

An exploration for quantification of overdiagnosis and its effect for breast cancer screening

Lei Yang^{1*}, Shengfeng Wang^{2*}, Yubei Huang³

¹Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Beijing Office for Cancer Prevention and Control, Peking University Cancer Hospital & Institute, Beijing 100142, China; ²Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing 100191, China; ³Department of Epidemiology and Biostatistics, Key Laboratory of Cancer Prevention and Therapy, Tianjin, Key Laboratory of Breast Cancer Prevention and Therapy, Ministry of Education, National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China

*These authors contributed equally to this work.

Correspondence to: Yubei Huang, PhD. Department of Epidemiology and Biostatistics, Key Laboratory of Cancer Prevention and Therapy, Tianjin, Key Laboratory of Breast Cancer Prevention and Therapy, Ministry of Education, National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China. Email: yubei_huang@163.com.

Abstract

Objective: To redefine overdiagnosis and reestimate the proportion of overdiagnosis of breast cancer caused by screening based on the Surveillance, Epidemiology, and End Results (SEER, 1973–2015) Program data.

Methods: The breast cancer diagnosed before 1977 was defined as the no-screening cohort since America had initiated breast cancer screening from 1977. The breast cancer diagnosed in 1999 was defined as the screening cohort due to no increases in both the proportion of early-stage breast cancer until 1999 and the overall survival of early-stage breast cancer diagnosed over the three years since 1999. The magnitude of overdiagnosis was calculated as the difference in the proportions of early-stage breast cancer patients with long-time (15-year) survival to all breast cancer patients between two cohorts.

Results: Over 23 years before and after widespread screening in America, the proportion of early-stage breast cancer patients increased from 52.1% (16,891/32,443) to 72.7% (16,021/22,025) ($P < 0.001$). The 15-year survival rate of early-stage breast cancer patients increased from 51.1% to 61.5% ($P < 0.001$), while the proportions of early-stage breast cancer patients with long-time survival to all breast cancer patients increased from 26.6% ($52.1\% \times 51.1\%$) to 44.7% ($72.7\% \times 61.5\%$). Assuming no improvements in cancer screening technology and treatment technology, 18.1% (44.7%–26.6%) of breast cancer patients were overdiagnosed associated with screening. The age-specific overdiagnosis rates were 18.9%, 24.7%, 24.5%, 20.5%, and 8.3% for breast cancer patients aged 40–49, 50–59, 60–69, 70–74, and ≥ 75 years old, respectively.

Conclusions: Overdiagnosis caused by mammographic screening is probably overestimated in current screening practices. Further trials with more sophisticated designs and analyses are needed to validate our findings in the future.

Keywords: Breast cancer; screening; overdiagnosis

Submitted Nov 19, 2019. Accepted for publication Jan 17, 2020.

doi: 10.21147/j.issn.1000-9604.2020.01.04

View this article at: <https://doi.org/10.21147/j.issn.1000-9604.2020.01.04>

Introduction

Cancer screening has always been considered a double-edged sword. Females who participate in mammographic

screening often expect that screening would result in a health benefit and minimal risk of harm. Moreover, physicians believe that cancers detected by screening have favorable biologic characteristics and would have better

prognoses than those diagnosed in the clinic. However, the benefits of screening always coexist with the risks of screening, like two sides of a coin. Therefore, before organizing a population-based screening, the key questions under consideration for policy makers are whether the benefits outweigh the risks, and vice versa.

Several previous studies have investigated the risk of breast cancer screening, especially the magnitude of overdiagnosis. However, highly variable estimates of overdiagnosis have been reported by different studies (1-16). As reported by Welch using data from the U.S. Surveillance, Epidemiology, and End Results (SEER) program, only 30 of the 162 additional small cancers detected per 100,000 women were expected to progress to large cancers, and the remaining 132 cases were classified as overdiagnosed cancers (5). More recent studies from Denmark and the Netherlands showed that nearly one in three and half of screening-detected breast cancers represent overdiagnosis, respectively (14,15). Based on these results, distinctly different standpoints for mammographic screening can be found from different organizations. The Swiss Medical Board even suggests that mammography screening should be abolished due to overwhelming overdiagnosis (17). However, the International Agency for Research on Cancer (IARC) Handbook Working Group concluded that there is a net benefit for women aged 50–69 years who were invited to undergo screening (18).

Several reasons could lead to these highly variable estimates of overdiagnosis. The definitions and methodologies used to calculate overdiagnosis are the primary reasons (19). Even among cohort studies, selection of different proximate indicators of overdiagnosis [e.g., the excessive incidence of smaller tumors, non-advanced cancers, or carcinoma *in situ* before and after widespread screening, or between no screening and screening areas (5,14,15)] could also lead to huge differences in the estimates of overdiagnosis. However, whether all these excessive early-stage cancers detected by screening should be considered overdiagnosed remains controversial. This study aimed to assess the basic conditions of overdiagnosed cancers and quantify overdiagnosis of breast cancer screening using the latest SEER data.

Materials and methods

Redefinition of overdiagnosed cancer

As defined by Welch, an overdiagnosed cancer is a cancer

that would never present clinically or from which the patient would not die during his or her lifetime in the absence of screening (20). According to this definition, we identified two basic conditions of overdiagnosed cancers. First, overdiagnosed cancers must be early-stage cancers detected by screening due to occult symptoms and signs, whereas most advanced-stage cancers, even in the absence of screening, would also be diagnosed in clinics due to obvious lumps or other clinical symptoms. Therefore, any advanced-stage cancers should not be classified as overdiagnosed. In contrast, should all early-stage cancers detected by screening be considered overdiagnosed because of the lack of obvious clinical symptoms and signs? According to Welch's definition, if early-stage breast cancer patients die of cancer itself or other diseases (such as cardiovascular diseases) in their life expectancy, these early-stage cancers should not always be classified as overdiagnosed. Therefore, the second condition is that an overdiagnosed breast cancer patient would not die early in their life expectancy. Namely, overdiagnosed breast cancer patients should survive longer than clinic-diagnosed breast cancer patients within the same life expectancy. In summary, if we classify all cancers into three categories: symptomatic advanced-stage cancer, asymptomatic early-stage cancer with short-term survival, and asymptomatic early-stage cancer with long-term survival, the asymptomatic early-stage cancer with long-term survival should be considered as overdiagnosed cancers.

Based on the two basic conditions, overdiagnosed cancers caused by screening should be redefined as asymptomatic early-stage cancers that are excessively detected by screening and have a longer survival time than asymptomatic early-stage cancers in the absence of screening.

Data source

All analyses were based on the latest incidence data of breast cancer from the SEER database (SEER9) (21). The SEER program is a population-based registry program for incident cancers in the United States, and SEER9, which include 9 registries, represents approximately 10% of the population of the United States (21). SEER-based estimates of breast-cancer mortality are virtually identical to those ascertained from U.S. mortality data (22). Moreover, the SEER program has had virtually complete case ascertainment and reporting for decades (23). More importantly, the study period (1973–2015) spans the

periods before and after the widespread dissemination of mammographic screening.

The analyses were performed in 756,636 female breast cancer patients with complete information of age at diagnosis [≥ 40 years old, which was the youngest starting age for mammographic screening suggested by the American Cancer Society (ACS) and U.S. Preventive Services Task Force (USPSTF), *Supplementary Table S1, S2*], year of diagnosis, SEER historic stage (*in situ*, localized, regional or distant), vital status (alive or dead), and survival months. Although American Joint Committee on Cancer (AJCC) TNM stage is a better predictor of cancer prognosis, data of AJCC TNM stage were only available since 1988. Tumor size, as suggested by Welch (5), could be considered as a direct proximate indicator of screening effect. However, the SEER historic stage is a more comprehensive and direct proximate indicator of screening effect and includes comprehensive information of the involvement in the limits of the original organ, lymph-node involvement, and metastasis. Moreover, SEER historic stage has been collected since 1973, and only 2.7% of patients in the database lack this information. Therefore, SEER historic stage, rather than AJCC TNM stage or tumor size, was used in our final analyses. Moreover, in this study, early-stage breast cancers were defined as carcinomas *in situ* and localized carcinomas, and the remaining cancers (regional and distant carcinomas) were defined as advanced-stage cancers.

Screening and no screening cancer cohorts

Guidelines for mammographic screening were first suggested by the ACS in 1977 (24), and the use of mammographic screening was rare before 1977. Therefore, the breast cancer patients diagnosed before 1977 were defined as the no-screening cohort. Due to the limited number of early-stage breast cancers diagnosed before 1977, breast cancer patients diagnosed from 1973 to 1976 (32,443 cancers; 16,891 early-stage cancers and 15,552 advanced-stage cancers) were incorporated into the no-screening cohort.

As shown in *Figure 1*, the proportion of early-stage breast cancer seemed to remain steady at a high plateau since 1999, which may suggest that the effectiveness of screening also reached a potential high plateau in 1999. In order to validate whether the screening effectiveness reached a stable high plateau since 1999, a series of Cox regression models were conducted to compare the long-

term cumulative survival rates of cancer cohorts diagnosed in different years. To reduce the potential bias, age at diagnosis was adjusted in all Cox regression analyses. As shown in *Supplementary Table S3*, there was a significant increasing trend in the 15-year cumulative survival rates for the early-stage breast cancer cohort diagnosed from 1973 to 1999 (P value for trend < 0.001). Similar increasing trends were also observed for survival rates associated with less than 15 years' follow-up among early-stage breast cancer diagnosed before 1999. However, no significant improvement in survival rates was found in the 3 years since 1999.

If there were any improvements in cancer survival, these improvements may come from screening or treatment. Conversely, if no improvements in cancer survival were found after an increasing trend, it meant that cancer screening technology and treatment technology had not changed significantly during this period. It also meant that the impact of cancer screening and treatment technology on cancer survival had arrived at a relatively stable high-plateau period in 1999. Therefore, breast cancer patients diagnosed at 1999 (22,025 cancers; 16,021 early-stage cancers and 6,004 advanced-stage cancers) were defined as the screening cohort.

Statistical analysis

In order to estimate the magnitude of overdiagnosis, we first calculated the proportions of early-stage breast cancer patients to all breast cancer patients and the proportions of early-stage breast cancer patients with long-term (15 years) survival to all early-stage breast cancer patients in the screening and no screening cohorts, respectively. Second, we calculated the proportions of early-stage breast cancer patients with long-term survival to all breast cancer patients in the two cohorts, which were calculated as the proportion of early-stage breast cancer patients multiplied by the proportion of early-stage breast cancer patients with long-term survival. Finally, assuming no obvious improvements in cancer screening and treatment skills from 1977 to 1999 in America, the magnitude of overdiagnosis is the difference in the proportions of early-stage breast cancer patients with long-time survival to all cancer patients between the two cohorts.

In theory, the magnitudes of overdiagnosis would differ among patients diagnosed at different ages. Based on starting and ending ages for mammographic screening suggested by the ACS and USPSTF at different periods

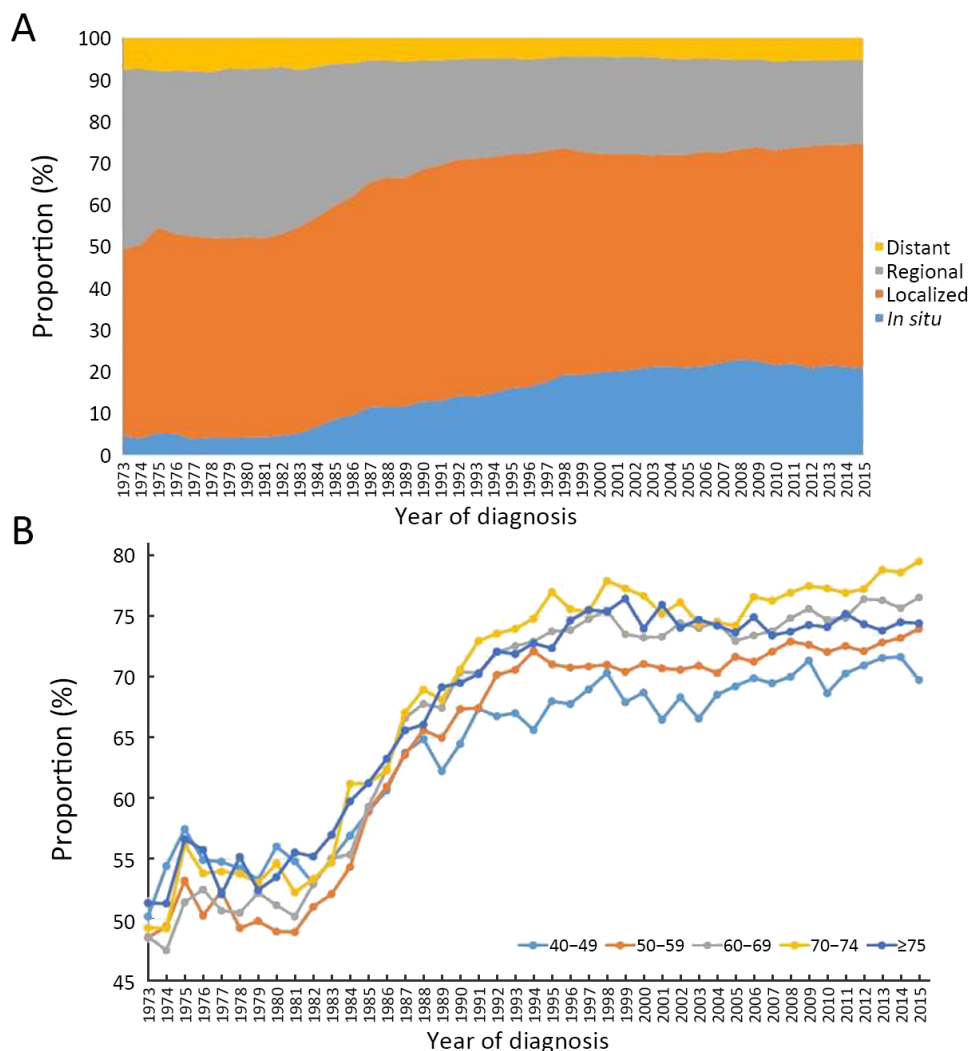


Figure 1 Stage shifts of breast cancer from 1973 to 2015. Overall SEER stage shifts (A) and age-specific early-stage shifts (B). SEER, the Surveillance, Epidemiology and End Results.

(Supplementary Table S1,S2) (24-32), subgroup analyses on the magnitudes of overdiagnosis were conducted according to five age groups: 40–49, 50–59, 60–69, 70–74, and ≥75 years old. Moreover, both the ACS and USPSTF guidelines recommend regular screening in females aged 50–74 years old until now. Therefore, further sensitivity analyses were conducted among breast cancer patients diagnosed at 50–74 years old to investigate the magnitude of overdiagnosis associated with regular screening.

We used Pearson’s Chi-square test to compare the proportion of early-stage cancers between the screening and no screening cohorts. Kaplan-Meier analysis was used to plot survival curves and estimate cumulative survival at specific follow-up times. The Cox regression analysis was used to assess whether there were significant differences in

the long-term cumulative survival rates at specific time points between screening cohort and no screening cohort after adjusting for age at diagnosis. All reported P values were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Trends in early-stage breast cancer diagnosis

The shift in the SEER historic stage of breast cancers is shown in Figure 1. From 1973 to 2015, the proportion of early-stage breast cancers increased from 49.5% to 74.8%. There seemed to be three distinct periods: a low-plateau period before 1977, a continuous-climb period from 1977

to 1999, and a new high-plateau period after 1999 (Figure 1A). A similar three-period trend was also observed in the five age subgroups (Figure 1B). These three periods may correspond to three different periods for mammographic screening in the United States: before, during, and after the implementation of widespread mammographic screening.

Mammographic screening led to an increase in early-stage cancer diagnosis

As shown in Table 1, the overall proportion of early-stage breast cancer increased from 52.1% (16,891/32,443) in the 1973–1976 cohort (no screening) to 72.7% (16,021/22,025) in the 1999 cohort (screening), with 20.7% excessive proportion of breast cancer patients being diagnosed at an early stage in the screening cohort. The age-specific excessive proportion of early-stage breast cancers after the implementation of widespread screening was lowest for patients aged 40–49 years old [13.3% (54.6% vs. 67.9%)]

and highest for patients aged 70–74 years old [24.9% (52.4% vs. 77.3%)], then decreased to 22.4% (54.1% vs. 76.4%) for patients aged ≥ 75 years old (all $P < 0.001$). These results suggested not only that mammographic screening caused obvious stage shifts, but also that the stage shifts were different for patients diagnosed at different ages. Sensitivity analyses showed that the excessive proportion of early-stage breast cancers among patients aged 50–74 years old was 22.2% (50.7% vs. 72.9%) after regular screening ($P < 0.001$).

Mammographic screening led to an increase in survival of early-stage cancer patients

As shown in Table 2, Figure 2, after 15-year follow-up, the overall cumulative survival rate of early-stage breast cancer patients increased from 51.1% in the 1973–1976 cohort (no screening) to 61.5% in the 1999 cohort (screening), with an overall 10.4% improvement in survival ($P < 0.001$).

Table 1 Comparison on proportions of early-stage breast cancer between cancer cohorts with and without mammographic screening

Age (year)	1973–1976 cancer cohort (no screening)			1999 cancer cohort (screening)			Excessive proportion of early-stage cancer (%)	P
	Number of cancers	Number of early-stage cancers	Proportion of early-stage cancers (%)	Number of cancers	Number of early-stage cancers	Proportion of early-stage cancers (%)		
40–49	6,401	3,493	54.6	4,201	2,853	67.9	13.3	<0.001
50–59	9,018	4,561	50.6	5,618	3,957	70.4	19.9	<0.001
60–69	8,055	4,041	50.2	4,801	3,529	73.5	23.3	<0.001
70–74	3,261	1,709	52.4	2,575	1,990	77.3	24.9	<0.001
≥ 75	5,708	3,087	54.1	4,830	3,692	76.4	22.4	<0.001
Total	32,443	16,891	52.1	22,025	16,021	72.7	20.7	<0.001
50–74 in total	20,334	10,311	50.7	12,994	9,476	72.9	22.2	<0.001

Table 2 Comparison on 15-year cumulative survival rates between early-stage breast cancer cohorts with and without mammographic screening

Age (year)	1973–1976 early-stage cancer cohort (no screening)			1999 early-stage cancer cohort (screening)			Improvement on survival rate (%)	P
	Total number	Number of death	Cumulatively survival rate (%)	Total number	Number of death	Cumulatively survival rate (%)		
40–49	3,493	854	75.4	2,853	325	88.4	13.0	<0.001
50–59	4,561	1,451	68.0	3,957	632	83.9	15.8	<0.001
60–69	4,041	2,000	50.4	3,529	1,125	67.8	17.4	<0.001
70–74	1,709	1,179	30.9	1,990	1,038	47.5	16.6	<0.001
≥ 75	3,087	2,747	10.9	3,692	2,992	18.6	7.7	<0.001
Total	16,891	8,231	51.1	16,021	6,112	61.5	10.4	<0.001
50–74 in total	10,311	4,630	55.0	9,476	2,795	70.2	15.3	<0.001

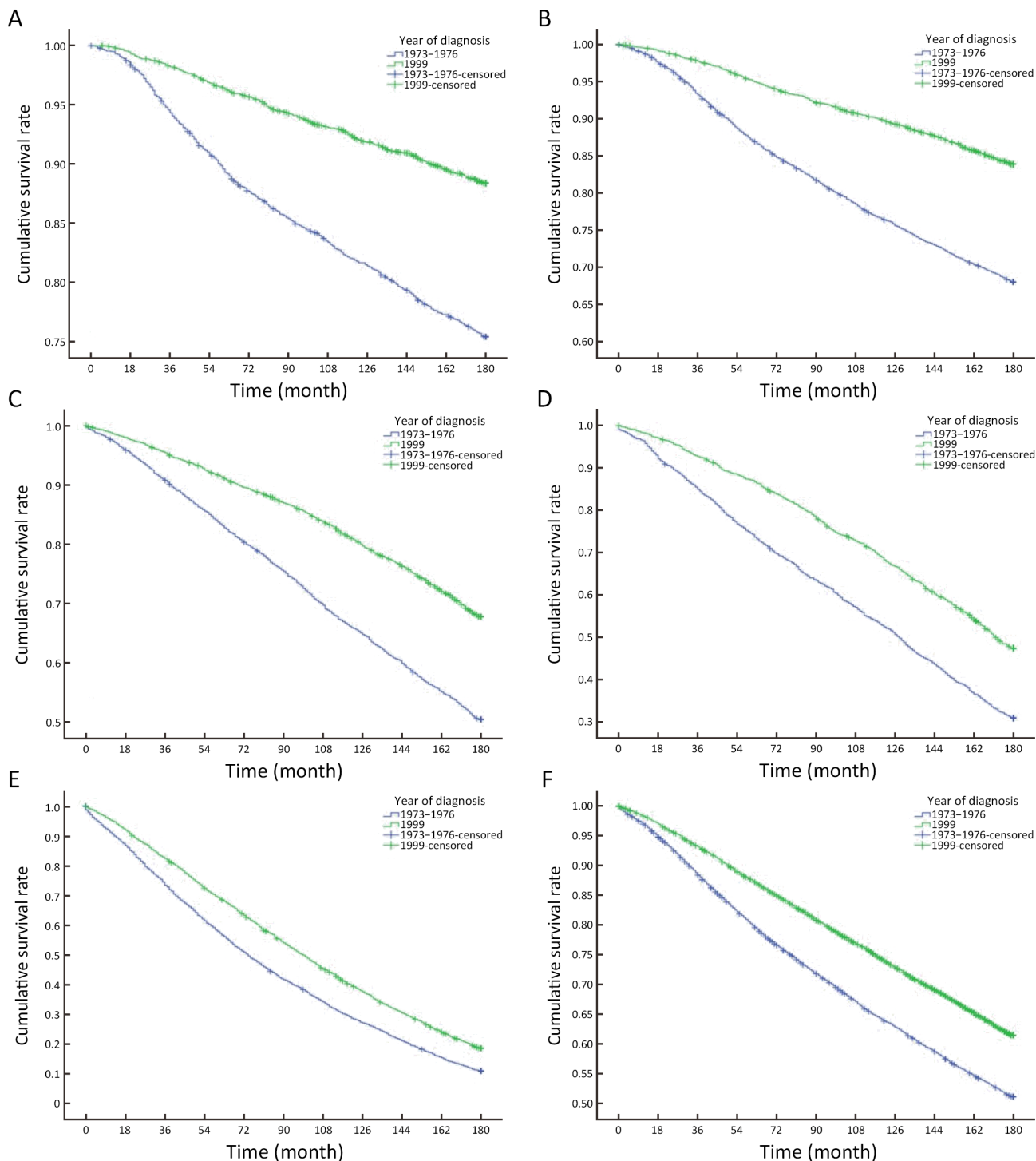


Figure 2 Overall and subgroups 15-year cumulative survival curves of early-stage breast cancers cohorts with and without screening. The 15-year cumulative survival rates of early-stage cancers aged 40–49 (A), 50–59 (B), 60–69 (C), 70–74 (D), ≥75 (E), and 40–74 (F) years old.

Moreover, the age-specific improvement in survival of early-stage breast cancer patients gradually increased from 13.0% (75.4% vs. 88.4%) for patients aged 40–49 years old

to 17.4% (50.4% vs. 67.8%) for patients aged 60–69 years old, and then gradually decreased to 7.7% (10.9% vs. 18.6%) for patients aged ≥75 years old (all $P < 0.001$).

Sensitivity analyses showed that the improvement in survival of early-stage cancer patients aged 50–74 years old was 15.3% (55.0% vs. 70.2%, $P < 0.001$) after regular screening.

Magnitudes of overdiagnosis

As shown in *Figure 3*, based on the proportion of early-stage breast cancers and the proportion of early-stage breast cancer with long-time survival, the proportions of early-stage cancer patients with long-time survival to all breast cancer in the no screening and screening cohorts were 26.6% and 44.7%, respectively. Under the assumption that there were no obvious improvements in the cancer screening and treatment skills before and after widespread screening, 18.1% of breast cancers were overdiagnosed, which would be associated with mammographic screening. The proportions of age-specific overdiagnosis were 18.9%, 24.7%, 24.5%, 20.5%, and 8.3% for breast cancer patients aged 40–49, 50–59, 60–69, 70–74, and ≥ 75 years old, respectively. Further analyses showed that 23.3% of breast cancer patients diagnosed at 50–74 years old were estimated to be overdiagnosed cancers associated with regular screening.

Discussion

Our analysis shows that, over 23 years (from 1977 to 1999) after implementation of widespread mammographic screening in the United States, less than 18.1% of breast cancers were overdiagnosed, since the improvement of

screening and therapy skills cannot be ignored. This result was similar to that reported by the Independent UK Panel on Breast Cancer Screening (33). However, it was obviously lower than those reported by recent studies in Denmark and the Netherlands. Overdiagnosis is probably overestimated in the current screening practice.

The major argument against screening mammography is that excessive detection of early-stage cancers (or smaller tumors/non-advanced cancers/carcinomas *in situ*) through screening does not reduce the incidence of advanced-stage cancers (or larger tumors/advanced cancers/invasive tumors) (5,14–16). Therefore, almost all excessive early-stage cancers detected through screening should be considered as overdiagnosed cancers. However, most of these excessive early-stage cancers have very similar prognoses as early-stage cancers diagnosed without screening, and more than half of these patients diagnosed before 1977 died within 15 years after diagnosis. Therefore, according to Welch's definition, these patients with early-stage cancers who died early after diagnosis due to the natural progression or potentially fatal conditions of the cancer should not be classified as overdiagnosed. Moreover, if we extended the follow-up time, we would have observed more patients with early-stage cancers who died early.

Another concern regarding screening mammography is that the incidence of early-stage cancers (or smaller tumors/non-advanced cancers/carcinomas *in situ*) will continue to increase as screening programs continue (5,14–

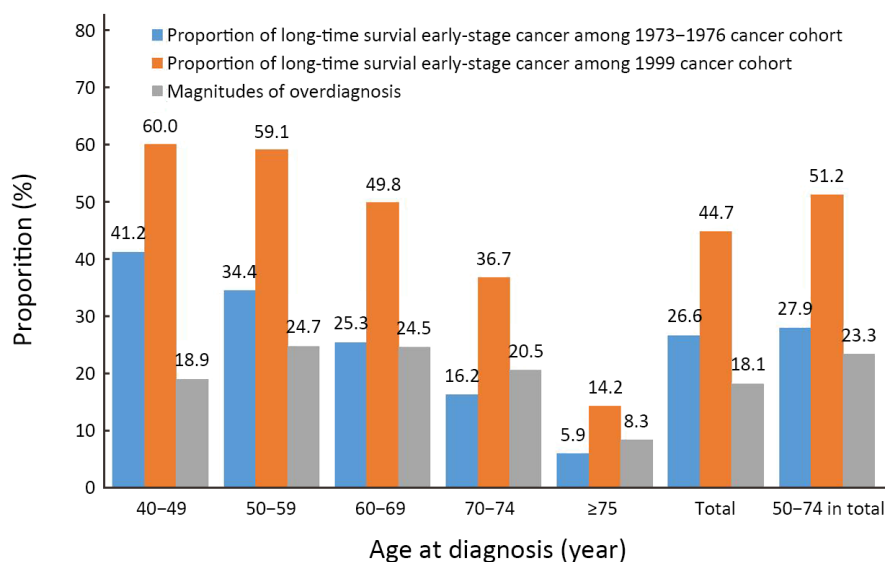


Figure 3 Potential magnitudes of overdiagnosis associated with screening.

16). Therefore, the magnitude of overdiagnosed cancers will also increase whereas the incidence of metastatic breast cancer is essentially unchanged in the United States (34). However, as shown in *Figure 1A*, the proportion of early-stage breast cancers seemed to plateau in 1999. Based on this high-plateau period, it is rational to assume that it will become increasingly more difficult to achieve an additional increase in the proportion of early-stage breast cancers, even though the number of newly diagnosed early-stage cancers continue to increase. Therefore, although screening programs continue, it appears that the proportion of early-stage cancers will not continue to increase in the future, but will remain at a high plateau, as observed in *Figure 1A*.

In addition, as argued by Welch (5), cancer stage may be subject to “upstaging” over time as technology and practice change. Therefore, tumor size would be a better proximate indicator of screening effect than cancer stage because of its relative stability over time. However, the most needed improvement in screening technology is the ability to detect more occult and smaller tumors. If the largest dimension of a newly diagnosed tumor was 20 mm by new technology and 19 mm by old technology, and there was no finding of lymph-node involvement or metastasis by either technology. The tumor would be reclassified from a small tumor (<20 mm) into a large tumor (≥ 20 mm) according to Welch’s classification. However, according to the SEER historic stage standard, if the tumor was confined entirely to the organ of origin, it is classified as localized cancer, regardless of tumor size. Namely, although new technology would bring more improvements in the classifications of the SEER summary stage or modern staging (such as the AJCC TNM stage) (35), the overall simplified four-item SEER historic stage classifications (*in situ*, localized, regional, and distant) are more stable than tumor size over time.

Moreover, we should emphasize that the magnitude of overdiagnosis in our study is likely to be overestimated, since the improvement in treatment skills between 1977 and 1999 cannot be ignored. However, in fact, if there was no improvement in treatment skills, more early-stage breast cancers would die early during their life expectancy, which would lead to less number of early-stage breast cancer with long-term survival. Furthermore, as mentioned by Welch (5), it is not rational to ignore the possibility that women would seek medical care earlier in the course of their disease. Both increased breast cancer awareness and support from family members would promote women to

seek medical service earlier than they did in the past (36,37). As shown in *Supplementary Figure S1*, the therapy pattern for early-stage breast cancer patients changed during the past two decades, especially during the period before 1999. Therefore, the magnitude of overdiagnosis would be no more than 18.1% after excluding the influences of improvement of therapy and earlier treatment seeking.

Our findings are limited for several reasons. First, the underlying true contribution of therapy to the survival of early-stage cancer patients is unobservable. However, as mentioned above, even though this contribution of therapy has been ignored, the magnitude of overdiagnosis would be no more than 18.1% after the implementation of widespread mammographic screening in the United States. Second, the SEER historic stage is a crude stage classification method than the AJCC TNM stage standard; therefore, the estimated magnitude of overdiagnosis would be biased in this study. Third, we have not considered the proportion of the people who received mammography screening actively during the follow-up period among no screening group, which would underestimate our results. Fourth, we did not investigate the influences of the change in risk factors (including environmental, behavioral, and dietary factors), advances in screening methods, or the lead-time bias on the stage shift in the distribution of early-stage breast cancers before and after widespread screening. However, ignorance of these influences would also bias the final estimate of overdiagnosis.

Conclusions

Over 23 years (from 1977 to 1999) after implementation of widespread mammographic screening in the United States, less than 18.1% of breast cancers were overdiagnosed, since survival improvement of early-stage cancer patients because of therapy cannot be ignored. Moreover, overdiagnosis is probably overestimated in the current screening practice. Further RCTs with more sophisticated design and analyses are needed to validate our findings in the future.

Acknowledgements

This work was supported by the Natural Science Foundation of Tianjin (No. 18JCQNJC80300); Chinese National Key Research and Development Project (No. 2018YFC1315600); National Natural Science Foundation of China (No. 81502476) and the Beijing Young Talent

Program (No. 2016000021469G189).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

- Seigneurin A, Francois O, Labarere J, et al. Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data. *BMJ* 2011;343:d7017.
- Jacklyn G, Glasziou P, Macaskill P, et al. Meta-analysis of breast cancer mortality benefit and overdiagnosis adjusted for adherence: improving information on the effects of attending screening mammography. *Br J Cancer* 2016;114:1269-76.
- Miller AB, Wall C, Baines CJ, et al. 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.
- Zackrisson S, Andersson I, Janzon L, et al. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ* 2006;332:689-92.
- Welch HG, Prorok PC, O'Malley AJ, et al. Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness. *N Engl J Med* 2016;375:1438-47.
- Morrell S, Barratt A, Irwig L, et al. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. *Cancer Causes Control* 2010;21:275-82.
- Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;339:b2587.
- Njor SH, Olsen AH, Blichert-Toft M, et al. Overdiagnosis in screening mammography in Denmark: population based cohort study. *BMJ* 2013;346:f1064.
- Hellquist BN, Duffy SW, Nyström L, et al. Overdiagnosis in the population-based service screening programme with mammography for women aged 40 to 49 years in Sweden. *J Med Screen* 2012;19:14-9.
- Kalager M, Adami HO, Bretthauer M, et al. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. *Ann Intern Med* 2012;156:491-9.
- Jørgensen KJ, Zahl PH, Gøtzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. *BMC Womens Health* 2009;9:36.
- Peeters PH, Verbeek AL, Straatman H, et al. Evaluation of overdiagnosis of breast cancer in screening with mammography: results of the Nijmegen programme. *Int J Epidemiol* 1989;18:295-9.
- Etzioni R, Gulati R. Recognizing the limitations of cancer overdiagnosis studies: A first step towards overcoming them. *J Natl Cancer Inst* 2016:108.
- Autier P, Boniol M, Koechlin A, et al. Effectiveness of and overdiagnosis from mammography screening in the Netherlands: population based study. *BMJ* 2017;359:j5224.
- Jørgensen KJ, Gøtzsche PC, Kalager M, et al. Breast cancer screening in Denmark: A cohort study of tumor size and overdiagnosis. *Ann Intern Med* 2017;166:313-23.
- Harding C, Pompei F, Burmistrov D, et al. Breast cancer screening, incidence, and mortality across US counties. *JAMA Intern Med* 2015;175:1483-9.
- Biller-Andorno N, Jüni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. *N Engl J Med* 2014;370:1965-7.
- Lauby-Secretan B, Scoccianti C, Loomis D, et al. Breast-cancer screening — Viewpoint of the IARC Working Group. *N Engl J Med* 2015;372:2353-8.
- Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. *BMJ* 2015;350:g7773.
- Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;102:605-13.
- Surveillance, Epidemiology, and End Results (SEER) Program. Research Data (1973-2015), National Cancer Institute, DCCPS, Surveillance Research Program, released April 2018, based on the November 2017 submission. Available online: <https://seer.cancer.gov/data/>
- Merrill RM, Dearden KA. How representative are the surveillance, epidemiology, and end results (SEER) program cancer data of the United States? *Cancer*

- Causes Control 2004;15:1027-34.
23. Zippin C, Lum D, Hankey BF. Completeness of hospital cancer case reporting from the SEER Program of the National Cancer Institute. *Cancer* 1995;76:2343-50.
 24. Dodd GD. American Cancer Society guidelines from the past to the present. *Cancer* 1993;72:1429-32.
 25. Mammography guidelines 1983: background statement and update of cancer-related checkup guidelines for breast cancer detection in asymptomatic women age 40 to 49. *CA Cancer J Clin* 1983;33:255.
 26. Mettlin C, Smart CR. Breast cancer detection guidelines for women aged 40 to 49 years: rationale for the American Cancer Society reaffirmation of recommendations. American Cancer Society. *CA Cancer J Clin* 1994;44:248-55.
 27. Leitch AM, Dodd GD, Costanza M, et al. American Cancer Society guidelines for the early detection of breast cancer: update 1997. *CA Cancer J Clin* 1997;47:150-3.
 28. Smith RA, Saslow D, Sawyer KA, et al. American Cancer Society guidelines for breast cancer screening: update 2003. *CA Cancer J Clin* 2003;53:141-69.
 29. Oeffinger KC, Fontham ET, Etzioni R, et al. Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update from the American Cancer Society. *JAMA* 2015;314:1599-614.
 30. U.S. Preventive Services Task Force Screening for breast cancer: recommendations and rationale. *Ann Intern Med* 2002;137:344-6.
 31. Nelson HD, Tyne K, Naik A, et al. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009;151:727-37, W237-42.
 32. Siu AL, U.S. Preventive Services Task Force Screening for breast cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* 2016;164:279-96.
 33. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;380:1778-86.
 34. Welch HG, Gorski DH, Albertsen PC. Trends in metastatic breast and prostate cancer — Lessons in cancer dynamics. *N Engl J Med* 2015;373:1685-7.
 35. Young JL Jr, Roffers SD, Ries LAG, et al. SEER Summary Staging Manual - 2000: Codes and Coding Instructions. Bethesda: National Institute of Health, NIH Pub, 2001.
 36. Rostgaard K, Vaeth M, Rootzén H, et al. Why did the breast cancer lymph node status distribution improve in Denmark in the pre-mammography screening period of 1978-1994? *Acta Oncol* 2010;49:313-21.
 37. Stamatovic L, Vasovic S, Trifunovic J, et al. Factors influencing time to seeking medical advice and onset of treatment in women who are diagnosed with breast cancer in Serbia. *Psychooncology* 2018;27:576-82.

Cite this article as: Yang L, Wang S, Huang Y. An exploration for quantification of overdiagnosis and its effect for breast cancer screening. *Chin J Cancer Res* 2020;32(1):26-35. doi: 10.21147/j.issn.1000-9604.2020.01.04

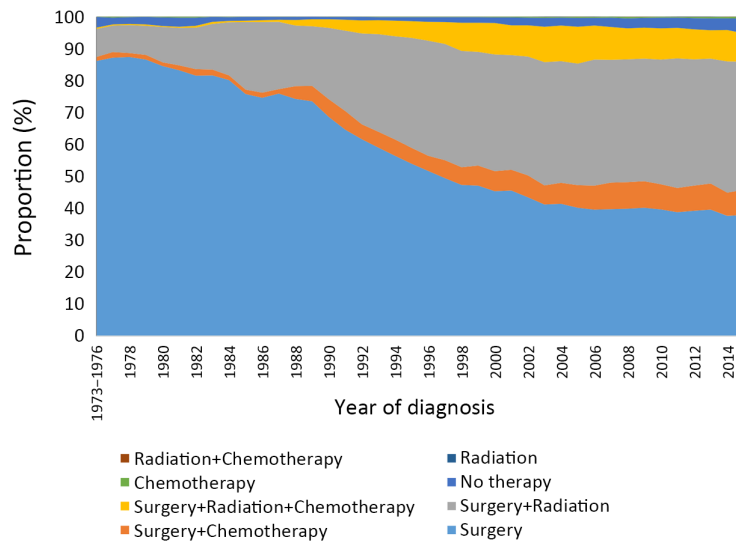


Figure S1 Trend of therapy patterns for early-stage breast cancer patients during the past two decades.

Table S1 Guidelines for breast cancer screening from American Cancer Society

Years	40–49 years old	≥50 years old	References
1977	If women themselves, or their mothers or sisters, had breast cancer	Annual screening	(24). <i>Cancer</i> 1993;72:1429-32.
1983	Screening with intervals of 1–2 years	Annual screening	(25). <i>CA Cancer J Clin</i> 1983;33:255.
1994	Screening with intervals of 1–2 years	Annual screening as long as a woman is in good health	(26). <i>CA Cancer J Clin</i> 1994;44:248-55.
1997	Annual screening	Annual screening as long as a woman is in good health	(27). <i>CA Cancer J Clin</i> 1997;47:150-3.
2003	Annual screening	Annual screening as long as a woman is in good health	(28). <i>CA Cancer J Clin</i> 2003;53:141-69.
2015	40–44 years old: should have the opportunity to begin annual screening; 45–49 years old: annual screening	50–54 years old: annual screening; ≥55 years old: biennial screening or have the opportunity to continue screening annually as long as their overall health is good and they have a life expectancy of 10 years or longer	(29). <i>JAMA</i> 2015;314:1599-614.

Table S2 Guidelines for breast cancer screening from U.S. Preventive Services Task Force

Years	40–49 years old	50–74 years old	≥75 years old	References
2002	Screening with intervals of 1–2 years	Screening with intervals of 1–2 years	Screening with intervals of 1–2 years	(30). <i>Ann Intern Med</i> 2002;137:344-6.
2009	Do not screen routinely. Individualize decision to begin biennial screening according to the patient’s context and values	Screen every 2 years	No recommendation	(31). <i>Ann Intern Med</i> 2009;151:727-37.
2016	The decision to start screening should be an individual one	Screen every 2 years	No recommendation	(32). <i>Ann Intern Med</i> 2016;164:279-96.

Table S3 Comparisons on cumulative survival rates of early-stage breast cancers cohort diagnosed at different years with Cox regression after adjusting age at diagnosis

Year at diagnosis	Number of cancers	Cumulatively survival rates [% (P values for comparison to 1999 cohort)]										
		5-year	6-year	7-year	8-year	9-year	10-year	11-year	12-year	13-year	14-year	15-year
1973–1976	16,891	80.6 (<0.001)	76.8 (<0.001)	73.6 (<0.001)	70.5 (<0.001)	67.4 (<0.001)	64.3 (<0.001)	61.6 (<0.001)	58.9 (<0.001)	56.0 (<0.001)	53.6 (<0.001)	51.1 (<0.001)
1977	4,602	80.0 (<0.001)	76.3 (<0.001)	72.6 (<0.001)	69.1 (<0.001)	65.8 (<0.001)	62.5 (<0.001)	59.3 (<0.001)	56.5 (<0.001)	54.0 (<0.001)	51.6 (<0.001)	49.1 (<0.001)
1978	4,650	79.2 (<0.001)	75.4 (<0.001)	71.4 (<0.001)	68.1 (<0.001)	64.4 (<0.001)	61.8 (<0.001)	58.7 (<0.001)	55.9 (<0.001)	53.3 (<0.001)	50.7 (<0.001)	48.2 (<0.001)
1979	4,832	79.3 (<0.001)	75.4 (<0.001)	71.6 (<0.001)	68.6 (<0.001)	64.8 (<0.001)	61.6 (<0.001)	59.0 (<0.001)	56.4 (<0.001)	53.9 (<0.001)	51.1 (<0.001)	48.7 (<0.001)
1980	4,948	78.7 (<0.001)	74.6 (<0.001)	70.8 (<0.001)	67.3 (<0.001)	64.3 (<0.001)	61.3 (<0.001)	58.4 (<0.001)	55.5 (<0.001)	53.3 (<0.001)	50.7 (<0.001)	48.2 (<0.001)
1981	5,189	79.1 (<0.001)	75.2 (<0.001)	71.2 (<0.001)	67.6 (<0.001)	64.4 (<0.001)	60.6 (<0.001)	57.7 (<0.001)	55.0 (<0.001)	51.9 (<0.001)	49.2 (<0.001)	46.6 (<0.001)
1982	5,337	78.9 (<0.001)	75.1 (<0.001)	71.7 (<0.001)	68.4 (<0.001)	65.6 (<0.001)	62.6 (<0.001)	59.1 (<0.001)	56.0 (<0.001)	53.2 (<0.001)	50.7 (<0.001)	48.6 (<0.001)
1983	5,871	80.6 (<0.001)	76.6 (<0.001)	73.0 (<0.001)	69.6 (<0.001)	66.7 (<0.001)	63.6 (<0.001)	60.2 (<0.001)	57.1 (<0.001)	54.4 (<0.001)	51.5 (<0.001)	49.2 (<0.001)
1984	6,594	82.3 (<0.001)	78.5 (<0.001)	74.9 (<0.001)	71.6 (<0.001)	68.0 (<0.001)	64.3 (<0.001)	61.3 (<0.001)	59.0 (<0.001)	56.1 (<0.001)	53.6 (<0.001)	50.9 (<0.001)
1985	7,673	82.2 (<0.001)	78.7 (<0.001)	75.3 (<0.001)	72.1 (<0.001)	68.8 (<0.001)	65.9 (<0.001)	62.9 (<0.001)	60.5 (<0.001)	57.8 (<0.001)	55.0 (<0.001)	52.4 (<0.001)
1986	8,438	83.4 (<0.001)	80.3 (<0.001)	77.0 (<0.001)	73.8 (<0.001)	70.7 (<0.001)	67.5 (<0.001)	64.6 (<0.001)	61.8 (<0.001)	58.8 (<0.001)	56.2 (<0.001)	52.9 (<0.001)
1987	9,895	84.7 (<0.001)	81.1 (<0.001)	77.9 (<0.001)	74.9 (<0.001)	71.6 (<0.001)	68.7 (<0.001)	65.5 (<0.001)	62.7 (<0.001)	59.5 (<0.001)	57.0 (<0.001)	54.2 (<0.001)
1988	10,113	85.7 (<0.001)	82.3 (<0.001)	79.2 (<0.001)	76.2 (<0.001)	72.9 (<0.001)	70.0 (<0.001)	66.8 (<0.001)	63.6 (<0.001)	60.9 (<0.001)	58.0 (<0.001)	55.4 (<0.001)
1989	9,849	85.7 (<0.001)	82.2 (<0.001)	79.3 (<0.001)	76.2 (<0.001)	73.2 (<0.001)	70.2 (<0.001)	67.4 (<0.001)	64.0 (<0.001)	61.3 (<0.001)	58.4 (<0.001)	55.7 (<0.001)
1990	10,847	85.2 (<0.001)	81.9 (<0.001)	78.5 (<0.001)	75.5 (<0.001)	72.5 (<0.001)	69.6 (<0.001)	66.5 (<0.001)	63.7 (<0.001)	60.8 (<0.001)	57.7 (<0.001)	54.8 (<0.001)
1991	11,369	85.5 (<0.001)	82.5 (<0.001)	79.4 (<0.001)	76.1 (<0.001)	73.0 (<0.001)	69.8 (<0.001)	67.2 (<0.001)	64.3 (<0.001)	61.5 (<0.001)	58.8 (<0.001)	56.3 (<0.001)
1992	11,927	85.1 (<0.001)	82.0 (<0.001)	78.8 (<0.001)	76.0 (<0.001)	72.8 (<0.001)	70.0 (<0.001)	67.2 (<0.001)	64.3 (<0.001)	61.4 (<0.001)	58.9 (<0.001)	56.0 (<0.001)
1993	11,917	85.4 (<0.001)	82.4 (<0.001)	79.2 (<0.001)	76.1 (<0.001)	73.2 (<0.001)	70.4 (<0.001)	67.3 (<0.001)	64.6 (<0.001)	61.9 (<0.001)	59.3 (<0.001)	56.7 (<0.001)
1994	12,527	86.5 (0.002)	83.7 (0.001)	80.8 (0.001)	78.0 (0.001)	75.1 (0.001)	72.1 (<0.001)	69.0 (<0.001)	66.1 (<0.001)	63.3 (<0.001)	60.6 (<0.001)	58.1 (<0.001)
1995	13,208	86.7 (0.005)	83.7 (0.001)	80.6 (<0.001)	77.6 (<0.001)	74.6 (<0.001)	71.8 (<0.001)	69.0 (<0.001)	66.1 (<0.001)	63.3 (<0.001)	60.8 (<0.001)	58.4 (<0.001)
1996	13,608	86.2 (<0.001)	83.1 (<0.001)	80.5 (<0.001)	77.6 (<0.001)	75.0 (<0.001)	72.3 (<0.001)	69.6 (<0.001)	67.1 (<0.001)	64.3 (<0.001)	61.8 (<0.001)	59.2 (<0.001)
1997	14,637	87.3 (0.181)	84.5 (0.127)	81.8 (0.122)	79.2 (0.240)	76.2 (0.117)	73.5 (0.108)	70.6 (0.072)	67.8 (0.013)	65.0 (0.007)	62.6 (0.022)	60.0 (0.022)
1998	15,905	87.5 (0.296)	84.8 (0.421)	82.0 (0.333)	79.1 (0.201)	76.4 (0.223)	73.5 (0.095)	70.8 (0.139)	68.3 (0.090)	65.5 (0.039)	62.8 (0.030)	59.8 (0.007)
1999	16,021	87.8 (1.000)	85.2 (1.000)	82.5 (1.000)	79.8 (1.000)	77.1 (1.000)	74.4 (1.000)	71.7 (1.000)	69.4 (1.000)	66.8 (1.000)	64.1 (1.000)	61.5 (1.000)

Table S3 (continued)

Table S3 (continued)

Year at diagnosis	Number of cancers	Cumulatively survival rates [% (P values for comparison to 1999 cohort)]										
		5-year	6-year	7-year	8-year	9-year	10-year	11-year	12-year	13-year	14-year	15-year
2000	15,786	88.1 (0.957)	85.5 (0.850)	83.1 (0.558)	80.4 (0.508)	77.8 (0.459)	75.3 (0.334)	72.4 (0.609)	69.5 (0.743)	66.9 (0.555)	64.2 (0.485)	
2001	16,282	87.9 (0.712)	85.3 (0.493)	82.7 (0.375)	80.0 (0.352)	77.5 (0.159)	74.8 (0.218)	72.2 (0.161)	69.3 (0.579)	66.7 (0.614)		
2002	16,448	88.8 (0.147)	86.1 (0.221)	83.5 (0.200)	81.0 (0.160)	78.4 (0.122)	75.9 (0.132)	73.4 (0.068)	70.7 (0.262)			
2003	15,529	88.4 (0.433)	86.0 (0.289)	83.5 (0.204)	81.0 (0.106)	78.8 (0.016)	76.1 (0.034)	73.5 (0.023)				
2004	16,042	88.9 (0.077)	86.4 (0.058)	84.0 (0.019)	81.5 (0.009)	79.0 (0.009)	76.5 (0.006)					
2005	16,182	89.3 (0.005)	86.9 (0.005)	84.3 (0.006)	82.1 (0.001)	79.7 (<0.001)						
2006	16,641	89.5 (0.001)	87.1 (<0.001)	84.7 (<0.001)	82.3 (<0.001)							
2007	17,312	90.4 (<0.001)	88.0 (<0.001)	85.8 (<0.001)								
2008	18,023	90.2 (<0.001)	87.8 (<0.001)									
2009	18,878	90.1 (<0.001)										