Population-Based Breast Cancer Screening With Risk-Based and Universal Mammography Screening Compared With Clinical Breast Examination
A Propensity Score Analysis of 1429890 Taiwanese Women

Amy Ming-Fang Yen, PhD; Huei-Shian Tsau, PhD; Jean Ching-Yuan Fann, PhD; Sam Li-Sheng Chen, PhD; Sherry Yueh-Hsia Chiu, PhD; Yi-Chia Lee, PhD; Shin-Liang Pan, PhD; Han-Mo Chiu, PhD; Wen-Hong Kuo, PhD; King-Jen Chang, PhD; Yi-Ying Wu, PhD; Shu-Lin Chiang, PhD; Chien-Yen Wu, PhD; Tsai-Jung Lin, MD; Jui-Ping Liu, MD; Ching-Lung Hung, MD; Shu-Lih Chia, MS; Mei-Ju Chen, MS; Hsiu-Hsi Chen, PhD; Shu-Ti Chiou, PhD

IMPORTANCE Different screening strategies for breast cancer are available but have not been researched in quantitative detail.

OBJECTIVE To assess the benefits and the harms of risk-based and universal mammography screening in comparison with annual clinical breast examination (CBE).

DESIGN Population-based cohort study comparing incidences of stage II+ disease and death from breast cancer across 3 breast cancer screening strategies, with adjustment for a propensity score for participation based on risk factors for breast cancer and comparing the 3 strategies for overdetection between January 1999 and December 2009. Asymptomatic women attending outreach screening in the community or undergoing mammography in hospitals were enrolled in the 3 screening programs.

INTERVENTIONS Risk-based biennial mammography, universal biennial mammography, and annual CBE.

MAIN OUTCOMES AND MEASURES Detection rates, stage II+ disease incidence, mortality from breast cancer, and overdiagnosis were compared using a time-dependent Cox proportional hazards regression model.

RESULTS A total of 1429890 asymptomatic women attending outreach screening in the community or undergoing mammography in hospitals were enrolled in the 3 screening programs. Detection rates (prevalent screen and subsequent screens per 1000) were the highest for universal biennial mammography (4.86 and 2.98, respectively), followed by risk-based mammography (2.80 and 2.77, respectively), and lowest for annual CBE (0.97 and 0.70, respectively). Universal biennial mammography screening, compared with annual CBE, was associated with a 41% mortality reduction (risk ratio, 0.59; 95% CI, 0.48-0.73) and a 30% reduction of stage II+ breast cancer (RR, 0.70; 95% CI, 0.66-0.74). Risk-based mammography screening was associated with an 8% reduction of stage II+ breast cancer (RR, 0.92; 95% CI, 0.86-0.99) but was not associated with a statistically significant mortality reduction (risk ratio [RR], 0.86; 95% CI, 0.73-1.02). Estimates of overdiagnosis were no different from CBE for risk-based screening and 13% higher than CBE for universal mammography.

CONCLUSIONS AND RELEVANCE Compared with population-based screening for breast cancer with annual CBE, universal biennial mammography resulted in a substantial reduction in breast cancer deaths, whereas risk-based biennial mammography resulted in only a modest benefit. Compared with annual CBE, risk-based and universal mammography screening did not result in significant overdiagnosis of breast cancer.

Published online March 31, 2016.

Copyright 2016 American Medical Association. All rights reserved.
Population-based organized breast cancer screening has increased worldwide based in part on supporting evidence of the efficacy of mammography screening observed in randomized clinical trials conducted between 1970 and 1990.\(^1,3\) and in part on the significant global increase in breast cancer incidence. In Taiwan, the incidence rate of breast cancer doubled from 30 per 100,000 in 1995 to 59 per 100,000 in 2003.\(^4\) Since population-based, organized service screening programs are influenced by cultural-specific factors, the infrastructure of the health care system, and economic concerns over limited resources, different approaches to breast cancer screening have been implemented, including clinical breast examination (CBE), risk-based mammography screening, and universal mammography screening.

Most evaluations of mass screening for breast cancer have focused on a single detection method, and very few studies have been conducted to address comparisons of benefits and harms across different detection methods, especially in the same population. Although universal mammography screening is the most widely adopted approach to population-based screening, comparative effectiveness research focused on different screening strategies is still important owing to the enduring debate over the value of mammography screening. While some systematic reviews have concluded that mammography screening is associated with a 20% to 25% reduction in breast cancer deaths,\(^5-10\) other reviews claim that there is no benefit\(^11,12\) and also express concerns that universal mammography screening leads to significant harms, specifically overdiagnosis, sometimes called overdetection of breast cancer.

These claims have caused health care decision makers to face some uncertainty regarding whether universal mammography screening is the best way to screen for breast cancer. We conduct a large, population-based cohort study to clarify concerns about limited benefits and excessive harms and to explore the possibility of alternative approaches to reduce avoidable deaths from breast cancer and also to help health care policy makers to determine the best screening strategy. The aim of this study is to present basic background information, screening performance data, and outcomes of each screening program, and to compare effectiveness in reducing advanced stages of breast cancer, death from breast cancer, and the estimates of overdiagnosis associated with each screening strategy.

### Methods

The study was certified as exempt from institutional review board review by the Research Ethics Committee of the National Taiwan University Hospital, also waiving written informed consent.

#### Participants Experiencing 1, 2, or All 3 Screening Programs

To evaluate and compare the effectiveness of 3 screening programs (annual CBE, risk-based biennial mammography based on qualifying questionnaire results, and universal biennial mammography), we classified 1,429,890 women involved with the Taiwanese breast cancer mass screening programs in chronological order from 1999 to 2009, identifying 7 types of screening history (Figure). Three patient groups underwent only a single screening method, including only CBE (group 1, \(n = 641,735\)); only risk-based mammography (group 2, \(n = 126,698\)); and only universal mammography (group 3, \(n = 350,514\)). Three groups underwent 2 screening methods, including risk-based and universal mammography (group 4, \(n = 56,082\)), CBE and universal mammography (group 5, \(n = 139,307\)); and CBE and risk-based mammography (group 6, \(n = 67,112\)). Finally, 1 group underwent all 3 screening methods (group 7, \(n = 48,442\)). These groups are all illustrated in the Figure.

#### Annual CBE for Breast Cancer

The first large-scale outreach screening program for breast cancer used annual CBE. The detailed results have been published previously.\(^18\) In brief, among 4,944,715 eligible women 35 years or older, 896,596 (18.13%) were screened between 1999 and 2001. Women with preexisting clinical signs or symptoms related to breast cancer (such as tenderness, nipple discharge, and mass) and those who had a history of breast cancer were excluded from this screening program. Among the 896,596 women who underwent screening, approximately 12% (108,046) underwent at least 1 repeated screen over a period of 3 consecutive years, for a total of 115,640 subsequent screening episodes. Women with suspected cancer after CBE (58,085 [6.48%] and 11,666 [10.09%] at prevalent and subsequent screenings, respectively), were referred to receive routine clinical diagnostic workup. Data on breast cancers, including in situ and invasive cases, were obtained from linkage of this cohort with the national cancer registry and are summarized in Table 1.

It should be noted that women participating in the CBE screening program could subsequently undergo risk-based and/or universal biennial mammography screening (subgroups A, B, and D in the Figure). They would, of course, be older when participating in the later programs. We therefore included women aged 40 to 69 years at the time of study entry but made adjustments for year of birth and propensity score in the subsequent analysis to render the comparison of effectiveness across the 3 screening methods as valid as possible.

#### Risk-Based Biennial Mammography Screening

The second population-based screening program was a risk-based biennial mammography screening program.\(^19\) The reasons for adopting a risk-based strategy were the low breast cancer incidence rate in Taiwan at that time (2002-2004) and
concerns about the clinical capacity to deliver population-based screening. Risk-based mammography screening was conducted between January 2002 and June 2004, with a 2-year interscreening interval.

The first stage of the program was to identify the high-risk group using a median risk score as a cutoff. The risk score was computed from conventional risk factors, specifically reproductive and menstrual history, and family history, derived from the questionnaire administered to women who took part in the CBE program between 1999 and 2001. Among 1934981 eligible women aged 50 to 69 years, 298334 women were screened in the first stage based on questionnaire results. The subcohorts of these participants illustrated in the Figure consist only of women who were screened in the risk-based screen-

Table 1. Characteristics of Mass Breast Cancer Screening Programs and Patients in Taiwan Between 1999 and 2009

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Annual Clinical Breast Examination</th>
<th>Risk-Based Biennial Mammography</th>
<th>Universal Biennial Mammography</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prevalent</td>
<td>Subsequent</td>
<td>Prevalent</td>
</tr>
<tr>
<td>Eligible populationa</td>
<td>4944715</td>
<td>ND</td>
<td>1934981</td>
</tr>
<tr>
<td>Screened women</td>
<td>896596</td>
<td>115640</td>
<td>298334</td>
</tr>
<tr>
<td>Screening rate, %b</td>
<td>18.13</td>
<td>ND</td>
<td>15.42</td>
</tr>
<tr>
<td>Positive ratec</td>
<td>6.48</td>
<td>10.09</td>
<td>44.5</td>
</tr>
<tr>
<td>Recall rate 1, %d</td>
<td>ND</td>
<td>ND</td>
<td>8.7</td>
</tr>
<tr>
<td>Recall rate 2, %e</td>
<td>ND</td>
<td>ND</td>
<td>BI-RADS 0 7.23 7.38</td>
</tr>
<tr>
<td>BI-RADS 4</td>
<td>ND</td>
<td>ND</td>
<td>1.24</td>
</tr>
<tr>
<td>BI-RADS 5</td>
<td>ND</td>
<td>ND</td>
<td>0.23</td>
</tr>
<tr>
<td>Biopsy rate, %f</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Breast cancer cases, No.</td>
<td>867</td>
<td>81</td>
<td>216</td>
</tr>
<tr>
<td>PPV, %g</td>
<td>1.49</td>
<td>0.69</td>
<td>3.22</td>
</tr>
<tr>
<td>Detection rate, ‰h</td>
<td>0.97</td>
<td>0.70</td>
<td>2.80</td>
</tr>
</tbody>
</table>

Abbreviations: NA, not applicable; ND, no data; PPV, positive predictive value.

a Average population between 1999 and 2009 reported by national census data.
b Screening rate was calculated as the number of screening examinations performed divided by the eligible population.
c Positive rate was calculated as the number of positive findings on clinical breast examination or risk scores above the cutoff value for risk-based screening divided by the number of participants.
d Recall rate 1 was calculated as the number of cases with BI-RADS scores of 0, 4, or 5 divided by the number of participants.
e Recall rate 2 was calculated similarly to recall rate 1 but using the respective individual BI-RADS scores.
f Biopsy rate was calculated as the total number of cases from final assessment with BI-RADS scores 4 and 5 that underwent pathological diagnosis divided by the number of screening examinations.
g The PPV was calculated as the number of screen-detected breast cancers divided by number of positive findings (ie, abnormal finding in clinical breast examination screening or a BI-RADS score of 0, 4, or 5 in risk-based and universal biennial mammography screening).
h Detection rate was calculated as the number of screen-detected breast cancers divided by number of participants.
Risk-Based and Universal Mammography vs Clinical Breast Examination

Overdiagnosis Associated With Mammography Screening

The detection rates (prevalent screenings and subsequent screenings per 1000) were highest for universal biennial mammography (4.86 and 2.98, respectively), followed by risk-based biennial mammography (2.80 and 2.77, respectively) and lowest for annual CBE (0.97 and 0.70, respectively). Similar rankings were noted for positive predictive value (Table 1).

Stage Shifting Resulting From Screening

Table 2 summarizes the TNM stage distribution for the 3 screening programs. There was a clear downstaging shift for universal biennial mammography, but only a negligible shift was observed in the risk-based screening program compared with CBE.

Results

Screening Findings

Table 1 summarizes the basic findings from the 3 screening programs. The details of each indicator, with the exception of the detection rate, are described in the eAppendix in the Supplement. Age-specific findings with 5-year brackets across the 3 screening programs were similar and are reported in eTable 1 in the Supplement.

Overall, the detection rates (prevalent screenings and subsequent screenings per 1000) were highest for universal biennial mammography (4.86 and 2.98, respectively), followed by risk-based biennial mammography (2.80 and 2.77, respectively) and lowest for annual CBE (0.97 and 0.70, respectively). Similar rankings were noted for positive predictive value (Table 1).

Stage Shifting Resulting From Screening

Table 2 summarizes the TNM stage distribution for the 3 screening programs. There was a clear downstaging shift for universal biennial mammography, but only a negligible shift was observed in the risk-based screening program compared with CBE.

Overdiagnosis Associated With Mammography Screening

We compared the cumulative incidence of breast cancer cases in the 3 screening programs. Compared with those who par-
Risk-Based and Universal Mammography vs Clinical Breast Examination

Table 2. Stage Distribution of Breast Cancer by Three Screening Programs*  

<table>
<thead>
<tr>
<th>TNM Stage</th>
<th>Annual Clinical Breast Examination</th>
<th>Risk-Based Biennial Mammography</th>
<th>Universal Biennial Mammography</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>569 (11.1)</td>
<td>199 (12.1)</td>
<td>870 (18.8)</td>
<td>1638</td>
</tr>
<tr>
<td>I</td>
<td>1641 (31.9)</td>
<td>520 (31.5)</td>
<td>1537 (33.2)</td>
<td>3698</td>
</tr>
<tr>
<td>II</td>
<td>2027 (39.4)</td>
<td>631 (38.2)</td>
<td>1498 (32.4)</td>
<td>4156</td>
</tr>
<tr>
<td>III</td>
<td>816 (15.9)</td>
<td>275 (16.7)</td>
<td>612 (13.2)</td>
<td>1703</td>
</tr>
<tr>
<td>IV</td>
<td>90 (1.7)</td>
<td>26 (1.6)</td>
<td>107 (2.3)</td>
<td>223</td>
</tr>
<tr>
<td>II+, %</td>
<td>57.7</td>
<td>56.5</td>
<td>47.9</td>
<td>NA</td>
</tr>
<tr>
<td>Total No.</td>
<td>5143</td>
<td>1651</td>
<td>4624</td>
<td>11 418</td>
</tr>
</tbody>
</table>

Abbreviations: BC, breast cancer; CBE, clinical breast examination.

*Unless otherwise indicated, data are reported as number (percentage) of all breast cancers.

**Because pathological staging in the cancer registry during the period of clinical breast examination was not based on the entire cohort, only 5143 samples were abstracted for study, no information on TNM stage was available for 641 breast cancers in the risk-based biennial mammography screening group and 520 breast cancers in universal biennial mammography screening group.

Table 3. Time-Dependent Cox Regression Models*  

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Hazard Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
</tr>
<tr>
<td>Breast cancer death</td>
<td></td>
</tr>
<tr>
<td>Annual clinical breast exam.</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Risk-based biennial mammag.</td>
<td>0.91 (0.77-1.07)</td>
</tr>
<tr>
<td>Universal biennial mammag.</td>
<td>0.67 (0.54-0.82)</td>
</tr>
<tr>
<td>Stage II+ breast cancer</td>
<td></td>
</tr>
<tr>
<td>Annual clinical breast exam.</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Risk-based biennial mammag.</td>
<td>0.99 (0.93-1.06)</td>
</tr>
<tr>
<td>Universal biennial mammag.</td>
<td>0.82 (0.78-0.87)</td>
</tr>
</tbody>
</table>

Abbreviation: CBE, clinical breast examination.

*Model 1 was adjusted for birth year; model 2 was adjusted for both birth year and propensity score; model 3 was adjusted for both birth year and propensity score by decile.

Mortality and Stage II+ Breast Cancer

Table 3 lists the hazard ratios (HRs) of universal biennial mammography and the risk-based screening compared with CBE by a time-dependent Cox regression model with and without adjustment for propensity score (probability/decile) and year of birth. The crude HR of breast cancer deaths for the universal biennial mammography screening after adjusting for birth year was only 0.67 (95% CI, 0.54-0.82), indicating a 33% mortality reduction. The mortality reduction was 41% (95% CI, 27%-52%) for the universal biennial mammography screening after adjusting for propensity score and year of birth. A similar finding (38%; 95% CI, 24%-50%) was noted when the decile of propensity score was used. The mortality reduction was not significant for the risk-based biennial mammography screening compared with CBE. Model selection with the Akaike information criterion (AIC)25 shows that the model including the continuous propensity score (AIC = 25.301) was superior to the one that did not adjust for propensity score (AIC = 25.329) and to the one that adjusted the propensity score in deciles (AIC = 25.311).

The crude HR for incidence of stage II+ breast cancer was 0.82 (95% CI, 0.78-0.87) for universal biennial mammography screening compared with CBE. The figures with adjustment for propensity score were 0.70 (95% CI, 0.66-0.74) and 0.73 (95% CI, 0.69-0.77) for the continuous probability and decile of propensity score, respectively. A smaller reduction in stage II+ was found for risk-based mammography screening (Table 3).

Discussion

Our study provides evidence on the relative benefits and harms of 3 screening strategies for breast cancer in Taiwan. Our findings suggest that universal biennial screening yielded the largest benefit, reducing mortality from breast cancer by 41%, and that risk-based screening failed to yield a significant benefit compared with CBE. Overdiagnosis from mammography screening was, at worst, modest.

The significant mortality reduction revealed in this study supports the use of mammography for mass breast cancer screening in a country with a low but increasing trend in breast cancer incidence. Our findings are in contrast to the view advanced by some that there is little benefit from mammography screening13-16 and are consistent with the supporting view of the benefits of mammography screening reported in most systematic reviews and meta-analyses.6,7,9,26,27

The conflicting opinions in this enduring debate may not be as relevant to Asian countries because the context is dif-
different in Asian countries than it is in Western countries. The evaluation of the effectiveness of any screening method, including mammography, is highly dependent on background characteristics, attendance rates, awareness, quality assurance, and the evolution of mammography screening. Western countries have already experienced the benefits of increased awareness and the downstreaming of breast cancer as mammography use has evolved and spread, and so the relative benefit of mammography screening in these settings today may be smaller, but this is not the case for Asian countries. For example, in the Taiwanese program, the proportion of larger tumors identified (≥2 cm in diameter) was 68% in the initial period and slowly dropped to 59% by 2009, compared with the correspondingly lower figures during a comparable time frame in Western countries (39.5% in Ticino, Switzerland and 26% in Finland between 2006 and 2010).

While concerns of overdiagnosis have been raised when small breast tumors are detected, this problem is not serious in the current findings, with overdiagnosis rates estimated at only 13% for the universal biennial screening program, with an adjustment of only 3 years for lead time. Moreover, breast cancer awareness is low and downstreaming of breast cancer is still limited in Asian countries, which may account for why universal biennial mammography screenings resulted in a substantial mortality reduction from breast cancer. The marginal benefit is more easily achieved when mammography screening is new and relatively uncommon.

This analysis has 2 important limitations. First, follow-up time, particularly for the universal biennial mammography screening program, may be too short to yield an unbiased estimate. Long-term follow-up is required for the confirmation of the estimate of the effectiveness of biennial mammography screening in the current study. Second, it is reasonable to infer that some of the mortality reduction may be attributed to improvements in treatment, but we did not have individual information on treatment and therapy. However, to address this possibility, we compared the survival of 2 refuser groups (ie, those who did not attend CBE or universal biennial mammography screening) and found the survival of the refuser group in universal biennial mammography screening (83%) was better than that of the CBE (77%). Based on these survival figures, we translated the improvement in survival during this period likely attributable to improvements in treatment to the outcomes used in the time-dependent Cox proportional hazard regression model. In doing so, we found that the improvement in treatment accounted for 5% of the observed mortality reduction (adjusted HR inflated from 0.59 to 0.64 after adjusting for treatment effect). Thus, while improvements in treatment account for some of the benefit we observed, most is attributable to the benefit of mammography in downstaging breast cancer.

The implications for the generalizability of our findings to the countries worldwide are 2-fold. First, our finding on universal biennial mammography screening are similar to breast cancer mortality reductions observed in some Western randomized clinical trials. In addition, recent evaluations of breast cancer service screening programs such as the Pan-Canadian study reported findings similar to our estimated results for universal biennial mammography screening.

Our findings also confirm that mammography screening with good quality control is reproducible in organized breast cancer service screening programs implemented in countries with low or intermediate incidence but increasing incidence trends. Our findings on overdiagnosis also indicate that little harm is caused by universal biennial mammography screening. These findings are consistent with previous studies with adjustment for lead time but at odds with the highly elevated estimates of overdiagnosis attributable to mammography screening when there has been no adjustment for lead time.

Conclusions

In conclusion, our evaluation of benefits and harms of breast cancer screening with 3 screening methods in a large population-based cohort in Taiwan shows that universal biennial mammography was the most effective strategy for detecting breast cancer early: it achieved a 40% mortality reduction through the reduction in stage II+ disease. In contrast, implementation of risk-based mammography screening, while seemingly pragmatic in the presence of limited resources, resulted in only a small benefit compared with CBE. Universal biennial mammography screening also was associated with only a modest level of overdiagnosis. The results of this comparison of different screening approaches applied to the same population should be informative to health policy makers seeking to determine if and how they might initiate breast cancer screening programs.


Conflict of Interest Disclosures: None reported.

Funding/Support: This study was supported by Health Promotion Administration, Ministry of Health and Welfare, Taiwan.

Role of the Funder/Sponsor: The supporting institution had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: The authors are obliged to the following professional scientists in the area of breast cancer screening for their contributions to editing the final manuscript: Stephen W. Duffy, MS, Wolfson Institute of Preventive Medicine, Queen Mary University of London; Robert A. Smith, PhD, Cancer Control Science Department, American Cancer Society; and Paula Gordon, PhD, BC Women’s Hospital, Vancouver, British Columbia, Canada. They received no compensation for their contributions.

REFERENCES


