Effects of Mammography Screening Under Different Screening Schedules: Model Estimates of Potential Benefits and Harms

Jeanne S. Mandelblatt, MD, MPH; Kathleen A. Cronin, PhD; Stephanie Bailey, PhD; Donald A. Berry, PhD; Harry J. de Koning, MD, PhD; Gerrit Draisma, PhD; Hui Huang, MS; Sandra J. Lee, DSc; Mark Munsell, MS; Sylvia K. Plevritis, PhD; Peter Ravdin, MD, PhD; Clyde B. Schechter, MD, MA; Bronislava Sigal, PhD; Michael A. Stoto, PhD; Natasha K. Stout, PhD; Nicolien T. van Ravesteyn, MSc; John Venier, MS; Marvin Zelen, PhD; and Eric J. Feuer, PhD; for the Breast Cancer Working Group of the Cancer Intervention and Surveillance Modeling Network (CISNET)*

Background: Despite trials of mammography and widespread use, optimal screening policy is controversial.

Objective: To evaluate U.S. breast cancer screening strategies.

Design: 6 models using common data elements.

Data Sources: National data on age-specific incidence, competing mortality, mammography characteristics, and treatment effects.

Target Population: A contemporary population cohort.

Time Horizon: Lifetime.

Perspective: Societal.

Interventions: 20 screening strategies with varying initiation and cessation ages applied annually or biennially.

Outcome Measures: Number of mammograms, reduction in deaths from breast cancer or life-years gained (vs. no screening), false-positive results, unnecessary biopsies, and overdiagnosis.

Results of Base-Case Analysis: The 6 models produced consistent rankings of screening strategies. Screening biennially maintained an average of 81% (range across strategies and models, 67% to 99%) of the benefit of annual screening with almost half the number of false-positive results. Screening biennially from ages 50 to 69 years achieved a median 16.5% (range, 15% to 23%) reduction in breast cancer deaths versus no screening. Initiating biennial screening at age 40 years (vs. 50 years) reduced mortality by an additional 3% (range, 1% to 6%), consumed more resources, and yielded more false-positive results. Biennial screening after age 69 years yielded some additional mortality reduction in all models, but overdiagnosis increased most substantially at older ages.

Results of Sensitivity Analysis: Varying test sensitivity or treatment patterns did not change conclusions.

Limitation: Results do not include morbidity from false-positive results, patient knowledge of earlier diagnosis, or unnecessary treatment.

Conclusion: Biennial screening achieves most of the benefit of annual screening with less harm. Decisions about the best strategy depend on program and individual objectives and the weight placed on benefits, harms, and resource considerations.

Primary Funding Source: National Cancer Institute.

Ann Intern Med. 2009;151:738-747. www.annals.org
For author affiliations, see end of text.

In 2009, an estimated 193 370 women in the United States will develop invasive breast cancer, and about 40 170 of them will die of this disease (1). Randomized trials of mammography (2–4) have demonstrated reductions in breast cancer mortality associated with screening from ages 50 to 74 years. Trial results for women aged 40 to 49 years and women aged 74 years or older were not conclusive, and the trials (4, 5) had some problems with design, conduct, and interpretation. However, it is not feasible to conduct additional trials to get more precise estimates of the mortality benefits from extending screening to women younger than 50 years or older than 74 years or to test different screening schedules.

We developed models of breast cancer incidence and mortality in the United States. These models are ideally suited for estimating the effect of screening under a variety of policies (6, 7). Modeling has the advantage of being able to hold selected conditions (for example, screening intervals or test sensitivity) constant, which facilitates comparison of strategies. Because all models make assumptions about unobservable events, use of several models provides a

See also:

Print
Editorial comment. 750
Related articles. 703, 716, 727
Summary for Patients. I-44

Web-Only
Appendix Tables
Appendix Figure
Conversion of graphics into slides

* This work was done by 6 independent modeling teams from Dana-Farber Cancer Institute (Dr. Lee, principal investigator); Erasmus University (Dr. de Koning, principal investigator); Georgetown University Medical Center, Lombardi Comprehensive Cancer Center (Dr. Mandelblatt, principal investigator); Harvard School of Public Health, Harvard Medical School, Harvard Pilgrim Health Care/University of Wisconsin (Dr. Stout, principal investigator); M.D. Anderson Comprehensive Cancer Center (Dr. Berry, principal investigator); and Stanford University (Dr. Plevritis, principal investigator). Drs. Mandelblatt and Cronin were the writing and coordinating committee for the project; all other collaborators are listed in alphabetical order. Dr. Feuer was responsible for overall CISNET project direction.

range of plausible effects and can illustrate the effects of differences in model assumptions (7).

We used 6 established models to estimate the outcomes across 20 mammography screening strategies that vary by age of initiation and cessation and by screening interval among a cohort of U.S. women. The results are intended to contribute to practice and guideline policy debates.

METHODS

The 6 models were developed independently within the Cancer Intervention and Surveillance Modeling Network (CISNET) of the National Cancer Institute (NCI) (7, 8) and were exempt from institutional review board approval. The models have been described elsewhere (7, 9–15). Briefly, they share common features and inputs but differ in some ways (**Appendix Table 1**, available at www.annals.org). Model E (Erasmus Medical Center, Rotterdam, the Netherlands), model G (Georgetown University Medical Center, Washington, DC, and Albert Einstein College of Medicine, Bronx, New York), model M (M.D. Anderson Cancer Center, Houston, Texas), and model W (University of Wisconsin, Madison, Wisconsin, and Harvard Medical School, Boston, Massachusetts) include ductal carcinoma in situ (DCIS). Models E and W specifically assume that some portions of DCIS are nonprogressive and do not result in death. Model W also assumes that some cases of small invasive cancer are nonprogressive. Model S (Stanford University, Palo Alto, California) and model D (Dana-Farber Cancer Institute, Boston, Massachusetts) include only invasive cancer. Some groups model breast cancer in stages, but 3 (models E, S, and W) use tumor size and tumor growth. The models also differ by whether treatment affects the hazard for death from breast cancer (models G, S, and D), results in a cure for some fraction of cases (models E and W), or both (model M). Despite these differences, in previous collaborations (7) all the models came to similar qualitative estimates of the relative contributions of screening and treatment to observed decreases in deaths from breast cancer.

Model Overview

We used the 6 models to estimate the benefits, resource use (as measured by number of mammograms), and harms of 20 alternative screening strategies varying by starting and stopping age and by interval (annual and biennial) (**Table 1**). The models begin with estimates of breast cancer incidence and mortality trends without screening and treatment and then overlay screening use and improvements in survival associated with treatment (7). We use a cohort of women born in 1960 and follow them beginning at age 25 years for their entire lives. Breast cancer is generally depicted as having a preclinical, screening-detectable period (sojourn time) and a clinical detection point. On the basis of mammography sensitivity (or thresholds of detection), screening identifies disease in the preclinical screening-detection period and results in the identification of earlier-stage or smaller tumors than might be identified by clinical detection, resulting in reduction in breast cancer mortality. Age, estrogen receptor status, and tumor size– or stage–specific treatment have independent effects on mortality. Women can die of breast cancer or of other causes.

Table 1. **Breast Cancer Screening Strategies***

No screening
Screen from age 40 to 69 y
Screen from age 40 to 79 y
Screen from age 40 to 84 y
Screen from age 45 to 69 y
Screen from age 50 to 69 y
Screen from age 50 to 74 y
Screen from age 50 to 79 y
Screen from age 50 to 84 y
Screen from age 55 to 69 y
Screen from age 60 to 69 y

* Each strategy was evaluated by using an annual or biennial schedule, for a total of 20 strategies; we include no screening for comparison.

Model Data Variables

All 6 modeling groups use a common set of age-specific variables for breast cancer incidence, mammography test characteristics, treatment algorithms and effects, and nonbreast cancer competing causes of death (**Appendix Table 2**, available at www.annals.org). In addition to these common variables, each model includes model-specific inputs (or intermediate outputs) to represent preclinical detectable times, lead time, dwell time within stages of disease, and stage distribution in unscreened versus screened women on the basis of their specific model structure (7, 9–15).

We use an age–period–cohort model to estimate what breast cancer incidence rates would have been without screening (16). This approach considers the effect of age, temporal trends in risk by cohort, and time period. Because we do not have data on future incidence of breast cancer, we extrapolate forward assuming that future age-specific incidence increases as women age, as observed in 2000. To isolate the effect of technical effectiveness of screening and to assess the effect of screening on mortality while holding treatment constant, models assume 100% adherence to screening and indicated treatment.

Three groups use the age-specific mammography sensitivity (and specificity) values observed in the Breast Cancer Surveillance Consortium (BCSC) program for detection of all cases of breast cancer (invasive and in situ). Separate values are used for initial and subsequent mammography performed at either annual or biennial intervals (17). Two of the models (D and G) use these data directly as input variables (10, 14), and 1 model (S) uses the data to calibrate the model (13). The other 3 models (E, M, and W) use the BCSC data as a guide and to fit sensitivity estimates from this and other sources (9, 11, 15).

All women who have estrogen receptor–positive invasive tumors receive hormonal treatment (tamoxifen if women aged <50 years at diagnosis and anastrozole if ≥50 years) and nonhormonal treatment with an anthracycline-based regimen. Women with estrogen receptor–negative invasive tumors receive nonhormonal therapy only. Women with DCIS who have estrogen receptor–positive tumors receive hormonal therapy only (18). Treatment effectiveness is based on a synthesis of recent clinical trials and is modeled as a proportionate reduction in mortality risk or the proportion cured (19, 20).

Benefits

We estimated the cumulative probability of unscreened women dying of breast cancer from age 40 years to death. Screening benefit is then calculated as the percentage of reduction in breast cancer mortality (vs. no screening). We also examined life-years gained because of averted or delayed breast cancer death. Benefits are cumulated over the lifetime of the cohort to capture reductions in breast cancer mortality (or life-years gained) occurring years after the start of screening, after considering non-breast cancer mortality (21, 22).

Harms

As measures of the burden that a regular screening program imposes on a population, 3 different potential screening harms were examined: false-positive mammograms, unnecessary biopsies, and overdiagnosis. We define the rate of false-positive mammograms as the number of mammograms read as abnormal or needing further follow-up in women without cancer divided by the total number of positive screening mammograms based on the specificity reported in the BCSC (17). We define unnecessary biopsies post hoc as the proportion of women with false-positive screening results who receive a biopsy (23). We define overdiagnosis as the proportion of cases in each strategy that would not have clinically surfaced in a woman's lifetime (because of lack of progressive potential or death from another cause) among all cases arising from age 40 years onward.

Base-Case Analysis

We compared model results for the 20 strategies to select the most efficient approach. In a decision analysis, we considered a new intervention more efficient than a comparison intervention if it results in gains in health outcomes, such as life-years gained or deaths averted, while consuming fewer resources (or costs). If the new intervention results in worse outcomes and requires a greater investment, it is inefficient and would not be considered for further use. In economic analysis, inefficient strategies are said to be "dominated" when this occurs. To rank the screening strategies, we first look at the results of each model independently. For a particular model, a strategy that requires more mammographies (our measure of resource use) but has a lower relative percentage of mortality reduction (or life-years gained) is considered inefficient or dominated by other strategies. To evaluate strategies on the basis of results from all 6 models together, we classify them as follows: If a strategy is dominated in all or in 5 of 6 of the models, we considered it dominated overall. If a strategy is not dominated in any of the models, we classified it as efficient. For a strategy with mixed results across the models, we classified it as borderline.

After all dominated strategies were eliminated, the remaining strategies were represented as points on a graph plotting the average number of mammograms versus the percentage of mortality reduction (or life-years gained) for each model. We obtained the efficiency frontier for each graph by identifying the sequence of points that represent the largest incremental gain in percentage of mortality reduction (or life-years gained) per additional screening mammography. Screening strategies that fall on this frontier are the most efficient (that is, no alternative exists that provides more benefit for fewer mammographies performed).

Sensitivity Analysis

We conducted a sensitivity analysis to see whether our conclusions about the ranking of strategies change when we vary input variables. First, we investigate the effect of assuming that mammography sensitivity for a given age, screening round, and screening interval is 10 percentage points less than that observed. Second, we examine whether ranking of strategies varies if treatment includes newer hormonal and nonhormonal adjuvant regimens (for example, taxanes). Third, because adjuvant therapy is unlikely to reach 100% of women as modeled in our base-case analysis, we reassess the ranking of strategies if we assume that actual observed current treatment patterns apply to the cohort (24).

Model Validation and Uncertainty

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 gold standard. For instance, because some models used results from screening trials (or SEER [Surveillance, Epidemiology and End Results] data) for calibration or as input variables, we cannot use comparisons of projected mortality reductions to trial results to validate all of the models. In addition, we cannot directly compare the results of this analysis, which uses 100% actual screening for all women at specified intervals, with screening trial results in which invitation to screening and participation varied. In our previous work (7, 9–11, 13–15), results of each model accurately projected independently estimated trends in the absence of intervention and closely approximated modern stage distributions and observed mortality trends. Overall, using 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.

Role of the Funding Source

This work was done under contracts from the Agency for Healthcare Research and Quality (AHRQ) and NCI and grants from the NCI. Staff from the NCI provided some data and technical assistance, and AHRQ staff reviewed the manuscript. Model results are the sole responsibility of the investigators.

Results

In an unscreened population, the models predict a cumulative probability of breast cancer developing over a woman's lifetime starting at age 40 years ranging from 12% to 15%. Without screening, the median probability of dying of breast cancer after age 40 years is 3.0% across the 6 models. Thus, if a particular screening strategy leads to a 10% reduction in breast cancer mortality, then the probability of breast cancer mortality would be reduced from 3.0% to 2.7%, or 3 deaths averted per 1000 women screened.

Benefits

The 6 models produce consistent results on the ranking of the strategies (**Appendix Table 3**, available at www.annals.org). Eight approaches are "efficient" in all models (that is, not dominated, because they provide additional mortality reductions for added use of mammography); 7 of these have a biennial interval, and all but 2 start at age 50 years. The **Figure** shows these results, and again we see that most strategies on the efficiency frontier have a biennial interval. Screening every other year from ages 50 to 69 years is an efficient strategy for reducing breast cancer mortality in all models. In all models, biennial screening starting at age 50 years and continuing through ages 74, 79, or 84 years are of fairly similar efficiency.

In examining benefits in terms of life-years gained (**Appendix Table 4**, available at www.annals.org), 6 of the 8 consistently nondominated strategies have a biennial interval. In contrast to results for mortality reduction, half of the nondominated strategies include screening initiation at age 40 years. Annual screening strategies that include screening until age 79 or 84 years are on the efficiency frontier (**Appendix Figure**, available at www.annals.org), but are less resource-efficient than biennial approaches for increasing life-years gained.

As another way to examine the effect of screening interval, we calculated for each screening strategy and model the proportion of the annual benefit (in terms of mortality reduction) that could be achieved by biennial screening (**Table 2**). Biennial screening maintains an average of 81% (range across strategies and models, 67% to 99%) of the benefits achieved by annual screening.

We also examined the incremental benefits gained by extending screening from ages 50 to 69 years to either earlier or later ages of initiation and cessation (**Table 3**). Continuing screening to age 79 years (vs. 69 years) results in a median increase in percentage of mortality reduction of 8% (range, 7% to 11%) and 7% (range, 6% to 10%) under annual and biennial intervals, respectively. If screening begins at age 40 years (vs. 50 years) and continues to age 69 years, all models project additional, albeit small, reductions in breast cancer mortality (3% median reduction with either annual or biennial intervals) (**Table 3**). This translates into a median of 1 additional breast cancer death averted (range, 1 to 2 deaths) per 1000 women screened under a strategy of annual screening from age 40 to 69 years (vs. 50 to 69 years). Thus, greater mortality reductions could be achieved by stopping screening at an older age than by initiating screening at an earlier age.

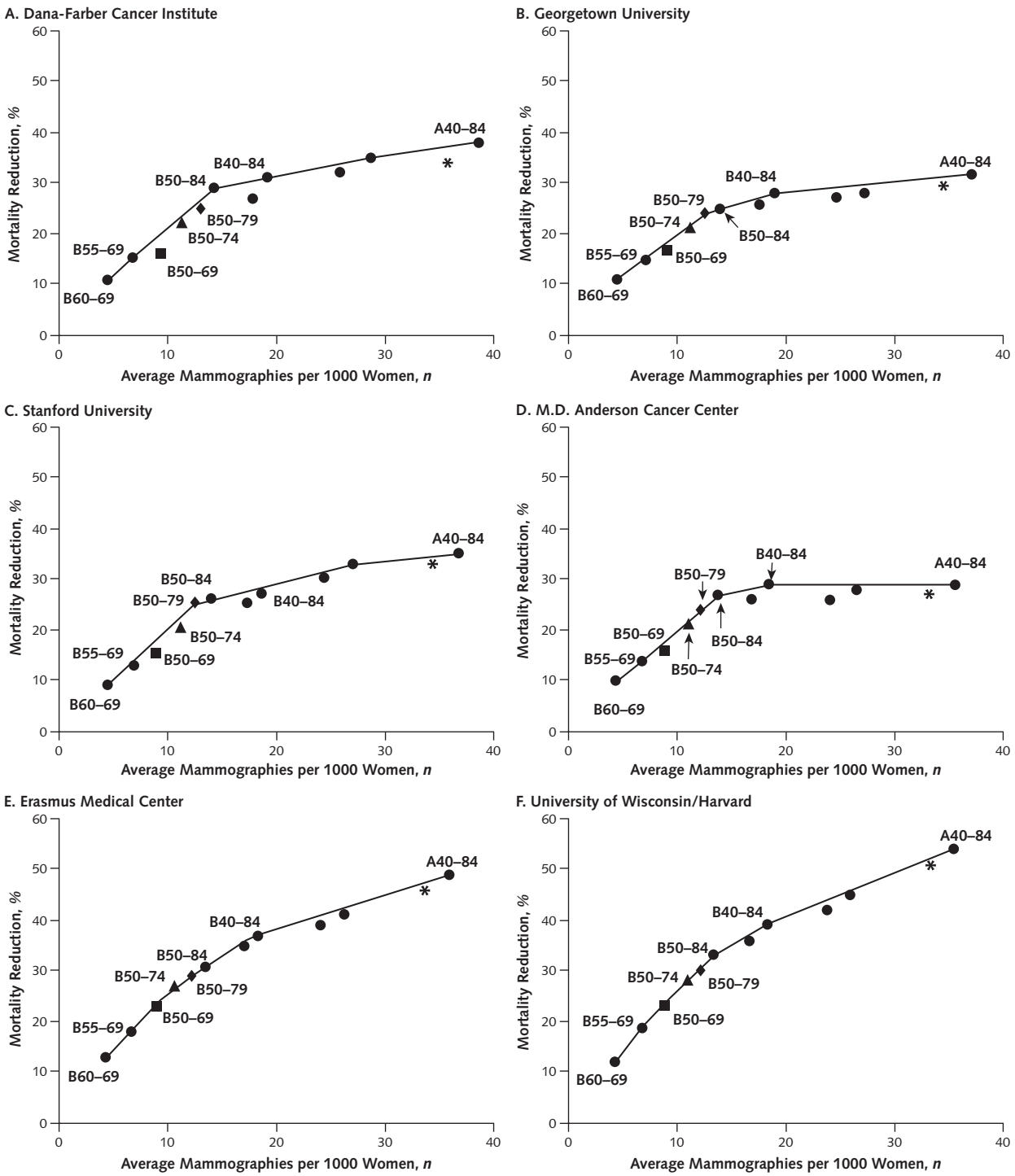
However, when life-years gained is the outcome measure, 3 of the models conclude that benefits are greater from extending screening to the younger rather than the older age group (**Table 3**). For instance, starting annual screening at age 40 years (vs. 50 years) and continuing annually to age 69 years yields a median of 33 (range, 11 to 58) life-years gained per 1000 women screened, whereas extending annual screening to age 79 years (vs. 69 years) yields a median of only 24 (range, 18 to 38) life-years gained per 1000 women screened.

Harms

All the models project similar rates of false-positive mammograms over the lifetime of screened women across the screening strategies; **Table 4** summarizes results for an exemplar model. More false-positive results occur in strategies that include screening from ages 40 to 49 years than in those that initiate screening at age 50 years or later and those that include annual screening rather than biennial screening. For instance, annual screening from ages 40 to 69 years yields 2250 false-positive results for every 1000 women screened over this period, almost twice as many as that of biennial screening in this age group. The proportion of biopsies that occur because of these false-positive results that are retrospectively deemed unnecessary (that is, the woman did not have cancer) is about 7%; therefore, many more women will undergo unnecessary biopsies under annual screening than biennial screening.

Of the 6 models, 5 estimated rates of overdiagnosis. They showed an increase in the risk for overdiagnosis as age increases (data not shown). Although the increase with age occurs over the entire age range considered in the different screening strategies, the rate of increase accelerates in the older age groups, mostly because of increasing rates of competing causes of mortality. Rates of overdiagnosis were higher for DCIS than for invasive disease, proportionately affecting younger women more because more cases of DCIS are diagnosed at younger ages. However, overall, initiating screening at age 40 years (vs. 50 years) had a smaller effect on overdiagnosis than did extending screening beyond age 69 years. Biennial strategies decrease the rate of overdiagnosis, but by much less than one half. The absolute estimate of overdiagnosis varied between models depending on whether DCIS was or was not included and

Figure. **Percentage of breast cancer mortality reduction versus number of mammographies performed per 1000 women, by model and screening strategy.**



The panels show an efficiency frontier graph for each model. The graph plots the average number of mammographies performed per 1000 women against the percentage of mortality reduction for each screening strategy (vs. no screening). Strategies are denoted as annual (*A*) or biennial (*B*) with starting and stopping ages. We plot efficient strategies (that is, those in which increases in use of mammography resources result in greater mortality reduction than the next least-intensive strategy) in all 6 models. We also plot "borderline" strategies (approaches that are efficient in some models but not others). The line between strategies represents the "efficiency frontier." Strategies on this line would be considered efficient because they achieve the greatest gain per use of mammography resources compared with the point (or strategy) immediately below it. Points that fall below the line are not considered as efficient as those on the line. When the slope in the efficiency frontier plot levels off, the additional reductions in mortality per unit increase in use of mammography are small relative to the previous strategies and could indicate a point at which additional investment (use of screening) might be considered as having a low return (benefit).

Table 2. **Percentage of Reduction in Breast Cancer Mortality Maintained When Moving From an Annual Screening Interval to a Biennial Interval, by Screening Strategy and Model**

Model*	Maintained Reduction in Breast Cancer Mortality, by Screening Strategy, %†									
	Ages 50–69 y	Ages 40–69 y	Ages 45–69 y	Ages 40–79 y	Ages 40–84 y	Ages 55–69 y	Ages 60–69 y	Ages 50–74 y	Ages 50–79 y	Ages 50–84 y
D	76	75	78	79	82	83	79	81	78	83
E	75	73	74	75	75	75	73	76	75	76
G	85	86	91	87	88	91	86	89	88	89
M	90	96	97	97	99	92	84	95	93	95
S	74	73	78	76	77	80	74	79	85	79
W	68	67	70	70	71	71	70	72	70	73

* Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S = Stanford University; W = University of Wisconsin/Harvard.
† Differences in the range of results reflect differences in modeling approaches. For example, the benefit of screening in model M is modeled through stage shift, as with most other models, but also includes a "beyond stage shift" factor based on a cure fraction for small tumors. However, because many of these "cures" occur among women with invasive cancer that is not fatal, finding such cancer 1 year earlier confers very little mortality advantage to annual (vs. biennial) screening.

on the assumptions related to progression of DCIS and invasive disease, reflecting the uncertainty in the current knowledge base.

Sensitivity Analysis

The overall conclusions are robust across the 6 models under different assumptions about mammography sensitivity, treatment patterns, and treatment effectiveness (data not shown).

DISCUSSION

This study uses 6 established models that use common inputs but different approaches and assumptions to extend previous randomized mammography screening trial results to the U.S. population and to age groups in whom trial results are less conclusive. All 6 modeling groups concluded that the most efficient screening strategies are those that include a biennial screening interval. Conclusions about the optimal starting ages for screening depend more on the measure chosen for evaluating outcomes. If the goal of a national screening program is to reduce mortality in the most efficient manner, then programs that screen biennially from age 50 years to age 69, 74, or 79 years are among the most efficient on the basis of the ratio of benefits to the number of screening examinations. If the goal of a screening program is to efficiently maximize the number of life-years gained, then the preferred strategy would be to screen biennially starting at age 40 years. Decisions about the best starting and stopping ages also depend on tolerance for false-positive results and rates of overdiagnosis.

The conclusion of this modeling analysis—that biennial intervals are more efficient and provide a better balance of benefits and harms than annual intervals—is contrary to some current practices in the United States (25–27). However, our result that biennial screening is

Table 3. **Incremental Changes in Percentage of Reduction in Breast Cancer Mortality and Life-Years Gained per 1000 Women, by Age of Screening Initiation and Cessation**

Model*	Start at Age 40 y vs. 50 y†						Stop at Age 79 y vs. 69 y‡					
	Difference in Percentage of Reduction in Breast Cancer Mortality		Difference in Breast Cancer Deaths Averted per 1000 Women		Difference in Life-Years Gained per 1000 Women		Difference in Percentage of Reduction in Breast Cancer Mortality		Difference in Breast Cancer Deaths Averted per 1000 Women		Difference in Life-Years Gained per 1000 Women	
	Annual	Biennial	Annual	Biennial	Annual	Biennial	Annual	Biennial	Annual	Biennial	Annual	Biennial
D	3	2	1	1	25	20	11	9	3	3	28	26
E	8	5	2	1	58	40	8	6	2	2	18	15
G	3	3	1	1	34	29	7	7	2	2	27	25
M	2	3	1	1	11	18	7	7	2	2	21	21
S	2	1	1	1	32	21	10	10	4	4	38	31
W	10	6	2	1	57	37	8	6	2	1	19	15
Median across models	3	3	1	1	33	25	8	7	2	2	24	23.5

* Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S = Stanford University; W = University of Wisconsin/Harvard.
† Incremental difference between screening from 40 to 69 y versus 50 to 69 y.
‡ Incremental difference between screening from 50 to 79 y versus 50 to 69 y.

Table 4. **Benefits and Harms Comparison of Different Starting and Stopping Ages Using the Exemplar Model***

Strategy	Average Screenings per 1000 Women	Potential Benefits (vs. No Screening)			Potential Harms (vs. No Screening)†	
		Percentage of Mortality Reduction	Cancer Deaths Averted per 1000 Women	Life-Years Gained per 1000 Women	False-Positive Results per 1000 Women	Unnecessary Biopsies per 1000 Women
Comparison of different starting ages						
Biennial screening						
40–69 y	13 865	16‡	6.1	120‡	1250	88
45–69 y	11 771	17‡	6.2	116‡	1050	74
50–69 y	8944	15	5.4	99	780	55
55–69 y	6941	13	4.9	80	590	41
60–69 y	4246	9	3.4	52	340	24
Annual screening						
40–69 y	27 583	22‡	8.3	164‡	2250	158
45–69 y	22 623	22‡	8.0	152‡	1800	126
50–69 y	17 759	20‡	7.3	132‡	1350	95
55–69 y	13 003	16‡	6.1	102‡	950	67
60–69 y	8406	12‡	4.6	69‡	600	42
Comparison of different stopping ages						
Biennial						
50–69 y	8944	15	5.4	99	780	55
50–74 y	11 109	20	7.5	121	940	66
50–79 y	12 347	25	9.4	130	1020	71
50–84 y	13 836	26	9.6	138	1130	79
Annual						
50–69 y	17 759	20‡	7.3	132‡	1350	95
50–74 y	21 357	26‡	9.5	156‡	1570	110
50–79 y	24 439	30	11.1	170	1740	122
50–84 y	26 913	33	12.2	178	1880	132

* Results are from model S (Stanford University). Model S was chosen as an exemplar model to summarize the balance of benefits and harms associated with screening 1000 women under a particular screening strategy.
† Overdiagnosis is another significant harm associated with screening. However, given the uncertainty in the knowledge base about ductal carcinoma in situ and small invasive tumors, we felt that the absolute estimates are not reliable. In general, overdiagnosis increases with age across all age groups but increases more sharply for women who are screened in their 70s and 80s.
‡ Strategy is dominated by other strategies; the strategy that dominates may not be in this table.

more efficient than annual screening is consistent with previous modeling research (28–32) and screening trials, most of which used 2-year intervals (2–5). The model results also agree with reports showing similar intermediate cancer outcomes (for example, stage distribution) between programs using annual and biennial screening, especially among women aged 50 years or older (33–37). In addition, we demonstrated substantial increases in false-positive results and unnecessary biopsies associated with annual intervals, and these harms are reduced by almost 50% with biennial intervals. Our results are also consistent with current knowledge of disease biology. Slow-growing tumors are much more common than fast-growing tumors, and the ratio of slow- to fast-growing tumors increases with age, (38) so that little survival benefit is lost between screening every year versus every other year. For the small subset of women with aggressive, fast-growing tumors, even annual screening is not likely to confer a survival advantage. Guidelines in other countries (4) include biennial screening. However, whether it will be practical or acceptable to change the existing U.S. practice of annual screening cannot be addressed by our models.

In all models, some reductions in breast cancer mortality, albeit small, were seen with strategies that started screening at age 40 years versus 50 years. Because models can represent millions of observations, they are well-suited to detect small differences in a group over time that might not be seen in even the largest clinical trial with a 10- to 15-year follow-up (4, 39–42). If program benefits are measured in life-years, the measure most commonly used in cost-effectiveness analysis, then our results suggest that initiating screening at age 40 years saves more life-years than extending screening past age 69 years (albeit at the cost of increasing the number of false-positive mammograms).

Previous recommendations on breast cancer screening have suggested an upper age limit for screening cessation because of decreasing program efficiency due to competing mortality (26, 43). Our result that screening strategies that include an upper age limit beyond age 69 years remain on the efficiency frontier (albeit with low incremental gains over strategies that stop screening at earlier ages and with greater harms) is consistent with previously reported results of screening benefit from observational and modeled data (31, 32, 44–47). However, the observational data reports may have been confounded by the inability to capture lead time and length biases (48–50). Any benefits of screening older women must be balanced against possible harms. For instance, the probability of overdiagnosis increases with age

and increases more dramatically for the oldest age groups. Model estimates for the oldest age groups also have more uncertainty compared with estimates for ages 50 to 74 years because of the lack of primary data on natural history of breast cancer and the absence of screening trial data after age 74 years. With the demographic pressure of an aging society, more research will be needed to fully understand the natural history of this disease and the balance of risks and benefits of screening and treatment in the older age groups (38, 50).

Our results also highlight the need for better primary data on the natural history of DCIS and small invasive cancer to draw reliable conclusions on the absolute magnitude of overdiagnosis associated with different screening schedules (37, 51). Clinical investigation (52), follow-up in screening trials (53), epidemiologic trends in incidence (54), and previous modeling efforts (9, 55) all indicated that some DCIS cases will not progress (56, 57), but how many is not known.

The collaboration of 6 groups with different modeling philosophies and approaches to estimate the same end points by using a common set of data provides an excellent opportunity to cross-replicate data generated from modeling, represent uncertainty related to modeling assumptions and structure, and give insight into which results are consistent across modeling approaches and which are dependent on model assumptions. The resulting conclusions about the ranking of screening strategies were very robust and should provide greater credibility than inferences based on 1 model alone.

Despite our consistent results, our study had some limitations (58). First, our models provide estimates of the average benefits and harms expected across a cohort of women and do not reflect personal data for individual women. Also, although our models project mortality reductions similar to those observed in clinical trials, the range of results includes higher mortality reductions than that achieved in the trials because we model lifetime screening and assume adherence to all screening and treatment. The trials followed women for limited numbers of years and have some nonadherence. The models also do not capture differences in outcomes among certain risk subgroups, such as women with *BRCA1* or *BRCA2* genetic susceptibility mutations, women who are healthier or sicker than average, or black women who seem to have more disease at younger ages than white women (59).

Second, the outcomes considered do not capture morbidity associated with surgery for screening-detected disease (60) or decrements in quality of life associated with false-positive results, living with earlier knowledge of a cancer diagnosis, or overdiagnosis (61).

Third, in estimating lifetime results, we projected breast cancer trends from background incidence rates of a 1960 birth cohort extrapolated forward in time. However, future background incidence (and mortality) may change as the result of several different forces, such as changes in patterns of reproduction; less use of hormone replacement therapy after 2002 or prescription of tamoxifen or other agents for primary disease prevention; increasing rates of obesity; and further advances in treatment (for example, trastuzumab) (62). Although most models portray known differences in biology by age (for example, distribution of estrogen receptor–positive tumors, sensitivity of screening, and length of the preclinical sojourn times), some aspects of the natural history of disease are not known or cannot be fully captured.

We assumed 100% adherence to screening and treatment to evaluate program efficacy. Benefits will always fall short of the projected results because adherence is not perfect. If actual adherence varies systematically by age or other factors, the ranking of strategies could change. In addition, we did not consider "mixed" strategies (for example, screening annually from age 40 to 49 years and then biennially from age 50 to 79 years) as was done in some trials (5) and other analyses (36, 63). We found that the benefits of screening from ages 40 to 49 years were small. Benefits in this age group were also associated with harms in terms of false-positive results and unnecessary biopsies. Thus, although strategies that include annual screening from ages 40 to 49 years might be efficient, this would be largely driven by the more favorable balance of benefits and harms after age 50 years. In addition, we judged that mixed strategies are very difficult to communicate to consumers and implement in public health practice.

Finally, we did not discount benefits or include costs in our analysis, although the average number of mammograms per woman (and false-positive results) provides some proxy of resource consumption. Even with these acknowledged limitations, the models demonstrate meaningful, qualitatively similar outcomes despite variations in structure and assumptions.

Overall, the evaluation of screening strategies by the 6 models suggests that optimal program design is based on biennial intervals. Choices about optimal ages of initiation and cessation will ultimately depend on program goals, resources, weight attached to the presence of trial data, the balance of harms and benefits, and considerations of efficiency and equity.

From the Georgetown University Medical Center and Lombardi Comprehensive Cancer Center, Washington, DC; National Institutes of Health, Bethesda, Maryland; University of Texas M.D. Anderson Cancer Center, Houston, Texas; Erasmus University Medical Center, Rotterdam, the Netherlands; Dana-Farber Cancer Institute, Harvard School of Public Health, Harvard Medical School, and Harvard Pilgrim Health Care, Boston, Massachusetts; Stanford University, Palo Alto, California; and Albert Einstein College of Medicine, New York, New York.

Acknowledgment: The authors thank the BCSC investigators, participating mammography facilities, and radiologists for the data they provided that were used to inform some of our model data input variables. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes is at http://breastscreening.cancer.gov/. The

authors also thank Mary Barton, MD, MPP, and William Lawrence, MD, MSc, from AHRQ; members of the U.S. Preventive Services Task Force; the Oregon Evidence-based Practice Center; Ann Zauber, PhD; and Karla Kerlikowske, MD, for helpful comments and review of earlier versions of this article. The authors thank Jackie Ford and Aimee Near for manuscript preparation.

Grant Support: By NCI cooperative agreements (2U01CA088270, 2U01CA088283, 2U01CA088248, and F32 CA125984). Portions of this work were performed under contract HHSN261200800769P. Data collection in the BCSC was supported by NCI-funded BCSC cooperative agreements (U01CA63740, U01CA86076, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, and U01CA70040) and several U.S. state public health departments and cancer registries. CISNET data management and Web site support were provided by Cornerstone Systems Northwest (NCI contract HHSN261200800002C).

Potential Conflicts of Interest: None disclosed.

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

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

References

1. **Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ.** Cancer statistics, 2009. CA Cancer J Clin. 2009;59:225-49. [PMID: 19474385]
2. **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]
3. **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]
4. **Vainio H, Bianchini F, eds.** Breast Cancer Screening. International Agency for Research on Cancer Handbook on Cancer Prevention, Report No. 7. Lyon, France: International Agency for Research on Cancer; 2002.
5. **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]
6. **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]
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. **Cancer Intervention and Surveillance Modeling Network.** Accessed at http://cisnet.cancer.gov/breast/profiles.html on 15 September 2008.
9. **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]
10. **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]
11. **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]
12. **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]
13. **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]
14. **Lee S, Zelen M.** A stochastic model for predicting the mortality of breast cancer. J Natl Cancer Inst Monogr. 2006:79-86. [PMID: 17032897]
15. **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]
16. **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]
17. **Breast Cancer Surveillance Consortium.** Performance Measures for 3,884,059 Screening Mammography Examinations from 1996 to 2007 by Age & Time (Months) Since Previous Mammography. Accessed at http://breastscreening.cancer.gov/data/performance/screening/perf_age_time.html on 7 October 2009.
18. **National Comprehensive Cancer Network.** NCCN Clinical Practice guidelines in oncology v.2.2008. Accessed at www.nccn.org/professionals/physician_gls/f_guidelines.asp on 22 September 2009.
19. **Clarke M, Coates AS, Darby SC, Davies C, Gelber RD, Godwin J, et al; Early Breast Cancer Trialists' Collaborative Group (EBCTCG).** Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials. Lancet. 2008;371:29-40. [PMID: 18177773]
20. **Early Breast Cancer Trialists' Collaborative Group (EBCTCG).** Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;365:1687-717. [PMID: 15894097]
21. **Rosenberg MA.** Competing risks to breast cancer mortality. J Natl Cancer Inst Monogr. 2006:15-9. [PMID: 17032889]
22. **Cronin KA, Feuer EJ, Clarke LD, Plevritis SK.** Impact of adjuvant therapy and mammography on U.S. mortality from 1975 to 2000: comparison of mortality results from the CISNET breast cancer base case analysis. J Natl Cancer Inst Monogr. 2006:112-21. [PMID: 17032901]
23. **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]
24. **Mariotto AB, Feuer EJ, Harlan LC, Abrams J.** Dissemination of adjuvant multiagent chemotherapy and tamoxifen for breast cancer in the United States using estrogen receptor information: 1975-1999. J Natl Cancer Inst Monogr. 2006:7-15. [PMID: 17032888]
25. **Smith RA, Saslow D, Sawyer KA, Burke W, Costanza ME, Evans WP 3rd, et al; American Cancer Society High-Risk Work Group.** American Cancer Society guidelines for breast cancer screening: update 2003. CA Cancer J Clin. 2003;53:141-69. [PMID: 12809408]
26. **National Cancer Institute.** NCI Statement on Mammography Screening [press release]. Bethesda, MD: National Cancer Institute; 31 January 2002. Accessed at www.cancer.gov/newscenter/mammstatement31jan02 on 22 September 2009.
27. Preventive Services: Breast Cancer Screening. Accessed at www.medicare.gov/Health/Mammography.asp on 22 September 2009.
28. **Salzmann P, Kerlikowske K, Phillips K.** Cost-effectiveness of extending screening mammography guidelines to include women 40 to 49 years of age. Ann Intern Med. 1997;127:955-65. [PMID: 9412300]
29. **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]
30. **Lee S, Huang H, Zelen M.** Early detection of disease and scheduling of screening examinations. Stat Methods Med Res. 2004;13:443-56. [PMID: 15587433]
31. **Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Extermann M, et al; Breast Cancer in Older Women Research Consortium.** Toward optimal screening strategies for older women. Costs, benefits, and harms of breast cancer screening by age, biology, and health status. J Gen Intern Med. 2005;20:487-96. [PMID: 15987322]
32. **Kerlikowske K, Salzmann P, Phillips KA, Cauley JA, Cummings SR.** Con-

tinuing screening mammography in women aged 70 to 79 years: impact on life expectancy and cost-effectiveness. JAMA. 1999;282:2156-63. [PMID: 10591338]

33. **Hofvind S, Vacek PM, Skelly J, Weaver DL, Geller BM.** Comparing screening mammography for early breast cancer detection in Vermont and Norway. J Natl Cancer Inst. 2008;100:1082-91. [PMID: 18664650]

34. **Smith-Bindman R, Chu PW, Miglioretti DL, Sickles EA, Blanks R, Ballard-Barbash R, et al.** Comparison of screening mammography in the United States and the United kingdom. JAMA. 2003;290:2129-37. [PMID: 14570948]

35. **Smith-Bindman R, Ballard-Barbash R, Miglioretti DL, Patnick J, Kerlikowske K.** Comparing the performance of mammography screening in the USA and the UK. J Med Screen. 2005;12:50-4. [PMID: 15814020]

36. **White E, Miglioretti DL, Yankaskas BC, Geller BM, Rosenberg RD, Kerlikowske K, et al.** Biennial versus annual mammography and the risk of late-stage breast cancer. J Natl Cancer Inst. 2004;96:1832-9. [PMID: 15601639]

37. **Wai ES, D'yachkova Y, Olivotto IA, Tyldesley S, Phillips N, Warren LJ, et al.** Comparison of 1- and 2-year screening intervals for women undergoing screening mammography. Br J Cancer. 2005;92:961-6. [PMID: 15714210]

38. **Fracheboud J, Groenewoud JH, Boer R, Draisma G, de Bruijn AE, Verbeek AL, et al.** Seventy-five years is an appropriate upper age limit for population-based mammography screening. Int J Cancer. 2006;118:2020-5. [PMID: 16287064]

39. **Miller AB, To T, Baines CJ, Wall C.** The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. Ann Intern Med. 2002;137:305-12. [PMID: 12204013]

40. **Elmore JG, Armstrong K, Lehman CD, Fletcher SW.** Screening for breast cancer. JAMA. 2005;293:1245-56. [PMID: 15755947]

41. **Elmore JG, Reisch LM, Barton MB, Barlow WE, Rolnick S, Harris EL, et al.** Efficacy of breast cancer screening in the community according to risk level. J Natl Cancer Inst. 2005;97:1035-43. [PMID: 16030301]

42. **Norman SA, Russell Localio A, Weber AL, Coates RJ, Zhou L, Bernstein L, et al.** Protection of mammography screening against death from breast cancer in women aged 40-64 years. Cancer Causes Control. 2007;18:909-18. [PMID: 17665313]

43. **U.S. Preventive Services Task Force.** Screening for breast cancer: recommendations and rationale. Ann Intern Med. 2002;137:344-6. [PMID: 12204019]

44. **McCarthy EP, Burns RB, Freund KM, Ash AS, Shwartz M, Marwill SL, et al.** Mammography use, breast cancer stage at diagnosis, and survival among older women. J Am Geriatr Soc. 2000;48:1226-33. [PMID: 11037009]

45. **Lash TL, Fox MP, Buist DS, Wei F, Field TS, Frost FJ, et al.** Mammography surveillance and mortality in older breast cancer survivors. J Clin Oncol. 2007;25:3001-6. [PMID: 17548838]

46. **Badgwell BD, Giordano SH, Duan ZZ, Fang S, Bedrosian I, Kuerer HM, et al.** Mammography before diagnosis among women age 80 years and older with breast cancer. J Clin Oncol. 2008;26:2482-8. [PMID: 18427152]

47. **Boer R, de Koning HJ, van Oortmarssen GJ, van der Maas PJ.** In search of the best upper age limit for breast cancer screening [Abstract]. Eur J Cancer. 1995;31A:2040-3. [PMID: 8562162]

48. **Berry DA, Baines CJ, Baum M, Dickersin K, Fletcher SW, Gøtzsche PC, et al.** Flawed inferences about screening mammography's benefit based on observational data [Letter]. J Clin Oncol. 2009;27:639-40; author reply 641-2. [PMID: 19075270]

49. **Schonberg MA, McCarthy EP.** Mammography screening among women age 80 years and older: consider the risks [Letter]. J Clin Oncol. 2009;27:640-1; author reply 641-2. [PMID: 19075269]

50. **Mandelblatt JS, Silliman R.** Hanging in the balance: making decisions about the benefits and harms of breast cancer screening among the oldest old without a safety net of scientific evidence [Editorial]. J Clin Oncol. 2009;27:487-90. [PMID: 19075258]

51. **Bryan BB, Schnitt SJ, Collins LC.** Ductal carcinoma in situ with basal-like phenotype: a possible precursor to invasive basal-like breast cancer. Mod Pathol. 2006;19:617-21. [PMID: 16528377]

52. **Kerlikowske K, Molinaro A, Cha I, Ljung BM, Ernster VL, Stewart K, et al.** Characteristics associated with recurrence among women with ductal carcinoma in situ treated by lumpectomy. J Natl Cancer Inst. 2003;95:1692-702. [PMID: 14625260]

53. **Moss S.** Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled trials of breast cancer screening. Breast Cancer Res. 2005; 7:230-4. [PMID: 16168145]

54. **Feuer EJ, Etzioni R, Cronin KA, Mariotto A.** The use of modeling to understand the impact of screening on U.S. mortality: examples from mammography and PSA testing. Stat Methods Med Res. 2004;13:421-42. [PMID: 15587432]

55. **de Koning HJ, Draisma G, Fracheboud J, de Bruijn A.** Overdiagnosis and overtreatment of breast cancer: microsimulation modelling estimates based on observed screen and clinical data. Breast Cancer Res. 2006;8:202. [PMID: 16524452]

56. **Burstein HJ, Polyak K, Wong JS, Lester SC, Kaelin CM.** Ductal carcinoma in situ of the breast. N Engl J Med. 2004;350:1430-41. [PMID: 15070793]

57. **Jones JL.** Overdiagnosis and overtreatment of breast cancer: progression of ductal carcinoma in situ: the pathological perspective. Breast Cancer Res. 2006; 8:204. [PMID: 16677423]

58. **Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al; ISPOR Task Force on Good Research Practices—Modeling Studies.** Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value Health. 2003;6:9-17. [PMID: 12535234]

59. **Mandelblatt JS, Liang W, Sheppard VB, Wang J, Isaacs C.** Breast cancer in minority women. In: Harris J, Lippman M, Morrow M, Osborne CK, eds. Diseases of the Breast. 4th ed. Philadelphia: Lippincott Williams & Wilkin; 2009.

60. **El-Tamer MB, Ward BM, Schifftner T, Neumayer L, Khuri S, Henderson W.** Morbidity and mortality following breast cancer surgery in women: national benchmarks for standards of care. Ann Surg. 2007;245:665-71. [PMID: 17457156]

61. **Bonomi AE, Boudreau DM, Fishman PA, Ludman E, Mohelnitzky A, Cannon EA, et al.** Quality of life valuations of mammography screening. Qual Life Res. 2008;17:801-14. [PMID: 18491217]

62. **Ravdin PM, Cronin KA, Howlader N, Berg CD, Chlebowski RT, Feuer EJ, et al.** The decrease in breast-cancer incidence in 2003 in the United States. N Engl J Med. 2007;356:1670-4. [PMID: 17442911]

63. **Buist DS, Porter PL, Lehman C, Taplin SH, White E.** Factors contributing to mammography failure in women aged 40-49 years. J Natl Cancer Inst. 2004; 96:1432-40. [PMID: 15467032]

Current Author Addresses: Dr. Mandelblatt: Lombardi Comprehensive Cancer Center, 3300 Whitehaven Street, Northwest, Suite 4100, Washington, DC 20007.
Dr. Cronin: National Cancer Institute, 6116 Executive Boulevard, Suite 504, Bethesda, MD 20892.
Dr. Bailey: 8666 Macawa Avenue, San Diego, CA 92123.
Dr. Berry: The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 1409, Houston, TX 77030.
Dr. de Koning: Erasmus University Medical Center, Dr. Molewaterplein 50, Rotterdam 3015 GE, the Netherlands.
Dr. Draisma: Department of Public Health, Room AE-235, Erasmus University Medical Center, PO Box 2040, 3000 CA Rotterdam, the Netherlands.
Mr. Huang: Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 01720.
Dr. Lee: Dana-Farber Cancer Institute, 3 Blackfan Circle, Boston, MA 02115.
Mr. Munsell and Mr. Venier: The University of Texas M.D. Anderson Cancer Center, PO Box 301402, Houston, TX 77230-1402.
Dr. Plevritis: Stanford Univeristy, 1201 Welch Road, Room P267, Stanford, CA 94305-5488.
Dr. Ravdin: 19931 Encino Royale, San Antonio, TX 78259.
Dr. Schechter: Albert Einstein College of Medicine of Yeshiva University, 1300 Morris Park Avenue, Mazer Building 110, Bronx, NY 10461.
Dr. Sigal: Department of Radiology, Stanford University, Lucas Center for Imaging, 1201 Welch Road, Stanford, CA 94305-5488.
Dr. Stoto: Georgetown University School of Nursing & Health Studies, 3700 Reservoir Road, Northwest, Room 235, Washington, DC 20057-1107.
Dr. Stout: Harvard Medical School, 133 Brookline Avenue, 6th Floor, Boston, MA 02215.
Mr. van Ravesteyn: Erasmus University Medical Center, PO Box 2040, Rotterdam 3000 CA, the Netherlands.
Dr. Zelen: Dana-Farber Cancer Institute, 3 Blackfan Circle, 11th Floor, Center for Life Sciences B, Boston, MA 02115.
Dr. Feuer: National Cancer Institute, 6116 Executive Boulevard, Room 5041, Mail, Stop Code 8317, Bethesda, MD 20892-8317.

Author Contributions: Conception and design: J.S. Mandelblatt, K.A. Cronin, D.A. Berry, H.J. de Koning, S.J. Lee, S.K. Plevritis, C.B. Schechter, M.A. Stoto, N.K. Stout, M. Zelen, E.J. Feuer.
Analysis and interpretation of the data: J.S. Mandelblatt, K.A. Cronin, S. Bailey, D.A. Berry, H.J. de Koning, G. Draisma, H. Huang, S.J. Lee, M. Munsell, S.K. Plevritis, C.B. Schechter, M.A. Stoto, N.K. Stout, N.T. van Ravesteyn, M. Zelen, E.J. Feuer.
Drafting of the article: J.S. Mandelblatt, H.J. de Koning, C.B. Schechter.
Critical revision of the article for important intellectual content: J.S. Mandelblatt, K.A. Cronin, S. Bailey, D.A. Berry, H.J. de Koning, G. Draisma, S.K. Plevritis, C.B. Schechter, M.A. Stoto, N.K. Stout, N.T. van Ravesteyn, M. Zelen, E.J. Feuer.
Final approval of the article: J.S. Mandelblatt, K.A. Cronin, S. Bailey, D.A. Berry, H.J. de Koning, G. Draisma, S.J. Lee, M. Munsell, S.K. Plevritis, C.B. Schechter, M.A. Stoto, N.K. Stout, N.T. van Ravesteyn, M. Zelen, E.J. Feuer.
Statistical expertise: D.A. Berry, G. Draisma, S.J. Lee, M. Munsell, S.K. Plevritis, C.B. Schechter, M.A. Stoto, M. Zelen, E.J. Feuer.
Obtaining of funding: J.S. Mandelblatt, D.A. Berry, S.K. Plevritis.
Administrative, technical, or logistic support: J.S. Mandelblatt, K.A. Cronin, S.K. Plevritis, J. Venier, E.J. Feuer.
Collection and assembly of data: K.A. Cronin, S. Bailey, D.A. Berry, H.J. de Koning, S.K. Plevritis, M. Munsell, C.B. Schechter, N.K. Stout, J. Venier.

Appendix Table 1. Summary of Model Features

Feature	Model*					
	D	E	G	M	S	W
Includes DCIS	No	Yes	Yes	Yes	No	Yes
Includes ER status	Yes	Yes	Yes	Yes	Yes	Yes
How treatment affects mortality	Hazard reduction	Cure fraction	Hazard reduction	Hazard reduction and cure fraction based on mode of diagnosis†	Hazard reduction	Cure fraction
Calibrated to mortality?	No	No	No	Yes	No	Yes‡
Calibrated to incidence?	No	Yes	Yes	Yes	Yes	Yes
Factors affecting screening benefits§	Stage shift, age shift	Size (larger or smaller than fatal diameter)	Stage shift, age shift	Stage shift, age shift	Stage shift, size within stage, age shift	Effectiveness of treatment by stage and age shifts
Factors affecting treatment benefits (independent of screening)	ER status, age, calendar year	ER status, age	ER status, age	ER status, age, calendar year (and improvements in care)	ER status, age	ER status, age, calendar year (which affect cure probability)

DCIS = ductal carcinoma in situ; ER = estrogen receptor.
* Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S = Stanford University; W = University of Wisconsin/Harvard.
† If cancer is clinically detected in model M, a hazard reduction is applied to the survival function. If cancer is detected by screening, then a cure fraction is applied for cases diagnosed in stages 1 and 2a. If cancer is detected by screening in stages 2b, 3, or 4, a similar hazard reduction is applied as for the clinically detected cases. This results in screening benefits due to stage shift and better prognosis for screening-detected versus clinically detected cases within early-stage disease. The use of a cure fraction for early-stage screening-detected cancer is a modification of the model published elsewhere (7, 11).
‡ Model W is calibrated only to mortality for a subset of the cure fraction variables after the natural history model was calibrated to incidence.
§ Note that all models use age-specific inputs for sensitivity of mammography screening. Sensitivity, in turn, has a small effect on screening benefits.

Appendix Table 2. Summary of Base-Case Input Data Sources*

Model Inputs	Data Sets			
	BCSC	SEER 9 Registry	Connecticut Tumor Registry	Berkeley Mortality Database
Secular breast cancer incidence	No	Yes	Yes	No
Mammography test characteristics	Yes	No	No	No
Other cause of death	No	No	No	Yes
Breast cancer survival in 1975	No	Yes	No	No
Breast cancer prevalence in 1975	No	Yes	Yes	No

BCSC = Breast Cancer Surveillance Consortium; SEER 9 = Surveillance, Epidemiology, and End Results 9.
* For this analysis, we assume that 100% of women are screened and that all women detected with cancer are treated as per current practice guidelines.

Appendix Table 3. Average Number of Screening Examinations and Percentage of Reduction in Breast Cancer Mortality, by Screening Strategy

Screening Strategy	Average Screenings per 1000 Women*	Reduction in Breast Cancer Mortality (vs. No Screening), by Model, %†					
		D	E	G	M	S	W
Efficient strategies (not dominated in 6 of 6 models)							
Biennial screening, ages 60–69 y	4263	11	13	11	10	9	12
Biennial screening, ages 55–69 y	6890	15	18	15	14	13	19
Biennial screening, ages 50–69 y	8947	16	23	17	16	15	23
Biennial screening, ages 50–74 y	11 066	22	27	21	21	20	28
Biennial screening, ages 50–79 y	12 366	25	29	24	24	25	30
Biennial screening, ages 50–84 y	13 837	29	31	25	27	26	33
Biennial screening, ages 40–84 y	18 708	31	37	28	29	27	39
Annual screening, ages 40–84 y	36 550	38	49	32	29‡	35	54
Borderline strategies (dominated in 2–3 of 6 models)							
Biennial screening, ages 40–79 y	17 241	27§	35	26	26§	25§	36
Annual screening, ages 50–79 y	24 419	32	39	27§	26§	30	42
Annual screening, ages 50–84 y	26 905	35	41	28§	28§	33	45
Annual screening, ages 40–79 y	34 078	34§	46	30	27§	33§	51
Inefficient/dominated strategies (dominated in all 6 models)							
Annual screening, ages 60–69 y	8438	14§	18§	13§	12§	12§	17§
Biennial screening, ages 45–69 y	11 694	18§	26§	20§	19§	17§	27§
Annual screening, ages 55–69 y	13 009	18§	25§	17§	15§	16§	26§
Biennial screening, ages 40–69 y	13 831	18§	28§	20§	19§	16§	29§
Annual screening, ages 50–69 y	17 733	21§	31§	20§	18§	20§	33§
Annual screening, ages 50–74 y	21 330	27§	35§	24§	22§	26§	38§
Annual screening, ages 45–69 y	22 546	23§	35§	22§	20§	22§	39§
Annual screening, ages 40–69 y	27 428	24§	39§	23§	20§	22§	43§

* Average number of mammograms across models. Not all possible mammograms in the age group are obtained in strategies that continue to the oldest age groups, because many women die of other causes before screening would occur.
† Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S = Stanford University; W = University of Wisconsin/Harvard.
‡ Because of rounding, this strategy seems to be dominated, but the actual result is 29.4.
§ Strategy is dominated ("inefficient") within the specific model. A strategy is classified as dominated if another strategy (from the efficient, borderline, or inefficient/dominated category) results in an equal or higher percentage of mortality reduction with fewer average screening examinations.

Appendix Table 4. Average Number of Screening Examinations and Life-Years Gained, by Screening Strategy

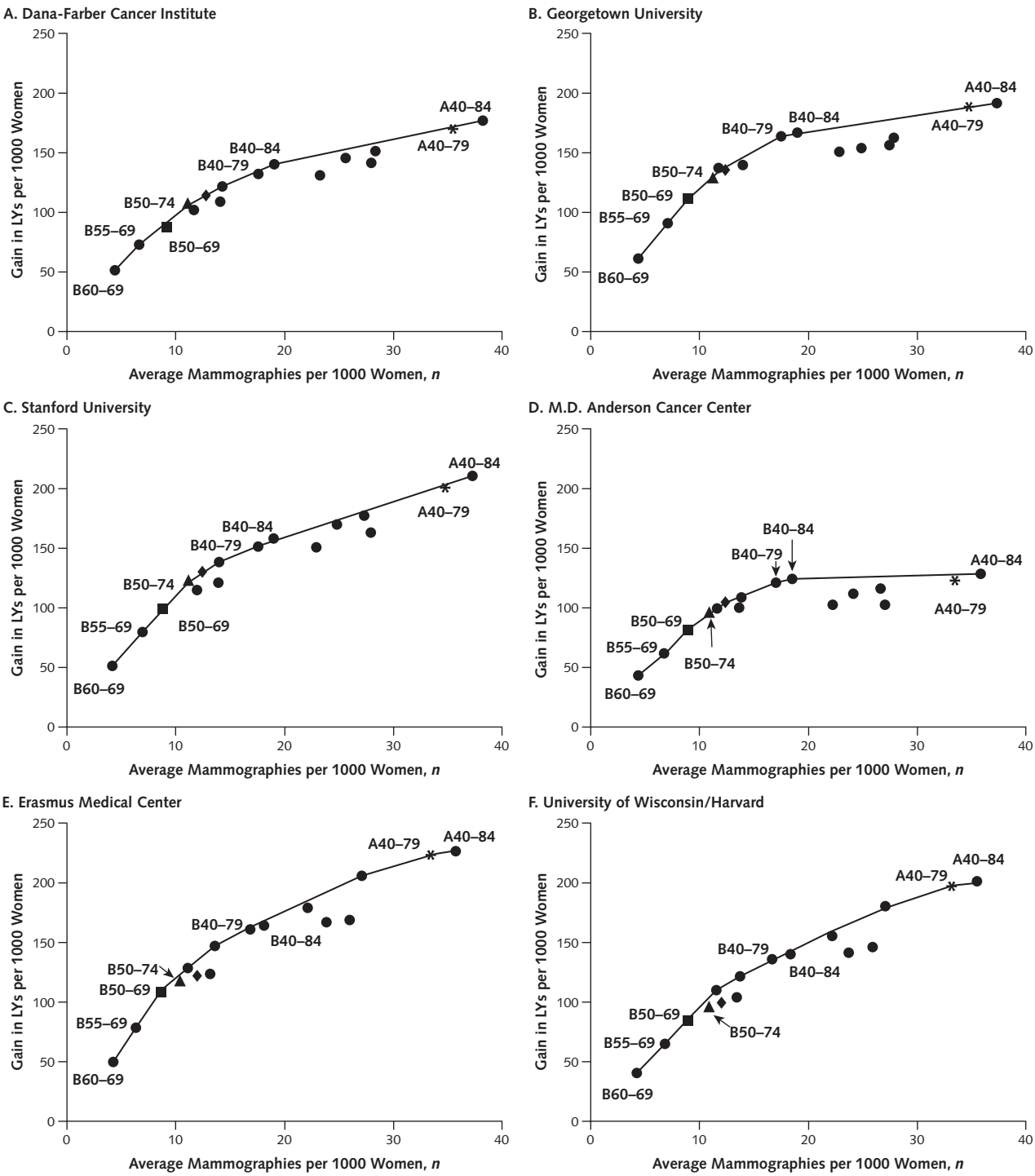
Screening Strategy	Average Screenings per 1000 Women*	Life-Years Gained per 1000 Women (vs. No Screening), by Model†					
		D	E	G	M	S	W
Efficient strategies (not dominated in 5 or 6 of 6 models)							
Biennial screening, ages 60–69 y	4263	51	49	61	43	52	39
Biennial screening, ages 55–69 y	6890	73	78	91	62	80	64
Biennial screening, ages 50–69 y	8947	88	107	111	82	99	84
Biennial screening, ages 50–74 y	11 066	106	116	128	96	121	95
Biennial screening, ages 40–79 y	17 241	133	161	164	122	151	136
Biennial screening, ages 40–84 y	18 708	140	164	167	126	158	140
Annual screening, ages 40–79 y	34 078	170	224	188	123‡	202	198
Annual screening, ages 40–84 y	36 550	177	227	192	128	210	202
Borderline strategies (dominated in 2–4 of 6 models)							
Biennial screening, ages 45-69 y	11 694	102‡	129	136	99	116‡	109
Biennial screening, ages 50–79 y	12 366	114	122‡	136‡	103	130	99
Biennial screening, ages 50–84 y	13 837	121	124‡	139‡	108	138	103
Biennial screening, ages 40–69 y	13 831	108‡	147	140	101‡	120‡	121
Annual screening, ages 45–69 y	22 546	131‡	179	152‡	103‡	152‡	155
Annual screening, ages 50–79 y	24 419	145	166‡	154‡	112‡	170	142‡
Annual screening, ages 50–84 y	26 905	152	169‡	157‡	116‡	178	146‡
Annual screening, ages 40–69 y	27 428	142‡	206	162‡	103‡	164‡	180
Inefficient or dominated strategies (dominated in all 6 models)							
Annual screening, ages 60–69 y	8438	65‡	69‡	71‡	53‡	69‡	56‡
Annual screening, ages 55–69 y	13 009	91‡	107‡	100‡	68‡	102‡	90‡
Annual screening, ages 50–69 y	17 733	117‡	148‡	128‡	91‡	132‡	123‡
Annual screening, ages 50–74 y	21 330	134‡	160‡	144‡	104‡	156‡	135‡

* Average number of mammograms across models. Not all possible mammograms in the age group are obtained in strategies that continue to the oldest age groups, because many women die of other causes before screening would occur.

† Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S = Stanford University; W = University of Wisconsin/Harvard.

‡ Strategy is dominated within a specific model. Strategy is classified as dominated if another strategy (from the efficient, borderline or inefficient/dominated category) results in an equal or higher gain in life-years with fewer average screening examinations.

Appendix Figure. **Life-years gained versus number of mammographies performed per 1000 women, by model and screening strategy.**



The panels show an efficiency frontier graph for each model. The graph plots the average number of mammographies performed per 1000 women against LYs gained for each screening strategy (vs. no screening). Strategies are denoted as annual (*A*) or biennial (*B*) with starting and stopping ages. We plot efficient strategies (that is, those in which increases in use of mammography resources result in greater LYs gained than the next least-intensive strategy) in all 6 models. We also plot "borderline" strategies (approaches that are efficient in some models but not others). The line between strategies represents the "efficiency frontier." Strategies on this line would be considered efficient because they achieve the greatest gain per use of mammography resources compared with the point (or strategy) immediately below it. Points that fall below the line are not considered as efficient as those on the line. When the slope in the efficiency frontier plot levels off, the additional LYs gained per unit increase in use of mammography are small relative to the previous strategies and could indicate a point at which additional investment (use of screening) might be considered as having a low return (benefit). LY = life-year.