Using Risk Stratification to Optimize Mammography Screening in Chinese Women

Kathy Leung1,6†, PhD, Joseph T Wu1,6†*, PhD, Irene OL Wong1†, PhD, Xiao-Ou Shu2, MD, PhD, Wei Zheng2, MD, PhD, Wanqing Wen2, PhD, Ui-Soon Khoo3, MD, Roger Ngan4, MD, Ava Kwong5, MD, PhD, Gabriel M Leung1,6, MD, PhD

1 Division of Epidemiology and Biostatistics, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong

2 Division of Epidemiology, Department of Medicine, and Vanderbilt Epidemiology Center, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America

3 Department of Pathology, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong

4 Department of Clinical Oncology, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong

5 Department of Surgery, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong

6 Laboratory of Data Discovery for Health (D24H), Hong Kong Science Park, New Territories, Hong Kong SAR, China

* Corresponding author: Prof Joseph T. Wu, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, 1/F North Wing, Patrick Manson Building, 7 Sassoon Road, Pok Fu Lam, Hong Kong. Tel: + 852 3917 6709; Email: joewu@hku.hk; ORCID: 0000-0002-3155-5987

† These authors contributed equally.
Abstract

Background

The cost-effectiveness of mammography screening among Chinese women remains contentious. Here we characterized breast cancer (BC) epidemiology in Hong Kong and evaluated the cost-effectiveness of personalized risk-based screening.

Methods

We used the Hong Kong Breast Cancer Study (a case-control study with 3,501 cases and 3,610 controls) and Hong Kong Cancer Registry to develop a risk stratification model based on well-documented risk factors. We used the Shanghai Breast Cancer Study to validate the model. We considered risk-based programs with different screening age ranges and risk thresholds under which women were eligible to join if their remaining BC risk at the starting age exceeded the threshold.

Results

The lifetime risk (15-99 years) of BC ranged from 1.8% to 26.6% with a mean of 6.8%. Biennial screening was most cost-effective when the starting age was 44 years and screening from age 44 to 69 years would reduce breast cancer mortality by 25.4% (95% credible interval [CrI] = 20.5-29.4%) for all risk strata. If the risk threshold for this screening program was 8.4% (the average remaining BC risk among US women at their recommended starting age of 50 years), the coverage was 25.8% and the incremental cost-effectiveness ratio (ICER) was US$18,151 (95% CrI = $10,408-$27,663) per quality-of-life-year (QALY) compared to no screening. The ICER of universal screening was $34,953 (95% CrI = $22,820-$50,268) and $48,303 (95% CrI = $32,210-$68,000) per QALY compared to no screening and risk-based screening with 8.4% threshold, respectively.
Conclusion

Organized BC screening in Chinese women should commence as risk-based programs. Outcome data (e.g., QALY loss due to false-positive mammograms) should be systemically collected for optimizing the risk threshold.

Keywords

Breast cancer; personalized mammography screening; cost-effectiveness analysis; Hong Kong Breast Cancer Study; Shanghai Breast Cancer Study; cancer epidemiology
Globally breast cancer (BC) is the most common malignancy in women, accounting for an estimated one-quarter of all malignancies [1]. Although BC is the top female cancer among Chinese populations and the incidence has been increasing [2], the lifetime risk of developing BC in Hong Kong, Shanghai, Singapore, Taiwan and elsewhere in mainland China remains 32-82% lower than that in western populations [2-4]. The epidemiology of BC is different between Chinese and western women: both the age-specific BC incidence and mortality are different (e.g., BC incidence increases earlier in Chinese women and plateau at the age of menopause before decreasing around 70); the effects of breast density and other risk factors of BC are different (e.g., Chinese women have denser breasts which would make mammography less sensitive); and most importantly BC is susceptible to life-course and contemporaneous risks and major epidemiologic differences are anticipated given the different stage and trajectory over time between China and the West. Therefore, wholesale adoption of inferences drawn from the West that have so far dominated the literature would be inappropriate [5]. Secondary prevention by mass screening mammography in Chinese women remains controversial with limited direct evidence of benefit supporting its population-based deployment [2, 6]. Despite such an empirical vacuum, haphazard opportunistic screening in women at average risk has substantially increased in mainland China [7].

Nevertheless, at the individual level, it is important to offer women an informed choice, especially in places where the private sector thrives but with no organized BC screening programs, including the prosperous Chinese coastal cities. Individual variation in risk is substantial within any given population. For example, although the average lifetime risk of developing BC was 4.5% for women in Shanghai, the lifetime risk for women at the 90th risk percentile was 9.5% which was comparable to the population average in the UK and US [8, 9]. As such, compared to universal
screening conventionally adopted in western populations, risk-based screening that aims to stratify the female population by their remaining lifetime risk and targets only high-risk women for organized screening would be more in keeping with precision preventive care. Indeed, some western countries have already begun to assess the potential benefits of switching from universal to risk-based screening [10]. Recent studies in the UK and US suggested that personalized screening tailored to individuals’ risks and preferences could improve the efficiency and effectiveness of BC screening [11, 12].

Here we report the development and validation of a risk stratification model for screening Chinese well women in Hong Kong and Shanghai and evaluation of the comparative cost-effectiveness of risk-based and universal screening.

Methods

Data

We characterized the epidemiology of BC in Hong Kong using: 1) data from the Hong Kong Breast Cancer Study (HKBCS) which is a hospital-based case-control study (Supplementary Figure 1, available online) that we conducted in Sep 2016–Aug 2019 to elucidate the risk factors of BC cases and effectiveness of mammography screening in reducing BC mortality; and 2) 2014-2016 data on BC incidence and mortality from the Hong Kong Cancer Registry (HKCR) [4]. Briefly, HKBCS comprised 400 and 3,101 women who were newly diagnosed with ductal carcinomas in situ (DCIS) and invasive breast cancer (IBC) respectively between September 2016 and June 2019 (response rate = 75%). We recruited 3,610 control subjects with similar ages who were diagnosed with diseases unrelated to BC in other hospital departments during the same period at the recruitment sites with no history of cancer (response rate = 57%).
All subjects gave written informed consent. The study was approved by the Institutional Review Board (IRB) of The University of Hong Kong and hospital clusters of HKW, HKE, KC/KE, KW, NTW, NTE under the Hospital Authority in the public sector and the relevant institutions in the private sector in Hong Kong.

The model

We developed a proportional hazard model with parameters \( \theta \) (Table 1; Supplementary Tables 1 and 2 and Supplementary Figure 2) to emulate the HKBCS and HKCR data simultaneously and simulated the development of BC in a hypothetical risk-stratified birth cohort over their lifetime. In addition to age, we assumed the risk of developing BC depended on a woman’s: 1) family history of BC among first degree relatives, 2) prior benign breast disease diagnosis, 3) age of menarche, 4) age at first live birth, 5) body mass index (BMI) and 6) physical activity level [8].

We partitioned the cohort into 288 risk strata which corresponded to all combinations of risk factor levels. The case-fatality rate (and hence survival probability) of IBC depended only on the age and stage at diagnosis whereas the stage-specific relative 5-yr survival probabilities were constant [13].

We stratified HKBCS subjects into screenees and nonscreenees based on their screening history and assumed that the average behavior of the screenees corresponded to biennial screening (Supplementary Figure 3). We assumed that screening had no effect on the inherent biological risk of BC [14-16]. However, compared to nonscreenees, screenees would be diagnosed earlier with less advanced stages (i.e., higher survival probabilities) if they developed BC (Supplementary Tables 3 and 4). We estimated the model parameters \( \theta \) using Markov Chain Monte Carlo (MCMC) methods (see the Supplementary Methods for details) [17].
To test whether our framework was applicable to other populations in China, we ran it with data from the Shanghai Breast Cancer Study (SBCS) and Shanghai Cancer Registry and then compared the inferred BC epidemiology in Shanghai to that reported in the original SBCS publication [8]. SBCS is a population-based case control study of 3,039 patients with invasive breast cancer and 3,082 age- and frequency-matched controls who were randomly selected the general population through the Shanghai Resident Registry. We also compared the inferred effects of the risk factors on the risk of BC in Hong Kong and Shanghai.

The effectiveness and cost-effectiveness of screening

We compared no screening to biennial screening with starting age 40-60 years and stopping age 69 or 74 years. The specificity of mammography is typically below 90%, thus when applied to whole populations mass screening would lead to a substantial number of false-positive mammograms and consequently unnecessary breast tissue biopsies [18, 19]. As such, numerous previous studies have emphasized that the impact of screening on quality-adjusted life years (QALYs) depended strongly on the quality-of-life (QoL) detriment associated with positive-mammograms (e.g., due to anxiety) and invasive diagnostic procedures [20, 21]. However, a recent study [22] reported that although women with false-positive mammograms suffered from increased short-term anxiety, there was no measurable health utility decrement compared to women with negative mammograms. We assumed that sensitivity and specificity of mammography were constant across all risk strata. To avoid underestimating the cost-effectiveness of screening in the base case, we assumed no QALY loss for all screens including positive screens but accounted for the QALY loss due to confirmatory tissue biopsy arising from positive mammograms (Supplementary Table 1).
We calculated the differential cost and QALY associated with screening for each of the 288 risk strata and the corresponding incremental cost-effectiveness ratios (ICER) at an annual discount rate of 3%. The ICER of universal screening was calculated from the aggregated differential costs and QALYs for all risk strata. Under a risk-based program with starting age \( a^* \) and risk threshold \( r^* \), women would be eligible to join the program if their remaining lifetime risk of BC at age \( a^* \) exceeded \( r^* \). All costs were converted to US dollars based on the exchange rate in 2018 (1 US$ = 7.8 HK$).

**Results**

**The epidemiology of breast cancer**

The calibrated/fitted model was congruent with the data (Figure 1A-C) with a c-statistic of 0.60 (95% credible interval [CrI] = 0.54-0.65) for HKBCS. We estimated that the lifetime risk (i.e., 15-99 years old assuming competing mortality by age in the model) of IBC ranged from 1.8% to 26.6% among all Hong Kong Chinese women with mean 6.8% (Figure 1D), whereas the lifetime risk of BC mortality ranged from 0.2% to 3.0% with mean of 1.1% (Figure 1E). Women in the top 30% risk strata accounted for around 50% of the BC cases in the general population. Family history of BC, history of benign breast disease, obesity and lack of physical activity increased the risk by 96% (RR = 1.96, 95% CrI = 1.68-2.25), 61% (RR = 1.61, 95% CrI = 1.43-1.79), 36% (RR = 1.36, 95% CrI = 1.30-1.45) and 8% (for regular physical activity, RR = 0.92, 95% CrI = 0.85-0.98), respectively (Table 1). Compared with women whose age of menarche was 12-14 years, those who began menstruation at younger and older ages were 1.19 (95% CrI = 1.11-1.30) and 0.66 (95% CrI = 0.57-0.75) times more likely to develop BC, respectively. Compared with women whose age at first live birth was 30 years or younger, those who gave their first live birth at older
ages and were nulliparous were 1.50 (95% CrI = 1.33-1.71) and 1.64 (95% CrI = 1.44-1.79) times more likely to develop BC, respectively. The estimated five-year survival probabilities at diagnosis for Stage 1-4 were 99.9% (95% CrI = 99.6-100%), 94.7% (95% CrI = 93.2-96.0%), 77.7% (95% CrI = 75.1-80.2%), 27.7% (95% CrI = 24.9-30.4%), respectively; the corresponding ten-year survival probabilities were 99.7% (95% CrI = 99.4-100%), 89.7% (95% CrI = 87.7-91.5%), 60.3% (95% CrI = 57.3-63.4%) and 7.7% (95% CrI = 6.1-9.4%) (Figure 1F). Overall, the average five- and ten-year survival probability of IBC were 89% (95% CrI = 88-91%) and 84% (95% CrI = 81-88%) respectively (Supplementary Table 2).

Applying our framework to the Shanghai SBCS data yielded relative hazards that were similar to the odds ratios estimated in the original SBCS publication (Supplementary Table 5) with comparable c-statistic (0.62 for SBCS from our model vs 0.63 in the original SBCS publication [8]; the former had slightly lower discrimination power because the model required it to converge with the population-level BC incidence and mortality statistics as well). The inferred relative hazards were similar to that in Hong Kong (Table 1). However, the inferred lifetime risk of BC and survival probability were lower in Shanghai. This was unsurprising because SBCS was conducted during 1998-2005 (i.e., 15-20 years earlier than HKBCS) when BC incidence was lower and access, quality and affordability of BC screening and treatments were substantially inferior compared to the present. These results suggested that our framework for BC risk stratification and disease progression was likely generalizable to other urban populations in China.

The uptake and effectiveness of screening

We estimated that screening uptake was 5.1% (95% CrI = 4.6-5.5%), 12.0% (95% CrI = 10.2-13.9%) and 17.2% (95% CrI = 15.6-18.7%) at age 40, 55 and 70 years, respectively.
(Supplementary Table 2). We estimated that biennial screening would allow BC to be diagnosed 0.45 (95% CrI = 0.34-0.58) years earlier on average with substantial stage shift to the left or downstaging (see Supplementary Table 2 and Figure 2A). Regardless of risk stratum, the hazard ratio of BC mortality between screenees and nonscreenees was almost constant at 0.76 (95% CrI = 0.60-0.90) between age 40 and 80 years (Figure 2B). Consequently, a woman who screened biennially from age 50-69 years would reduce lifetime risk of BC mortality by 21% (95% CrI = 17-24%) with 0.09 (95% CrI = 0.07-0.11) probability of experiencing one or more episodes of unnecessary tissue biopsy due to false-positive mammograms (Figure 2C). Extending the starting age to 40 years and stopping age to 74 years would increase BC mortality risk reduction to 28% (95% CrI = 22-32%) and 22% (95% CrI = 18-26%), respectively.

The cost-effectiveness of screening

Unlike relative reduction in BC mortality risk, the cost, QALY gained, and ICER of screening strongly depended on a woman’s risk of BC (Figure 3A-F). The ICER was minimized for all risk strata if screening started at age 44 years (Figure 3E-F), though the difference in ICER was marginal compared with starting age of 40 years. As such, we set the default starting age at 44 years in what follows. Compared to no screening, biennial screening from age 44 to 69 years had an ICER of $153,983 (95% CrI = $109,193-$215,108) and $6,718 (95% CrI = $2,067-$12,631) per QALY for women in the 1st and 99th risk percentile (whose lifetime risk was 0.34 and 3.62 times the population average), respectively. The ICERs would increase by 15% (95% CrI = 11-18%) under the most pessimistic assumption regarding QALY loss attributed to positive screening mammograms. If status quo opportunistic screening was used as the comparator instead of no screening, the ICERs would increase by 4.2% (95% CrI = 2.3-6.1%).
Universal screening

Biennial screening from age 44 to 69 years for all women would provide 0.020 (95% CrI = 0.017-0.024) QALY gain at a net cost of $709 (95% CrI = $505-$908) per woman which corresponded to an ICER of $34,953 (95% CrI = $22,820-$50,268) per QALY (Figure 3G). Extending the starting age to 40 years and stopping age to 74 years would increase the ICER by 3.4% (95% CrI = 2.7-5.8%) and 12.9% (95% CrI = 11.4-14.7%), respectively (Figure 3G-H). Therefore, lowering the starting age was a more cost-effective way than extending the stopping age for maximizing the health benefits of screening.

Risk-based screening

The proportion of women eligible for screening increased markedly as the risk threshold decreased (Figure 4). If the risk threshold for biennial screening from age 44 to 69 was set at 8.4%, then the average remaining lifetime risk of BC among eligible screenees at the starting age of 44 years was 11.1%, which would be equivalent to the US national average at their recommended screening starting age of 50 years (Figure 4A). Under this risk-based screening program, 25.8% of the cohort would be eligible for screening (Figure 4C). Compared to no screening, this risk-based program (with risk threshold at 8.4%) provided a health gain of 0.009 (95% CrI = 0.007-0.011) QALY at a net cost of $159 (95% CrI = $98-$224) per woman, respectively, which corresponded to an ICER of $18,151 (95% CrI = $10,408-$27,663) per QALY (Figure 5A). Expanding this risk-based program into universal screening from age 44 to 69 years would incur an ICER of $48,303 (95% CrI = $32,210-$68,000) per QALY (Figure 5A). The ICERs among no screening, risk-based
screening and universal screening increased by 12-15% if the stopping age was extended to 74 years (Figure 5B).

**Discussion**

We have developed a generic and robust inference framework for characterizing the epidemiology of BC and effectiveness of screening in Hong Kong Chinese women, which could serve as a reliable sentinel for the rest of China and the overseas diaspora given its relatively advanced trajectory and stage of socioeconomic development (as we have illustrated using Shanghai as a comparator). The validity of our framework and results can be further assessed against findings from other local and overseas studies. Our findings concerning risk factors among Chinese women accord with the established literature [8]. The inferred BC stage distribution of screenees and nonscreenees are very similar to that reported for women in Taiwan which has implemented organized biennial screening for women aged 45-69 since 2004 (Supplementary Table 3); this supports our assumption that the average screening behaviour of the screenees in HKBCS corresponded to biennial screening. The age-specific five-year survival probabilities of IBC and the relative reduction in BC mortality inferred in our model are consistent with that reported in the UK and other high-income countries [23-25]. Taken together, these comparisons lend added credence to the reliability of our calibrated model for BC and estimated cost-effectiveness of mammography screening.

We conclude that risk-based and conventional universal BC screening would provide similar relative reduction in BC mortality among screenees but the former would be far more cost-effective, at different screening starting ages. However, a recent cohort study in Taiwan [26] reported that compared with annual clinical breast examination, risk-based biennial
mammography screening only provided a modest reduction in BC mortality (hazard ratio [HR] = 0.89, 0.75, 1.06) compared to its universal counterpart (HR = 0.62, 95% CI = 0.50-0.76). This should be interpreted in the context where 45-49% of women enrolled in their risk-based screening were assessed as high-risk and referred for mammography and the proportion adherent to these referrals was 58-62%. That is, 26% of the women enrolled in their risk-based screening underwent mammography. Because they used the initial number of enrolled women (i.e., before risk assessment was done) as the denominator for calculating HR, the HR reduction in risk-based screening (0.11) was approximately 0.26 times that in universal screening (0.38). Of additional note, we recommend caution in understanding their propensity-score based findings which concluded a relative risk reduction of 38% for universal biennial screening contrasted with virtually the entire corpus of past work consisting of both RCT and empirical evidence suggesting about only half that quantum at around 20% [24]. The correct interpretation reconciling the apparent contradiction however requires careful teasing out of the underlying data reporting structures. Moreover, we assessed biennial mammography screening strategies following the recommendations in most of western populations. The screening interval could be adjusted with more data on the pathology of breast cancer cases detected after the risk-based screening program will have been implemented.

Risk-based strategies optimize BC screening by reducing unnecessary mammography and tissue biopsy among low-risk women. The probability of a biopsy arising from an initial false-positive screen over 10 years of biennial mammography was 6-10% in the US [18] and breast biopsy had a complication rate of 8-15% [6, 19], i.e., 1.5%-4.5% of low- and average-risk women would experience at least one episode of complications due to unnecessary biopsy for biennial screening from age 44 to 69 years. Given that the cost-effectiveness of BC screening strongly
depends on QALY loss brought about by false-positive mammograms and consequent tissue biopsy, jurisdictions that consider commencing organized population-wide screening could take the following risk-based approach: 1) select an initial risk threshold that could be accommodated by current screening capacity (e.g. based on the national average in the US or UK); 2) measure the QALY loss among screenees with false-positive mammograms and tissue biopsy during a pilot phase of the program; and 3) re-evaluate the screening parameters in light of the additional data generated from the pilot phase of the program and adjust the risk threshold (in accordance with other existing cancer prevention programs, e.g. colorectal screening in Hong Kong).

Enhancing the discriminative power of the BC risk prediction model is key to improve the performance of risk-based screening (via more accurate risk stratification). After the risk-based BC screening program is rolled out, our model could be further improved with data on mammographic breast density that has been shown to be an additional useful factor for BC risk stratification [27]. Previous studies have also shown that polygenic risk profiles based on single nucleotide polymorphisms (SNPs) are strong predictors of BC [28]. In a recent polygenic risk score model based on 313 SNPs [28], 35% of all breast cancer would be expected to occur in women in the highest 20% of the risk distribution. Although we did not consider SNPs in our model because no data were available in HKBCS and only limited data on a few SNPs were available in SBCS [8], similarly we estimated 40% of breast cancer would be expected to occur in women in the highest 25.8% of the risk distribution in the Hong Kong population. Given that sequencing for personalized medicine is becoming more accessible, albeit still struggling to overcome challenges associated with direct application of the hitherto predominantly Caucasian-derived evidence to other racial or ethnic groups, and that the cost will continue to fall in the future, genetic risk profiles should be included as a core predictor for next-generation BC screening.
Minimizing unnecessary tissue biopsy is another key to improve the compliance and cost-effectiveness of BC screening. There is a need to develop innovative and effective methods to replace mammography or supplement it with non-invasive and accurate reflex testing to improve the positive predictive value of screening.

Finally, while the ICER for universal screening at $34,953 falls within the acceptable range as one would expect in most western developed countries, we should highlight that China officially remains a middle-income country with large disparities at the subnational level. The risk-based screening ICER of $18,151, that is almost half that of universal screening, represents much better value-based care. As a first step coming from no organized screening, to expand the opportunity to be screened as fairly and as efficiently as possible, it would be prudent to commence risk-based screening before further considering a universal strategy. Such a decision for the population will of course be controversial and subject to debate at the policy level. In future, we would also recommend that individual choice be taken into account, based on the ethos of personal preference and tailored preventive care, where our model could further incorporate an additional module that takes into account women’s own risk appetite [29, 30].

**Funding**

This study was supported by a commissioned grant from the Health and Medical Research Fund (HMRF) from The Government of the Hong Kong Special Administrative Region (BC-HKU) and Award Number U54GM088558 from the National Institute of General Medical Sciences (NIGMS). Dr Irene OL Wong was supported by The University of Hong Kong/China Medical Board Grants (HKU/CMB) 2016/2017 (second round) to visit Vanderbilt University Medical Center, Nashville, Tennessee, United States.
Notes

Role of the funders: The funders had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Disclosures: No authors have any conflicts of interest to disclose.

Author contributions: Conceptualization: KL, JTW, IOLW, GML. Data collection: IOLW, XS, WZ, WW, KL, USK, RN, AK, GML. Model development: KL, JTW. Formal analysis: KL, IOLW, JTW. Supervision: GML, JTW. Writing – Original Draft: KL, JTW. Writing – Review and Editing: all authors.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health, USA.

Acknowledgement: We thank the patients who participated the Hong Kong Breast Cancer Study (HKBCS). Special thanks are due to the study-site collaborators in public and private sectors and fieldworkers in HKBCS. We would also acknowledge research support from Hong Kong Cancer Registry and its breast cancer team for providing population-based cancer figures and facilitating data validation.

Data Availability

Request for data should be made to the corresponding author. Request of codes should be made to the corresponding author.
References

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International Journal of Cancer 2015;136(5):E359-386.
2. Fan L, Strasser-Weippl K, Li J-J, et al. Breast cancer in China. The lancet oncology 2014;15(7):e279-e289.
3. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians 2018;68(6):394-424.
4. Hong Kong Hospital Authority. Hong Kong Cancer Registry. http://www3.ha.org.hk/cancereg/.
5. Wong IOL, Schooling CM, Cowling BJ, et al. Breast cancer incidence and mortality in a transitioning Chinese population: current and future trends. British Journal of Cancer 2015;112:167-170.
6. Leung GM, Lam TH, Thach TQ, et al. Will screening mammography in the East do more harm than good? American Journal of Public Health 2002;92(11):1841-6.
7. Mei Z, Yijing Z, Heling B, et al. Breast Cancer Screening Rates Among Women Aged 20 Years and Above — China, 2015. China CDC Weekly 2021;3(13):267-273.
8. Zheng W, Wen W, Gao Y-T, et al. Genetic and clinical predictors for breast cancer risk assessment and stratification among Chinese women. Journal of the National Cancer Institute 2010;102(13):972-981.
9. Sun L, Legood R, Sadique Z, et al. Cost-effectiveness of risk-based breast cancer screening programme, China. Bulletin of the World Health Organization 2018;96(8):568-577.
10. Onega T, Beaber EF, Sprague BL, et al. Breast cancer screening in an era of personalized regimens: A conceptual model and National Cancer Institute initiative for risk-based and preference-based approaches at a population level. Cancer 2014;120(19):2955-2964.
11. Pashayan N, Morris S, Gilbert FJ, et al. Cost-effectiveness and Benefit-to-Harm Ratio of Risk-Stratified Screening for Breast Cancer: A Life-Table Model. JAMA Oncology 2018;4(11):1504-1510.

12. Ayer T, Alagoz O, Stout NK. OR Forum—A POMDP Approach to Personalize Mammography Screening Decisions. Operations Research 2012;60(5):1019-1034.

13. United States National Cancer Institute. Surveillance, Epidemiology and End Results Program. https://seer.cancer.gov/.

14. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. New England Journal of Medicine 2005;353(17):1784-1792.

15. Plevritis SK, Munoz D, Kurian AW, et al. Association of screening and treatment with breast cancer mortality by molecular subtype in US women, 2000-2012. Jama 2018;319(2):154-164.

16. Huang X, Li Y, Song J, et al. A Bayesian simulation model for breast cancer screening, incidence, treatment, and mortality. Medical Decision Making 2018;38(1_suppl):78S-88S.

17. George EI, McCulloch RE. Variable selection via Gibbs sampling. Journal of the American Statistical Association 1993;88(423):881-889.

18. Nelson HD, Pappas M, Cantor A, et al. Harms of breast cancer screening: systematic review to update the 2009 US Preventive Services Task Force recommendation. Annals of internal medicine 2016;164(4):256-267.

19. Hofvind S, Ponti A, Patnick J, et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. Journal of medical screening 2012;19(1_suppl):57-66.

20. van Ravesteyn NT, Stout NK, Schechter CB, et al. Benefits and harms of mammography screening after age 74 years: model estimates of overdiagnosis. J Natl Cancer Inst 2015;107(7).

21. Schousboe JT, Kerlikowske K, Loh A, et al. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. Ann Intern Med 2011;155(1):10-20.
22. Tosteson AN, Fryback DG, Hammond CS, et al. Consequences of false-positive screening mammograms. JAMA internal medicine 2014;174(6):954-961.

23. Weedon-Fekjær H, Romundstad PR, Vatten LJ. Modern mammography screening and breast cancer mortality: population study. BMJ : British Medical Journal 2014;348:g3701.

24. Marmot MG, Altman D, Cameron D, et al. The benefits and harms of breast cancer screening: an independent review. British journal of cancer 2013;108(11):2205.

25. Nelson HD, Fu R, Cantor A, et al. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 US Preventive Services Task Force recommendation. Annals of internal medicine 2016;164(4):244-255.

26. Yen AM-F, Tsau H-S, Fann JC-Y, et al. Population-Based Breast Cancer Screening With Risk-Based and Universal Mammography Screening Compared With Clinical Breast Examination: A Propensity Score Analysis of 1 429 890 Taiwanese WomenRisk-Based and Universal Mammography vs Clinical Breast ExaminationRisk-Based and Universal Mammography vs Clinical Breast Examination. JAMA Oncology 2016;2(7):915-921.

27. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Annals of internal medicine 2003;138(3):168-175.

28. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. The American Journal of Human Genetics 2019;104(1):21-34.

29. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018;359(6378):926-930.

30. Puzhko S, Gagnon J, Simard J, et al. Health professionals’ perspectives on breast cancer risk stratification: understanding evaluation of risk versus screening for disease. Public Health Reviews 2019;40(1):2.
Table 1. Relative hazards in Hong Kong and Shanghai

| Risk factor                                      | Relative hazard (95% CrI) | No. of controls | Overall proportion of women (95% CrI) |
|-------------------------------------------------|---------------------------|-----------------|--------------------------------------|
|                                                 | Hong Kong 2016-2019 | Shanghai Early 2000s | Hong Kong 2016-2019 | Shanghai Early 2000s | Hong Kong 2016-2019 | Shanghai Early 2000s |
| Age of menarche, years                          |                           |                  |                        |                        |                        |                        |
| ≥ 15                                            | 0.66 (0.57-0.75)         | 0.73 (0.64-0.81) | 535                    | 673                    | 0.153 (0.144-0.162)   | 0.330 (0.314-0.345)   |
| 12–14                                          | 1                         | 1                | 1569                   | 1212                   | 0.448 (0.438-0.462)   | 0.594 (0.577-0.609)   |
| ≤ 11                                           | 1.19 (1.11-1.30)         | 1.23 (1.01-1.49) | 1439                   | 159                    | 0.398 (0.386-0.410)   | 0.076 (0.068-0.087)   |
| Age at first live birth, years                  |                           |                  |                        |                        |                        |                        |
| < 25                                           | 1                         | 1                | 974                    | 558                    | 0.277 (0.267-0.288)   | 0.288 (0.277-0.305)   |
| 25–29                                          | 1.00 (0.89-1.13)         | 1.12 (1.01-1.24) | 1066                   | 1097                   | 0.300 (0.289-0.309)   | 0.553 (0.544-0.572)   |
| ≥ 30                                           | 1.50 (1.33-1.71)         | 1.84 (1.73-2.11) | 745                    | 349                    | 0.208 (0.197-0.218)   | 0.159 (0.147-0.167)   |
| Nulliparous                                    | 1.64 (1.44-1.79)         |                  | 794                    |                        | 0.215 (0.203-0.226)   |                        |
| Family history of breast cancer among first degree relatives |
| No                                              | 1                         | 1                | 3357                   | 1985                   | 0.929 (0.921-0.934)   | 0.970 (0.965-0.975)   |
| Yes                                             | 1.96 (1.68-2.25)         | 1.55 (1.13-1.91) | 253                    | 61                     | 0.071 (0.066-0.079)   | 0.030 (0.026-0.035)   |
| Prior benign breast disease diagnosis           |                           |                  |                        |                        |                        |                        |
| No                                              | 1                         | 1                | 3045                   | 1464                   | 0.848 (0.840-0.856)   | 0.730 (0.719-0.743)   |
| Yes                                             | 1.61 (1.43-1.79)         | 1.77 (1.72-1.81) | 557                    | 582                    | 0.152 (0.144-0.161)   | 0.270 (0.257-0.280)   |
| Body mass index, kg/m²                          |                           |                  |                        |                        |                        |                        |
| < 18.5                                         | 0.95 (0.83-1.00)         | 0.72 (0.58-0.95) | 208                    | 110                    | 0.067 (0.060-0.072)   | 0.055 (0.048-0.064)   |
| 18.5–23                                        | 1                         | 1                | 1381                   | 887                    | 0.429 (0.418-0.440)   | 0.436 (0.421-0.451)   |
| > 23                                           | 1.36 (1.30-1.45)         | 1.27 (1.15-1.42) | 1777                   | 1047                   | 0.505 (0.492-0.516)   | 0.509 (0.492-0.519)   |
| Physical activity                               |                           |                  |                        |                        |                        |                        |
| No                                              | 1                         | 1                | 2828                   | 1430                   | 0.784 (0.776-0.793)   | 0.686 (0.674-0.698)   |
| Yes                                             | 0.92 (0.85-0.98)         | 0.92 (0.86-0.99) | 767                    | 615                    | 0.216 (0.208-0.224)   | 0.314 (0.299-0.325)   |

a The c-statistic (which is the same as the Area Under the Receiver Operating Characteristic Curve or AUC) is 0.60 for Hong Kong Breast Cancer Study (HKBCS) and 0.62 for Shanghai Breast Cancer Study (SBCS) from our model.

b The proportion of females were estimated jointly in the model based on the number of participants in the control group accounting for missing data.
Less than 2% of Shanghai women were nulliparous and they were grouped with the women whose age at first live birth ≥ 30. Physical activity refers to exercising intensively (e.g., lifting heavy objects, cardiovascular exercise, riding fast on bicycle etc.) at least once a week on average, in the last 10 years.
Figure Legends

**Figure 1. Inferred breast cancer epidemiology.** Black dots and vertical bars indicate point estimates and 95% confidence intervals from the data (A, B and F). Lines and shades indicate posterior means and 95% credible intervals (CrIs) from the model (A, B and C). (A) The calibrated model was congruent with the observed incidence of ductal carcinoma in situ (DCIS) and invasive breast cancer (IBC) in Hong Kong Cancer Registry (HKCR). (B) The calibrated model was congruent with the observed breast cancer mortality in HKCR. (C) The calibrated model was congruent with the observed age distribution of the cases in Hong Kong Breast Cancer Study (HKBCS). (D-E) Probability density function (PDF) of lifetime risk of IBC and breast cancer (BC) mortality among women in the hypothetical birth cohort comprising 288 risk strata. (F) Inferred average five-year and ten-year survival probability by the stage at diagnosis.

**Figure 2. Effectiveness of biennial mammography screening.** (A) Inferred stage distribution of breast cancer (BC) cases among screenees and nonscreenees. Bars indicate posterior means. Vertical lines indicate 95% credible intervals (CrIs). (B) Hazard ratio of BC mortality between screenees and nonscreenees. Lines indicate posterior means. Shades indicate 95% CrIs. (C) Relative reduction in BC mortality risk conferred by biennial screening with different starting and stopping ages. Lines indicate posterior means. Shades indicate 95% CrIs.
Figure 3. Cost-effectiveness of biennial mammography screening compared to no screening in the hypothetical cohort with stopping age 69 and 74 years. (A-F) Cost, quality-adjusted life years (QALYs) gained and incremental cost-effectiveness ratio (ICER) as a function of starting age for women at the 1st, 25th, 50th, 75th, and 99th risk percentile. (G-H) ICER as a function of starting age for the entire cohort. Lines and shades indicate means and 95% prediction intervals. The circles indicate that the ICERs were minimized when the starting age was 44 years.

Figure 4. Screening coverage, effectiveness and cost-effectiveness of risk-based screening as a function of risk threshold. Biennial screening started at age 44 and stopped at age 69 or 74 years. (A-B) Average remaining lifetime risk among women eligible for risk-based screening. The red circles corresponding to setting the risk threshold such that the remaining lifetime risk of breast cancer (BC) among eligible screenes was the same as the United States national average at age 50 when their screening program starts (i.e., 11.1%). (C-D) Screening coverage, i.e., proportion of the birth cohort eligible for risk-based screening. (E-H) Relative reduction in BC mortality for the cohort and the associated incremental cost-effectiveness ratio (ICER) conferred by risk-based screening.

Figure 5. Comparative cost-effectiveness of universal and risk-based screening. Circles indicate universal screening. Squares indicate risk-based screening under which the average remaining lifetime risk of eligible screenees at the starting age was the same as the United States national average at the age of 50 (see Figure 4). Dashed lines indicate the risk-based screening at different risk thresholds. (A) The cost-effectiveness planes show the increase in cost and quality-
adjusted life years (QALYs) compared to no screening per birth cohort when biennial screening started at age 44 years and stopped at age 69 years. (B) The cost-effectiveness planes show the increase in cost and quality-adjusted life years (QALYs) compared to no screening per birth cohort when biennial screening started at age 44 years and stopped at age 74 years.
Figure 1

A. Incidence per 100,000

B. BC mortality per 100,000

C. CDF

D. PDF

E. Lifetime risk of BC mortality

F. Survival probability
Figure 2

A

Proportion

Nonscreened

Screened

Stage

DCIS
1
2
3
4

B

Hazard ratio of BC mortality

Age at diagnosis (years)

40
50
60
70
80

C

Relative reduction in BC mortality risk

Starting age (years)

40
45
50
55
60

A

B

C

Figure 2
Figure 3

Stopping age = 69 years

A

Stopping age = 74 years

B

C

D

E

F

G

H

Starting age (years)

The entire cohort
Figure 4

Biennial screening starts at age 44 and ends at age 69

Average risk of eligible females

Screening coverage

BC mortality reduction

ICER (US$10,000 per QALY)

Risk threshold

Biennial screening starts at age 44 and ends at age 74
Figure 5

A

Biennial screening starts at age 44 and ends at age 69

Universal

Risk-based (at the average risk of US females)

$34,553/QALY

$37,226/QALY

QALYs gained per 10,000

Cost per 10,000 (US$)

10 × 10^6

B

Biennial screening starts at age 44 and ends at age 74

Universal

Risk-based (at the average risk of US females)

$39,469/QALY

$52,802/QALY

QALYs gained per 10,000

10 × 10^6

$17,225/QALY

$19,703/QALY

$46,838/QALY

$52,692/QALY