Individually tailored screening of breast cancer with genes, tumour phenotypes, clinical attributes, and conventional risk factors

Y-Y Wu¹, M-F Yen², C-P Yu³,4 and H-H Chen*,¹

¹Graduate Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, 5F. No. 17, Hsu Chow Road, Zhongzheng District, Taipei City 100, Taiwan, ROC; ²School of Oral Hygiene, Oral Medicine of College, Taipei Medical University, No. 250, Wu-Hsing Street, Sinyi District, Taipei City 110, Taiwan, ROC; ³Department of Pathology, Tri-Service General Hospital, No. 325, Sec. 2, Chenggong Road, Neihu District, Taipei City 114, Taiwan, ROC and ⁴Department of Pathology, and Graduate Institute of Pathology and Parasitology, National Defense Medical Center, No. 161, Sec. 6, Minquan E. Road, Neihu District, Taipei City 114, Taiwan, ROC

Background: We demonstrated how to comprehensively translate the existing and updated scientific evidence on genomic discovery, tumour phenotype, clinical features, and conventional risk factors in association with breast cancer to facilitate individually tailored screening for breast cancer.

Methods: We proposed an individual-risk-score-based approach that translates state-of-the-art scientific evidence into the initiators and promoters affecting onset and subsequent progression of breast tumour underpinning a novel multi-variable three-state temporal natural history model. We applied such a quantitative approach to a population-based Taiwanese women periodical screening cohort.

Results: Risk prediction for pre-clinical detectable and clinical-detected breast cancer was made by the two risk scores to stratify the underlying population to assess the optimal age to begin screening and the inter-screening interval for each category and to ascertain which high-risk group requires an alternative image technique. The risk-score-based approach significantly reduced the interval cancer rate as a percentage of the expected rate in the absence of screening by 30% and also reduced 8.2% false positive cases compared with triennial universal screening.

Conclusion: We developed a novel quantitative approach following the principle of translational research to provide a roadmap with state-of-the-art genomic discovery and clinical parameters to facilitate individually tailored breast cancer screening.
With the advent of genetic and biological markers for breast cancer, individually tailored screening for breast cancer can now be achieved by making use of information on genes, conventional risk factors, clinical attributes, and relevant tumour phenotypes such as HER-2/neu. However, it is very rare to focus on how to link these scientific discoveries with temporal natural history to assess which factors (initiators) affect the onset of breast cancer, that is, the transition from free-of-breast-cancer to the PCDP and which ones (promoters) responsible for the subsequent progression from the PCDP to the CP by comparing the distribution of each factor between screen-detected cases and clinically detected ones with the premise that the distribution of initiators should be identical between both the detection modes.

To this end, we conducted a literature review on state-of-the-art genomic discovery, tumour phenotypes, clinical correlates, and conventional risk factors. These factors were further superimposed into the three-state temporal natural history of breast cancer in the light of the roles of initiators and promoters mentioned above to build a multi-variable and three-state Markov regression model. We then estimated the effects of initiators and promoters to construct composite risk scores by the application of the proposed Markov regression model to the data set generated above. We applied this risk-score-based approach to emulating a scenario of Taiwanese population-based periodic screening programme for breast cancer to demonstrate how such a novel quantitative approach can be used to develop an individually tailored screening strategy that is efficient in terms of inter-screening interval, age at the start of screening and the expedient use of costly but accurate alternative image techniques.

**MATERIALS AND METHODS**

**Study framework on initiators and promoters.** As shown in Figure 1, our study design was to quantify the effects of initiators on the occurrence of breast cancer in the PCDP and those of promoters on the transition from the PCDP to CP. The definitions of these three-state outcomes (free of breast cancer (state 1), breast cancer in the PCDP (state 2, sometimes called pre-symptomatic cases), and breast cancer in the CP (state 3, sometimes called symptomatic cases)) are realised by the following three corresponding outcomes, free of breast cancer, screen-detected cases and interval cancer obtained from population-based screening data as defined in the previous studies (Shapiro et al., 1974; Prorok, 1976; Walter and Day, 1983; van Oortmarssen et al., 1990; Duffy et al., 1995; Chen et al., 1997a, 1997b). The initiators in Figure 1 included BRCA1, BRCA2 (Peto et al., 1999; Chen and Parmigiani, 2007), single-nucleotide polymorphisms (Cox et al., 2007; Easton et al., 2007; Hunter et al., 2007; Stacey et al., 2007), breast density (del Carmen et al., 2007; Martin and Boyd, 2008; Cummings et al., 2009), and conventional risk factors (body mass index (BMI) and age at first full-term pregnancy (AP)) (Hsieh et al., 2002; Chen et al., 2004). The promoters were identified by comparing the distributions of relevant tumour phenotypes between screen-detected cases and interval cancer without making any assumption because the initiators affecting the occurrence of breast cancer would not lead

![Figure 1. Study design ascertaining outcomes from three rounds of screens based on the temporal natural history of breast cancer. BC, breast cancer; Ki-67, Ki-67 proliferation; HER-2/neu, HER-2/neu immunohistochemistry score.](image-url)
to a disparity between screen-detected cases and interval cancer (Dong et al., 2008). According to the literatures of previous studies (Hsieh et al., 2002; Dong et al., 2008), the promoters responsible for the progression from the PCDP to the CP in the current model included BMI, AP, oestrogen receptor (ER), HER-2 immunochimical score, and Ki-67 expression.

**Data sources.** We emulated an organised service periodical screening of 1 million Taiwanese women aged 50 years or older assuming the distributions of covariates by various detection methods are identical to the first box in Figure 1. To fit in with this scenario, we tuned baseline pre-clinical incidence rate following age-specific incidence rate, ranging from 0.0017 to 0.021 per women-year during 2001–2005 in Taiwan (Taiwan Cancer Registry Annual Report, 2009). Since it is impossible to have a comprehensive empirical data to directly estimate the parameters underpinning the model, we turned, by using simulation technique, these empirical tabular data obtained from literatures into individual data as if they were collected according to the concept of sufficient statistics (Casella and Berger, 2001). For example, the sufficient statistics of generating data on breast density are the total numbers of each subgroup, which is equivalent to the collection of individual data when a multi-nominal distribution following the percentage of each category. The data on other covariates were spawned in a similar manner. After generating data set with covariates, we further applied a three-state Markov regression model to add information on various detection methods (including screen-detected cases and interval cancer) to capture time to enter the PCDP and time to the CP by using relative risk as specified in Tables 1 and 2. Data on three rounds of screen were generated as illustrated in Figure 1. This completes the creation of data set for parameter estimation (the second box of Figure 2).

**Statistical analysis**

*Estimation of parameters.* The data created above were used to build up a series of likelihood functions based on our proposed Markov regression model (the third box of Figure 2). These details on the formula of likelihood function (the fourth box of Figure 2) are similar to those used in the previous studies (Hsieh et al., 2002).

**Composite risk scores.** The maximum likelihood estimation method was used to calculate risk scores (Chen et al., 2000; Hsieh et al., 2002) (the fourth box of Figure 2). The risk scores are the natural logarithms of the estimated transition rates for initiators and promoters. The risk stratification by deciles of risk scores is constructed to demonstrate the following individually tailored screening strategies. The median value of 10-year breast cancer risk for the average-risk population at age 50 was used as a criterion to recommend the suitable starting screening age for other deciles, similar to Pharoah et al. (2008).

**Model validation.** Figure 2 also shows model validation including cross-validation and external validation. The former was to divide full data set into the two parts, 2/3 trained data for estimating parameters and 1/3 validated data that were calculated by the application of parameters estimated from 2/3 trained data to the distributions of covariates of the other 1/3 validated data. We then compared the observed with the predicted number of breast cancer by different detection modes (first screen, subsequent screen, and interval cancer).

As far as external validation is concerned, we applied the parameters estimated from full data set we simulated from the empirical tabular data published in literatures to Taiwanese multi-centre screening (TAMCAS). The study design and results have been elaborated in previous studies (Lai et al., 1998; Chen et al., 2000; Wu et al., 2008). 4867 women with family history of breast cancer were invited and attended the first screen. We ascertainment 76 breast cancer cases out of 4867 women. The observed breast cancer cases and number of women free of breast cancer identified at subsequent screen and interval cancers are listed in Table 3. Since different populations have different underlying incidence rates of breast cancer, the baseline pre-clinical incidence rate (in the absence of characteristics of relevant covariates) should be tuned to accommodate the underlying incidence rate of breast cancer for the target population of interest when the proposed Markov regression model is attempted. In our TAMCAS study, we let the pre-clinical incidence rate follow the Weibull distribution

### Table 1. Percentage and relative risks for initiators affecting the transition rate from free-of-breast cancer to the pre-clinical screen-detectable phase (PCDP)

| Variables                  | Percentage (%) | Relative risk of being breast cancer in the PCDP (pre-symptomatic cases) | References                  |
|----------------------------|----------------|------------------------------------------------------------------------|-----------------------------|
| Gene mutation              |                |                                                                        |                             |
| BRCA1 gene mutation        | 0.11           | 20.26                                                                  | Peto et al, 1999; Chen and Parmigiani, 2007 |
| BRCA2 gene mutation        | 0.12           | 15.44                                                                  |                             |
| Breast density (BI-RADS)   |                |                                                                        |                             |
| Almost entirely fat         | 0.92           | 1.00                                                                   | del Carmen et al, 2007; Cumming et al, 2009 |
| Scattered fibroglandular densities | 15.86         | 2.03                                                                   |                             |
| Heterogeneously dense      | 56.09          | 2.95                                                                   |                             |
| Extremely dense            | 27.13          | 4.03                                                                   |                             |
| Polygenes (dbSNP no.) per allele |             |                                                                        |                             |
| rs2981582                  | 38             | 1.26                                                                   | Easton et al, 2007; Hunter et al, 2007 |
| rs3803662                  | 25             | 1.20                                                                   | Easton et al, 2007          |
| rs889312                   | 28             | 1.13                                                                   | Easton et al, 2007          |
| rs3817198                  | 30             | 1.07                                                                   | Easton et al, 2007          |
| rs13281615                 | 40             | 1.08                                                                   | Stacey et al, 2007          |
| rs13387042                 | 50             | 1.20                                                                   | Cox et al, 2007             |
| rs1045485                  | 86             | 1.13                                                                   | Easton et al, 2007          |
| BMI (kg m⁻²)               |                |                                                                        |                             |
| ≤ 23                      | 37.11          | 1.00                                                                   | Chen et al, 2004            |
| > 23                      | 62.89          | 2.59                                                                   | Hsieh et al, 2002           |
| Age at first full-term pregnancy (years) |             |                                                                        |                             |
| ≤ 25                      | 66.53          | 1.00                                                                   | Chen et al, 2004            |
| > 25                      | 33.47          | 1.99                                                                   | Hsieh et al, 2002           |
According to Berg of high-risk categories requiring alternative image techniques. Ultrasound or 66.7% when it was combined with MRI in the light improved 14.8% when mammography was combined with ultrasound. The sensitivity by using a combination of ultrasound or magnetic based approach. We also applied estimates for improving relative risks of those across studies would not vary with various attributes were based on Table 1 obtained from literatures, whereas those of other frequencies of the risk factors across free of breast cancer, screen-detected cases, and clinically detected cases (such as interval cancers). The percentage (prevalence) and relative risk of having breast cancer in the PCDP (pre-symptomatic cases) for each initiator, which was extracted from the literature, are shown in Table 1. The transition rates from free-of-breast-cancer to the PCDP among carriers were estimated from the results of a meta-analysis (Chen and Parmigiani, 2007). The total law of probability was used to calculate the baseline effect of non-carriers. The estimated results for non-carrier, BRCA1 carrier and BRCA2 carrier were equal to $1.04 \times 10^{-6}$, $2.10 \times 10^{-5}$, and $1.60 \times 10^{-5}$. The corresponding relative risks are equal to 20.26 and 15.44. The percentages listed in Table 2 for three tumour-specific risk factors (ER, Ki-67, and HER-2) are indicative of the distribution of breast cancer in the PCDP (pre-symptomatic cases). The relative risk for the progression from the PCDP to the CP (symptomatic cases) listed accompanied with the percentage of pre-symptomatic cases was used to simulate the corresponding percentage of interval cancers (symptomatic cases). Note that the percentages, of BMI and AP applied to pre-symptomatic (screen-detected) cases were simulated following the distribution from the normal (disease-free) women in Table 1, as these two attributes have been carried for the general women during the lifetime.

The estimated clinical weights of the two-step progression to breast cancer were expressed by the two risk scores

$\text{Risk score 1} = \log(1.04 \times 10^{-6}) + \log(1.94) + 0.69 \times \text{scattered fibroglandular densities breast}$

$+ 1.20 \times \text{heterogeneously dense breast} + 1.38 \times \text{extremely dense breast}$

$+ 0.25 \times (\text{no. of rs2981582 risk alleles}) + 0.20 \times (\text{no. of rs803662 risk alleles})$

$+ 0.11 \times (\text{no. of rs889312 risk alleles}) + 0.07 \times (\text{no. of rs817198 risk alleles})$

$+ 0.08 \times (\text{no. of rs13281615 risk alleles}) + 0.18 \times (\text{no. of rs1338794 risk alleles})$

$+ 0.11 \times (\text{no. of rs1045485 risk alleles}) + 0.96 \times (\text{BMI} > 23) + 0.68 \times (\text{AP} > 25)$

$\text{Risk score 2} = \log(1.06) + 0.72 \times (\text{BMI} > 23) + 0.42 \times (\text{AP} > 25) + 0.34 \times (\text{ER}^{-})$

$+ 0.32 \times (\text{Ki-67} 10–30% \text{proliferation}) + 0.11 \times (\text{Ki-67} > 30 \text{% proliferation})$

$+ 0.24 \times (\text{HER-2} > 2.5) + 0.08 \times (\text{HER-2} > 3.3)$

Information on risk score 1 provided the basis for risk stratification of the underlying general population. The 10-year risk of having breast cancer in each category is diagrammed in Figure 3. The 90th percentile of risk score of 10-year follow-up was

| Variables | Percentage (%) of breast cancer in the PCDP (pre-symptomatic cases) | Relative risk of the progression from the PCDP to the CP (symptomatic cases) | References |
|-----------|---------------------------------------------------------------|---------------------------------------------------------------|-------------|
| BMI (kg m$^{-2}$) | | | |
| $\leq 23$ | 26.20* | 1.00 | Hsieh et al, 2002 |
| $> 23$ | 73.80* | 2.00 | Chen et al, 2004 |
| Age at full-term first pregnancy (years) | | | |
| $\leq 25$ | 57.62* | 1.00 | Hsieh et al, 2002 |
| $> 25$ | 42.38* | 1.56 | Chen et al, 2004 |
| ER status | | | |
| Positive | 81.00 | 1.00 | Dong et al, 2008 |
| Negative | 19.00 | 1.35 | |
| Ki-67 proliferation | | | |
| $< 10\%$ | 30.1 | 1.00 | Dong et al, 2008 |
| 10–30% | 50.7 | 1.40 | |
| $> 30\%$ | 19.2 | 2.11 | |
| HER-2/neu immunohistochemistry score | | | |
| 0 or 1+ | 75.6 | 1.00 | Dong et al, 2008 |
| 2+ | 11.9 | 1.28 | |
| 3+ | 12.5 | 1.07 | |

*These values are based on the simulated PCDP (pre-symptomatic cases) simulated from free of breast cancer to the PCDP following the distribution of women free of breast cancer in Table 1.

Applications to individually tailored screening. To determine the optimal inter-screening interval in different high-risk groups, we adopted a triennial screening interval as a reference to identify the number of interval cancer for other risk score percentiles that did not exceed the median-risk score based on a triennial programme for high-risk screening. This is called a risk-score-based approach. We also applied estimates for improving sensitivity by using a combination of ultrasound or magnetic resonance imaging (MRI) with mammography compared with mammography alone borrowing information from previous studies (Berg et al, 2008; Kuhl et al, 2010) to identify the threshold of high-risk categories requiring alternative image techniques. According to Berg et al (2008), sensitivity improved 27.5% when mammography was combined with ultrasound. The sensitivity improved 14.8% when mammography was combined with ultrasound or 66.7% when it was combined with MRI in the light of Kuhl et al (2010). We calculated the incidence rate of interval cancer at which the threshold of risk percentiles combining ultrasound with mammography would be the same as the median population rate using mammography only.

We further extended the study period until 12 years. We calculated the proportion of interval cancer and the incidence rate of interval cancer as a percentage of the expected incidence rate (0.00185 per year) in the absence of screening ($I/E$ ratio), a reflection of the sensitivity of the screening method, by different inter-screening intervals and risk-score-based strategy. The higher $I/E$ ratio is, the lower the screening programme performance is. This also allowed us to assess how a risk-score-based approach contributes to reducing false positive cases resulting from the low-risk group and false negative cases from the high-risk group compared with universal screening policy. The sensitivity for each risk decile was estimated from the $1 – I/E$ ratio, and each corresponding false-positive fraction from Berg et al (2008) was applied for calculating the false-positive fraction.

### RESULTS

**Composite score for multi-state risk prediction.** To identify the role of initiators and promoters, we therefore compared the frequencies of the risk factors across free of breast cancer, screen-detected cases, and clinically detected cases (such as interval cancers). The percentage (prevalence) and relative risk of having breast cancer in the PCDP (pre-symptomatic cases) for each initiator, which was extracted from the literature, are shown in Table 1. The transition rates from free-of-breast-cancer to the PCDP among carriers were estimated from the results of a meta-analysis (Chen and Parmigiani, 2007). The total law of probability was used to calculate the baseline effect of non-carriers. The estimated results for non-carrier, BRCA1 carrier and BRCA2 carrier were equal to $1.04 \times 10^{-6}$, $2.10 \times 10^{-5}$, and $1.60 \times 10^{-5}$. The corresponding relative risks are equal to 20.26 and 15.44. The percentages listed in Table 2 for three tumour-specific risk factors (ER, Ki-67, and HER-2) are indicative of the distribution of breast cancer in the PCDP (pre-symptomatic cases). The relative risk for the progression from the PCDP to the CP (symptomatic cases) listed accompanied with the percentage of pre-symptomatic cases was used to simulate the corresponding percentage of interval cancers (symptomatic cases). Note that the percentages, of BMI and AP applied to pre-symptomatic (screen-detected) cases were simulated following the distribution from the normal (disease-free) women in Table 1, as these two attributes have been carried for the general women during the lifetime.

The estimated clinical weights of the two-step progression to breast cancer were expressed by the two risk scores

Risk score 1 = $\log(1.04 \times 10^{-6}) + \log(1.94) + 0.69 \times \text{scattered fibroglandular densities breast}$

$+ 1.20 \times \text{heterogeneously dense breast} + 1.38 \times \text{extremely dense breast}$

$+ 0.25 \times (\text{no. of rs2981582 risk alleles}) + 0.20 \times (\text{no. of rs803662 risk alleles})$

$+ 0.11 \times (\text{no. of rs889312 risk alleles}) + 0.07 \times (\text{no. of rs817198 risk alleles})$

$+ 0.08 \times (\text{no. of rs13281615 risk alleles}) + 0.18 \times (\text{no. of rs1338794 risk alleles})$

$+ 0.11 \times (\text{no. of rs1045485 risk alleles}) + 0.96 \times (\text{BMI} > 23) + 0.68 \times (\text{AP} > 25)$

and

Risk score 2 = $\log(1.06) + 0.72 \times (\text{BMI} > 23) + 0.42 \times (\text{AP} > 25) + 0.34 \times (\text{ER}^{-})$

$+ 0.32 \times (\text{Ki-67} 10–30\% \text{proliferation}) + 0.11 \times (\text{Ki-67} > 30 \text{\% proliferation})$

$+ 0.24 \times (\text{HER-2} > 2.5) + 0.08 \times (\text{HER-2} > 3.3)$

Information on risk score 1 provided the basis for risk stratification of the underlying general population. The 10-year risk of having breast cancer in each category is diagrammed in Figure 3. The 90th percentile of risk score of 10-year follow-up was
2.3%. The corresponding figure of lifetime risk was 10.2%. The corresponding relative risk compared with the median value (50th percentile) was 2.24.

**Model validation.** Table 3 shows the results of comparisons between the observed and the predicted number of breast cancer by detection modes for cross-validation and external validation. The proposed regression model may be adequate for risk score 1 and risk score 2, as there was lacking of statistical significance for cross-validation ($\chi^2(7) = 8.08$, $P = 0.33$) and external validation ($\chi^2(4) = 5.30$, $P = 0.26$).

**Individually tailored screening policies.** Table 4 shows the recommended policies by age at the start of screening and inter-screening intervals by different decile categories. The recommended age for the start of screening was suggested by identifying the equivalent age-specific risk for other percentiles, which is commensurate with the 10-year risk of the 50th percentile of the population and was estimated as 1% or so.

Table 4 also shows the higher the risk of entering the PCDP and further progressing to the CP becomes, the more intensive is the inter-screening interval demanded. In addition to shortening the inter-screening interval to reduce the incidence rate of interval cancer, a higher sensitivity tool could also be considered. When triennial mammography plus ultrasound was applied, the incidence rate for interval cancer of women in the 85th risk percentile was reduced to the same rate as the median-risk women when the triennial mammography screening alone was applied in the light of Berg et al (2008). The thresholds for using ultrasound and MRI based on Kuhl et al (2010) (the sensitivity improved 14.8% when mammography was combined with ultrasound or 66.7% when it was combined with MRI) were at the 76th and 94th percentiles, respectively.

Following the spirit of translational research, we showed two cases to demonstrate how to implement individually tailored screening as suggested in Table 4. For a woman having average risk, she would be classified into the 40th percentile by calculating her risk score 1. By using Table 4, the recommend age to start screening is around 57 years and the recommended inter-screening interval is 4 years. For a woman in the 90th~100th percentile, age at start of screening should not be older than 30 years, and the recommended inter-screening interval should not be over 6 months and the screening tool may consider using alternative image technique in addition to mammography.

The number of screenings, incidence of screen-detected cases, interval cancer, the proportion of interval cancer among breast cancer cases, and the $I/E$ ratio for 12 years are simulated and listed in Table 5, which compares the risk-score-based policies with traditional universal screening policies. In terms of an indicator of $I/E$ ratio, the risk-score-based approach (the optimal inter-screening interval was determined from Table 4) was equivalent to a universal biennial screen. The risk-score-based approach significantly reduced the interval caner rate as a percentage of...
A risk-stratification approach truly reflects the priority for targeting individuals invited to be screened. However, a false positive cases than the universal triennial screening based approach was more effective for reducing interval cancer and determination of an inter-screening interval using the risk-score-based approach was more effective for reducing interval cancer and false positive cases than the universal triennial screening programme. A risk stratification approach truly reflects the priority for targeting individuals invited to be screened. However, denser breast tissues in younger women may lead to rapidly growing malignancies due to the low sensitivity of mammography.

**DISCUSSION**

An individually tailored screening strategy can be developed by applying empirical evidence on genomic discovery, tumour phenotype, clinical correlates, and epidemiological profiles to a periodical population-based screening for breast cancer with a multi-variable and three-state Markov regression model. Developing this model is potentially helpful to ascertain the roles of different risk factors responsible for different phases in breast cancer progression, and also for targeting individually based preventive or surveillance interventions.

**Figure 3. Ten-year risk and lifetime risk of developing breast cancer by risk score percentile.**

The optimal age was determined by the age at which the 10-year risk equivalent to 1% of the 10-year risk for the 50th percentile group at age 50 years.

**Table 4. The recommend age to start screening and inter-screening interval for two types of screening at different percentiles of risk score 1**

| Percentile | Start screening age | Inter-screening interval (year) | Alternative imaging technique (threshold/study) |
|------------|---------------------|---------------------------------|-----------------------------------------------|
| 90–100     | 29                  | 0.4 (4.8 months)                | Mammography + MRI  
(94th/Kuhl et al [2010]) |
| 80–90      | 34                  | 1                              | Mammography + Ultrasound  
(85th/Berg et al, 2008) |
| 70–80      | 39                  | 1.5                            | Mammography + Ultrasound  
(76th/Kuhl et al [2010]) |
| 60–70      | 44                  | 2                              |                                              |
| 50–60      | 50d                 | 3d                             |                                              |
| 40–50      | 57                  | 4                              |                                              |
| 30–40      | 69                  | 8                              |                                              |
| 20–30      | NA                  | NA                             |                                              |
| 10–20      | NA                  | NA                             |                                              |
| 0–10       | NA                  | NA                             |                                              |

**Table 5. A comparison of different screening policies for 1 million women during 12 years**

| Screening policies | Number of Screening | Incident screen-detected cases (A) | Interval cancers (B) | Interval cancers proportion (B) | I/E ratio |
|--------------------|---------------------|------------------------------------|-----------------------|-------------------------------|-----------|
| Universal          |                     |                                    |                       |                               |           |
| Annual             | 12 895 635          | 11 444                             | 2989                  | 20.71                         | 0.14      |
| Biennial           | 6 947 795           | 9344                               | 5089                  | 35.26                         | 0.23      |
| Triennial          | 4 964 306           | 7339                               | 6694                  | 46.38                         | 0.30      |
| Risk-score-based   | 7 297 357           | 9017                               | 4670                  | 34.12                         | 0.21      |
Although more intensive screening can reduce interval cancer, an alternative image technique may be a better solution to reduce interval cancer for some high-risk women. Our risk assessment model can be also applied to ascertaining women with the risk higher than 75% of the underlying population that may require mammography in combination with ultrasound and 90% that may need mammography combined with MRI, which were consistent with the previous findings (Kriege et al, 2004; Boyd et al, 2007).

In terms of translational research, our study is the first to demonstrate how to translate genetic markers, tumour phenotypes, clinical attributes, and conventional risk factors to facilitate individually tailored screening for breast cancer. Our approach can help women who have the BRCA1 or BRCA2 mutations in conjunction with other risk factors to reduce the breast cancer risk by intensive surveillance, chemotherapy (tamoxifen), oral contraceptives, prophylactic mastectomy, and prophylactic oophorectomy (Anderson et al, 2006). In addition, the lower risk with less-intensive screening or the lower likelihood of using alternative image techniques can also reduce false positive cases as demonstrated in our study. A cost-effectiveness study on testing breast cancer susceptibility genes showed that test and no-test strategies were very similar in cost and effectiveness, even with a small mutation probability (Holland et al, 2009).

In our current model, there are five promoters responsible for the progression from the PCDP (pre-symptomatic disease) to the CP (symptomatic disease) (including BMI, age at first full-term pregnancy (AP), ER, Ki-67, and HER-2). Although BMI and AP have been reported to have the role of promoters (Hsieh et al, 2002), the other three tumour-specific risk factors (ER, Ki-67, and HER-2) associated with the role of promoter have not been validated yet. It is therefore very interesting to assess whether and how the three tumour-specific factors (ER, Ki-67, HER-2) make contributions to the progression from the PCDP to the CP. By using the likelihood ratio test, we compared the full model with the reduced model (excluding three biomarkers) and found the contribution of three tumour-specific factors was statistically significant ($\chi^2 = 137.06, P < 0.0001$). This confirms that these three biomarkers have an important role of promoter.

Promoters identified from screen-detected cases and interval cancer can be used as prognostic or predictive factors for personalised early treatment or prophylactic intervention. Other than the interval cancer resulting from false negatives, rapid tumour progression also accounts for the failure of detection by screening. The identification of promoters affecting the transition rate between the PCDP and the CP leading to interval cancer with rapid progression has an important role in personalised treatment and therapy (Dowsett and Dunbier, 2008). Perou et al (2000) proposed four main breast cancer molecular classes categorised by histological grade and gene-expression profile, including ER, PR, and HER-2/new statuses, all of which may be defined as promoters and provide a better prediction of prognosis for optimal treatment (Dowsett and Dunbier, 2008). For example, trastuzumab is a first-line treatment for HER-2/new-positive metastatic breast cancer (Viani et al, 2007). The other new high-throughput genomic technologies, such as Oncotype DX and Mammaprint, are already commercialised as prognostic tests (Dowsett and Dunbier, 2008). By identifying different risk groups at different stages, our risk assessment model provides guidance for implementing all these new therapies to achieve efficacy and efficiency.

In conclusion, this is the first to demonstrate how to translate state-of-the-art biological information on genetic, tumour, and clinical attributes as well as conventional risk factors into the realisation of the concept of an individually tailored screening strategy. The proposed risk-score-based approach can be readily applied to other screening programs in different countries and areas by tuning their own genetic susceptibility factors, tumour phenotypes, clinical attributes, and risk factors.

ACKNOWLEDGEMENTS

This study was funded by National Science Council, Taiwan (NSC 96-2314-B-002-002 and NSC 97-2314-B-002-019-MY3).

CONFLICT OF INTEREST

The authors declare no conflict of interest.
