intermediate outputs) to represent preclinical detectable times, lead time, and age- and ER/HER2-specific stage distribution in screen- versus non–screen-detected women on the basis of their specific model structure (1, 7, 23–30). These model-specific parameters were based on assumptions about combinations of values that reproduced U.S. trends in incidence and breast cancer–specific mortality, including proportions of DCIS that were nonprogressive and would not be detected without screening. Models M and W also assumed some small nonprogressive invasive cancers.

The models adopted an age–period–cohort modeling approach to project incidence rates of breast cancer in the absence of screening (31, 36); model M used 1975–1979 rates from the Surveillance, Epidemiology, and End Results program. The models assumed 100% adherence to screening and receipt of the most effective treatment to isolate the effect of varying screening strategies.

Four models used age-specific sensitivity values for digital mammography that were observed in the Breast Cancer Surveillance Consortium (BCSC) for detection of invasive and DCIS cancers combined (model S uses data for invasive cancers only). Separate values were used for initial and subsequent mammography by screening interval, using standard BCSC definitions: "Annual" includes data from screens occurring within 9 to 18 months of the prior screen, and "biennial" includes data on screens within 19 to 30 months (37, 38). Model D used these data as input variables (28), and models GE, S, and W used the data for calibration (24, 25, 27). Models E and M fit estimates from the BCSC and other data (26, 29).

Women with ER-positive tumors received 5 years of hormone therapy and an anthracycline-based regimen accompanied by a taxane. Women with ER-negative invasive tumors received anthracycline-based regimens with a taxane. Those with HER2-positive tumors also received trastuzumab. Women with ER-positive DCIS received hormonal therapy (16). Treatment effectiveness was based on clinical trials and was modeled as a reduction in breast cancer–specific mortality risk or increase in the proportion cured compared with ER/HER2-specific survival in the absence of adjuvant treatment (32).

Benefits

Screening benefits (vs. no screening or incremental to other strategies) included percentage of reduction in breast cancer mortality, breast cancer deaths averted, and life-years and quality-adjusted life-years (QALYs) gained because of averted or delayed breast cancer death. Benefits (and harms) were accumulated from age 40 to 100 years to capture the lifetime effect of screening.

We considered preferences, or utilities, to account for morbidity from screening and treatment. A disutility for age- and sex-specific general population health was first applied to quality-adjust the life-years (39). These were further adjusted to account for additional decrements in life-years related to undergoing screening (–0.006 for 1 week, or –1 hour), evaluating a positive screen (–0.105 for 5 weeks, or –3.7 days), undergoing initial treatment by stage (for the first 2 years after diagnosis), and having distant disease (for the last year of life for all women who die of breast cancer) (**Appendix Table 1**, available at www.annals.org) (33, 40, 41).

Use and Harms

Use of services focused on the number of mammograms required for the screening strategy. Harms included false-positive mammograms, benign biopsies, and overdiagnosis. Rates of false-positive mammograms were calculated as 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 biopsies among women with false-positive screening results; we assumed 100% compliance with biopsy recommendations (42). Overdiagnosis was defined as all cases that would not have been clinically detected in the absence of screening because of lack of progressive potential or death from competing non–breast cancer mortality. The effect of overdiagnosis on QALYs was captured by the disutility of being treated for cancer but dying of other causes.

Statistical Analysis

For each model, strategies were ranked by the number of mammograms performed. We report the median use, benefits, and harms and range across models. We also obtained an efficiency frontier by plotting the sequence of points that represent the largest incremental percentage of reduction in breast cancer mortality (or life-years or QALYs) per mammogram performed or harm entailed. Screening strategies that fell on this frontier were considered the most efficient (that is, they have the steepest slope such that no alternative exists that provides more benefit with less use or fewer harms).

Three models (E, GE, and W) also evaluated results based on combinations of breast cancer risk and density. Risk levels included 1.3 times higher than average (for example, nulliparity or age at first live birth >30 years) (18, 43); 2.0 times higher than average (for example, family history of 1 first-degree relative) (18); or 4.0 times higher than average (for example, family history of 2 or more first-degree relatives) (18, 44). Greater risk levels, such as those seen with *BRCA1* and *BRCA2* mutations, were not considered because such groups have specific screening guidelines. We made the simplifying assumption that risk affected incidence, but not other aspects of disease.

Breast density was modeled as entirely fatty ("a"), scattered density ("b"), heterogeneously dense ("c") and extremely dense ("d"). On the basis of observed age-specific prevalence rates, density was assigned at age 40 years and remained the same or decreased by 1 level at age 50 years and again at age 65 years (45). Density modified the sensitivity and specificity of mammography on the basis of age, interval, and first versus subsequent screening (33). Density also modified the age group–specific (40 to 49 years, 50 to 64, and ≥65 years) risk for breast cancer, using age-group-specific risk among those with average population density as the referent category (BCSC. Unpublished data) (44, 46). Density was assumed not to affect molecular subtype or disease natural history. Results for density were grouped into low ("a and b") or high ("c and d") for presentation. The risk- and density-specific results were also compared with those for screening average-risk and average-density groups biennially from 50 to 74

years, because many guideline groups accept the latter.

In other analyses, 2 models (model E and GE) examined the effect of comorbidity on screening cessation by using comorbidity-specific life expectancy groups. Examples of conditions that placed women in severe and moderate comorbidity groups included congestive heart failure and diabetes, respectively; the specific conditions and their associated life expectancies are reported elsewhere (20, 35, 47). We compared results of continuing to screen biennially past age 74 years among women with no or low comorbidity or stopping earlier than 74 years for those with moderate or high comorbidity. These analyses included women who survived and were free of breast cancer up until the point where screening was to be extended or stopped.

Four models evaluated whether high disutility values would eliminate screening benefits. Finally, we evaluated the ability of the models to independently predict external trends and results (Appendix Figure 1 and Appendix Table 2, available at www.annals.org).

Role of the Funding Source

We worked with the USPSTF and Agency for Healthcare Research and Quality (AHRQ) to develop the scope and key questions for this research. National Cancer Institute investigators (K.C., E.F.) collaborated in their role as scientific project officers. AHRQ staff distributed earlier versions of the draft for peer review. The agencies had no role in the study conduct or decision to submit the manuscript for publication. The investigators are solely responsible for the content and the decision to submit the manuscript for publication.

RESULTS

Benefits in the Average-Risk Population

The models produced consistent rankings of the screening strategies (Table 1). For instance, biennial screening from age 50 to 74 years yielded a median 25.8% reduction in breast cancer mortality compared with no screening (range, 24.1% to 31.8%). Annual screening led to slightly greater reductions in mortality than biennial strategies. However, biennial strategies maintained a median of 79.8% to 81.3% of the breast cancer–specific mortality reduction of annual screening (range, 68.3% to 98.9%) (Appendix Table 3, available at www.annals.org).

Biennial screening also maintained the most annual benefits for life-years and QALYs, and quality adjustment did not change the ranking of strategies. Across all strategies, the largest decrement from quality adjustment to life-years was related to declines in general health as women aged; smaller decrements occurred owing to the disutility of undergoing diagnostic evaluation of an abnormal screening examination and for having cancer. The disutility associated with screening itself had a minimal effect on QALYs (33).

The incremental benefits of initiating screening at age 40 years were slightly greater than those of starting at age 50 years in terms of breast cancer deaths averted with both annual and biennial screening (median, 1.3 [range, 1.1 to 1.7] and 1.0 [0.8 to 1.7] per 1000 women screened, respectively) (Table 2). Initiating screening at age 45 years yielded benefits intermediate between beginning at 40 and 50 years, although there were slightly greater incremental benefits when starting at age 45 years (vs. 50 years) than starting at age 40 years (vs. 45 years) (for example, 10.6 vs. 8.0 and 15.4 vs. 7.9 QALYs for biennial and annual strategies, respectively) (Table 1).

Harms in the Average-Risk Population

All models projected more harms (false-positive results, benign biopsies, and overdiagnosed cases) under annual versus biennial schedules and by starting earlier than age 50 years versus at age 50 years (Table 3). For instance, if biennial screening began at age 40 years instead of age 50 years, for every 1000 women screened there would be a median of 1 more death averted, but 576 more false-positive results, 67 more benign biopsies, and 2 additional overdiagnosed cases. Compared with screening initiation at age 45 years, starting screening at age 40 years had 1 or fewer added deaths averted depending on interval, but more incremental harms.

Efficiency Frontiers for Average-Risk Populations

Efficiency frontier plots were used to graphically depict the balance between the number of mammograms and benefits (life-years gained) of screening strategies. Biennial strategies starting at age 40, 45, or 50 years were all efficient (Figure and Appendix Figure 2, available at www.annals.org). Points that were close to but fell below the frontier were less efficient than those on the frontier line. For example, compared with the point on the efficient frontier for biennial screening at age 45 years, the hybrid strategy of annual screening at 45 years was less efficient than biennial screening starting at 40 years. This is because the hy-brid strategy at 45 years would require 405.8 more mammograms to gain an additional life-year for every 1000 women screened compared with biennial screen-ing at 45 years, whereas biennial screening starting at 40 years would require only 189.5 extra mammograms to gain an additional life-year.

Finally, annual screening from age 50 to 74 years was consistently inferior to other strategies (that is, it was inefficient, or dominated) because it yielded the same or fewer benefits than the next less resource intensive efficient strategy but required more mammograms or entailed more harms (for example, A50 vs. B40 in Figure 1). These patterns were generally seen with other harm and benefit metrics (Appendix Figure 2).

Sensitivity Analyses for Average-Risk Populations

Varying the disutilities for usual health, screening, diagnosis, and treatment did not affect strategy rankings for average-risk populations, and QALY gains per-

Table 1. Ranking of Benefits per 1000 Women Screened, by Model and Screening Strategy

Strategy*	Screens, *n*†	Model						Median (Range Across Models), %
		D	E	GE	M	S	W	
Reduction in Breast Cancer Mortality per 1000 Women Screened vs. No Screening, %*‡								
Biennial, 50–74	11 127	25.6	26.0	31.8	26.8	24.1	25.4	25.8 (24.1–31.8)
Biennial, 45–74	13 212	26.6	27.6	33.9	28.4	25.9	26.7	27.2 (25.9–33.9)
Hybrid, 45–74	15 966	27.7	29.7	35.9	29.2	27.3	30.1	29.5 (27.3–35.9)
Biennial, 40–74	16 013	28.3	30.3	35.9	31.9	28.2	30.5	30.4 (28.2–35.9)
Hybrid, 40–74	20 884	29.0	32.3	37.9	31.7§	29.3	32.8	32.0 (29.0–37.9)
Annual, 50–74	21 318	32.1	33.9	37.6§	27.1§	29.1§	35.3	33.0 (27.1–37.6)
Annual, 45–74	26 136	34.2	37.6	41.6	29.4§	32.3	39.1	35.9 (29.4–41.6)
Annual, 40–74	31 037	35.5	40.1	43.6	32.5	34.4	42.6	37.8 (32.5–43.6)
Years of Life Gained per 1000 Women Screened vs. No Screening*								
Biennial, 50–74	11 127	153.8	94.0	140.5	146.5	104.2	74.6	122.4 (74.6–153.8)
Biennial, 45–74	13 212	168.4	107.7	161.2	171.3	115.2	84.0	138.2 (84.0–171.3)
Hybrid, 45–74	15 966	175.3	117.9	170.2	171.4	125.1	95.7	147.7 (95.7–175.3)
Biennial, 40–74	16 013	183.7	123.7	172.4	194.8	131.6	98.8	152.0 (98.8–194.8)
Hybrid, 40–74	20 884	191.1	137.6	187.2	211.5	141.0	110.9	164.1 (110.9–211.5)
Annual, 50–74	21 318§	180.0§	125.9§	167.3§	156.3§	133.3§	104.3§	144.8 (104.3–180.0)§
Annual, 45–74	26 136	201.3	149.3	196.7	177.8§	154.2	123.0	166.0 (123.0–201.3)
Annual, 40–74	31 037	217.1	168.8	213.5	218.1	170.1	140.5	191.8 (140.5–218.1)
QALYs Gained per 1000 Women Screened vs. No Screening*‖								
Biennial, 50–74	11 127	114.5	67.3	100.1	109.6	71.9	47.1	86.0 (47.1–114.5)
Biennial, 45–74	13 212	123.8	75.6	114.4	129.4	78.8	51.9	96.6 (51.9–129.4)
Hybrid, 45–74	15 966	126.6	80.9	118.3	128.5	84.5	58.3	101.4 (58.3–128.5)
Biennial, 40–74	16 013	133.7	85.4	120.1	148.1	89.1	60.4	104.6 (60.4–148.1)
Hybrid, 40–74	20 884	134.2	91.0	126.1	159.4	92.5	64.8	109.3 (64.8–159.4)
Annual, 50–74	21 318	127.0§	84.1§	111.4§	113.2§	87.5§	62.4§	99.5 (62.4–127.0)§
Annual, 45–74	26 136	138.9	97.3	129.5	129.4§	99.5	71.7	114.5 (71.7–138.9)
Annual, 40–74	31 037	146.6	107.3	137.2	160.6	107.6	80.0	122.4 (80.0–160.6)

D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; GE = Georgetown University Medical Center and Albert Einstein College of Medicine; M = MD Anderson Cancer Center; QALY = quality-adjusted life-year; S = Stanford University; W = University of Wisconsin and Harvard Medical School.

* 100% of women receive adjuvant systemic therapy based on recommended stage and estrogen receptor/human epidermal growth factor-2 receptor–specific adjuvant therapy, and age (i.e., estrogen receptor–positive women older than 50 years were assigned tamoxifen, and those 50 years or older were assigned tamoxifen or an aromatase inhibitor).

† Strategies are ranked from the least to the most mammograms, where the number of mammographies is the median across models. Not all possible mammographies in the age interval are obtained because some women die of other causes before screening would occur.

‡ 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, 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. **Table 3** includes the absolute reduction in breast cancer deaths (i.e., deaths averted) versus no screening for each strategy.

§ Strategies that are inferior or inefficient ("dominated") within a specific model; a strategy is classified as such if another strategy results in an equal or higher benefit (percentage of decline in mortality, life-years gained, or QALYs) with fewer harms (e.g., average screening examinations).

‖ Adjusted for general health, diagnosis, screening, and treatment.

sisted under all screening strategies, although their magnitude decreased with increasing disutility (33).

Harms and Benefits, by Risk Level

The balance of harms and benefits differed by risk group, with women at higher risk having fewer false-positive results per 1000 women screened and higher gains from screening than lower-risk groups. Screening higher-risk women also yielded a lower ratio of over-diagnosed cases per breast cancer death averted than screening average-risk women. However, annual screening from age 50 to 74 years had the same or less benefit and more harms than other strategies at all risk levels (33).

For women with a 2- to 4-fold increase in risk, annual screening starting at age 40 or 45 years had similar or more favorable harm–benefit ratios (on the basis of false-positive results) as biennial screening of average-risk women from 50 to 74 years of age. If the harm–benefit ratio of 185.8 (range, 169.5 to 268.0) false-positive results per death averted for average-risk women screened biennially from 50 to 74 years of age is considered to be acceptable, then for women with a 2- to 4-fold increase in risk, annual screening starting at age 40 or 45 years had similar or more favorable harm–benefit ratios. For instance, women with a 2-fold increase in risk undergoing annual screening at age 40 years have a slightly more favorable corresponding ratio, at 182.5 (range, 177.4 to 231.9) false-positive results per death averted. For women with a 1.3-fold increase in risk, biennial screening starting at age 40

Table 2. Incremental Changes in Breast Cancer Deaths Averted, by Interval, Age of Screening Initiation, and Model*

Model	Breast Cancer Deaths Averted per 1000 Women, *n* (Breast Cancer Mortality Reduction, %)			
	Annual Screening		Biennial Screening	
	Start at Age 40 vs. 50 y	Start at Age 45 vs. 50 y	Start at Age 40 vs. 50 y	Start at Age 45 vs. 50 y
D	1.1 (3.4)	0.6 (2.1)	0.9 (2.7)	0.3 (1.0)
E	1.5 (6.2)	0.9 (3.6)	1.0 (4.3)	0.4 (1.6)
GE	1.5 (6.0)	1.0 (4.0)	1.0 (4.1)	0.5 (2.2)
M	1.7 (5.3)	0.7 (2.3)	1.7 (5.1)	0.5 (1.6)
S	1.1 (5.2)	0.7 (3.1)	0.9 (4.1)	0.4 (1.7)
W	1.1 (7.3)	0.6 (3.8)	0.8 (5.1)	0.2 (1.3)
Median	1.3 (5.7)	0.7 (3.4)	1.0 (4.2)	0.4 (1.6)

D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; GE = Georgetown University Medical Center and Albert Einstein College of Medicine; M = MD Anderson Cancer Center; S = Stanford University; W = University of Wisconsin and Harvard Medical School.
* Incremental difference between starting at age 40 or 45 versus 50. Annual is comparing A 40–74 (or A 45–74) with A 50–74; biennial is comparing B 40–74 (or B 45–74) with B 50–74. Hybrid strategies are compared with B 50–74; therefore, for those incremental comparisons, the hybrid results are the same as the annual results and are not shown separately here.

years had similar harm–benefit ratios as biennial screening of average-risk women from age 50 to 74 years.

Benefits and Harms, by Breast Density Group

Breast density (low vs. high) changed absolute benefits, but annual screening from 50 to 74 years remained inefficient across breast density groups. Women in the low-density group had a greater proportion of their cancers detected owing to greater sensitivity of digital mammography, and therefore a greater reduction in breast cancer–specific mortality than the high-density group. However, women in the high-density group had a greater absolute number of cancer cases detected because their risk for cancer was higher, leading to more life-years saved among these women than in the low-density group (33).

Benefits and Harms, by Comorbidity

For women with no comorbidity, biennial screening could continue to age 78 or 80 years and still have similar harm–benefit ratios as screening women with average non–breast cancer mortality biennially from 50 to 74 years. However, for women with moderate to se-

Table 3. Lifetime Benefits and Harms of Screening Strategies, Based on Starting Ages and Screening Intervals

Strategy and Age Group	Median Value (Range Across Models) per 1000 Women Screened vs. No Screening*					
	Screens, *n*	Breast Cancer Deaths Averted, *n*	False-Positive Screens, *n*	Benign Breast Biopsies, *n*	Overdiagnosed Cases (Invasive and DCIS), *n*†‡	All Overdiagnosed Cases, %†‡
Biennial						
50–74	11 127	7 (4–9)	953 (830–1325)	146 (121–205)	19 (11–34)	12 (8–22)
45–74	13 212	8 (4–9)	1220 (930–1599)	176 (131–232)	19 (11–34)	12 (8–22)
40–74	16 013	8 (5–10)	1529 (1100–1976)	213 (153–276)	21 (12–38)	13 (9–24)
Hybrid						
45–74	15 966	8 (5–9)	1520 (1160–1968)	202 (154–266)	21 (12–40)	13 (8–25)
40–74	20 884	9 (5–10)	2106 (1480–2623)	256 (184–325)	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)	283 (265–376)	28 (12–74)	17 (9–38)
40–74	31 037	10 (6–11)	2941 (2550–3742)	338 (296–435)	30 (13–77)	18 (9–39)

DCIS = ductal carcinoma in situ.
* In all scenarios, 100% of women receive adjuvant systemic therapy based on recommended stage and estrogen receptor/human epidermal growth factor-2 receptor–specific adjuvant therapy, and age (i.e., estrogen receptor–positive women older than 50 years were assigned tamoxifen, and those 50 years or older were assigned tamoxifen or an aromatase inhibitor).
† Overdiagnosed cases are those that would not have been clinically detected in the absence of screening (i.e., cases in which the patient does not die of breast cancer owing to lack of progressive potential or death from a competing non–breast cancer cause). The result includes overdiagnosis of DCIS and invasive disease. Overdiagnosis is calculated by comparing cases detected in the screening scenario with those detected in the nonscreened scenario. Model S (Stanford University) is excluded because it does not include DCIS. The percentage of overdiagnosis is calculated as the proportion of all cases detected in the screening strategy that are overdiagnosis.
‡ The upper range for all estimates of overdiagnosis is based on results from model M (MD Anderson Cancer Center). Model M generates very high overdiagnosis on the basis of the assumption that incidence in the absence of screening has essentially remained flat since 1975–1979, with almost 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. In general, models that assume higher proportions of DCIS or invasive cancer to be nonprogressive generate higher estimates of overdiagnosis than models that assume less nonprogressive disease. 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.

vere comorbidity, the comparable ratios were equivalent at about age 66 to 68 years (33).

DISCUSSION

We used 6 established models to estimate the potential efficacy of different breast cancer screening strategies in the United States. All 6 models demonstrated that screening initiation at age 40 years has some benefits for average-risk populations, but also higher levels of harms, than strategies starting at age 50 years. The findings also suggest that comorbidity levels could be used to tailor the age of screening cessation. Biennial screening strategies were the most efficient, but annual screening could be considered from age 40 to 74 years in groups with a 2- to 4-fold higher-than-average risk.

Results from all models indicated that digital mammography screening of average-risk women in their 40s modestly lowers breast cancer–specific mortality and extends the length and quality of life, even after disutilities related to the screening process are considered. The absolute benefits of starting screening in the 40s varied somewhat on the basis of model structure and assumptions, but were consistent with observations from randomized trials (6). However, starting at age 40 versus 45 years was associated with increasing incremental harms relative to the increase in benefits. Thus, decisions about initiating screening before age 50 years may depend on the weight attached to screening benefits and harms.

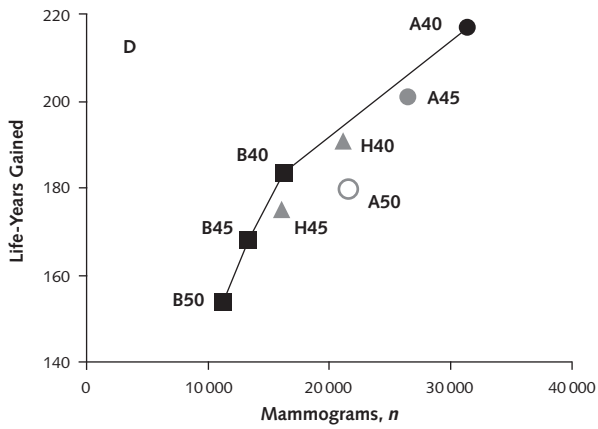
Consistent with other analyses of upper age limits for screening (20, 48–50) and other recommendations (12, 51), our results suggested that the balance of harms and benefits of screening was affected by competing non–breast cancer mortality, so that age at screening cessation could be tailored by comorbidity levels.

Similar to our 2009 analysis (1), biennial strategies are most consistently efficient. Screening annually from age 50 to 74 years had the same or fewer benefits for any given harm for all population groups in virtually all models and would be considered inefficient. However, annual screening in the 40s followed by biennial screening at age 50 years, and the most intensive schedule evaluated (annual screening from 40 to 74 years), were also efficient or close to being efficient. In addition, annual screening of women with a 2- to 4-fold increased risk (for example, due to non–*BRCA*-related family history) from age 40 to 74 years had harm–benefit ratios similar to those of biennial screening from age 50 to 74 years in average-risk populations.

The results also suggest that benefits of screening vary by breast density, at least when grouped into low and high categories. Women with dense breasts have a higher risk for cancer and a higher absolute detection rate, but a lower relative detection rate (19, 52). This is because the sensitivity of digital mammography, although optimized for density, is still lower in women with dense breasts than those with nondense breasts (53–56). Improving outcomes for women with dense breasts (55) may require new innovations in imaging (57–60) or identification of risk biomarkers (61, 62).

This analysis extends our prior work by explicitly considering overdiagnosis as a screening harm. Depending on screening strategy, the models estimated that 2% to 12% of invasive disease cases and 30% to 50% of DCIS cases might represent overdiagnosis. Although the models differed in absolute estimates, they agreed on how overdiagnosis affected the ranking of strategies and the finding that most overdiagnosed cases were DCIS. The model results for overdiagnosis are not directly comparable with other published estimates (8, 63) because the models followed women for their entire lives. The models also made assumptions about unobservable input parameters related to natural history. Although there is no agreement on methods to estimate overdiagnosis (64) or on its true rate (65, 66), there is agreement that it is an important harm. Active surveillance for DCIS with a low risk for progression is one potential future approach to reduce harms

Figure. Efficiency frontier for life-years gained versus mammograms performed per 1000 women in model D (Dana-Farber Cancer Institute).



Efficiency frontier graphs for all models are shown in **Appendix Figure 2** (available at www.annals.org). This graph plots the average gain in life-years per additional mammogram performed per 1000 women for each screening strategy (vs. no screening) in model D. Biennial strategies are indicated with a square; hybrid strategies (annual in the 40s followed by biennial from 50 to 74 years of age) with a triangle; and annual strategies with a circle. Efficient strategies were plotted (i.e., those in which increases in mammography use resulted in greater life-years gained than the next less intensive strategy). The line represents the "efficiency frontier" by joining efficient strategies in which increases in mammography use resulted in greater life-years gained than the next less intensive efficient strategy. Strategies on this line would be considered efficient because they achieve the greatest gain in benefit (life-years gained) per harm or use of mammograms. Strategies that use more mammograms but still have small benefits (i.e., a shallower slope than the next best strategy) are considered to be less efficient (i.e., weakly dominated). When and if the slope in the efficiency frontier plot levels off, it means that 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 (or additional benefit). There is no definitive inflection point across the models for the strategies or metrics evaluated. Black strategies are efficient; gray strategies close to the efficiency frontier are less efficient; and open gray strategies are inefficient (inferior, or dominated). Reference 33 provides efficiency frontiers for other harm and benefit metrics.

from DCIS overdiagnosis. More information is also needed on consumer knowledge of and willingness to risk overdiagnosis (67).

Overall, this study has several important strengths, including collaboration of 6 long-established, independent modeling groups; use of well-calibrated models that reproduce temporal epidemiologic trends and a screening trial result; inclusion of digital technology; incorporation of increasingly effective treatments; and consideration of quality of life, risk factors, breast density, and comorbidity (68). The conclusions about the ranking of screening strategies are robust and should provide greater credibility than inferences based on 1 model alone.

Our study also had limitations. First, to evaluate program efficacy, we assumed 100% adherence to screening, prompt evaluation of abnormal results, and full use of optimal treatment. Actual benefits will fall short of our projected results because adherence is not perfect.

Second, we focused only on hybrid strategies for women in their 40s. Alternative hybrid strategies may be important to examine in future research.

Third, the analysis did not consider other imaging technologies for average-risk populations or for groups with high breast density; such technologies include ultrasonography (69), computer-aided detection (70), tomosynthesis, or magnetic resonance imaging. Data on tomosynthesis performance and needs for radiologist retraining are still emerging (58).

Fourth, we did not model any radiation-induced breast cancers, owing to more intensive mammography schedules (71). Fifth, we assumed that risk factors influenced the incidence of disease, but not its natural history.

Sixth, certain risk factors, such as family history, are age-dependent in their effects (18, 72). Because we held relative risk levels constant over age, our benefit estimates could be over- or underestimated for specific risk factors (17).

Seventh, we did not consider polygenic risk (73, 74) or explicitly model menopausal status; we used age 50 years as a proxy for the average age of menopause. Eighth, the analysis did not include screening program costs or utility estimates specific to some of the newest treatments.

Finally, compared with our earlier research (1), the models all estimated similar, but somewhat greater, reductions in breast cancer–specific mortality (for example, a median 22% vs. 25.8% reduction with biennial screening from 50 to 74 years in 2009 vs. current models, respectively). The primary reasons for this modeled improvement relate to the increased sensitivity of digital compared with film mammography, advances in molecular-targeted therapies, and changes in underlying breast cancer trends.

Overall, the 6 models conclude that biennial screening strategies are the most efficient. Choices about the optimal age at initiation (and cessation) and screening intervals will ultimately depend on program goals, the weight attached by the decision maker to screening harms and benefits (75), and considerations of efficiency.

From Georgetown University Medical Center and Cancer Prevention and Control Program, Georgetown-Lombardi Comprehensive Cancer Center, Washington, DC; Harvard Pilgrim Health Care Institute, Dana-Farber Cancer Institute, and Harvard Medical School, Boston, Massachusetts; Albert Einstein College of Medicine, Bronx, New York; Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands; UC Davis School of Medicine, Davis, Stanford University, Stanford, and University of California, San Francisco, San Francisco, California; Group Health Research Institute, Seattle, Washington; Information Management Services, Calverton, and National Cancer Institute, Bethesda, Maryland; Carbone Cancer Center, University of Wisconsin, Madison, Wisconsin; University of Texas MD Anderson Cancer Center, Houston, Texas; Norris Cotton Cancer Center and Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire; and College of Medicine, and University of Vermont, Burlington, Vermont.

Portions of earlier versions of this research were presented on the USPSTF Web site as part of a Technical Report (AHRQ Publication No. 14-05201-EF-4) (33).

Note: This work was done by 6 independent modeling teams, for the Breast Cancer Working Group of the Cancer Intervention and Surveillance Modeling Network (CISNET) and the Breast Cancer Surveillance Consortium (BCSC). The teams were from Dana-Farber Cancer Institute (principal investigator [PI], Dr. Lee); Erasmus MC, University Medical Center Rotterdam (PI, Dr. de Koning); Georgetown University Medical Center, Lombardi Comprehensive Cancer Center (PI, Dr. Mandelblatt) and Albert Einstein College of Medicine (PI, Dr. Schechter); Harvard Medical School and Harvard Pilgrim Health Care (PI, Dr. Stout) and University of Wisconsin (PIs, Drs. Trentham-Dietz and Alagoz); University of Texas MD Anderson Cancer Center (PI, Dr. Berry); and Stanford University (PI, Dr. Plevritis). Drs. Mandelblatt, Cronin, de Koning, Miglioretti, Schechter, and Stout were the coordinating committee for the project; Drs. Mandelblatt, Stout, Schechter, and Cronin were the writing committee; Drs. de Koning and Cronin served as dual senior authors; and Dr. Feuer was responsible for overall CISNET project direction. The BCSC collaborated and also provided key data inputs to the modeling teams. Members of the BCSC are listed at http://breastscreening.cancer.gov.

Disclaimer: The investigators worked with members of the USPSTF and AHRQ staff to develop the scope and key questions for this research. The USPSTF, 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.

Acknowledgment: 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 AHRQ for assistance with analysis and review of

results; members of the USPSTF and the Oregon Evidence-based Practice Center for comments on earlier versions of this research; and Jessica Garshell for data processing. They also thank Adrienne Ryans for manuscript preparation and John Wong, Stuart Baker, William Lawrence, and Tom Trikalinos for helpful suggestions on earlier versions of this paper.

Grant Support: By the National Institutes of Health (National Cancer Institute grant U01 CA152958 and National Cancer Institute–funded BCSC grant P01 CA154292, contract HSN261201100031C, and grant U54CA163303), and in part by American Cancer Society grant MRSG 14-027-01 CPHPS (Dr. Chang). 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 United States. A full list of BCSC investigators and a description of these sources is available at http://breastscreening.cancer.gov/work/acknowledgement.html.

Disclosures: Dr. Stout reports grants from the National Cancer Institute during the conduct of the study. Dr. Schechter reports grants from the National Cancer Institute during the conduct of the study. Dr. Miglioretti reports grants from AHRQ and the National Cancer Institute during the conduct of the study. Dr. Trentham-Dietz reports grants from the National Cancer Institute during the conduct of the study. Dr. van Ravesteyn reports grants from the National Cancer Institute during the conduct of the study. Dr. Alagoz reports grants from the National Cancer Institute during the conduct of the study. Dr. Tosteson reports grants from the National Cancer Institute during the conduct of the study. Dr. Heijnsdijk reports grants from SCOR and personal fees from the American Society of Breast Surgeons outside the submitted work. Dr. Gangnon reports grants from the National Cancer Institute during the conduct of the study. Dr. Sprague reports grants from the National Institutes of Health during the conduct of the study. Dr. de Koning reports grants from the National Cancer Institute during the conduct of the study. Authors not named here have disclosed no conflicts of interest. Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M15-1536.

Reproducible Research Statement: *Study protocol, statistical code, and data set:* Not available.

Requests for Single Reprints: Jeanne S. Mandelblatt, MD, MPH, Lombardi Comprehensive Cancer Center, 3300 Whitehaven Street, NW, Suite 4100, Washington, DC 20007; e-mail, mandelbj@georgetown.edu.

Current author addresses and author contributions are available at www.annals.org.

References

1. **Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al; Breast Cancer Working Group of the Cancer Intervention and Surveillance Modeling Network.** Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med. 2009;151:738-47. [PMID: 19920274] doi:10.7326/0003-4819-151-10-200911170-00010

2. **Gøtzsche PC, Jørgensen KJ.** Screening for breast cancer with mammography. Cochrane Database Syst Rev. 2013;6:CD001877. [PMID: 23737396] doi:10.1002/14651858.CD001877.pub5

3. **Biller-Andorno N, Jüni P.** Abolishing mammography screening programs? A view from the Swiss Medical Board. N Engl J Med. 2014;370:1965-7. [PMID: 24738641] doi:10.1056/NEJMp1401875

4. **Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE.** Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet. 2002;359:909-19. [PMID: 11918907]

5. **Tabár 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;38:625-51. [PMID: 10943268]

6. **Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L; Trial Management Group.** Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet. 2006;368:2053-60. [PMID: 17161727]

7. **Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al; Cancer Intervention and Surveillance Modeling Network (CISNET) Collaborators.** Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med. 2005;353:1784-92. [PMID: 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. [PMID: 24519768] doi:10.1136/bmj.g366

9. **Paci E, Broeders M, Hofvind S, Puliti D, Duffy SW; EUROSCREEN Working Group.** European breast cancer service screening outcomes: a first balance sheet of the benefits and harms. Cancer Epidemiol Biomarkers Prev. 2014;23:1159-63. [PMID: 24991022] doi:10.1158/1055-9965.EPI-13-0320

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;23:1139-46. [PMID: 24991021] doi:10.1158/1055-9965.EPI-13-0946

11. **Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, et al; International Agency for Research on Cancer Handbook Working Group.** Breast-cancer screening–viewpoint of the IARC Working Group. N Engl J Med. 2015;372:2353-8. [PMID: 26039523] doi:10.1056/NEJMsr1504363

12. **U.S. Preventive Services Task Force.** Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2009;151:716-26. [PMID: 19920272] doi:10.7326/0003-4819-151-10-200911170-00008

13. **Broeders M, Moss S, Nyström L, Njor S, Jonsson H, Paap E, et al; EUROSCREEN Working Group.** 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. [PMID: 22972807]

14. **Moss SM, Wale C, Smith R, Evans A, Cuckle H, Duffy SW.** Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. Lancet Oncol. 2015;16:1123-32. [PMID: 26206144] doi:10.1016/S1470-2045(15)00128-X

15. **U.S. Food and Drug Administration.** Mammography Quality Standards Act and Program. Silver Spring, MD: U.S. Department of Health and Human Services; 2013. Accessed at www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/default.htm on 12 January 2015.

16. **National Comprehensive Cancer Network.** NCCN clinical practice guidelines in oncology–breast cancer. 2014. Accessed at www.nccn.org/professionals/physician_gls/f_guidelines.asp on 12 January 2015.

17. **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;156:609-17. [PMID: 22547470] doi:10.7326/0003-4819-156-9-201205010-00002

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;156:635-48. [PMID: 22547473] doi:10.7326/0003-4819-156-9-201205010-00006
19. **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;155:10-20. [PMID: 21727289] doi:10.7326/0003-4819-155-1-201107050-00003
20. **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;161:104-12. [PMID: 25023249] doi:10.7326/M13-2867
21. **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;137:165-72. [PMID: 25430053] doi:10.1002/ijc.29364
22. **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;12:551-8. [PMID: 9294789]
23. **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. [PMID: 25255803] doi:10.1093/jnci/dju289
24. **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. 2006:37-47. [PMID: 17032893]
25. **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. 2006:47-55. [PMID: 17032894]
26. **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. 2006:30-6. [PMID: 17032892]
27. **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. 2006:86-95. [PMID: 17032898]
28. **Lee S, Zelen M.** A stochastic model for predicting the mortality of breast cancer. J Natl Cancer Inst Monogr. 2006:79-86. [PMID: 17032897]
29. **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. 2006:56-65. [PMID: 17032895]
30. **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. 2006:96-105. [PMID: 17032899]
31. **Gangnon RE, Sprague BL, Stout NK, Alagoz O, Weedon-Fekjær 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. Cancer Epidemiol Biomarkers Prev. 2015;24:905-12. [PMID: 25787716] doi:10.1158/1055-9965.EPI-14-1286
32. **Peto R, Davies C, Godwin J, Gray R, Pan HC, Clarke M, et al; Early Breast Cancer Trialists' Collaborative Group (EBCTCG).** 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;379:432-44. [PMID: 22152853] doi:10.1016/S0140-6736(11)61625-5
33. **Mandelblatt JS, Cronin K, de Koning H, Miglioretti DL, Schechter CS, Stout N.** Modeling Report: Collaborative Modeling of U.S. Breast Cancer Screening Strategies. AHRQ Publication No. 14-05201-EF-4. Rockville, MD: U.S. Preventive Services Task Force; 2015. Accessed at www.uspreventiveservicestaskforce.org/Page/Document/modeling-report-collaborative-modeling-of-us-breast-cancer-1/breast-cancer-screening1 on 4 May 2015.
34. **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;136:823-35. [PMID: 23104221] doi:10.1007/s10549-012-2274-3
35. **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;159:667-76. [PMID: 24247672] doi:10.7326/0003-4819-159-10-201311190-00005
36. **Holford TR, Cronin KA, Mariotto AB, Feuer EJ.** Changing patterns in breast cancer incidence trends. J Natl Cancer Inst Monogr. 2006:19-25. [PMID: 17032890]
37. **Kerlikowske K, Zhu W, Hubbard RA, Geller B, Dittus K, Braithwaite D, et al; Breast Cancer Surveillance Consortium.** Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. JAMA Intern Med. 2013;173:807-16. [PMID: 23552817] doi:10.1001/jamainternmed.2013.307
38. **Dittus K, Geller B, Weaver DL, Kerlikowske K, Zhu W, Hubbard R, et al; Breast Cancer Surveillance Consortium.** Impact of mammography screening interval on breast cancer diagnosis by menopausal status and BMI. J Gen Intern Med. 2013;28:1454-62. [PMID: 23760741] doi:10.1007/s11606-013-2507-0
39. **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;26:391-400. [PMID: 16855127]
40. **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;98:774-82. [PMID: 16757702]
41. **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;49:538-44. [PMID: 1917155]
42. **Rosenberg RD, Yankaskas BC, Abraham LA, Sickles EA, Lehman CD, Geller BM, et al.** Performance benchmarks for screening mammography. Radiology. 2006;241:55-66. [PMID: 16990671]
43. **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;100:1724-33. [PMID: 19033562] doi:10.1093/jnci/djn388
44. **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;98:1204-14. [PMID: 16954473]
45. **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. [PMID: 25217577] doi:10.1093/jnci/dju255
46. **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;105:1043-9. [PMID: 23744877] doi:10.1093/jnci/djt124
47. **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;66:1376-85. [PMID: 24035494] doi:10.1016/j.jclinepi.2013.07.002
48. **Walter LC, Covinsky KE.** Cancer screening in elderly patients: a framework for individualized decision making. JAMA. 2001;285:2750-6. [PMID: 11386931]
49. **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;59:1444-51. [PMID: 21797837] doi:10.1111/j.1532-5415.2011.03523.x

50. **van Ravesteyn NT, Stout NK, Schechter CB, Heijnsdijk EA, 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. [PMID: 25948872] doi:10.1093/jnci/djv103
51. **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;62:129-42. [PMID: 22261986] doi:10.3322/caac.20143
52. **Vilaprinyo E, Forné C, Carles M, Sala M, Pla R, Castells X, et al; Interval Cancer (INCA) Study Group.** Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. PLoS One. 2014;9:e86858. [PMID: 24498285] doi:10.1371/journal.pone.0086858
53. **Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al; Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group.** Diagnostic performance of digital versus film mammography for breast-cancer screening. N Engl J Med. 2005;353:1773-83. [PMID: 16169887]
54. **Kerlikowske K, Hubbard RA, Miglioretti DL, Geller BM, Yankaskas BC, Lehman CD, et al; Breast Cancer Surveillance Consortium.** Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. Ann Intern Med. 2011;155:493-502. [PMID: 22007043] doi:10.7326/0003-4819-155-8-201110180-00005
55. **Kerlikowske K, Zhu W, Tosteson AN, Sprague BL, Tice JA, Lehman CD, et al; Breast Cancer Surveillance Consortium.** Identifying women with dense breasts at high risk for interval cancer: a cohort study. Ann Intern Med. 2015;162:673-81. [PMID: 25984843] doi:10.7326/M14-1465
56. **Braithwaite D, Zhu W, Hubbard RA, O'Meara ES, Miglioretti DL, Geller B, et al; Breast Cancer Surveillance Consortium.** 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;105:334-41. [PMID: 23385442] doi:10.1093/jnci/djs645
57. **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;10:913-7. [PMID: 24295940] doi:10.1016/j.jacr.2013.09.010
58. **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;311:2499-507. [PMID: 25058084] doi:10.1001/jama.2014.6095
59. **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;203:223-8. [PMID: 24951219] doi:10.2214/AJR.13.11566
60. **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;202:293-8. [PMID: 24450668] doi:10.2214/AJR.13.11585
61. **Matamala N, Vargas MT, González-Cámpora R, Miñambres R, Arias JI, Menéndez P, et al.** Tumor microRNA expression profiling identifies circulating microRNAs for early breast cancer detection. Clin Chem. 2015;61:1098-106. [PMID: 26056355] doi:10.1373/clinchem.2015.238691
62. **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. [PMID: 25855707] doi:10.1093/jnci/djv036
63. **Welch HG, Passow HJ.** Quantifying the benefits and harms of screening mammography. JAMA Intern Med. 2014;174:448-54. [PMID: 24380095] doi:10.1001/jamainternmed.2013.13635
64. **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;158:831-8. [PMID: 23732716] doi:10.7326/0003-4819-158-11-201306040-00008
65. **Carter JL, Coletti RJ, Harris RP.** Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ. 2015;350:g7773. [PMID: 25569206] doi:10.1136/bmj.g7773
66. **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. [PMID: 25362701] doi:10.1093/jnci/dju315
67. **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:e0125165. [PMID: 25992887] doi:10.1371/journal.pone.0125165
68. **Elmore JG, Harris RP.** The harms and benefits of modern screening mammography [Editorial]. BMJ. 2014;348:g3824. [PMID: 24938686] doi:10.1136/bmj.g3824
69. **Sprague BL, Stout NK, Schechter C, van Ravesteyn NT, Cevik M, Alagoz O, et al.** Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. Ann Intern Med. 2015;162:157-66. [PMID: 25486550] doi:10.7326/M14-0692
70. **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;158:580-7. [PMID: 23588746] doi:10.7326/0003-4819-158-8-201304160-00002
71. **Miglioretti DL, Lange J, van Ravesteyn N, van den Broek JJ, Lee CI, Melnikow J, et al.** Radiation-Induced Breast Cancer and Breast Cancer Death from Mammography Screening [Abstract]. Rockville, MD: Agency for Healthcare Research and Quality; 2015.
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;145:165-75. [PMID: 24647890] doi:10.1007/s10549-014-2905-y
73. **Garcia-Closas M, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, Brook MN, et al; Gene ENvironmental Interaction and breast CAncer (GENICA) Network.** Genome-wide association studies identify four ER negative-specific breast cancer risk loci. Nat Genet. 2013;45:392-8, 398e1-2. [PMID: 23535733] doi:10.1038/ng.2561
74. **Stevens KN, Vachon CM, Couch FJ.** Genetic susceptibility to triple-negative breast cancer. Cancer Res. 2013;73:2025-30. [PMID: 23536562] doi:10.1158/0008-5472.CAN-12-1699
75. **Hoffmann TC, Del Mar C.** Patients' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. JAMA Intern Med. 2015;175:274-86. [PMID: 25531451] doi:10.1001/jamainternmed.2014.6016

Current Author Addresses: Dr. Mandelblatt: Lombardi Comprehensive Cancer Center, 3300 Whitehaven Street, NW, Suite 4100, Washington, DC 20007.
Dr. Stout: Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 133 Brookline Avenue, 6th Floor, Boston, MA 02215.
Dr. Schechter: Department of Family and Social Medicine, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Black Building 406, Bronx, NY 10461.
Mr. van den Broek and Drs. van Ravesteyn, Heijnsdijk, and de Koning: Erasmus MC, University Medical Center Rotterdam, P.O. Box 2040, 3000 CA, Rotterdam, the Netherlands.
Dr. Miglioretti: Department of Public Health Sciences, UC Davis School of Medicine, One Shields Avenue, Med Sci 1C, Room 145, Davis, CA 95616.
Mr. Krapcho: Information Management Services, Inc., 3901 Calverton Boulevard, Suite 200, Calverton, MD 20705.
Dr. Trentham-Dietz: University of Wisconsin Carbone Cancer Center, 610 Walnut Street, WARF Room 307, Madison, WI 53726.
Dr. Munoz: Stanford University, 318 Campus Drive, Room S255, Stanford, CA 94305
Dr. Lee: Dana-Farber Cancer Institute, Harvard Medical School, Mailstop CLS11007, 450 Brookline Avenue, Boston, MA 02215.
Dr. Berry: Department of Biostatistics, The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, 4-5062 Pickens Academic Tower, Houston, TX 77030-1402.
Dr. Alagoz: University of Wisconsin–Madison, 1513 University Avenue, Madison, WI 53706.
Dr. Kerlikowske: Veterans Affairs Medical Center (111A1), 4150 Clement Street, San Francisco, CA 94121.
Dr. Tosteson: Geisel School of Medicine at Dartmouth, One Medical Center Drive (HB7505), Lebanon, NH 03756.
Ms. Near, Ms. Hoeffken, and Dr. Chang: Lombardi Comprehensive Cancer Center, Georgetown University, 3300 Whitehaven Street NW, Suite 4100, Washington, DC 20007.
Mr. Chisholm: Department of Gynecologic Oncology and Reproductive Medicine, Unit 1362, The University of Texas MD Anderson Cancer Center, P.O. Box 301439, Houston, TX 77230-1439.
Dr. Huang: The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, Unit 1411, P.O. Box 301402, Houston, TX 77230-1402.
Ms. Huang: Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215.
Mr. Ergun: University of Wisconsin–Madison, 3233 Mechanical Engineering Building, 1513 University Avenue, Madison, WI, 53706.
Dr. Gangnon: University of Wisconsin–Madison, Department of Population Health Sciences, 610 Walnut Street, Madison, WI 53726.
Dr. Sprague: University of Vermont, Office of Health Promotion Research, 1 South Prospect Street, Burlington, VT 05401.
Dr. Plevritis: Department of Radiology, Stanford University, James H. Clark Center, Room S255, MC 5442, 318 Campus Drive, Stanford, CA 94305.
Drs. Feuer and Cronin: Division of Cancer Control and Population Sciences, National Cancer Institute, 9609 Medical Center Drive, MSC 9765, Room 4E534, Bethesda, MD 20892-9765.

Author Contributions: Conception and design: O. Alagoz, D.A. Berry, Y. Chang, K.A. Cronin, H.J. de Koning, J.S. Mandelblatt, C.B. Schechter, N.K. Stout.
Analysis and interpretation of the data: O. Alagoz, D.A. Berry, Y. Chang, Gary B.B. Chisholm, K.A. Cronin, H.J. de Koning, M.A. Ergun, E.J. Feuer, R.E. Gangnon, A.S. Hoeffken, X. Huang, H. Huang, K. Kerlikowske, M. Krapcho, S.J. Lee, J.S. Mandelblatt, D.L. Miglioretti, D.F. Munoz, A.M. Near, C.B. Schechter, B. Sprague, N.K. Stout, A.N. Tosteson, J.J. van den Broek, N.T. van Ravesteyn.
Drafting of the article: O. Alagoz, Gary B.B. Chisholm, K.A. Cronin, H.J. de Koning, A.S. Hoeffken, J.S. Mandelblatt, C.B. Schechter, N.K. Stout.
Critical revision of the article for important intellectual content: O. Alagoz, D.A. Berry, H.J. de Koning, R.E. Gangnon, A.S. Hoeffken, K. Kerlikowske, J.S. Mandelblatt, D.L. Miglioretti, A.M. Near, C.B. Schechter, B. Sprague, N.K. Stout, A.N. Tosteson, A. Trentham-Dietz, J.J. van den Broek, N.T. van Ravesteyn.
Final approval of the article: O. Alagoz, D.A. Berry, Y. Chang, Gary B.B. Chisholm, K.A. Cronin, H.J. de Koning, M.A. Ergun, E.J. Feuer, R.E. Gangnon, E.A.M. Heijnsdijk, A.S. Hoeffken, X. Huang, H. Huang, K. Kerlikowske, M. Krapcho, S.J. Lee, J.S. Mandelblatt, D.L. Miglioretti, D.F. Munoz, A.M. Near, S.K. Plevritis, C.B. Schechter, B. Sprague, N.K. Stout, A.N. Tosteson, A. Trentham-Dietz, J.J. van den Broek, N.T. van Ravesteyn.
Provision of study materials or patients: K. Kerlikowske, A. Trentham-Dietz.
Statistical expertise: D.A. Berry, Gary B.B. Chisholm, K.A. Cronin, H.J. de Koning, E.J. Feuer, R.E. Gangnon, A.S. Hoeffken, X. Huang, J.S. Mandelblatt, D.L. Miglioretti, D.F. Munoz, A.M. Near, C.B. Schechter, J.J. van den Broek.
Obtaining of funding: D.A. Berry, E.J. Feuer, J.S. Mandelblatt, D.L. Miglioretti, C.B. Schechter, B. Sprague, A. Trentham-Dietz.
Administrative, technical, or logistic support: M.A. Ergun, E.J. Feuer, A.S. Hoeffken, J.S. Mandelblatt.
Collection and assembly of data: D.A. Berry, H.J. de Koning, K. Kerlikowske, M. Krapcho, S.J. Lee, J.S. Mandelblatt, D.L. Miglioretti, C.B. Schechter, B. Sprague, N.K. Stout, A. Trentham-Dietz, J.J. van den Broek.

Appendix: Model Validation

Each model has a different structure and assumptions and some varying input variables, so no single method can be used to validate results against an external standard. Therefore, we used several approaches.

First, considering actual screening and treatment patterns instead of the efficacy strategies simulated in the base case, we compared model projections of incidence, breast cancer–specific mortality, and stage distribution with those reported by the Surveillance, Epidemiology, and End Results program for 1975 to 2010. In our previous work, results of each model accurately projected trends for incidence and breast cancer–specific mortality by ER status for 1975 to 2000 (23). Next, we approximated the Age screening trial (6), assuming perfect adherence to invitations for annual screening with 13-year follow-up of women aged 40 to 49 years

(6). Finally, we examined the consistency of results across models.

Using inputs for actual dissemination of screening and treatment in the United States, the models captured the major trends in incidence and the general shape of breast cancer–specific mortality decreases over time (Appendix Figure 1). They also closely matched current stage distribution (not shown) and the Age trial results (Appendix Table 2) (6, 33).

Thus, the models replicated patterns of observed US incidence and breast cancer–specific mortality over time. The models also estimated similar breast cancer–specific mortality reduction as that observed among women aged 40 to 49 years who actually attended screening in the Age trial, although the model results are slightly more optimistic than the trial because the models assume 100% screening and use of the most effective systemic regimens (6). Overall, use of 6 models to project a range of plausible screening outcomes provides implicit cross-validation, with the range of results from the models as a measure of uncertainty.

Appendix Table 1. Utility Input Parameter Values

Utilities for Cancer-Related States*

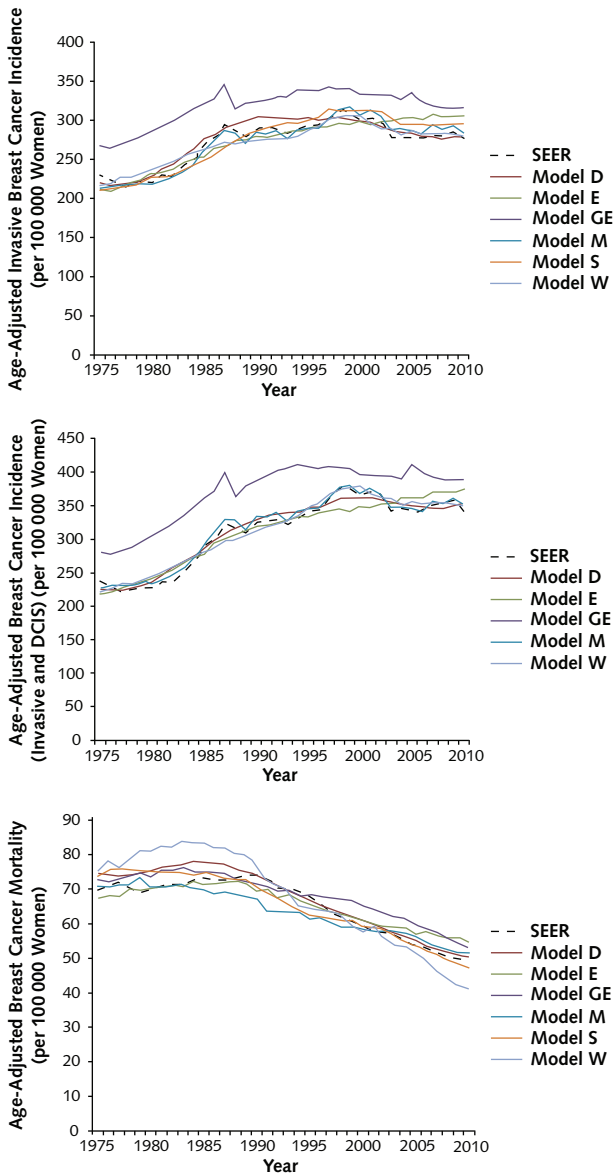
State	Utility	Disutility (Worst Case 150%, 200%)	Duration	Unit
Cancer treatment for local or DCIS	0.9	0.1 (0.15, 0.20)	2	Year
Cancer treatment for regional	0.75	0.25 (0.375, 0.50)	2	Year
Cancer treatment for distant	0.6	0.4 (0.6, 0.8)	Until death	–
Screening attendance (routine screening)	0.994	0.006 (0.009, 0.012)	1	Week
Diagnostic phase (evaluation of positive screen)	0.895	0.105 (0.158, 0.210)	5	Weeks

Age-Specific Utilities for General Health in U.S. Women†

Age	Healthy Base Value (Range)
20 y	0.913 (0.905–0.920)
25 y	0.913 (0.905–0.920)
30 y	0.893 (0.886–0.900)
35 y	0.893 (0.886–0.900)
40 y	0.863 (0.855–0.871)
45 y	0.863 (0.855–0.871)
50 y	0.837 (0.829–0.846)
55 y	0.837 (0.829–0.846)
60 y	0.811 (0.800–0.822)
65 y	0.811 (0.800–0.822)
70 y	0.771 (0.758–0.784)
75 y	0.771 (0.758–0.784)
80 y	0.724 (0.701–0.747)
85 y	0.724 (0.701–0.747)

DCIS = ductal carcinoma in situ.
* From references 40 and 41.
† Values from the EuroQoL-5D quality-of-life questionnaire (39).

Appendix Figure 1. Modeled versus observed incidence of breast cancer and breast cancer–specific mortality in women aged 40 to 100 years.



The models closely estimate observed U.S. trends in incidence of invasive disease (**top**), incidence of invasive disease and DCIS (**middle**)*, and breast cancer–specific mortality (**bottom**). Using inputs for actual dissemination of screening and treatment in the United States, the models all captured the major trends in incidence over time. Early increases with the advent of mammography in the mid-1980s are seen, followed by a downturn in the 2000s and then a leveling off. The models also captured the general shape of decreases in breast cancer–specific mortality over time. All models show an increase in incidence with the introduction of mammography screening. Model GE has a steep peak in incidence in 2005 owing to the specific method for capturing the transition from plain film to digital mammography, because digital mammography has higher sensitivity and detection of ductal carcinoma in situ than plain film mammography; other models include a more gradual transition surrounding this period. D = Dana-Farber Cancer Institute; DCIS = ductal carcinoma in situ; E = Erasmus Medical Center; GE = Georgetown University Medical Center and Albert Einstein College of Medicine; M = MD Anderson Cancer Center; QALY = quality-adjusted life-year; S = Stanford University; SEER = Surveillance, Epidemiology, and End Results; W = University of Wisconsin and Harvard Medical School.
* Model S does not include DCIS.

Appendix Table 2. Approximation of the Age Trial With 13-y Follow-up, by Model*

Model	Relative Risk for Breast Cancer Death With 100% Screening†
D	0.75
E	0.73
GE	0.65
M	0.72
S	0.69
W	0.71
Median (range)	0.72 (0.65–0.75)

D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; GE = Georgetown University Medical Center and Albert Einstein College of Medicine; M = MD Anderson Cancer Center; S = Stanford University; W = University of Wisconsin and Harvard Medical School.
* Projection of relative risk of breast cancer death with annual screening from age 40 to 49 y; biennial at age 50 and 52 y versus a control group with biennial screening at age 50 and 52 y. Because the models are estimating mortality reduction with actual screening, model estimates are most comparable to the Age trial results (6) among women who actually attended screening. Model results show more benefit than observed in the trial because the models assume that 100% of women complied with the trial-specified screening schedule. In reality, not all women who were invited attended screening, and among those who attended, many did not attend all scheduled screening rounds. In addition, the models assumed 100% receipt of the most effective treatments.
† Age trial invitation results (intention to treat): relative risk, 0.83 (95% CI, 0.66–1.04). Age trial results for women who actually were screened: relative risk, 0.76 (CI, 0.51–1.01).

Appendix Table 3. Annual Mortality Reduction Maintained by Biennial Screening, by Strategy and Model

Age at Screening	Mortality Reduction, %						
	Model D	Model E	Model GE	Model M*	Model S	Model 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

D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; GE = Georgetown University Medical Center and Albert Einstein College of Medicine; M = MD Anderson Cancer Center; S = Stanford University; W = University of Wisconsin and Harvard Medical School.

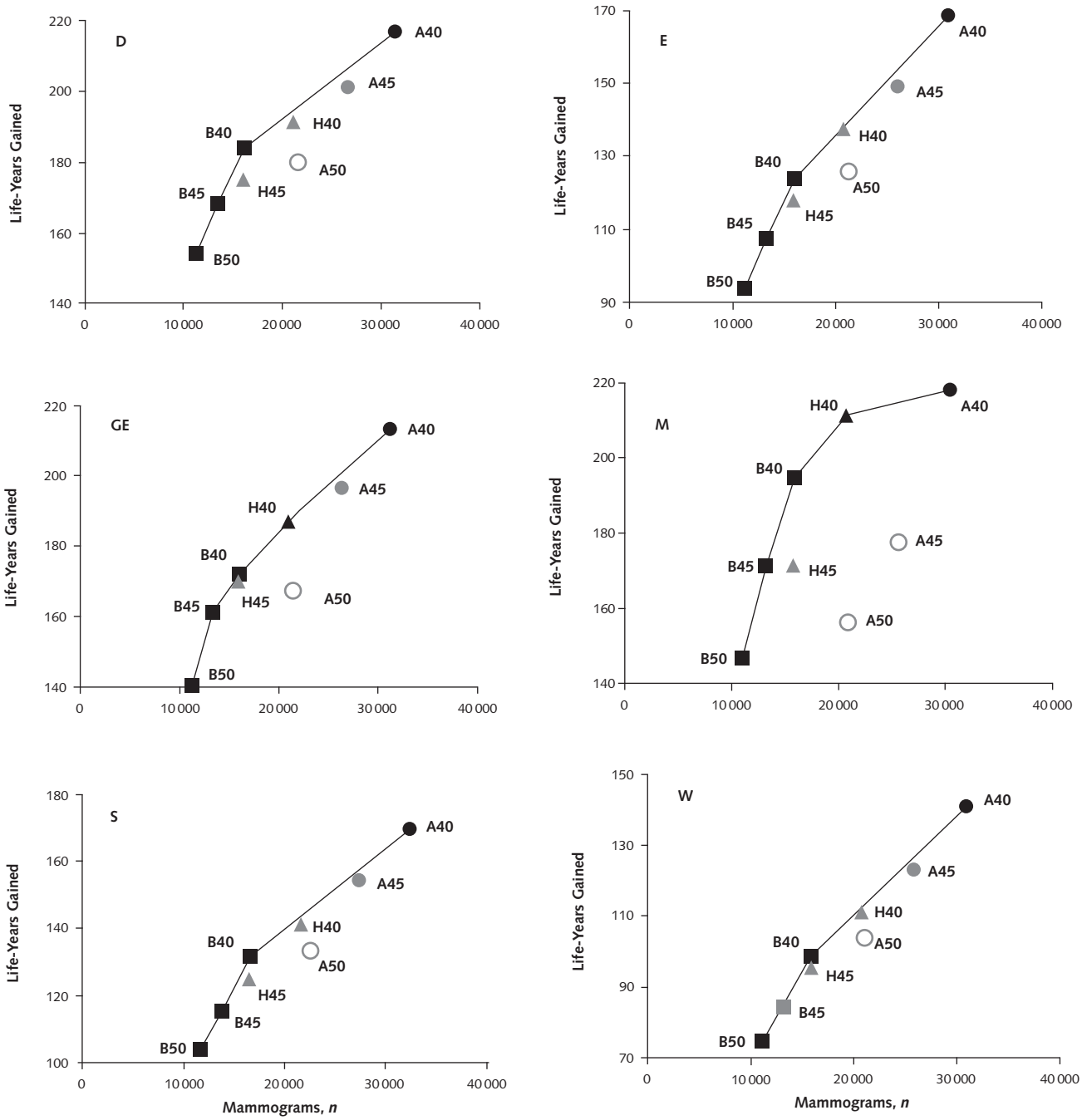
* Model M does not include a natural history component. On the basis of a combination of assumptions about underlying incidence trends in the absence of screening, it essentially yields a long lead time for invasive cancer; thus, all cancers found with annual screening can also be detected with biennial screening.

† Percentage of reduction with annual screening in women aged 50–74 y that is maintained by biennial screening in women aged 50–74 y is calculated as the percent mortality reduction with biennial screening in women aged 50–74 y divided by the percent mortality reduction with annual screening in women aged 50–74 y.

‡ Percentage of reduction with annual screening in women aged 45–74 y that is maintained by biennial screening in women aged 45–74 y is calculated as the percent mortality reduction with biennial screening in women aged 45–74 y divided by the percent mortality reduction with annual screening in women aged 45–74 y.

§ Percentage of reduction with annual screening in women aged 40–74 y that is maintained by biennial screening in women aged 40–74 y is calculated as the percent mortality reduction with biennial screening in women aged 40–74 y divided by the percent mortality reduction with annual screening in women aged 40–74 y.

Appendix Figure 2. Efficiency frontier for life-years gained versus mammograms performed for each screening strategy, by model.



The average gain in life-years per additional mammogram performed per 1000 women for each screening strategy (vs. no screening). Biennial strategies are indicated with a square; hybrid strategies (annual in the 40s followed by biennial from 50 to 74 years of age) with a triangle; and annual strategies with a circle. Efficient strategies were plotted (those in which increases in mammography use resulted in greater life-years gained than the next least-intensive strategy). The line represents the "efficiency frontier" by joining efficient strategies in which increases in mammography use resulted in greater life-years gained than the next less intensive efficient strategy. Strategies on this line would be considered efficient because they achieve the greatest gain in benefit (life years gained) per harm or use of mammograms. Strategies that use more mammograms but still have small benefits (i.e., a shallower slope than the next best strategy) are considered to be less efficient (i.e., weakly dominated). When and if the slope in the efficiency frontier plot levels off, it means that 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 (or additional benefit). There is no definitive inflection point across the models for the strategies or metrics evaluated. Black strategies are efficient; gray strategies close to the efficiency frontier are less efficient; and open gray strategies are inefficient (inferior, or dominated). Reference 33 provides efficiency frontiers for other harm and benefit metrics. D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; GE = Georgetown University Medical Center and Albert Einstein College of Medicine; M = MD Anderson Cancer Center; QALY = quality-adjusted life-year; S = Stanford University; W = University of Wisconsin and Harvard Medical School.