Supplementary Online Content


eAppendix 1. Analysis of Individual Tumor Sizes
eAppendix 2. Analysis of Lead Time Effects
eAppendix 3. Analyses for Different Regions
eAppendix 4. Analysis by County Population
eAppendix 5. Analysis of Duration of Follow-up
eAppendix 6. Definitions of Surgical Treatments

This supplementary material has been provided by the authors to give readers additional information about their work.
eAppendix 1. Analysis of Individual Tumor Sizes

**Breast cancer diagnoses in 2000, according to tumor size**

<table>
<thead>
<tr>
<th>Tumor Size</th>
<th>Rate (per 100,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0-0.9 cm</td>
<td></td>
</tr>
<tr>
<td>1.0-1.9 cm</td>
<td></td>
</tr>
<tr>
<td>2.0-2.9 cm</td>
<td></td>
</tr>
<tr>
<td>3.0-3.9 cm</td>
<td></td>
</tr>
<tr>
<td>4.0-4.9 cm</td>
<td></td>
</tr>
<tr>
<td>&gt;5.0 cm</td>
<td></td>
</tr>
</tbody>
</table>

Proportion with mammogram in past 2 years

Extent of screening and breast cancer incidence for individual centimeters of tumor size, women age 40 and older in 547 US counties

In the main text, Figure 2 presents results for ≤2 vs. >2 cm cancers. Here, we show analogous results for individual centimeters of tumor size. Screening-related increases in incidence are concentrated at small sizes of breast cancer. Increases were significant for 0.0–0.9 and 1.0–1.9 cm breast cancers, but not for larger sizes (see eTable). No screening-related reduction in breast cancer diagnoses was found in any strata of tumor size. **Notes:** Each circle presents data for a single county, and the circle area is proportional to the county population of women ≥40 years old (see legend). Each fitted curve is a smoothing spline with Bonferroni-corrected 95% confidence bands. In the regression analysis, screening extent and tumor size (measured in mm) were both treated as continuous variables and included in a 2-dimensional tensor smooth. After fitting, the model predictions were stratified by centimeter for ease of interpretation. This procedure was selected to minimize error from discretizing tumor size, which is a fundamentally continuous measurement, while avoiding the difficulty of interpreting 3-dimensional plots. (We would otherwise obtain a 3-dimensional plot from this analysis of incidence, screening, and tumor size.) For clarity, results are binned for all cancers ≥5.0 cm; findings are not significant for greater sizes.

© 2015 American Medical Association. All rights reserved.
eAppendix 2. Analysis of Lead Time Effects

Screening extent and breast cancer incidence for the years 2000–2003, as stratified by the size of recent changes to screening rates

Counties where the extent of screening was stable from 1997–1999 to 2000–2003 (±5 percentage points; n = 189) are shown in green. Counties with recently increased screening over the same period (>10 percentage points; n = 120) are shown in purple. If lead times were responsible for the association between the absolute extent of screening (x-axis) and incidence (y-axis), then incidence should be elevated in counties with recently increased screening and reduced in counties with stable screening. Yet, as this figure shows, no such pattern is observed—in fact, counties with recently increased screening have a slightly lower breast cancer incidence at all values of the x-axis. Notes: As in the main text, results are shown for women age 40 and older in the same US counties. However, the years of analysis have been switched from 2000 to 2000–2003. Screening data are only available for the periods 1997–1999 and 2000–2003. Therefore, this switch was necessary to examine changes in screening. As usual, each circle presents data for a single county, and the circle area is proportional to the county population of women ≥40 years old (see legend). Each fitted curve is a smoothing spline, presented with 95% confidence bands.
Discussion of lead time effects

The incidence results presented in Main Figures 2 and 3 are consistent with overdiagnosis. However, they could hypothetically result from earlier detection of true breast cancers, as follows:

Tumors in screened women are diagnosed earlier than tumors in unscreened women. It is not known exactly how much advanced notice (lead time) is provided by mammography screening. If mammography screening results in long lead times, then breast cancer incidence should increase when more women are screened. For example, suppose a woman receives her first screening today, and the screening detects a breast tumor that otherwise would not have been detected for another two years. Accordingly, the incidence of breast cancer would be increased today, and would also be decreased two years from now.

If lead times were responsible for our incidence results then, when comparing counties that currently have the same absolute extent of screening, the counties where screening has recently increased would have higher incidence than the counties where the extent of screening has been stable. For example, if we compare counties that currently have around 60% screening, then the incidence of breast cancer should be higher in counties where the extent of screening has increased substantially (>10 percentage point increase) during the past 7 years, as compared with counties where the extent of screening has remained stable during the same period (±5 percentage point change). Yet, as seen in the Figure, there is no evidence of this effect. Instead, breast cancer incidence is broadly similar in counties with recently increased screening (purple line) and in counties with stable screening (green line). In fact, breast cancer incidence was observed to be slightly lower where the extent of screening had recently increased. In summary, the county evidence suggests that lead times are not responsible for our results.

This lead time analysis can be formalized by including both the absolute extent of screening and recent increases in screening within a multivariable regression (with breast cancer incidence as the dependent variable). As in the main analyses, spline regression was performed under a negative binomial model, fit within the framework of a generalized additive model by restricted maximum likelihood. A two-dimensional tensor of thin plate splines was used to allow incidence to depend on essentially any smooth relationship between the absolute extent of screening and increases in screening that might be prompted by lead times (although it turned out that this interaction was non-significant).

If lead times were responsible for the incidence data, then recent increases in screening should be positively associated with breast cancer incidence. Yet, the regression demonstrated that the association between recent increases in screening and breast cancer incidence was negative (RR = 0.90 for a recent 10 percentage point increase in screening; p < 0.001). Further, the absolute extent of screening remained significantly and positively associated with breast cancer incidence in the multivariable regression (RR = 1.15 for every 10 percentage points difference in the absolute extent of screening; p < 0.001), consistent with the findings of our main analysis (RR = 1.16).

Altogether, we found no evidence that lead time could explain the relationship between the extent of screening and breast cancer incidence.
Appendix 3. Analysis for Different Regions

Separate log-linear relationships are fit for each state and metropolitan area. Nonetheless, each region shows broadly similar relationships between screening and incidence, without any significant relationships between screening and mortality. Accordingly, the patterns observed in the main text (Figures 2 & 3) do not result from regional differences, such as in insurance coverage or quality of mammography screening. 

Notes: Each fitted curve is from a generalized linear model with a negative binomial family, and is presented with 95% confidence bands (Bonferroni correction, 11 comparisons). Although splines were used for greater flexibility elsewhere in this paper, it was necessary to switch to parametric regression here because there was too little data for spline regression in some regions (e.g., Hawaii, which includes only 4 counties).
Discussion of regional results

As can be seen in the figure, generally similar relationships between screening and incidence are found throughout the states and metropolitan areas that report to the SEER cancer registries. Further, no significant screening-mortality relationship was found in any state or metropolitan area.

To illustrate the similarity of results across the US, we fit a model in which each region had its own relative rate of breast cancer incidence per change in screening (visible as the different slopes of the curves in the figure). Statistically speaking, however, this model was substantially inferior to a simpler model in which all regions shared the same relative rate of breast cancer incidence per change in screening (the same slopes).* Indeed, the former model had a relative likelihood of 0.017 as compared with the latter, simpler model. Consistent with the main analysis, the simpler model suggested that a 10 percentage point increase in screening was associated with a 15% increase in breast cancer incidence (RR = 1.15, 95% CI: 1.12–1.18), regardless of state or metropolitan area.

Further, a model of mortality with no screening term was preferable to any alternative that included a screening term (relative likelihood = 0.234).** Somewhat informally, this means that if we were trying to predict mortality in counties based on their extent of screening and the region that they belonged to, then ignoring the extent of screening would provide better predictions than trying to include it in our model.

Together, these findings further emphasize the strength and consistency of the relationship between screening and incidence, as well as the lack of a concomitant relationship between screening and mortality.

*Note that, in both models, each region had its own intercept term. The simpler model offered the minimum Akaike Information Criterion (AICmin = 3430.3, AICalt = 3438.5).

**AICmin = 2244.0, min(AICalt) = 2246.9.
Extent of screening, breast cancer incidence, and breast cancer mortality, according to the county population of women age 40 and older

Similar relationships between screening and incidence were found regardless of whether the counties included 5000, 50 000, or 500 000 women. Further, there was no significant relationship between screening and 10-year mortality from breast cancer at any population size. However, low statistical power was available for county populations <2500, with broad confidence bands making it difficult to interpret the relationships between screening and breast cancer rates. (These counties included only 0.9% and 1.3% of the total breast cancer diagnoses and deaths analyzed in this study, respectively.) Notes: Each circle presents data for a single county, and the circle area is proportional to the county population of women ≥40 years old (see legend). Each fitted curve is a smoothing spline with Bonferroni-corrected 95% confidence bands. In the regression analysis, screening extent and log(county population) were both treated as continuous variables and included in a 2-dimensional tensor smooth, allowing essentially any smooth relationship between them. After fitting, the model predictions were stratified by county population for ease of interpretation. This procedure was selected to minimize error related to coarsely discretizing county population, while avoiding the difficulty of interpreting 3-dimensional plots.
Discussion of analysis by county population

Ecological biases arise because values that are defined for individuals (screening history, breast cancer diagnosis) are aggregated before being analyzed, resulting in a loss of information.\textsuperscript{1} More information is usually lost when the aggregation takes place at a larger scale. For example, more information is lost when data are aggregated at the level of nations than when data are aggregated at the level of counties.

Ecological biases inherently cannot be ruled out without access to individual data.\textsuperscript{1,2} However, because the information lost through aggregation results in ecological bias, an ecological relationship that changes with the scale of aggregation provides a warning sign that can suggest the presence of ecological bias. For example, if the relationship between screening and breast cancer diagnosis differs significantly according to the county population, this would be considered a warning sign that our county-level analysis is affected by ecological bias (although the absence of that warning sign cannot exclude ecological bias).

If present, this warning sign would be especially evident in our study, because of the wide range of county populations. Yet, as can be seen in the figure, similar relationships between screening and incidence were found regardless of whether the counties included 5000, 50 000, or 500 000 women. Further, there was no significant relationship between screening and 10-year mortality from breast cancer in any of the strata. In summary, the ecological relationships did not change across 2 orders of magnitude in the scale of aggregation—the warning sign of ecological bias was not observed.

However, for counties with <2500 women in our analysis, it was not clear whether the trends were or were not similar to the main analysis—broad confidence intervals made it difficult to interpret the relationships between screening and breast cancer rates (RR for incidence = 0.97, 95% CI: 0.80–1.16; RR for mortality = 0.83; 95% CI: 0.66–1.04). Indeed, these counties included only 0.9% and 1.3% of the total breast cancer diagnoses and deaths analyzed in this study, respectively.
eAppendix 5. Analysis of Duration of Follow-up

Cumulative incidence-based mortality from breast cancer for diagnoses in 2000 (red line)

Cumulative mortality is shown by years since diagnosis. For context, the black line presents the conventional mortality rate for the year 2000. Notes: The mortality rates include 7729 deaths from breast cancer. All results are presented for women ≥40 years old in the SEER registries, and are age standardized to the total US population in that age range.

Cumulative mortality rates from breast cancer for diagnoses in 2000, stratified by county-level extent of screening in 2000

As previously, the black line presents the conventional mortality rate for the year 2000. Notes: For very low, low, medium, high, and very high screening, the mortality rates respectively include 547, 1936, 2089, 2645, and 522 deaths from breast cancer.
Discussion of the duration of follow-up

In the current study, the duration of follow-up was 10 years. It is important to consider whether 10 years is sufficient, both because some tumors are slow growing and because late recurrences play a substantial role in breast cancer. Further, it is believed that tumors detected by screening are especially likely to be slow growing, a phenomenon that is generally described as a length bias or the length-biased sampling of screening.3

The first figure shows cumulative incidence-based mortality from breast cancer, assessed monthly over the course of the 10-year follow-up. For context, the figure also shows the conventional mortality rate from breast cancer in 2000 (black horizontal line). As breast cancer mortality has been decreasing since 1990,4 there is no reason to suspect that breast cancer mortality in future years will exceed breast cancer mortality in 2000.

The second figure shows the same results stratified according to screening extent. Through the 10-year follow-up, the cumulative mortality rates remain very close to each other in strata with ≥60% screening. Since the mortality curves do not diverge for the first 10 years after diagnosis, there is no evidence to suggest that they would begin to do so >10 years after diagnosis.

In addition to supporting the 10-year follow-up, the similarity of these mortality curves provides a second, independent source of support for the arguments in eAppendix 2 — Lead Time. Briefly, if the observed increases in overall incidence (Figure 2) were explained by differences in lead time between counties with high and low screening, then lead times would modify the intervals between diagnosis and death that were observed in these counties. Yet, the incidence-based mortality curves are highly similar over the 10-year follow-up.

For counties with <60% screening, incidence-based mortality is slightly lower 6–10 year after diagnosis. We do not know exactly why this is the case, but it could simply be the relatively small sample size in this strata, which accounts for 7% of all women and 7% (n = 547) of all deaths recorded in the study. Indeed, while the mortality curve is visually distinct from the curves for other strata, it is not significantly different (log-rank tests, p > 0.2 for all comparisons with other strata).

Based on the findings shown in the figures, 10 years of follow-up appears sufficient to investigate overdiagnosis in the county data.
eAppendix 6. Definitions of Surgical Treatment

Breast-conserving surgery was defined as

- Partial mastectomy, not otherwise specified (NOS); less than total mastectomy, NOS
- Partial mastectomy WITH nipple resection
- Lumpectomy or excisional biopsy
- Reexcision of the biopsy site for gross or microscopic residual disease
- Segmental mastectomy (including wedge resection, quadrantectomy, tylectomy)

Non-breast-conserving surgeries are sometimes described as mastectomies. Non-breast-conserving surgery was defined as

- Subcutaneous mastectomy
- Total (simple) mastectomy, NOS
  - WITHOUT removal of uninvolved contralateral breast
    - Reconstruction, NOS
    - Tissue
    - Implant
    - Combined (tissue and implant)
  - WITH removal of uninvolved contralateral breast
    - Reconstruction, NOS
    - Tissue
    - Implant
    - Combined (tissue and implant)
- Bilateral mastectomy for a single tumor involving both breasts, as for bilateral inflammatory carcinoma.
- Modified radical mastectomy
  - WITHOUT removal of uninvolved contralateral breast
    - Reconstruction, NOS
    - Tissue
    - Implant
    - Combined (tissue and implant)
  - WITH removal of uninvolved contralateral breast
    - Reconstruction, NOS
    - Tissue
    - Implant
    - Combined (tissue and implant)
- Extended radical mastectomy
  - WITHOUT removal of uninvolved contralateral breast
  - WITH removal of uninvolved contralateral breast

Surgical treatment included breast-conserving and non-breast-conserving surgeries, as well as

- Local tumor destruction, NOS
- Mastectomy, NOS
- Surgery, NOS


© 2015 American Medical Association. All rights reserved.
References


