PLOS ONE

Developing and validating an individualized breast cancer risk prediction model for women attending breast cancer screening

Javier Louro[1,2,3,4], Marta Román[1,2,3]*, Margarita Posso[1,2,3], Ivonne Vázquez[5], Francina Saladié[6], Ana Rodriguez-Arana[7], M. Jesús Quintana[8,9], Laia Domingo[1,2,3], Marisa Baré[2,10], Rafael Marcos-Gragera[9,11], María Vernet-Tomas[12], Maria Sala[1,2,3], Xavier Castells[1,2,3], on behalf of the BELE and IRIS Study Groups[¶]

1 IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain, 2 Research Network on Health Services in Chronic Diseases (REDISSEC), Barcelona, Spain, 3 Servei d'Epidemiologia i Avaluació, Hospital del Mar, Barcelona, Spain, 4 European Higher Education Area (EHEA) Doctoral Programme in Methodology of Biomedical Research and Public Health in Department of Pediatrics, Obstetrics and Gynecology, Preventive Medicine and Public Health, Universitat Autónoma de Barcelona (UAB), Bellaterra, Barcelona, Spain, 5 Servei de Patologia, Hospital del Mar, Barcelona, Spain, 6 Cancer Epidemiology and Prevention Service, Hospital Universitari Sant Joan de Reus, Institut d'Investigació Sanitària Pere Virgili, Universitat Rovira i Virgili, Reus, Spain, 7 Servei de Diagnòstic per la imatge, Hospital del Mar, Barcelona, Spain, 8 Department of Clinical Epidemiology and Public Health, University Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Barcelona, Barcelona, Spain, 9 CIBER of Epidemiology and Public Health (CIBERESP), Barcelona, Spain, 10 Clinical Epidemiology and Cancer Screening, Parc Taulí University Hospital, Sabadell, Spain, 11 Department of Health, Epidemiology Unit and Girona Cancer Registry, Oncology Coordination Plan, Autonomous Government of Catalonia, Catalan Institute of Oncology, Girona, Spain, 12 Servei d'Obstetricia i Ginecologia, Hospital del Mar, Barcelona, Spain

¶ Complete membership of the author group can be found in the Acknowledgments
* mroman@parcdesalutmar.cat

Abstract

Background

Several studies have proposed personalized strategies based on women's individual breast cancer risk to improve the effectiveness of breast cancer screening. We designed and internally validated an individualized risk prediction model for women eligible for mammography screening.

Methods

Retrospective cohort study of 121,969 women aged 50 to 69 years, screened at the long-standing population-based screening program in Spain between 1995 and 2015 and followed up until 2017. We used partly conditional Cox proportional hazards regression to estimate the adjusted hazard ratios (aHR) and individual risks for age, family history of breast cancer, previous benign breast disease, and previous mammographic features. We internally validated our model with the expected-to-observed ratio and the area under the receiver operating characteristic curve.

Results

During a mean follow-up of 7.5 years, 2,058 women were diagnosed with breast cancer. All three risk factors were strongly associated with breast cancer risk, with the highest risk

being found among women with family history of breast cancer (aHR: 1.67), a proliferative benign breast disease (aHR: 3.02) and previous calcifications (aHR: 2.52). The model was well calibrated overall (expected-to-observed ratio ranging from 0.99 at 2 years to 1.02 at 20 years) but slightly overestimated the risk in women with proliferative benign breast disease. The area under the receiver operating characteristic curve ranged from 58.7% to 64.7%, depending of the time horizon selected.

Conclusions

We developed a risk prediction model to estimate the short- and long-term risk of breast cancer in women eligible for mammography screening using information routinely reported at screening participation. The model could help to guiding individualized screening strategies aimed at improving the risk-benefit balance of mammography screening programs.

Introduction

There is ongoing debate on the benefits and harms of breast cancer screening [1–3]. To improve this balance, current evidence supports personalized screening [4,5]. Modeling studies have shown that modifying the screening interval, screening modality, or age range of the target population based on women's individual risk yielded greater benefit than conventional standard strategies [5–7]. Several risk models have been designed to estimate women's individual breast cancer risk based on their personal characteristics [8–15]. However, most of these models have not been specifically developed to estimate the risk of women targeted for breast cancer screening in order to offer them personalized strategies.

A recent consensus statement of the European Conference on Personalized Early Detection and Prevention of Breast Cancer (ENVISION) [16] stated the need to develop breast cancer risk prediction models based on data from large screening cohorts and including risk factors easily obtainable at screening participation, such previous mammographic features and prior benign breast disease.

To date, only one model has specifically aimed to predict women's individual risk looking to personalize breast cancer screening strategies [17]. Although highly valuable, the model was based on short-term risk estimates and did not account for relevant characteristics of prospective studies such as internal time-dependent covariates. This model only estimates the two-year risk, which could lead to bias as one of the aims proposed in breast cancer screening personalization is to see which women are at a lower risk in order to extend their screening period to three or four years. Therefore, if new breast cancer risk models are developed with the aim of analyzing the possibilities offered by personalized screening strategies, it would be interesting to estimate the biennial risk of each woman, in other words, to obtain estimators not only at 2 years, but also every two years (2, 4, 6, 8. . . up to 20 years, which is the total time a woman is screened). This will help to better understand the different possibilities of screening strategies and will allow to observe the differences in the validation of the model estimators for the different time horizons. There is therefore a need for breast cancer risk prediction models, with risk estimates in the short- and long-term, and based on data from large screening cohorts. These new risk models should include a limited and feasible number of variables for the proposed objective, for example, detailed information on the type of previous benign

breast disease or previous mammographic characteristics, which existing risk models tend not to use.

We aimed to design and validate an individualized risk prediction model to estimate the biennial risk of breast cancer in women eligible for mammography screening by using data from the long-standing population-based screening program in Spain.

Materials and methods

Setting and study population

Breast cancer screening in Spain started in 1990 in a single setting and expanded until it became nationwide in 2006. This program follows the recommendations of the European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis [18]. Women aged 50 to 69 years are invited to biennial screening mammography by written letter. Screening mammograms are interpreted according to the Breast Imaging Reporting and Data System (BI-RADS) scale by trained breast radiologists [19]. Women with an abnormal mammographic feature are recalled for further assessments to confirm or rule out malignancy. Women without a breast cancer diagnosis are invited again for routine screening at 2 years. Overall, breast cancer screening in Spain has a recall rate of 43.0, a detection rate of 4.0, and an interval cancer rate of 1.1 per 1,000 mammographic examinations [20]. The positive predictive value is 9.8% for recalls and 38.9% for recalls involving invasive procedures. Overall, 16.8% of all screen-detected cancers are ductal carcinoma in situ (DCIS). More details of breast cancer screening in Spain are described elsewhere [21].

We analyzed data from two centers forming part of the Spanish breast cancer screening program in the Metropolitan Area of Barcelona. These centers routinely gather information on family history of breast cancer, previous benign breast disease (BBD), and previous mammographic features. The centers collect information on screening mammography examinations, recalls, further assessments, and diagnostic results in their defined catchment areas. The cohort included all 123,251 women screened at least once between 1995 and 2015 and followed-up until December 2017. We excluded 758 women diagnosed with breast cancer at the first screen, 210 women with missing information on family history, 213 women with missing information on previous BBD, and 101 women with missing information for both family history and previous BBD. The study population for the analysis consisted of 121,969 women who underwent 437,540 screening mammograms during the study period.

Definition of study variables

Information on family history and history of prior breast biopsies was self-reported and collected from face-to-face interviews conducted by trained professionals at the time of mammography screening. This information was consistently collected over the 20 years study period. A family history of breast cancer was defined as having at least one first-degree relative with a history of breast cancer.

Breast biopsy results were classified by a community pathologist at each center using SNOMED codes [22]. Pathological diagnoses were grouped following the benign breast disease classification proposed by Dupont and Page [23–25] into non-proliferative and proliferative disease. Proliferative lesions with and without atypia were combined into a single category due to the small number of subsequent breast cancer cases among those with a proliferative lesion with atypia. If women reported having had a biopsy before the start of the screening but no pathology results were available, the biopsy was classified as having a prior biopsy, unknown diagnosis.

A community radiologist routinely reported on mammographic features found at mammography screening interpretation. We classified as mammographic features any mass, calcification, asymmetry or architectural distortion reported by radiologists at mammographic interpretation. Findings were assigned to the category of multiple mammographic features if more than one of the previous mammographic features had been reported simultaneously at screening interpretation.

We included both invasive breast cancers and DCIS for the analysis.

Model design

We built the risk prediction model using a random sample of 60% of the study population (estimation subcohort). The remaining 40% was used for an internal validation (validation subcohort).

We estimated the age-adjusted hazard ratios (aHR) and the 95% confidence intervals (95%CI) for the breast cancer incidence for each category of family history, previous BBD, and previous mammographic features with the estimation subcohort. Age was included in the model as a continuous variable. We used partly conditional Cox proportional hazards regression, an extension of the standard Cox model, to incorporate changes in these risk factors over time. Robust standard errors were used to estimate 95% confidence intervals using the Huber sandwich estimator [26]. If a woman has had a diagnosis of cancer, she will contribute women-years at risk from the date of her first mammogram to the diagnosis of cancer. Since we can identify all interval cancers, a woman who has not had a diagnosis of cancer at the end of her follow-up will contribute women-years at risk from the first mammogram to the last mammogram plus 2 years of follow-up.

We tested whether family history, previous BBD, and previous mammographic features interacted among themselves or with age. The interaction terms were not significant and were therefore not included in the model. The proportional hazards assumption was assessed by plotting the log-minus-log of the survivor function against log time for each predictor variable. The proportional hazards assumption appeared to be reasonable for all predictors.

Model validation

We calculated the absolute breast cancer risk estimates for each 2-year interval over the 20-year lifespan covered by screening (ages 50 to 69 years) for each individual in the validation subcohort. As proposed by Zheng and Heagerty, we used a general hazard function to predict the absolute risk of breast cancer diagnosis based on length of follow-up, prediction time, and women's risk profile [27].

We conducted an internal validation of the model to evaluate its predictive performance by assessing its calibration and discrimination. To assess calibration, we calculated the ratio between the expected breast cancer rate in the validation subcohort versus the observed rate in the estimation subcohort. To account for censoring, the observed rate was estimated using the Kaplan-Meier estimator. The expected breast cancer rate was calculated as the average of the risk estimates in the validation subcohort. The expected breast cancer rate in a specific risk group was calculated as the average of the risk estimates for each woman in that risk group of the validation subcohort. The expected-to-observed (E/O) ratio assessed whether the number of women predicted to develop breast cancer from the model matched the actual number of breast cancers diagnosed in the validation subcohort. An E/O ratio of 1.0 indicates perfect calibration. We calculated the E/O ratio 95% confidence intervals (95% CI) using the formula of the standardized mortality ratio proposed by Breslow and Day [28]. The discriminatory accuracy of our model was assessed by estimating the area under the receiving operating characteristic curve (AUC) for each 2-year interval based on the predicted risks for each woman and

whether she developed breast cancer during the time interval or not [29]. The predicted risks were calculated using the model coefficient estimates at the baseline mammogram for those women in the validation cohort who have been followed for a time greater than or equal to the time horizon being estimated. The AUC measured the ability of the model to discriminate between women who will develop breast cancer from those who will not. We calculated the 95% CI using the approach proposed by Hanley and McNeil [30].

Statistical tests were two-sided and all p-values <0.05 were considered statistically significant. All analyses were performed using the statistical software R version 3.4.3 (Development Core Team, 2014).

The study was approved by the Clinical Research Ethics Committee of Hospital del Mar Medical Research Institute (2015/6189/I). The review boards of the institutions providing data granted approval for data analyses. This is an entirely registry-based study that used anonymized retrospective data and hence there was no requirement for written informed consent.

The authors declare that they have no conflicts of interest.

Results

During a mean follow-up of 7.52 years, breast cancer was diagnosed in 2,058 out of the 121,969 women in the study population. The mean follow-up was shorter in women with a breast cancer diagnosis than in those without (5.8 years vs 7.6 years, p-value < 0.05). Women with breast cancer were more likely to have a family history of breast cancer (18.32% vs 13.86%), biopsies with unknown diagnosis (23.76% vs 21.72%), non-proliferative and proliferative BBD (5.59% vs 3.20%, and 1.60% vs 0.45%, respectively), masses (20.51% vs 18.12%), and calcifications (6.85% vs 2.71%) (Table 1).

Breast cancer was strongly associated with previous benign breast disease, with the highest risk being found among women with a proliferative BBD (aHR, 3.02; 95% CI: 1.75, 5.21) compared with those without a BBD (Table 2). Family history was also associated with breast cancer (aHR, 1.67; 95% CI: 1.41, 1.98). Among women with previous mammographic features, the highest risks were found in calcifications (aHR, 2.52; 95% CI: 1.93, 3.29) and architectural distortions (aHR, 2.07; 95% CI: 1.27, 3.38).

Overall calibration of the model was accurate across all 2-year time horizons. The E/O ratio ranged from 0.99 at 2 years to 1.02 at 20 years and was never significantly different than 1 (Table 3). The AUC was lowest at the 4-year risk estimate (AUC, 58.7%; 95%CI: 55.9%-61.5%) and highest at the 18-year risk estimate (AUC, 64.7%; 95%CI: 62.5%-66.9%) and were significantly higher than 50% for all the time horizons.

Estimates for the 10-year time horizon showed that the model slightly overestimated breast cancer rates in women with masses (E/O ratio, 1.18; 95%CI: 1.02–1.37) and in women aged 55–59 years (E/O ratio, 1.15; 95%CI: (1.03–1.29) (Table 4). The model also underestimated breast cancer rates in women aged 50–54 years (E/O ratio, 0.83; 95%CI: 0.75–0.94). Because of the small number of breast cancer cases, calibration was overestimated among women with proliferative BBD (E/O ratio, 1.85; 95%CI: 1.00–3.40).

Distribution of the absolute cumulative risk estimates at 2-, 10- and 20-year time horizons are shown in Fig 1. The 10-year risk was between 1.5% and 2% in 60% of the women and was higher than 2% in 35%. The 20-year risk was lower than 3% in only 4% of the women, between 5% and 7% in 17% of the women, and was higher than 7% in approximately 9% of the women.

Discussion

We used individual-level data from a large cohort of women regularly screened in Spain to design and validate a risk prediction model to estimate the biennial risk of breast cancer in

Table 1. Baseline characteristics of the study population.

	No breast cancer (n = 119,911)	Breast cancer (n = 2,058)	p-value
Mean follow- up	7.6 years	5.8 years	<0.001
Age (years)			
50–54	63,507 (52.96%)	1,149 (55.83%)	0.010
55–59	25,738 (21.46%)	542 (26.34%)	<0.001
60–64	22,796 (19.01%)	325 (15.79%)	<0.001
65–69	7,870 (6.56%)	42 (2.04%)	<0.001
Family history of breast cancer			
No	103,296 (86.14%)	1,681 (81.68%)	<0.001
Yes	16,615 (13.86%)	377 (18.32%)	<0.001
Benign breast disease			
None	89,500 (74.64%)	1,421 (69.05%)	<0.001
Prior biopsy, unknown diagnosis	26,042 (21.72%)	489 (23.76%)	0.028
Non-proliferative	3,832 (3.20%)	115 (5.59%)	<0.001
Proliferative	537 (0.45%)	33 (1.60%)	<0.001
Mammographic features			
None	86,326 (71.99%)	1,283 (62.34%)	<0.001
Mass	21,728 (18.12%)	422 (20.51%)	<0.001
Calcifications	3,246 (2.71%)	141 (6.85%)	<0.001
Asymmetry	3,371 (2.81%)	56 (2.72%)	0.858
Architectural distortion	1,249 (1.04%)	29 (1.41%)	0.129
Multiple features	3,991 (3.33%)	127 (6.17%)	<0.001

Differences in mean of follow-up were tested by Mann–Whitney U test.

Differences in qualitative variables were tested by two-sided test of equality for column proportions (z-test). Tests adjusted for all pairwise comparisons within each tumor characteristic using the Bonferroni correction.

https://doi.org/10.1371/journal.pone.0248930.t001

Table 2. Partly conditional Cox proportional hazards model results showing the hazard ratios of the risk factors on breast cancer.

	Women-years	Breast cancer cases	aHR* (95%CI)
Family history of breast cancer			
No	471,552	976	Ref.
Yes	79,471	227	1.67 (1.41–1.98)
Benign breast disease			
No	408,883	832	Ref.
Prior biopsy, unknown diagnosis	118,010	286	1.36 (1.16–1.59)
Non-proliferative	21,123	67	1.41 (1.02–1.94)
Proliferative	3,007	18	3.02 (1.75–5.21)
Mammographic features			
No	380,314	752	Ref.
Mass	110,597	239	1.32 (1.11–1.57)
Calcifications	17,160	81	2.52 (1.93–3.29)
Asymmetry	17,526	38	1.66 (1.16–2.39)
Architectural distortion	6,287	20	2.07 (1.27–3.38)
Multiple features	19,140	73	1.86 (1.43–2.43)

aHR: Adjusted Hazard Ratio. 95%CI: 95% Confidence Interval.

*Model adjusted by age, family history, previous benign breast disease and previous mammographic features.

https://doi.org/10.1371/journal.pone.0248930.t002

Table 3. E/O ratio and area under the ROC curve of the model for each time horizon.

	Observed events	E/O ratio (CI95%)	AUC
2-year risk	188	0.99 (0.86–1.14)	63.0 (59.1–66.9)
4-year risk	455	1.01 (0.92–1.11)	58.7 (55.9–61.5)
6-year risk	685	1.00 (0.92–1.07)	59.5 (57.2–61.8)
8-year risk	853	1.02 (0.95–1.09)	61.0 (58.9–63.0)
10-year risk	1,000	1.01 (0.95–1.08)	60.9 (59.0–62.8)
12-year risk	1,092	1.03 (0.97–1.09)	60.5 (58.6–62.4)
14-year risk	1,165	1.01 (0.96–1.07)	62.4 (60.5–64.3)
16-year risk	1,195	1.00 (0.95–1.06)	64.3 (62.4–66.3)
18-year risk	1,201	1.01 (0.96–1.07)	64.7 (62.5–66.9)
20-year risk	1,203	1.02 (0.97–1.08)	63.8 (61.3–66.3)

E/O: Expected observed. 95%CI: 95% Confidence Interval.

https://doi.org/10.1371/journal.pone.0248930.t003

women aged 50 to 69 years eligible for mammography screening. We tested a model that uses only variables easily obtainable at screening participation. The model showed very good calibration but only modest discrimination.
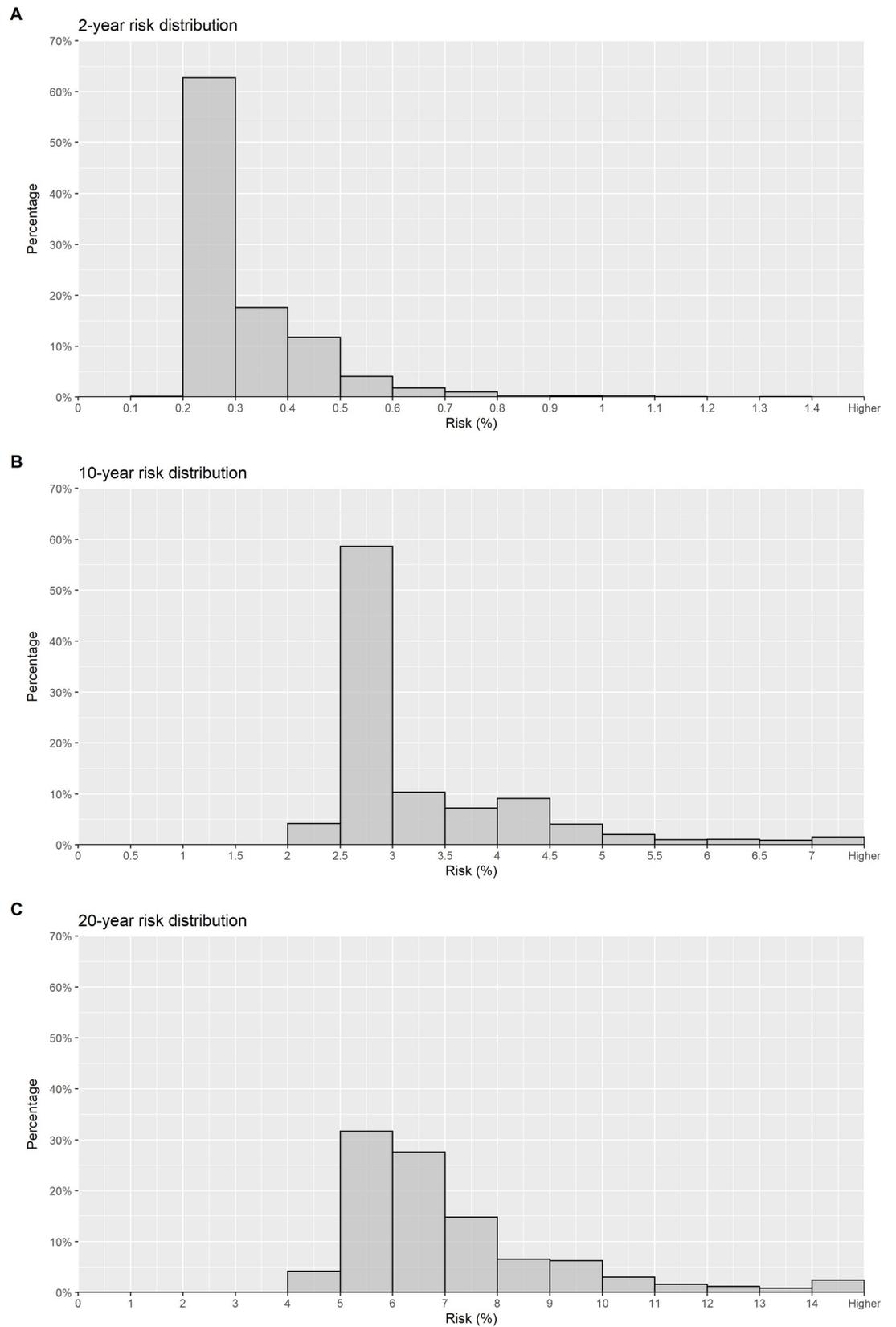
Our model calculates the risk of breast cancer for each 2-year time horizon during a woman's screening lifespan. Until now, the 5-year risk estimate has been the standard since the BCRAT model used a 5-year risk time horizon for decision making about chemoprevention. The BCRAT model was the basis for enrolment into the two major US prevention trials

Table 4. Calibration of the 10-year estimates from the model in risk factor subgroups.

	Observed events	E/O ratio (95%CI)
Overall	1,000	1.01 (0.95–1.08)
Family history		
No	824	1.00 (0.93–1.07)
Yes	176	1.10 (0.95–1.28)
Benign breast disease		
No	709	1.02 (0.95–1.10)
Prior biopsy, unknown diagnosis	238	1.02 (0.90–1.16)
Non-proliferative	43	1.17 (0.87–1.57)
Proliferative	10	1.85 (1.00–3.40)
Mammographic features		
No	661	1.03 (0.96–1.11)
Mass	175	1.18 (1.02–1.37)
Calcifications	60	1.07 (0.83–1.38)
Asymmetry	29	1.14 (0.79–1.63)
Architectural distortion	15	1.05 (0.63–1.73)
Multiple features	60	0.90 (0.70–1.16)
Age (years)		
50–54	296	0.83 (0.75–0.94)
55–59	296	1.15 (1.03–1.29)
60–64	278	0.96 (0.85–1.08)
65–69	130	1.02 (0.86–1.21)

E/O: Expected observed. 95%CI: 95% Confidence Interval.

https://doi.org/10.1371/journal.pone.0248930.t004

A

2-year risk distribution

B

10-year risk distribution

C

20-year risk distribution

Fig 1. Distribution of the absolute cumulative risk estimates.

https://doi.org/10.1371/journal.pone.0248930.g001

[31,32]. However, as stated in the statements of the last European Conference on Risk-Stratified Prevention and Early Detection of Breast Cancer, there is a need for risk models specifically designed for women eligible for breast cancer screening, based on data from large screening cohorts [16].

A previous model was designed to estimate the risk of breast cancer in women eligible for mammography screening [17]. The model used the Karma cohort from Sweden and included information on mammographic features. That study focused solely on estimating the short-term risk of breast cancer over the next mammographic examination. In addition, it used a case-control design to establish risk factors, which may bias the estimates of the short-term association with breast cancer risk. Our model adds to the breast cancer risk prediction models currently available and can be used to help guide personalized screening strategies by employing information easily obtained at screening participation. Additional useful information from our model is estimation of a woman's risk for breast cancer at 2-yearly intervals.

Our model was further developed by adding the effect of mammographic features, such as masses, calcifications, asymmetries, and architectural distortions. Previous studies have shown that mammographic features increase the subsequent risk of breast cancer [33]. In our model, the strongest influence on risk was conferred by calcifications. The biology behind calcifications is not well established. It has been suggested that mammary cells may acquire some mesenchymal characteristics, being able to contribute to the production of breast calcifications as a sign of carcinogenic transformation [34].

The role of BBD as a risk factor for breast cancer is well established [9,33,35]. However, its inclusion in breast cancer risk prediction models is rare, mainly because available information on BBD in large cohorts of women is uncommon. Only one previous risk model included different estimates for the different categories of the Dupont and Page BBD pathological classification [23–25]. The Breast Cancer Surveillance Consortium model was updated to include BBD, which led to only minimal improvement in discrimination [9]. This lack of significant improvement could be due to the absence of pathology results for most women who reported breast biopsies prior to their first screening round, as was also the case in our study. However, the addition of BBD to the model markedly increased the proportion of women identified as being at high risk for invasive breast cancer.

We assessed the internal validity of the model by means of its calibration and discriminatory accuracy. To perform internal validation we split our cohort in two sets, the estimation subcohort, to perform the analysis and development of the model and the validation subcohort, to perform the internal validation of the model. This technique known as split validation is common for this type of models [9] but cross validation or bootstrapping could also have been performed [36,37]. The model showed accurate calibration, neither overestimating nor underestimating the overall risk through the different years. In Table 4 we saw the calibration of the 10-year estimates from the model in risk factor subgroups. We also performed the E/O ratio estimates in risk factor subgroups for each one of the time horizons proposed. We only showed the 10-year estimates since showing all of them could be confusing. We showed the 10-year estimates since they have a good balance between the number of events observed (in the first time horizons some subcategories have a low number of observed events was lower) and the number of people observed (in the last time horizons we have some lost to follow–up, as the mean time of follow-up is 7.5 years). Nonetheless, the E/O ratio was overestimated for women with a proliferative BBD, due to the small number of cases among this subgroup.

The model showed modest discrimination with a maximum AUC of 64.7%. Discriminatory accuracy in breast cancer risk prediction models is usually low because a substantial proportion of cases are diagnosed in women with no known risk factors and the AUC of the different models vary between 60 and 70% [14]. This is clearly in contrast with prediction models for

other diseases, such as cardiovascular disease, which achieve good discrimination [38,39]. However, the model presented in this paper performed as well as other models that include many other risk factors that were not available in this study. As one of the reasons why the existing risk models have not been implemented for personalized screening is that it is difficult to collect all of the necessary risk factors in practice, a simpler model like the one we present could be useful. We tested other approaches to validate our model, such as the AUC estimation proposed by Li et al [40]. This estimation uses weights to calculate the contribution in the estimates of those women without a breast cancer diagnosis who were censored before reaching the time horizon. However, this approach produced no substantial differences in our validation.

A major strength of our model is that we used individual-level data from more than 120,000 women participating in a large, well-established, population-based screening program in Spain from 1995 to 2015, with a mean follow-up of more than 7.5 years and a maximum of 20 years. The program has a participation rate of 67% and a re-attendance rate of 91.2% [19].

This study also has some limitations. First, a major weakness is the lack of information on breast density, which was not systemically collected as part of screening data in the participating centers. Previous models estimating individual breast cancer risk have shown that the addition of breast density improved the discriminatory power of the models [9,17,41,42]. Dense breasts confer women a higher risk of breast cancer and are also associated with a higher risk of false-positive results, masking, and interval cancers [43]. In addition, we had no information on common genetic variants, which has been added to other breast cancer risk prediction models [44,45]. However, the discriminatory accuracy of the models was scarcely improved by the inclusion of information on single nucleotide polymorphisms (SNPs). This lack of both variables may be useful for some institutions where these risk factors are not available.

Second, the number of breast cancer cases among women with a proliferative BBD was small, which reduced our ability to accurately predict the expected number of cases across risk factor subgroups. Nevertheless, the overall calibration of the model across the time horizons assessed was highly accurate. Also, as a consequence of the small number of subsequent breast cancer cases among those women with a proliferative BBD with atypia, we merged proliferative BBD with and without atypia into a single category which might make the model less usable in practice.

Third, our model was based on a large set of representative data from the Breast Cancer Screening Program in Spain, which provides good generalizability. However, external validation of the results is needed to verify the predictive performance of our risk model.

Another limitation might be the reason for censoring. Over 52% of women in the cohort had their last mammogram in the last two years of the study follow-up and 17% of women had their last mammogram at ages 68 or 69 years. Most of the remaining 31% are women who did not participate in the 2014–2015 round or who have changed health areas and thus are not in our study population. The screening program does not have an exhaustive record of which women die and, therefore, we cannot differentiate them from non-participating women.

Finally, we were unable to analyze the association between the laterality of the BBD with the subsequent risk of breast cancer. In a previous analysis, we found that 40% of incident breast cancer cases in women with BBD were contralateral to the prior BBD, suggesting that a large proportion of benign lesions may be risk markers rather than precursors of subsequent cancer [46].

Conclusions

We designed and internally validated a risk prediction model to estimate the short- and long-term risk of breast cancer in women eligible for mammography screening based on their age,

family history, previous benign breast disease, and previous mammographic features. The model showed good calibration and modest discriminatory power, and could be improved by adding further variables such as breast density and polygenic risk scores. The model can be used biennially to predict a woman's breast cancer risk during her screening lifespan (age 50 to 69 years) using information easily obtained at screening participation. Risk prediction models specifically designed for women eligible for breast cancer screening are key to guide individualized screening strategies aiming to improve the risk-benefit balance of mammography screening programs.

Acknowledgments

Author Contributions

Conceptualization: Javier Louro, Marta Román, Margarita Posso, Xavier Castells.

References

1. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med. 2012; 367(21):1998–2005. https://doi.org/10.1056/NEJMoa1206809 PMID: 23171096

2. Paci E, Broeders M, Hofvind S, Puliti D, Duffy SW, Group EW. European breast cancer service screening outcomes: a first balance sheet of the benefits and harms. Cancer Epidemiol Biomarkers Prev. 2014; 23(7):1159–63. https://doi.org/10.1158/1055-9965.EPI-13-0320 PMID: 24991022

3. Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. JAMA Intern Med. 2014; 174(3):448–54. https://doi.org/10.1001/jamainternmed.2013.13635 PMID: 24380095

4. Onega T, Beaber EF, Sprague BL, Barlow WE, Haas JS, Tosteson AN, 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):2955–64. https://doi.org/10.1002/cncr.28771 PMID: 24830599

5. Vilaprinyo E, Forne C, Carles M, Sala M, Pla R, Castells X, et al. Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. PLoS One. 2014; 9(2):e86858. https://doi.org/10.1371/journal.pone.0086858 PMID: 24498285

6. Trentham-Dietz A, Kerlikowske K, Stout NK, Miglioretti DL, Schechter CB, Ergun MA, et al. Tailoring Breast Cancer Screening Intervals by Breast Density and Risk for Women Aged 50 Years or Older: Collaborative Modeling of Screening Outcomes. Ann Intern Med. 2016; 165(10):700–12. https://doi.org/10.7326/M16-0476 PMID: 27548583

7. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. 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. https://doi.org/10.7326/0003-4819-155-1-201107050-00003 PMID: 21727289

8. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst. 1989; 81(24):1879–86. https://doi.org/10.1093/jnci/81.24.1879 PMID: 2593165

9. Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast Density and Benign Breast Disease: Risk Assessment to Identify Women at High Risk of Breast Cancer. J Clin Oncol. 2015; 33 (28):3137–43. https://doi.org/10.1200/JCO.2015.60.8869 PMID: 26282663

10. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med. 2004; 23(7):1111–30. https://doi.org/10.1002/sim.1668 PMID: 15057881

11. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst. 2006; 98(17):1204–14. https://doi.org/10.1093/jnci/djj331 PMID: 16954473

12. Gail MH, Costantino JP, Pee D, Bondy M, Newman L, Selvan M, et al. Projecting individualized absolute invasive breast cancer risk in African American women. J Natl Cancer Inst. 2007; 99(23):1782–92. https://doi.org/10.1093/jnci/djm223 PMID: 18042936

13. Rosner B, Colditz GA, Iglehart JD, Hankinson SE. Risk prediction models with incomplete data with application to prediction of estrogen receptor-positive breast cancer: prospective data from the Nurses' Health Study. Breast Cancer Res. 2008; 10(4):R55. https://doi.org/10.1186/bcr2110 PMID: 18598349

14. Louro J, Posso M, Hilton Boon M, Román M, Domingo L, Castells X, et al. A systematic review and quality assessment of individualised breast cancer risk prediction models. British Journal of Cancer. 2019; 121(1):76–85. https://doi.org/10.1038/s41416-019-0476-8 PMID: 31114019

15. Antoniou AC, Pharoah PP, Smith P, Easton DF. THE BOADICEA model of genetic susceptibility to breast and ovarian cancer. Br J Cancer. 2004; 91(8):1580–90. https://doi.org/10.1038/sj.bjc.6602175 PMID: 15381934

16. Pashayan N., Antoniou A.C., Ivanus U. et al. Personalized early detection and prevention of breast cancer: ENVISION consensus statement. Nat Rev Clin Oncol (2020).

17. Eriksson M, Czene K, Pawitan Y, Leifland K, Darabi H, Hall P. A clinical model for identifying the short-term risk of breast cancer. Breast Cancer Res. 2017; 19(1):29. https://doi.org/10.1186/s13058-017-0820-y PMID: 28288659

18. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition—summary document. Ann Oncol. 2008; 19(4):614–22. https://doi.org/10.1093/annonc/mdm481 PMID: 18024988

19. Ascunce N, Salas D, Zubizarreta R, Almazan R, Ibanez J, Ederra M, et al. Cancer screening in Spain. Ann Oncol. 2010; 21 Suppl 3:iii43–51. https://doi.org/10.1093/annonc/mdq085 PMID: 20427360

20. Domingo L, Hofvind S, Hubbard RA, Roman M, Benkeser D, Sala M, et al. Cross-national comparison of screening mammography accuracy measures in U.S., Norway, and Spain. Eur Radiol. 2016; 26 (8):2520–8. https://doi.org/10.1007/s00330-015-4074-8 PMID: 26560729

21. Castells X, Sala M, Ascunce N, Salas D, Zubizarreta R, Casamijana M. Descripción del cribado del cáncer en España. Proyecto DESCRIC. Madrid: Plan de Calidad para el Sistema Nacional de Salud. Ministerio de Sanidad y Consumo: Agència d'Avaluació de Tecnologia i Recerca Medìques de Cataluña 2006 [Available from: http://www.cribadocancer.es/images/archivos/mama/documentacion/informes_tecnicos/archivos/InformeDESCRIC.pdf.

22. Ruch P, Gobeill J, Lovis C, Geissbuhler A. Automatic medical encoding with SNOMED categories. BMC Med Inform Decis Mak. 2008; 8 Suppl 1:S6. https://doi.org/10.1186/1472-6947-8-S1-S6 PMID: 19007443

23. Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. N Engl J Med. 1985; 312(3):146–51. https://doi.org/10.1056/NEJM198501173120303 PMID: 3965932

24. Page DL, Dupont WD, Rogers LW, Rados MS. Atypical hyperplastic lesions of the female breast. A long-term follow-up study. Cancer. 1985; 55(11):2698–708. https://doi.org/10.1002/1097-0142 (19850601)55:11<2698::aid-cncr2820551127>3.0.co;2-a PMID: 2986821

25. Page DL, Schuyler PA, Dupont WD, Jensen RA, Plummer WD Jr., Simpson JF. Atypical lobular hyperplasia as a unilateral predictor of breast cancer risk: a retrospective cohort study. Lancet. 2003; 361 (9352):125–9. https://doi.org/10.1016/S0140-6736(03)12230-1 PMID: 12531579

26. Freedman DA. On the So-Called 'Huber Sandwich Estimator' and 'Robust Standard Errors.' The American Statistician, vol. 60, no. 4, 2006, pp. 299–302. JSTOR, www.jstor.org/stable/27643806. Accessed 23 Feb. 2021.

27. Zheng Y, Heagerty PJ. Partly conditional survival models for longitudinal data. Biometrics. 2005; 61 (2):379–91. https://doi.org/10.1111/j.1541-0420.2005.00323.x PMID: 16011684

28. Breslow NE, Day NE. Statistical methods in cancer research. Volume II—The design and analysis of cohort studies. IARC Sci Publ. 1987(82):1–406.

29. Harrell FE Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996; 15(4):361–87. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4 PMID: 8668867

30. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982; 143(1):29–36. https://doi.org/10.1148/radiology.143.1.7063747 PMID: 7063747

31. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J Natl Cancer Inst. 1998; 90(18):1371–88. https://doi.org/10.1093/jnci/90.18.1371 PMID: 9747868

32. Vogel VG, Costantino JP, Wickerham DL, Cronin WM, Cecchini RS, Atkins JN, et al. Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes: the

NSABP Study of Tamoxifen and Raloxifene (STAR) P-2 trial. JAMA. 2006; 295(23):2727–41. https://doi.org/10.1001/jama.295.23.joc60074 PMID: 16754727

33. Castells X, Tora-Rocamora I, Posso M, Roman M, Vernet-Tomas M, Rodriguez-Arana A, et al. Risk of Breast Cancer in Women with False-Positive Results according to Mammographic Features. Radiology. 2016; 280(2):379–86. https://doi.org/10.1148/radiol.2016151174 PMID: 26878225

34. Scimeca M, Giannini E, Antonacci C, Pistolese CA, Spagnoli LG, Bonanno E. Microcalcifications in breast cancer: an active phenomenon mediated by epithelial cells with mesenchymal characteristics. BMC Cancer. 2014; 14:286. https://doi.org/10.1186/1471-2407-14-286 PMID: 24758513

35. Hartmann LC, Sellers TA, Frost MH, Lingle WL, Degnim AC, Ghosh K, et al. Benign breast disease and the risk of breast cancer. N Engl J Med. 2005; 353(3):229–37. https://doi.org/10.1056/NEJMoa044383 PMID: 16034008

36. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001; 54(8):774–781. https://doi.org/10.1016/s0895-4356(01)00341-9 PMID: 11470385

37. Efron B, Tibshirani R. Improvements on cross-validation: The .632+ bootstrap method. J Amer Statist Assoc. 1997; 92(438):548–560.

38. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998; 97(18):1837–47. https://doi.org/10.1161/01.cir.97.18.1837 PMID: 9603539

39. Marrugat J, D'Agostino R, Sullivan L, Elosua R, Wilson P, Ordovas J, et al. An adaptation of the Framingham coronary heart disease risk function to European Mediterranean areas. J Epidemiol Community Health. 2003; 57(8):634–8. https://doi.org/10.1136/jech.57.8.634 PMID: 12883073

40. Li L, Greene T, Hu B. A simple method to estimate the time-dependent receiver operating characteristic curve and the area under the curve with right censored data. Stat Methods Med Res. 2018; 27(8):2264–78. https://doi.org/10.1177/0962280216680239 PMID: 27895266

41. Kerlikowske K, Gard CC, Sprague BL, Tice JA, Miglioretti DL. One versus Two Breast Density Measures to Predict 5- and 10-Year Breast Cancer Risk. Cancer Epidemiol Biomarkers Prev. 2015; 24 (6):889–97. https://doi.org/10.1158/1055-9965.EPI-15-0035 PMID: 25824444

42. Armero C, Forne C, Rue M, Forte A, Perpinan H, Gomez G, et al. Bayesian joint ordinal and survival modeling for breast cancer risk assessment. Stat Med. 2016; 35(28):5267–82. https://doi.org/10.1002/sim.7065 PMID: 27523800

43. Posso M, Louro J, Sanchez M, Roman M, Vidal C, Sala M, et al. Mammographic breast density: How it affects performance indicators in screening programmes? Eur J Radiol. 2019; 110:81–7. https://doi.org/10.1016/j.ejrad.2018.11.012 PMID: 30599878

44. Zhang X, Rice M, Tworoger SS, Rosner BA, Eliassen AH, Tamimi RM, et al. Addition of a polygenic risk score, mammographic density, and endogenous hormones to existing breast cancer risk prediction models: A nested case-control study. PLoS Med. 2018; 15(9):e1002644. https://doi.org/10.1371/journal.pmed.1002644 PMID: 30180161

45. Shieh Y, Hu D, Ma L, Huntsman S, Gard CC, Leung JW, et al. Breast cancer risk prediction using a clinical risk model and polygenic risk score. Breast Cancer Res Treat. 2016; 159(3):513–25. https://doi.org/10.1007/s10549-016-3953-2 PMID: 27565998

46. Castells X, Domingo L, Corominas JM, Tora-Rocamora I, Quintana MJ, Bare M, et al. Breast cancer risk after diagnosis by screening mammography of nonproliferative or proliferative benign breast disease: a study from a population-based screening program. Breast Cancer Res Treat. 2015; 149 (1):237–44. https://doi.org/10.1007/s10549-014-3208-z PMID: 25503778