Universitat Autònoma de Barcelona

Individualized breast cancer risk prediction models applied to population-based screening mammography

Javier Louro Aldamiz-Echevarría

Universitat Autònoma de Barcelona

Individualized breast cancer risk prediction models applied to population-based screening mammography

PhD Thesis

Javier Louro Aldamiz-Echevarría

Directors: Dr. Marta Román Expósito, Dr. Xavier Castells Oliveres

Tutor: Dr. Xavier Castells Oliveres

PhD in Methodology of Biomedical Research and Public Health

Department of Pediatrics, Obstetrics & Gynecology, and Preventative Medicine and Public Health

Barcelona, April 2021

IMIM Parc de Salut MAR Barcelona

Institut Hospital del Mar d'Investigacions Mèdiques

PRBB

All models are wrong,
but some are useful

George E. P. Box

This PhD thesis has been carried out in the Epidemiology and Evaluation Group of Hospital del Mar-IMIM, in Barcelona, under the direction of Dr. Marta Román and Dr. Xavier Castells.

It is presented as a compendium of publications.

Barcelona, April 2021

Acknowledgments

I hope the reader will forgive me for writing this thesis acknowledgments in Spanish. English is the language of science, but I think I can express better in Spanish what I feel about all the people that have helped me through this travel.

"Deja los agradecimientos para el final". Este fue uno de los primeros consejos que me dieron el día que decidí sentarme a escribir la tesis. *"Avanza en lo importante; introducción, métodos, discusión, escribe un poco cada día, y deja las trivialidades como la portada, los agradecimientos, anexos, etc... para cuando la tengas terminada".*

Pues no. Aún estoy empezando la introducción (de hecho, llevo solo un par de frases) y no he podido evitar la tentación de desplazar el cursor hacia arriba y comenzar a escribir estas palabras. Esta sección está infravalorada y es tan importante como las demás. No, esta sección es más importante que las demás. Quizás no por su contenido científico. Quizás no para la mayoría de los lectores de esta tesis, si es que la lee alguien. Probablemente una persona ajena a estos 4 años no comprenderá porqué es importante, ya que la importancia de esta sección trasciende más allá de lo visible. Esta sección es apoyo, compañerismo, felicidad, sonrisas y celebraciones, pero también estrés, ansiedad, sudor y lágrimas. Esta sección es lo que no se ve en los artículos científicos, es la cara oculta de la investigación, es la metatesis.

Esta sección es cada paso en un viaje de 1560 días.

La verdad es que veo imposible resumir todo lo que quiero deciros en un texto breve. Me habéis acompañado durante más de cuatro años en este viaje. Algunos habéis caminado conmigo desde el principio hasta el final, otros habéis aparecido y aportado vuestro granito de arena en alguna parada del camino y otros simplemente, durante un breve instante, habéis conseguido hacer que este viaje sea mejor. Algunos habéis contribuido con consejos, ideas y correcciones, algunos me habéis enseñado, tanto en

lo personal como en lo profesional y otros me habéis escuchado cuando lo he necesitado. Pero todos tenéis algo en común, todos vosotros habéis contribuido a que mi doctorado sea una experiencia increíble.

Esta tesis es de todos vosotros:

Marta, no podría empezar esta sección con el nombre de otra persona. Al menos la mitad de esta tesis es tuya. Maestra, tutora y amiga. Cada momento que he pasado contigo me has dado la oportunidad de aprender y me has inspirado como investigadora, como estadística, como música y sobre todo como persona. He tenido mucha suerte de que volvieras al grupo en el momento exacto en el que iba a comenzar esta tesis. Jamás olvidaré lo que me has aportado estos años. Gracias por todo.

Xavier, siempre he creído que al mundo le faltan jefes como tú. Desde el primer momento me has dado responsabilidad y has confiado en mí cuando ni siquiera yo confiaba y con ello me has ayudado a formarme como profesional y persona. Siempre recordaré con cariño el tiempo invertido la primera media hora de las reuniones BELE hablando de música, cine y de los *Dunkerques*.

Anna, has conseguido que trabajar y divertirse sean dos cosas que van de la mano día a día. Tras más de 3 años de viernes musicales, *ratolins*, *yogi teas*, *AJs*, *rondinaires*, *gusteifs*, cotilleos del coronavirus y mil otras cosas, si hay alguna culpable de que todos los días me levantara con ganas de ir a trabajar, eres tú. Una parte muy grande de esta tesis es tuya.

Laia, el día que te fuiste del despacho 135 este perdió muchísimo. Por suerte, he podido seguir disfrutando de trabajar a tu lado como compañera de grupo. Gracias por toda la ayuda y consejos y sobre todo por tu optimismo y alegría que siempre nos contagias a todos.

Martiña, eres el ejemplo de que no hay que estar mucho tiempo en un sitio para formar parte de él. Eres también una parte muy importante de esta tesis, aun habiendo compartido, por desgracia, muy poco tiempo en el despacho.

Anabel, mi primera compi, que me enseñaste todo lo necesario para sobrevivir al primer contacto con el mundo laboral, gracias.

Gracias también a Isa Torá y a los *mesianos* con los que compartí el 135; Laura, Diego, Sergio y Stella por las risas y el buen ambiente creado en el despacho a lo largo de estos años.

Margarita, muchas gracias por todo lo que me has enseñado durante estos años y por lo mucho que has contribuido a esta tesis.

Mercè, mi primer contacto con el IMIM, tú eres la razón por la que hoy esté escribiendo estas líneas, gracias por ser la primera en creer en mí.

Y hablando de creer en mí, gracias Maria por darme la oportunidad, allá por septiembre de 2016, de formar parte de este increíble grupo. Me has enseñado mucho durante estos años y además has sido una jefa cercana, de las que se preocupan por nosotros y ayudan en todo lo que pueden. Una gran parte de esta tesis también es tuya.

Carlota, mi primera amiga en el hospital, gracias por todas las risas y el *buenrollismo* juntos a lo largo de estos años.

Gracias también a María Padilla, a Toño Gimeno y a toda la gente de aquel curso en Madrid que además de conocimientos nos dio nuevos amigos y compañeros.

Por supuesto quiero agradecer su participación en este viaje a toda la familia que me falta del *Servei*. Todos formáis parte de esta tesis: Andrea Burón, Esther Martínez, Irene Fernández, Cris Hernández, Cris González, Montse Bonilla, Inma Collet, Milagros Herranz, Anna Collado, Marta Blanqué, Francesc Macià, Mercè Esturi, Judith Sivilla, Alicia Noguera, Priscila Giraldo, Miriam Caracuel y a todos los residentes que han pasado por el Servei. Espero no olvidarme de nadie y de verdad que lo lamento si es así.

También a los investigadores de otros grupos del Hospital del Mar u otros hospitales que también han formado parte de este proyecto o con los que he trabajado estos años: Rodrigo Alcántara, Ana Rodríguez-Arana, Josep Corominas, Laia Serrano, Ivonne Vázquez, Laura Comerma, Mar Vernet, Ignasi Tusquets, Lidia Blay, Irene Zarcos, Michele Hilton Boon, Marisa Baré, Carmen Vidal, Joana Ferrer, Rafa Marcos, María Jesús Quintana, Mar Sánchez, Miguel Prieto, Francina Saladié, Jaume Galceran, Xavier Bargalló, Lupe Peñalva y Josep Alfons Espinàs.

I would also like to thank Solveig Hofvind and all the team from BreastScreen Norway for their support and help in making the international research component of this thesis possible. It is a pleasure to work with you.

Quiero agradecerle a Gail Craigie su profesionalidad y dedicación como revisora, tanto de los artículos que comprenden esta tesis como de la tesis misma.

Quiero aprovechar la oportunidad para agradecerle también su labor a Alexandra Elbakyan: sin su trabajo ni esta ni la mayoría de las tesis e investigaciones del mundo podrían realizarse. Thank you for *removing all barriers in the way of science*.

A las *matísticos* en general, y a Itxa y Sil en particular, por las largas conversaciones y risas en nuestros desayunos semanales (o casi) en la terraza.

A mi querido CED, que nació de una idea y un café y acabó como una gran familia: Lucho, Marta, Judith, Vanesa, Bea y Uxue.

A Natalia, y a nuestros desayunos y charlas, que nos dan energía para afrontar todo lo que el IMIM nos eche encima.

A Will y a los Incubakers.

A toda la URSS (la del IMIM, claro), con los que comer siempre me ha alegrado el día.

A mis alumnos durante todos estos años del Institut Bonanova y al resto de compañeros de este centro, con los que he descubierto una de las pasiones de mi vida: la docencia.

Y hablando de docencia: quiero hacer un agradecimiento especial a Rosa Crujeiras, profesora de la USC, que estuvo en el previaje, en la planificación, añadiendo a mi vida la pasión que padezco ahora: la estadística. Gracias Rosa, por inspirarme a mí y seguro que a muchos más para dedicarnos a la estadística.

A Barcelona, per ensenyar-me què és la felicitat. A los *Contrascados*, a los *Perros*, al *Gran Sol*, a la *Miau* y al *triángulo muletero,* que también vivieron esta tesis.

A mis padres y a mi hermana, por su apoyo incondicional, no solo con esta tesis si no durante toda una vida. Gracias.

Y a Helena, por todo. Es imposible enumerarlo, podría hacer otra tesis con todo lo que querría agradecerte, sobre todo estos últimos meses. Por acompañarme día a día, caída a caída, viaje a viaje y, espero, en muchos viajes más. Gracias.

Esta tesis es de todos vosotros.

Muchas gracias por todo.

This has been my PhD. I have found it worth living and would gladly live it again if the chance were offered me.

Funding

The completion of this thesis has been made possible by various grants:

- Research project funded by Instituto de Salud Carlos III FEDER (PI17/00047), "Individual breast cancer risk estimation according to known risk factors for the personalization of population-based screening".

- Research project funded by Instituto de Salud Carlos III FEDER (PI15/00098), "Benign breast disease, breast density and association with breast cancer in a population-based breast cancer screening program ".

- Funds from the Research Network on Health Services in Chronic Diseases (RD12/0001/0015).

- XV Alicia Llacer grant awarded by the Spanish Society of Epidemiology (SEE) for the best research by a young researcher.

Abstract

Keywords (MeSH terms): Breast neoplasms, breast density, breast disease, early detection of cancer, mammography, risk assessment, risk factors

Background: Mammographic screening has been shown to reduce mortality from breast cancer. Following the recommendations of the European Council, European countries have started population-based screening programs that offer biennial mammograms to women aged between 50 and 69 years. The results of the effectiveness of population-based screening are controversial in terms of the balance between mortality reduction and adverse effects. To improve this balance, current evidence supports personalized screening. Modeling studies have shown that modifying the screening interval, screening modality, or age range of the target population based on women's individual risk yields a greater benefit than conventional standard strategies. Several risk models have been designed to estimate women's individual breast cancer risk based on their personal characteristics. However, most of these models have not been specifically developed to estimate the risk of women targeted for breast cancer screening. There is therefore a need to broaden current information on risk factors for breast cancer and the estimation of individual risk prediction models through the analysis of large population-based databases.

Aims: The general objective of the thesis is to deepen the analysis of population-based breast cancer screening. Specifically, the aim of this thesis is to assess different breast cancer risk factors in order to develop and validate an individualized breast cancer risk prediction model. We evaluated how breast density affects screening performance indicators in a digital mammography context. Then, we assessed differences in breast cancer risk across benign breast disease diagnosed at prevalent or incident screens. To our knowledge, this is the first time that such an approach has been used. We also evaluated the interaction between breast density and benign breast disease. Subsequently, we performed a systematic review

to update the existing evidence, conduct a critical appraisal and risk of bias assessment and summarize the results of the individualized risk models that are used to estimate the risk of breast cancer in women in the general population. Finally, a breast cancer risk prediction model was designed and internally validated, based on information easily accessible at screening.

Methods: The study population included all women participating in 10 breast cancer screening programs of Spain from 1995 to 2015 and followed up until 2017. We analyzed 2,853,753 screening mammograms from 782,406 women aged 50-69 years. For the specific analysis of breast density and the different screening performance indicators, we used generalized estimating equations (GEE) models. For the main analysis of this thesis, Cox partly conditional models were used, an extension for repeated measures of the Cox proportional hazard model in which the hazard ratio of a given event is modeled over time. Using these models, we assess the difference of the risk factors in the various categories. We also designed an individual risk prediction model. The model was internally validated with the expected-to-observed ratio and with the area under the receiving operating characteristic curve.

Results: i) Mammogram sensitivity decreased from 89.2% in women with BI-RADS 1 to 67.9% for those in BI-RADS 4. Both the positive predictive value of recall and of invasive tests decreased from 10.4% to 5.7% and from 49.8% to 32.4% in women with BI-RADS 1 and BI-RADS 4, respectively. **ii)** Compared with women without benign breast disease, the risk of breast cancer was significantly higher in women with benign breast disease diagnosed in an incident screen (aHR, 2.67; 95%CI: 2.24-3.19) than in those with benign breast disease diagnosed in a prevalent screen (aHR, 1.87; 95%CI: 1.57-2.24). **iii)** The risk of breast cancer independently increased with the presence of benign breast disease and with greater breast density (p-value for interaction = 0.84). **iv)** The quality of the different existing breast cancer risk prediction models was moderate with some limitations in the discriminative power and data inputs. A maximum AUC value of 0.71 was reported in the study conducted in a screening context. **v)** We designed a risk prediction model based on family history, previous benign breast disease and previous mammographic features. All 3 risk factors were strongly associated with breast cancer risk, with the highest risk being found among women with a family history of

breast cancer (aHR: 1.67), 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 on the time horizon selected.

Conclusions: i) Performance screening measures are negatively affected by breast density, with sensitivity and positive predictive value decreasing as breast density increases. **ii)** The risk of breast cancer conferred by benign breast disease differed according to type of screen (prevalent or incident). To our knowledge, this is the first study to analyze the impact of screening type on the prognosis of benign breast disease. **iii)** The risk of breast cancer independently increased with the presence of benign breast disease and with greater breast density and remained elevated for over 15 years. **iv)** Individualized risk prediction models are promising tools for implementing risk-based screening policies. However, it is a challenge to recommend any of them since they need further improvement in their quality and discriminatory capacity. **v)** We designed and internally validated a risk prediction model able to estimate short- and long-term breast cancer risk using information routinely reported at screening participation. The model included age, family history of breast cancer, benign breast disease and previous mammographic findings, which were found to be related to an increase in breast cancer risk. The model should be externally validated and updated with new variables.

Resumen

Palabras clave: Cáncer de mama, densidad mamaria, lesión benigna de mama, cribado de cáncer, mamografía, modelos de predicción, factores de riesgo

Introducción: Se ha demostrado que el cribado mamográfico reduce la mortalidad por cáncer de mama. Siguiendo las recomendaciones de la Comisión Europea, los países europeos han establecido programas poblacionales de cribado que ofrecen mamografías bienales a mujeres de entre 50 y 69 años de edad. Sin embargo, el cribado de cáncer de mama no está libre de controversia ya que existe un debate en cuanto al equilibrio entre la reducción de la mortalidad y los efectos adversos. Para mejorar este equilibrio, la evidencia científica actual apoya el cribado personalizado. Los estudios de modelización han demostrado que modificar el intervalo de cribado, la prueba de cribado o el rango de edad de la población objetivo en función del riesgo individual de las mujeres produce un mayor beneficio que las estrategias convencionales. Se han diseñado varios modelos de riesgo para estimar el riesgo individual de cáncer de mama de las mujeres en función de sus características personales. Sin embargo, la mayoría de estos modelos no se han desarrollado específicamente para estimar el riesgo de las mujeres participantes en el cribado de cáncer de mama. Por lo tanto, es necesario ampliar la información actual sobre los factores de riesgo de esta enfermedad y crear modelos de predicción del riesgo individual mediante el análisis de grandes bases de datos poblacionales.

Objetivo: El objetivo general de esta tesis es profundizar en el análisis del cribado poblacional del cáncer de mama. En concreto, esta tesis pretende evaluar diferentes factores de riesgo de cáncer de mama para desarrollar y validar un modelo de predicción de riesgo individual de esta enfermedad. Se analizó cómo la densidad mamaria afecta a los distintos indicadores de cribado en el contexto de la mamografía digital. A continuación, se evaluaron las diferencias en el riesgo de cáncer de mama en función de si una lesión benigna de mama se diagnosticó en un cribado prevalente o un cribado incidente. También se analizó la interacción entre la densidad

mamaria y las lesiones benignas en el riesgo de cáncer de mama. Posteriormente, se realizó una revisión sistemática para actualizar la evidencia existente, llevar a cabo una valoración crítica y una evaluación del riesgo de sesgo y resumir los resultados de los modelos de riesgo individualizados que se utilizan para estimar el riesgo de cáncer de mama en las mujeres de la población general. Por último, se diseñó un modelo de predicción individual del riesgo de cáncer de mama y se validó internamente, basado en información fácilmente accesible en un episodio de cribado.

Métodos: La población de estudio incluyó a todas las mujeres que participaron en 10 programas de cribado de cáncer de mama españoles desde 1995 hasta 2015 y de las que se hizo seguimiento hasta 2017. Se analizaron 2,853,753 mamografías de cribado de 782,406 mujeres de entre 50 y 69 años. Para el análisis específico de la densidad mamaria y los diferentes indicadores del cribado se utilizaron modelos de ecuaciones de estimación generalizada (GEE). Para el análisis principal de esta tesis, se utilizaron modelos parcialmente condicionales de Cox, una extensión para medidas repetidas del modelo de riesgos proporcionales de Cox en el que el Hazard Ratio de un evento determinado se modela a lo largo del tiempo. Mediante estos modelos evaluamos la diferencia de los factores de riesgo en las distintas categorías de cada factor. También diseñamos un modelo de predicción del riesgo individual. El modelo se validó internamente con el ratio esperado-observado y con el área bajo la curva ROC.

Resultados: i) La sensibilidad de la mamografía disminuyó del 89.2% en las mujeres con BI-RADS 1 al 67.9% en las de BI-RADS 4. Tanto el valor predictivo positivo de la reconvocatoria como el de las pruebas invasivas disminuyeron del 10.4% al 5.7% y del 49.8% al 32.4% en las mujeres con BI-RADS 1 y BI-RADS 4, respectivamente. **ii)** En comparación con las mujeres sin un diagnóstico de lesión benigna, el riesgo de cáncer de mama fue significativamente mayor en las mujeres con una lesión benigna diagnosticada en un cribado incidente (aHR, 2.67; IC95%: 2.24-3.19) que en aquellas con una lesión benigna diagnosticada en un cribado prevalente (aHR, 1.87; IC95%: 1.57-2.24). **iii)** El riesgo de cáncer de mama aumentó de forma independiente con la presencia de lesión benigna y con una mayor densidad mamaria (p-valor para la interacción = 0,84). **iv)** La calidad de los diferentes modelos de predicción de riesgo de cáncer de mama existentes fue moderada, con algunas limitaciones en el poder discriminativo y en la forma de recoger los datos. En el estudio realizado en

un contexto de cribado se registró un valor máximo del área bajo la curva ROC de 0.71. **v)** Se diseñó un modelo de predicción del riesgo individual basado en los antecedentes familiares, los antecedentes de lesión benigna previa y la presencia de patrones mamográficos previos. Los 3 factores de riesgo se asociaron fuertemente con el riesgo de cáncer de mama, encontrándose el mayor riesgo entre las mujeres con antecedentes familiares de cáncer de mama (aHR: 1.67), lesión benigna proliferativa (aHR: 3.02) y calcificaciones previas (aHR: 2.52). El modelo está, en general, correctamente calibrado (el ratio esperado-observado oscila entre 0.99 a los 2 años y 1.02 a los 20 años), pero sobrestima ligeramente el riesgo en las mujeres con lesión benigna proliferativa. El área bajo la curva ROC oscila entre el 58,7% y el 64,7%, dependiendo del horizonte temporal seleccionado.

Conclusiones: i) Los distintos indicadores de cribado se ven afectados negativamente por la densidad mamaria, disminuyendo la sensibilidad y el valor predictivo positivo de la prueba a medida que aumenta la densidad mamaria. **ii)** El riesgo de cáncer de mama conferido por una lesión benigna difiere según el tipo de cribado (prevalente o incidente). Hasta donde sabemos, este es el primer estudio que analiza el impacto del tipo de cribado en el pronóstico de la lesión benigna. **iii)** El riesgo de cáncer de mama aumenta de forma independiente con la presencia de una lesión benigna y con una mayor densidad mamaria y se mantiene elevado durante más de 15 años. **iv)** Los modelos de predicción son herramientas prometedoras para implementar políticas de cribado basadas en el riesgo individualizado. Sin embargo, es un reto recomendar cualquiera de ellos para la personalización del cribado ya que necesitan mejorar su calidad y capacidad discriminatoria. **v)** Diseñamos y validamos internamente un modelo de predicción de riesgo capaz de estimar el riesgo de cáncer de mama a corto y largo plazo utilizando la información recogida de forma rutinaria en el cribado mamográfico. El modelo incluye edad, antecedentes familiares de cáncer de mama, antecedentes de lesión benigna y patrones mamográficos previos, que resultaron estar relacionados con un aumento del riesgo de cáncer de mama. El modelo debe ser validado externamente y actualizado con nuevas variables.

Content

List of figures, equations and abbreviations

List of figures

List of equations

List of abbreviations

95%CI: 95 percent confidence interval
aHR: adjusted Hazard Ratio
AUC: Area under the receiving operating characteristic curve
BELE-2: BEnign LEsion project
BCRAT: Breast Cancer Risk Assessment Tool
BCSC: Breast Cancer Surveillance Consortium
BI-RADS: Breast Imaging Reporting and Data System
COVID-19: Coronavirus disease 2019
E/O ratio: Expected-to-observed ratio
FPR: False-positive recall
GEE: Generalized estimating equation
HR: Hazard ratio
IBIS: International Breast Cancer Intervention Study
IRIS: Individualized RISk project
MyPeBS: My personal breast screening
PPV: Positive predictive value
PPV-1: Positive predictive value of recall
PPV-2: Positive predictive value of invasive tests
RR: Relative risk

Presentation

Breast cancer is the most common tumor worldwide and is also the leading cause of cancer-related death in women. Because of its wide scope, it is a major public health problem. For this reason, over the last 50 years, a huge part of cancer research has focused on understanding the risk factors of this disease and on promoting strategies to improve its prognosis.

Several clinical trials carried out in the last 3 decades of the past century showed that early detection of breast cancer through mammography reduces mortality and is cost-effective. Thus, throughout the 1980s and 1990s, most European countries launched breast cancer screening programs to promote the early detection of this disease. These programs offer free mammograms to women in the target age range to detect breast cancer as early as possible and, therefore, take advantage of the benefit of screening: the mortality reduction and improvement in possible treatment. Most of these programs offer biennial mammograms to women between the ages of 50 and 69 years.

However, the risk-benefit balance of breast cancer screening is not free of controversy. In recent years, several articles have emerged that criticize the low effectiveness of breast cancer screening due to adverse effects, especially false positives and overdiagnosis (unnecessary treatment). In response to this debate, the current strategy of population-based screening has begun to be rethought with proposals adapted to individual women's personal breast cancer risk. This would involve moving from a one-size-fits-all strategy, a universal approach with biennial coverage of all women aged 50-69 years, to screening where periodicities, age range and techniques are defined according to women's individual risk of breast cancer. The implementation of personalized strategies, however, poses several unresolved challenges. The first is how to effectively identify women at higher risk of developing breast cancer in the context of population-based screening. In 2019 a statement of the expert meeting of the European Conference on Risk-Stratified Prevention and Early

Detection of Breast Cancer posited 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.

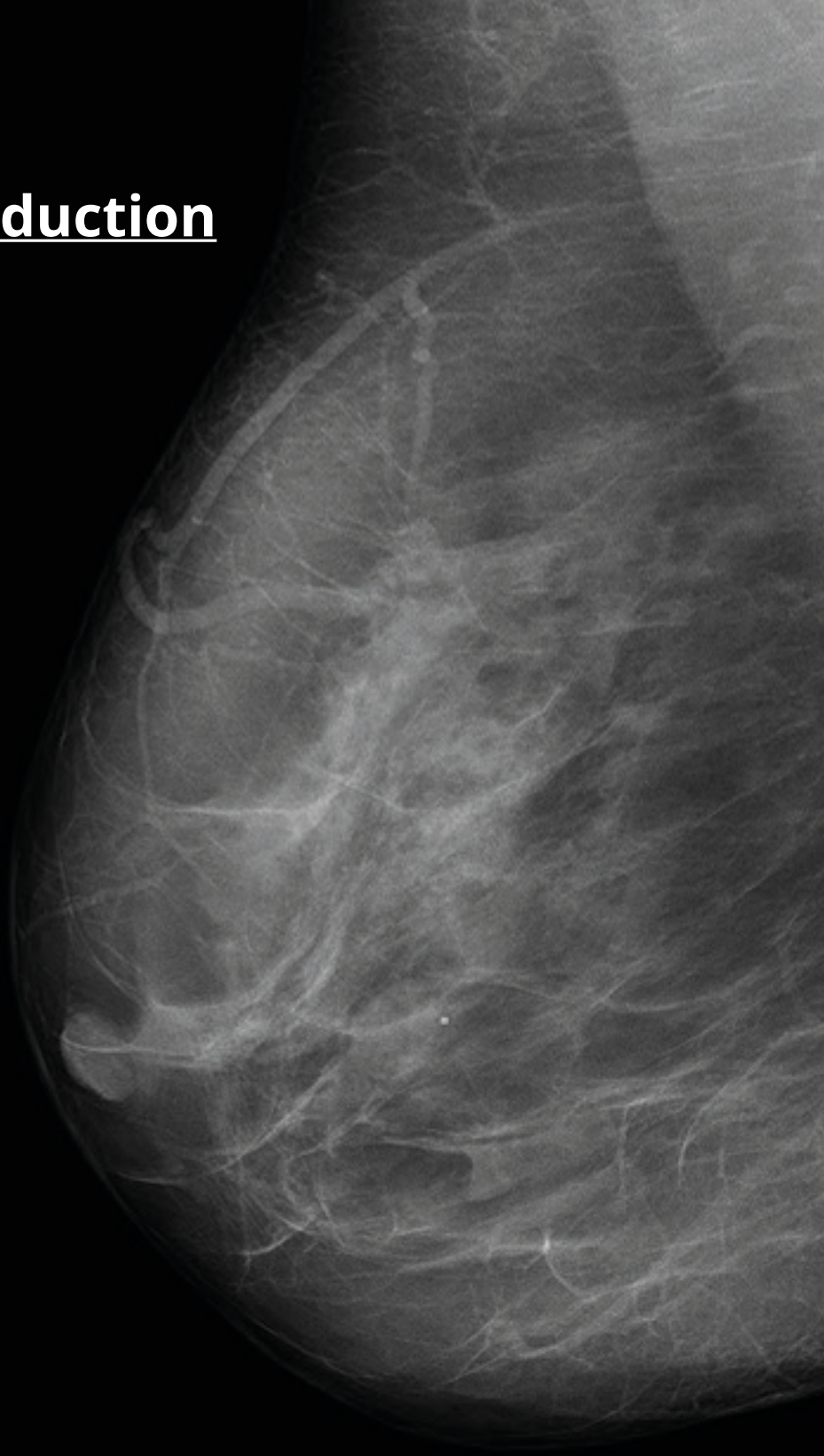
This thesis is part of the ongoing debate on personalizing breast cancer screening. Specifically, this thesis arises from the need to evaluate the different breast cancer risk factors with large longitudinal screening cohorts in order to develop an individualized breast cancer risk prediction model.

The first step of the thesis was to develop a retrospective cohort with information from the entire breast screening history of the population. The resulting database contains information from 10 screening programs in the Spanish context and includes over 2,800,000 screening mammograms from over 750,000 women, carried out between 1995 and 2015, with an average of 3 participations per woman. Then, the database was used to carry out different analyses to better understand the different breast cancer risk factors. To deepen the study, a systematic review of existing breast cancer prediction models was carried out to summarize the state of the art on the topic, and to identify needs for upcoming models. Finally, a breast cancer risk prediction model was designed, based on information easily accessible at screening.

This thesis is presented as a compendium of publications and is composed of 5 papers aiming to provide specific and unpublished answers to some of the questions described. It also attempts to highlight the complexity involved in the evaluation of mammographic screening in long-standing dynamic populations, as well as the need to use a longitudinal approach in order to correctly evaluate this procedure. The works presented were carried out at the Epidemiology and Evaluation Department of Hospital del Mar-IMIM, under the supervision of Dr. Marta Román and Dr. Xavier Castells. These tasks received specific funding from 2 projects of the Health Research Fund, FIS-ISCIII (PI15/00098, PI17/00047) and from the Alicia Llacer Award to the best research project by a young researcher awarded by the Spanish Society of Epidemiology.

I personally led the design, data collection, validation, and analysis of the database used in these projects.

Introduction



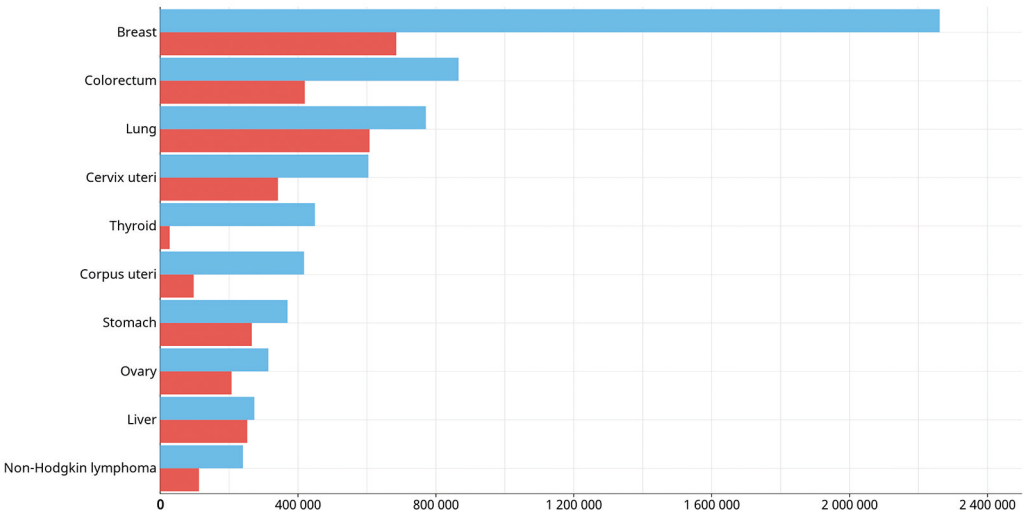


Figure 2. Estimated age-standardized incidence and mortality rates in 2020, worldwide, female, all ages.

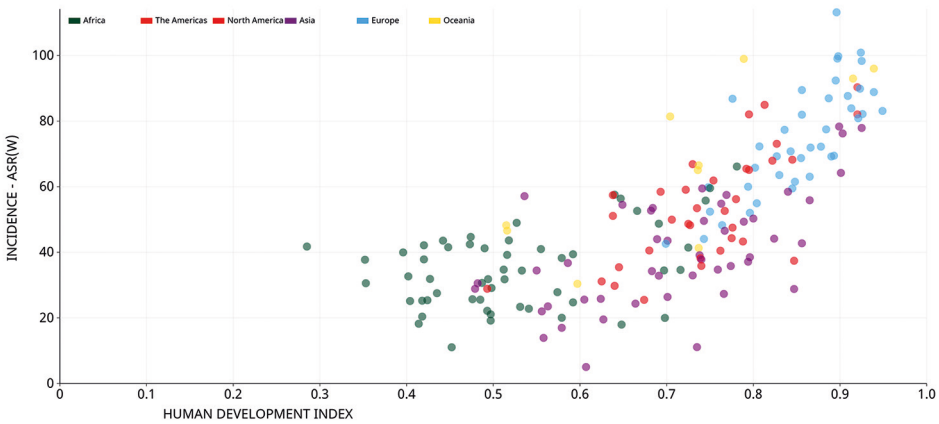


Figure 3. Incidence of breast cancer vs Human Development Index by country, in 2020, all ages.

A brief look at the epidemiology of breast cancer

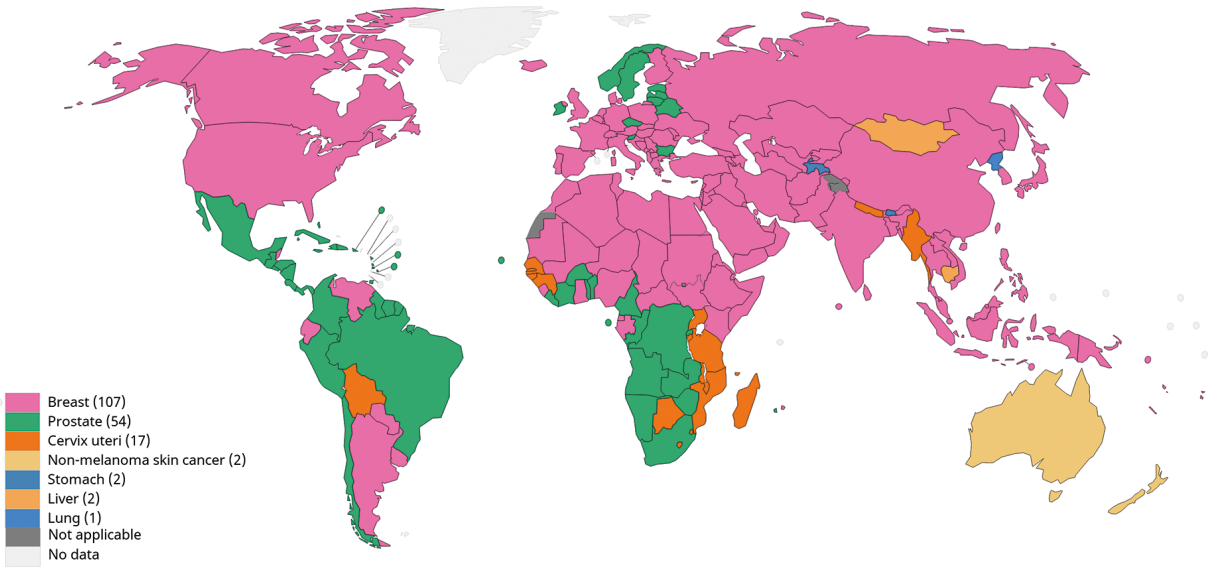


Figure 1. Most common cancer sites by country, estimated age-standardized incident rates in 2020, worldwide, both sexes, all ages.

In recent years, breast cancer has become the most common cancer in the world (Figure 1). It is estimated that 2,261,419 new cases of breast cancer were diagnosed in 2020 worldwide (1). With 684,996 deaths, it is also the leading cause of cancer-related death in women (1) (Figure 2).

Globally, breast cancer incidence varies widely around the world, with a clearly higher incidence in developed countries. This is evidenced by comparing breast cancer incidence and the Human Development Index across countries. This indicator was developed by the United Nations Development Program to analyze the average level of each country of the fundamental dimensions of human development (health, education and standard of living) (Figure 3).

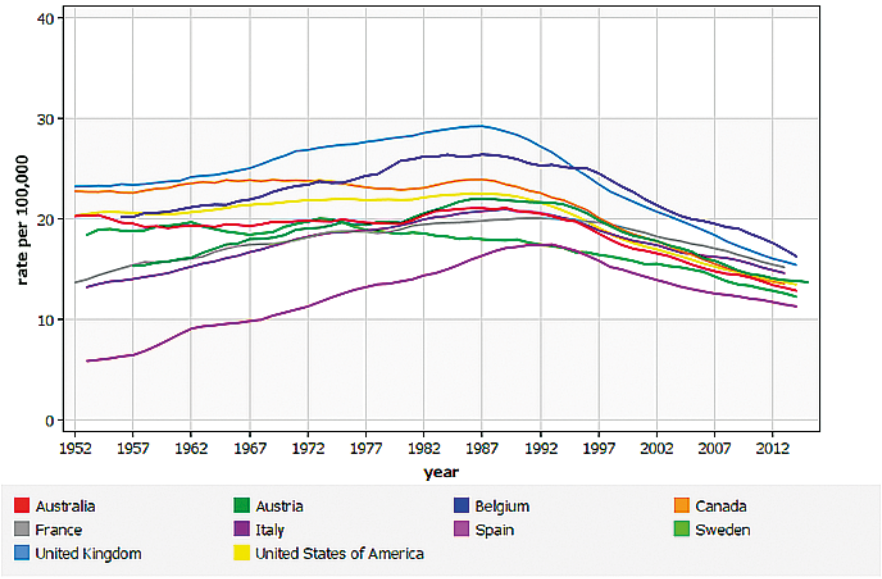


Figure 4. Age-standardized mortality rates for breast cancer in developed countries.

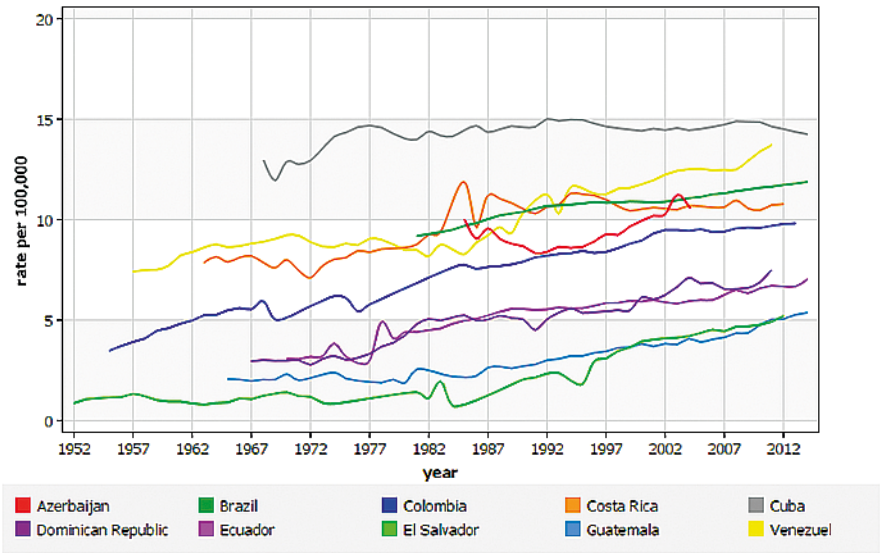


Figure 5. Age standardized mortality rates for breast cancer in undeveloped countries.

Source of Figures 1 to 5: World Health Organization International Agency for Research on Cancer (IARC). GLOBOCAN 2020: estimated cancer incidence, mortality, and prevalence worldwide (1).

Globally, breast cancer incidence is increasing. The World Health Organization estimates that by 2040 the annual number of cases will already exceed 3 million (1).

Despite the growing worldwide trend over the last 4 decades, since the 1990s most countries with a high per capita income, including Spain, have begun to see a decline in breast cancer incidence (1-10).

Similarly, over the last 2 decades, mortality from this disease has decreased substantially in developed countries and breast cancer is now one of the cancers with the highest survival rates, standing at over 80% at 5 years (1, 2, 11) (Figure 4).

However, due to the more favorable survival of breast cancer in developed countries, the range of mortality is much more homogeneous in developed regions (Figure 4) than in undeveloped countries (Figure 5).

This improvement in survival is attributed to improved treatment, the introduction of cancer functional units in hospitals and the implementation of early detection practices.

In Spain, more than 32,000 new breast cancer cases were estimated to be diagnosed in 2020 (12). The adjusted incidence rate for 2020 (using the world population as reference) was 77.5 new cases per 100,000 women, with the average for western European countries being 90.7 new cases per 100,000 women (1). Although Spain continues to be one of the European countries with the lowest incidence of breast cancer (1, 13), this disease is the most frequent cause of death from cancer in Spanish women, causing an estimated 6,000 deaths or more each year (12). The adjusted mortality rate in 2020 (using the world population as a reference) was 10.6 per 100,000 women, with the average for western European countries being 15.6 per 100,000 women per year (1).

Breast cancer screening

The genesis of screening programs

The main objective of early detection or screening is to reduce mortality from a disease by detecting it in the early stages. To achieve this aim, broad participation is essential. From a population perspective, it is expected that the implementation of screening will decrease mortality from the disease in the medium to long term. However, at an individual level, not all participates will benefit directly from such programs.

The genesis of screening programs dates back to the creation of a test for the detection of cervical cancer by the Greek physician Georgios Papanicolau in 1923, which was used in Canada in 1949 as the first known cancer screening test (14). Since then, numerous clinical trials have been conducted to identify the cancers that can be screened and the methods to do so. This practice is beginning to spread and currently there are screening programs for breast, cervix, prostate, lung, pancreas, mouth and colon and rectum cancers in different parts of the world.

The conditions necessary for the application of early detection tests for a disease were described at the end of the 1960s (15). First, the condition targeted should be an important health problem. Second, there should be an accepted treatment for patients diagnosed with the disease. Third, facilities for diagnosis and treatment should be available. Fourth, there should be a recognizable latent or early symptomatic stage. Fifth, there should be a suitable test or examination. Sixth, the test should be acceptable to the population. Seventh, the natural history of the condition, including development from latent to overt disease, should be adequately understood. Eighth, there should be an agreed policy about who should be treated as patients. Ninth, the cost of case-finding (including diagnosis and treatment of diagnosed patients) should be economically balanced in relation to possible expenditure on medical care as a whole. Finally, case-finding should be a continuing process and not a "once and for all" project. These criteria are considered classic and have been widely referenced. Subsequently, these criteria have been reviewed and updated.

Natural history of breast cancer

The natural history of breast cancer is not well understood, and there are various hypotheses on its origin and development. For many years the most widespread theory was the Wellings Jensen model (16, 17), which postulates that invasive breast cancer develops from the accumulation of multiple genetic alterations in benign breast disease. The various stages of this model are shown in Figure 6. However, only a low percentage of benign lesions eventually progress to invasive cancer.

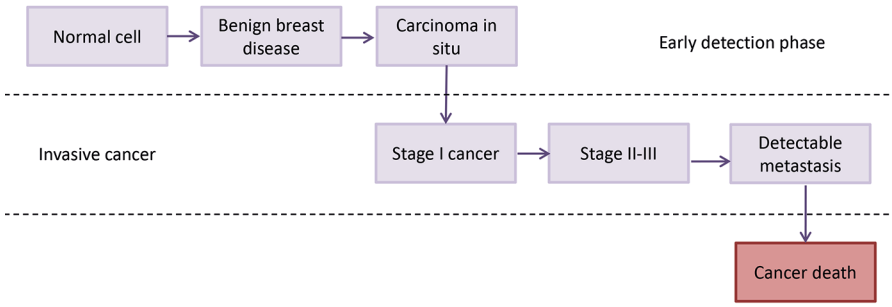


Figure 6. Lineal model of cancer progression.

On the other hand, there is evidence that not all lesions follow this linear pattern. More recent articles (18, 19) suggest that, although in situ carcinomas are possible precursors of invasive carcinomas, not all invasive carcinomas originate from an in situ carcinoma. It is postulated that some subtypes develop from the progression of progenitor cells, which would subsequently be affected by different mutations.

Rationale and evidence-base for screening mammography programs

Breast cancer meets all the criteria proposed by Wilson and Jungner (15), which favors the deployment of early detection programs through mammographic screening.

Since the early 1980s, several randomized controlled trials have studied the effect of mammography screening of breast cancer, assessing different

screening ages and mammogram frequencies (20-25) (Figure 7).
The evidence on the effectiveness of mammography screening in reducing mortality from the disease is widely accepted (26-39). Although the methods employed in some of these studies have been questioned, mainly due to the randomization criteria used, overall, the results show a 23% reduction in breast cancer mortality in women who have been screened (Figure 8).

	New York HIP	Malmö I and II	Swedish Two County	Canada I and II	Stockholm	Göteborg	UK Age trial	Edinburgh
Start date	1963	1976	1977	1980	1981	1982	1991	1978
Randomisation method	Individual	Individual	Cluster	Individual	Day of birth	Day of birth*	Individual	Cluster
Population of women								
Source	IC	P	P	Various†	P	P	PC	PC
Number of women‡ (clusters)	62 000	60 076	133 065 (45)	89 835	60 800	52 222	160 921	54 654 (87)
Age group (years)	40–64	45–69 and 43–49	38–75	40–49 and 50–59	39–65	39–59	39–41	45–64
Invited group intervention	M+PE	M	M+SE	M+PE+SE	M	M	M	M+PE
Mammography								
Number of views	2	2 then 1 or 2	1	2	1	2 then 1	2 then 1	2 then 1
Screening interval (months)	12	18–24	24–33	12	24–28	18	12	24
Number of screening rounds	4	6–8	2–4	4–5	2	4–5	8–10	2–4
Duration of screening (years)	3	12	7	5	4	7	8	6
Attendance	65%	74%	85%	88%	82%	84%	81%	65%
Control group intervention	None	None	None	PE+SE§	None	None	None	None
Follow-up								
Controls invited for screening¶	Not known	Never‖	After 7 years	Never	After 4 years	After 7 years	After 10 years	After 10 years
Cause of death determination	L	IEC, NS	L, IEC, NS	IEC, NS	IEC, NS	NS	NS	NS

Information is taken from various publications, but mainly the Cochrane Review;[5] Nyström and colleagues;[7] and Tabar and colleagues.[8] *These summaries are sometimes simplifications of characteristics that differ between subtrials or subgroups. Some discrepancies also exist between different publications. IC=insurance company register. P=population register. PC=primary care register. M=mammography. PE=physical examination. SE=self-examination. L=local. IEC=independent endpoint committee. NS=national statistics or register. †Day of birth, and later individual. ‡Includes P. IC, employee recruitment, and general publicity. Women were randomly assigned after initial PE, and evidence suggests that the women attending screening had a higher rate of breast cancer at that initial attendance than was expected from an age-matched population. §Some of these numbers are approximate, because the numbers vary in different publications. ¶After the initial assessment, only the women in Canada II underwent systematic PE during the screening period—in Canada I, they were taught how to do a physical examination. ‖Systematic invitation of all controls. ||Applies to Malmo I ages 55–69 years.

Table 1: Characteristics of the randomised trials of breast cancer screening

Figure 7. Main breast cancer screening trials.

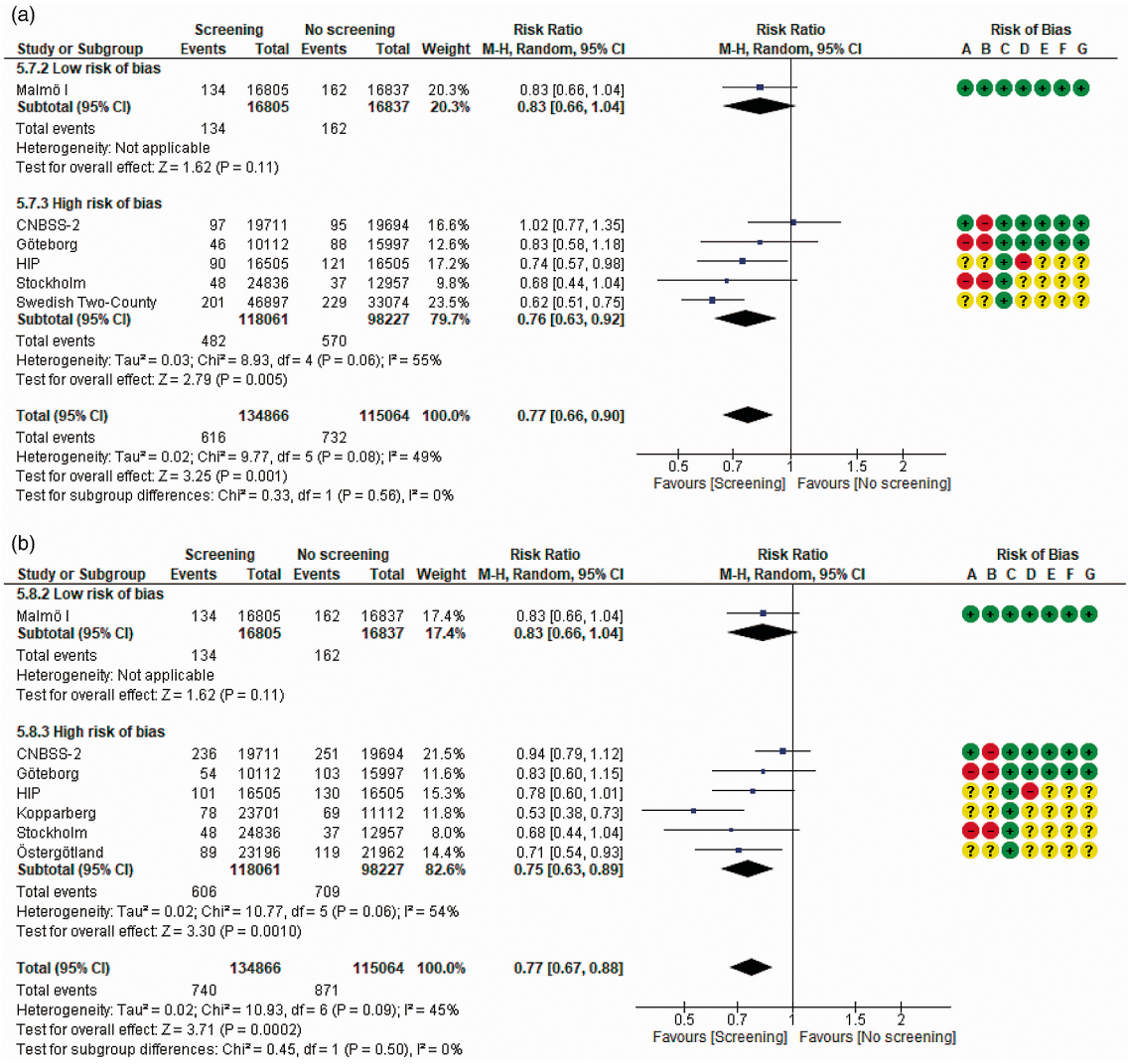


Figure 8. Effect of mammography screening on breast cancer mortality (women aged 50–69).

(a) short case accrual, mean follow-up across studies 17.6 years; (b) longest case accrual, mean follow-up across studies 15.5 years. Risk of bias legend: (A) random sequence generation, (B) allocation concealment, (C) blinding of participants and personnel, (D) blinding of outcome assessment, (E) incomplete outcome data, (F) selective reporting and (G) other bias.

Source Figure 6: Adapted from Esserman et al (18).
Source Figure 7: Adapted from The Independent UK Panel on Breast Cancer Screening (20).
Source Figure 8: Adapted from Canelo-Aybar C, et al. (36).

Breast cancer screening implementation in Europe

Following the publication of the first randomized trials showing favorable results (20, 21), in the 1990s, most scientific communities began to recommend mammography screening for the early detection of breast cancer. Many countries have progressively implemented population-based programs. Currently, the European Council recommends population-based breast cancer screening in women aged 50-69 years (40-42). The Council also recommends that programs follow the standards set out in the European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis (42, 43). By 2017, breast cancer screening programs had been implemented by all 28 member states of the European Union.

All programs use mammography as the screening test, and most countries have completely replaced analogic with digital mammography. Overall, it is estimated that about 25 million women between the ages of 50 and 69 years in the European Union are invited to participate in screening programs annually, out of an estimated total of 32 million women (invitation coverage 79.8%) and 16 million have been screened in these programs (participation coverage 49.2%) (41).

Although all member states follow the recommendations of the European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis (42, 43), which aim to standardize screening practices, there are organizational differences between screening programs across Europe. The greatest consensus can be found in the frequency of screening, which is 2 years in all countries, except in the United Kingdom and Malta, where it is 3 years. The age range of the target population is usually 50-69 years, but can be smaller (50-59 years, 50- 65 years) or larger (45-69 years, 40-70 years) in some countries. Most of the differences between programs, however, fall under the eligibility criteria and the screening practice protocol (41).

Breast cancer screening implementation in Spain

In Spain, the first population-based breast cancer screening program was launched in Navarre in 1990. The program is part of the National Health System strategy (44). Subsequently, programs were gradually initiated in the rest of the autonomous regions until reaching total coverage of the target population in 2006 (45). The organization of the programs falls to the health authorities of the various autonomous regions. Although the screening programs are independently managed, they are linked through a screening program network (46). The entities responsible for the different programs meet annually, and share the results of indicators related to organization, resources and other elements aimed at ensuring their quality based on specific and joint evaluation of the programs.

All programs in Spain have adopted the recommendations of the European guidelines for breast cancer screening (42, 43) and use biennial mammography as the screening test. The age group of the target population is 50-69 years. In addition, 6 autonomous regions also include the group aged 45 to 49 years. Data from 2016 indicate that screening coverage in Spain (defined as the proportion of the target population of the programs and the reference population according to the Spanish National Institute of Statistics data) is more than 90%, and that overall participation (defined as the number of participating women among the total number of women invited) is 75.7% (46).

However, there are substantial differences between programs, especially in the method of reading, single or double, and in the case of double reading, whether it is done with consensus or arbitration. In addition, regarding the type of mammography, since 2000, some programs have chosen to replace analogic with digital mammography. By 2017, 13 autonomous regions had introduced digital mammography in all or some of their screens (46).

The results of the screening process in Spain show that the overall percentage of additional tests in initial and subsequent screening was 12.3% and 4.1%, respectively, while the total detection rate was 4.1 tumors per 1,000 screening examinations (46).

Benefits and harms of breast cancer screening

Randomized clinical trials have demonstrated a clear effect of mammography screening on reducing breast cancer mortality under relatively controlled circumstances. The effectiveness of screening is expected to have improved since the publication of the first randomized trials. Quality assurance, training of specialist staff, and mammographic techniques have all improved over time (47), as has therapy through the widespread use of specific treatments (48).

In 2001, the Cochrane Collaboration published a meta-analysis of the results of the various randomized controlled trials, concluding that mammographic screening did not improve survival, and that the effects on breast cancer mortality were inconclusive (49, 50). The publication of that study was a turning point, sparking a debate about the appropriateness of breast cancer screening with current strategies. The results of the meta-analysis were disputed (51) and in 2006, the Cochrane Collaboration itself qualified its conclusions in a review of the article, showing a 15% reduction in breast cancer mortality due to mammographic screening (52).

In 2012, the EUROSCREEN Working Group published a systematic review of population-based studies conducted in Europe assessing the benefits and adverse effects of screening. The review estimated a 25%-31% reduction in breast cancer mortality (39). In parallel, a review of the benefits and adverse effects of screening by The Independent UK Panel on Breast Cancer Screening, based on the initial trials, estimated the mortality reduction to be 20% (28). At the end of 2014, the World Health Organization report reviewing the evidence on the benefits and adverse effects of screening concluded that population-based mammographic screening in women aged 50-69 years reduced mortality from this disease and was cost-effective in upper-middle income countries (38). However, these estimates vary widely from study to study, depending on their data and methodology, and other studies have estimated the mortality reduction to be 25% (31) or 43% (30). In 2020, a systematic review of European articles estimated that the reduction in breast cancer mortality in attenders versus non-attenders at screening ranged between 33% and 43% (northern Europe), 43%-45% (southern Europe) and 12%-58% (western Europe) (35). In 2021, a systematic review by the European

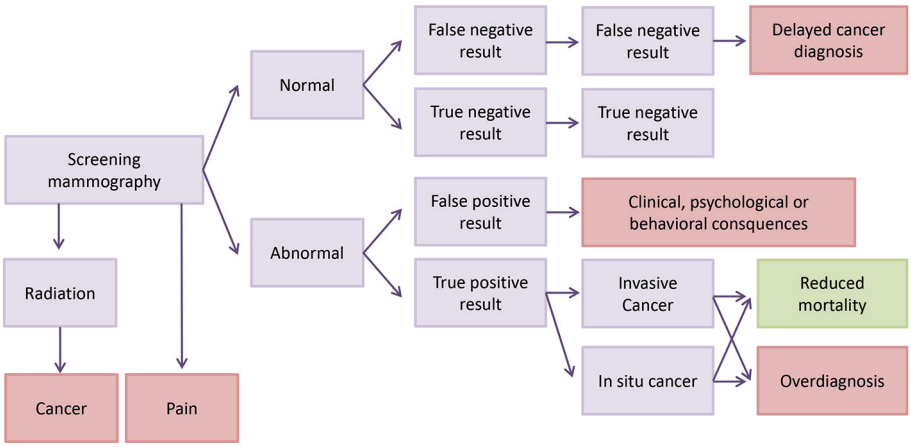


Figure 9. Risks and benefits of screening mammography.

Commission Initiative on Breast Cancer estimated a 23% reduction in breast cancer mortality in women who have been screened (36).

Like any health intervention, breast cancer screening has adverse effects (Figure 9). Their existence has been recognized for many years (54). Since mammographic screening is offered to a large, asymptomatic population, its adverse effects must be minimized, maintaining an acceptable risk-benefit balance. At the individual level, however, not all participating women will obtain the same benefits from screening nor will they experience the same adverse effects. The main adverse effects of screening are false positives, overdiagnosis and interval cancers, each with a different scope and consequences (53). False-positives are defined as results that incorrectly indicate the presence of cancer. To definitively diagnose or rule out malignancy, the patient is recommended to undergo further procedures (mammography, ultrasound, magnetic resonance imaging, fine needle biopsy, excisional biopsy, and/or surgical biopsy). False positive results cause concern and anxiety in affected women, as well as additional tests with an associated cost (55-58). In addition, several studies have shown that these women have a lower adherence to the program in successive calls and are at increased risk of developing breast cancer (59-63). At a population level, false-positives are the adverse effect with the strongest impact, both because of the number of women affected, and because of the resources and the care burden they entail. It is estimated that 1 out of every 5 women participating in screening will have

a false-positive result over 10 biennial participations in mammographic screening of the population aged between 50 and 69 years (64).

Overdiagnosis can be defined as the detection of histologically-confirmed breast cancer in screening that would never have been clinically diagnosed during a woman's lifetime (65). Because of the emotional impact generated by the diagnosis and treatment, it is considered to be the most serious adverse effect of early detection and is that generating the most controversy. On an individual level, it is impossible to distinguish which tumors have been overdiagnosed and which have been detected early, and both are treated equally.

At the population level, however, overdiagnosis involves treatment of a larger number of women, which increases costs as well as the number of women who will suffer from stress, anxiety, and the possible adverse effects of treatment due to the diagnosis of the disease. There are several mathematical approaches to estimate overdiagnosis without a standard methodology for its calculation. The main problem is to have a comparable reference population in the absence of screening. Estimates of overdiagnosis by the EUROSCREEN Working Group indicate rates of between 1% and 10%, with an estimated mean value of 6.5% (39), while the review by The Independent UK Panel on Breast Cancer Screening estimates that 19% of tumors diagnosed by screening are overdiagnosed (28). In 2021, the European Commission Initiative on Breast Cancer review estimated a 17.3% of overdiagnosis from an individual perspective (36). However, the estimates published to date vary widely and are highly susceptible to biased interpretation, ranging from 1% to 60% (66) (Figure 10).

In contrast, interval cancers can be considered a limitation of screening. Interval cancers are primary tumors diagnosed after negative screening and before the next invitation (42, 43). From a review of screening mammograms, they can be classified into true interval cancers (when the previous screening mammogram showed no suspicion of malignancy), false-negatives (when a retrospective review indicates that there was a suspicion of malignancy that was not detected either due to errors in radiological interpretation, or because of technical errors in performing the mammogram or additional tests), minimal signs (non-specific suspicion in the previous screening mammogram) and occult tumors (not visible by mammography). The incidence of interval cancers has

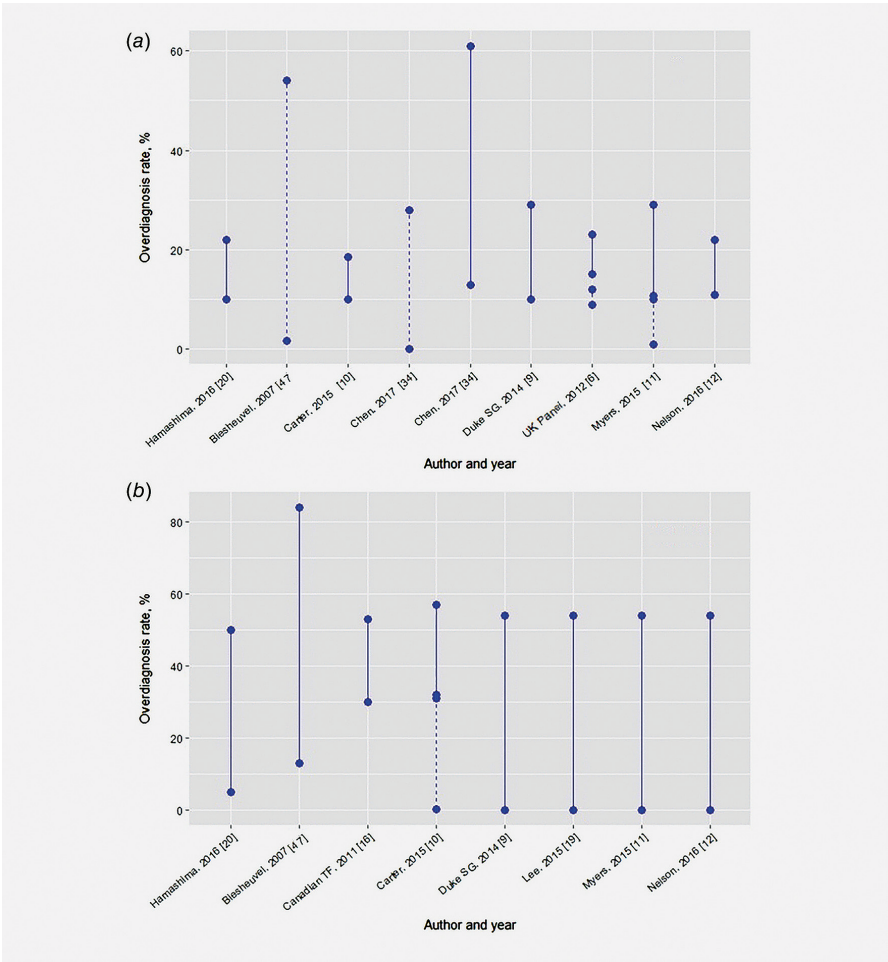


Figure 10. Overdiagnosis rates reported in systematic reviews.

(a) randomised controlled trials and (b) observational studies.

Source Figure 9: Armstrong K et al. (50).
Source Figure 10: Mandrik O et al. (63).

been estimated to range from 1.8 to 2.9 per 1,000 women screened (67). Interval cancers, and specifically false-negative tumors, delay diagnosis and are associated with a larger tumor size (68), greater nodal involvement (69) and shorter survival, with a 53% higher risk of dying from interval cancer compared to screen-detected breast cancer (70), thus reducing the effectiveness of screening. In addition, they create a false sense of security in women, making them less alert to possible signs of the disease.

Controversies in breast cancer screening: towards personalized screening

In the aforementioned EUROSCREEN Working Group article (39), the authors concluded that, with the current evidence generated from observational studies, breast cancer screening continues to be recommended and that the benefits of screening outweigh the adverse effects. That study estimated that for every 1,000 women screened biennially for 20 years between the ages of 50 and 69 years and followed-up until the age of 79 years, 71 breast cancers are diagnosed, 7-9 deaths are prevented, 4 tumors are overdiagnosed, and 200 women experience a false positive result. The Independent UK Panel on Breast Cancer Screening report concludes that for every 1,000 women aged 50-69 years screened for 20 years, and followed-up until the age 79 years, 68 breast cancers are diagnosed, 4 deaths are avoided, and 13 tumors are overdiagnosed (28).

The Independent UK Panel report highlights the need to review the information given to invited women, to make it as transparent as possible and allow informed decision-making. In the USA, the debate is ongoing with the publication of 2 articles that found that one in 3 cancers diagnosed by screening would be overdiagnosed and that, at best, screening would only have a small effect on reducing mortality from this disease (1, 72). In contrast, another publication argues that calculation of overdiagnosis in those articles is probably grossly overestimated (73). Because of the methodological limitations in the selection of comparison groups, especially the group of unscreened women, and the different methodological approaches used, evaluation of the benefits of population-based breast cancer screening remains a question with heterogeneous and inconclusive answers.

The likelihood that a woman will benefit from screening mammography depends on her risk of developing clinically significant breast cancer in her lifetime. Taking individual risk factors beyond age into account should enable the classification of women into groups at varying risks of breast cancer. One side of the current debate holds that personalized risk-based screening going beyond the current 'one-size fits all' recommendation may increase the effectiveness and benefit-harm balance of breast cancer screening (74-77). This approach proposes personalized screening strategies depending on the individual risk of each woman. This encompasses not only a more exhaustive screening of women at higher

risk or with more specific techniques such as ultrasound (78) or magnetic resonance imaging (79), but also a less exhaustive screening of women at lower risk (80). Thus, it is intended to maximize the benefits of screening and reduce the adverse effects. Some clinical trials are currently testing the effectiveness of personalization, but still without results (81, 82). If screening based on individual risk is accepted as the new screening paradigm, there will also be opportunities for primary prevention.

The implementation of personalized strategies, however, poses several unresolved challenges. The first is how to effectively identify women at greater risk of developing breast cancer.

Breast cancer risk factors

There are many characteristics that over the years have been proven to be risk factors for breast cancer. Analyzing all of them is beyond the scope of this thesis, which focuses on the risk factors relevant to screening mammography and useful for the development of personalized screening strategies.

Age

Breast cancer incidence and mortality increase proportionally with age. Only around 10% of cases are diagnosed in people younger than 45 years and less than 5% in people younger than 40 years (83) (Figure 11). Risk increases from the age of 50-55 years, associated with the onset of menopause and indicating the relevance of reproductive and hormonal factors in the aetiology of the disease (84) but then remains stable until the age of 65-70 years, when it declines.

Family history

Having a family history of breast cancer plays a significant role in a patient's lifetime breast cancer risk (85-89). A study published in 2001 by the Collaborative Group on Hormonal Factors in Breast Cancer (90) estimated that the risk ratios for breast cancer rose with increasing numbers of affected first-degree relatives: compared with women who had no affected relatives, the ratios were 1.80 (95% CI: 1.69-1.91), 2.93 (95% CI: 2.36-3.64), and 3.90 (95% CI: 2.03-7.49), respectively, for one, 2, and 3 or more affected first-degree relatives.

Benign breast disease

Benign breast disease is defined as non-malignant breast alterations that have also been proven to be associated with an increased risk of breast cancer both in the clinical setting (91-93) and in population-based screening (94). These lesions are usually classified as non-proliferative lesions, proliferative lesions without atypia and proliferative lesions with atypia (95, 96). A study by Hartmann in 2005 estimated a breast cancer risk ratio of 1.27 (95%CI: 1.15-1.41), 1.88 (95%CI: 1.66-2.12) and 4.24 (95%CI: 3.26-5.41) for people diagnosed with non-proliferative lesions, proliferative lesions without atypia and proliferative lesions with atypia, respectively, compared with woman who never had a benign breast disease diagnosis (92).

Suspicious mammographic findings

Despite being less well studied than benign breast disease, suspicious findings at mammogram reading such as masses, calcifications, distortions, or asymmetric density are also associated with an increased risk of subsequent breast cancer. A 2016 study by Castells et al (97) estimated the hazard ratio of subsequent breast cancer in women with suspicious findings to be 1.59, 2.24, 1.58, 2.09 and 2.73 for masses, calcifications, asymmetric density, distortion and calcifications associated with mass, respectively.

Breast density

Breast density reflects the amount of fibrous and glandular tissue compared with the amount of fatty tissue in a woman's breasts. Levels of density are commonly described from less to more dense using a results reporting system called the Breast Imaging Reporting and Data System (BI-RADS) (98) as: A: almost entirely fatty, B: scattered areas of fibroglandular density, C: heterogeneously dense, D: extremely dense. Denser breasts have a higher risk of breast cancer (99). A 2015 study (100) estimated that women with higher density were between 2.3 and 4.1 times more likely to develop breast cancer than those with lower density, depending on their age.

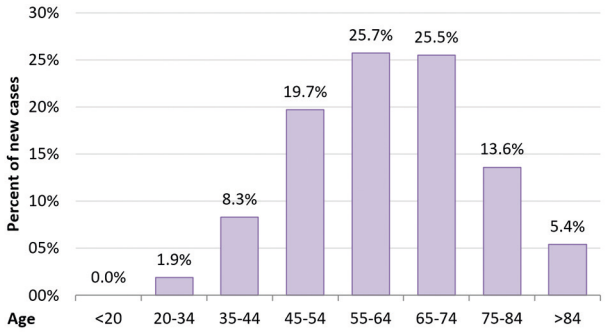


Figure 11. Percent of New Cases of Breast Cancer in US 2013-2017 by Age Group.

Source Figure 11: Adapted from National Cancer Institute (79).

Breast Cancer Risk prediction models

Breast cancer risk prediction models through history

Estimating women's individual risk of breast cancer based on their personal characteristics is not a recent idea. The first approach emerged in the early 1980s with an article that attempted to measure the incidence of breast cancer by the age of breast tissue (101). The first model as such

appeared at the end of that decade, when the first version of the model currently known as BCRAT (Breast Cancer Risk Assessment Tool) was developed, which aimed to estimate the probability of developing cancer based on age, family history of breast cancer, age at first birth, age at menarche, and previous biopsies to assist in medical counseling (102). Over the years, this model has been improved and has been joined by other widely used models, such as the BCSC (Breast Cancer Surveillance Consortium) model (100, 103), the IBIS (International Breast Cancer Intervention Study) model (104), the Rosner & Colditz model (105-108) and the BOADICEA model (109). Breast cancer has a large number of (possible) risk factors, and the number included in the different models varies widely (from 2 to 20). Some of those tested by the different models over the years are body mass index, hormone replacement therapy, alcohol consumption, physical activity, diet, breast density, benign breast disease, family history, parity, a polygenic risk score or hormone information, among others.

However, none of these models was developed for women participating in mammography screening but were rather created for the general population or for specific high-risk groups. Of note, a useful and practical risk model for screening populations would require a model that uses only variables that are readily and easily available at mammographic screening. To date, only one model has been specifically designed to predict individual risk among women participating in breast cancer screening (110). Although truly relevant, the model was based on short-term risk estimates, with only 3 years of follow-up, and it did not account for relevant characteristics of prospective studies such as internal time-dependent covariates.

The need to create an updated individualized breast cancer risk prediction model

As discussed above, evidence is beginning to emerge in support of the introduction of personalized screening based on women's individual risk rather than the current "one size fits all" strategy. In 2019 a statement of the European Conference on Risk-Stratified Prevention and Early Detection of Breast Cancer experts meeting (111), posed 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. The

power of calibration and discrimination of the model used to assign risk to each woman will be key to minimizing the adverse effects of screening.

Methodological challenges of creating an individualized risk model

To be able to build this model, information from a large number of women followed over a long period will be needed. Detailed and in-depth analysis of women's participation, beyond cross-sectional evaluation of the programs themselves, will require individualized information on each woman's participation in screening. Therefore, the different risk factors that will form part of the model will be needed for each participation, as well as long-term follow-up to check whether each woman received a breast cancer diagnosis or not. In general terms, this information is specific to each screening program and is not uniformly collected or coded.

The need to include a large volume of women in the model is justified by the frequency of the different events. The frequency of a breast cancer diagnosis in each participation in mammography screening is around 5.5‰ and the frequency of proliferative benign breast disease, which will be one of the risk factors, is around 1‰. This indicates that the group of women with proliferative benign breast disease diagnosed in a screening participation and a breast cancer diagnosed at any subsequent participation is extremely uncommon, and hence, that a large sample size is needed to allow performance of the pertinent analyses.

As population programs were deployed during the last 5 years of the 1990s and the early 2000s, various studies and approaches have emerged to evaluate their effectiveness. The current situation in which most programs have been running for almost 2 decades presents an updated scenario with the possibility of analyzing longitudinal data on screening practice over a long period of time. This scenario begs new questions about how to evaluate screening and how to work with this type of cohort.

An appropriate approach to the analysis of these data over the long term should consider the correlation between multiple observations of

women over the study period (repeat measurements). In addition, certain study characteristics, such as the presence of benign breast disease, the reading method or the screening technique used (digital or screen-film mammography) may change between different observations in the same woman, making it desirable to be able to include time changing variables, as opposed to assuming fixed information throughout the study period.

Finally, women in the target population are invited to undergo mammography every 2 years, which could substantially restrict our observations; however, but with the help of population and hospital records, information can be obtained on those cancers that are detected between 2 mammograms (known as interval cancers). This last peculiarity allowed us to perform the analyses from a continuous time perspective and therefore to use survival models such as Cox regression models.

Presentation of the articles constituting this thesis

Article 1

Mammographic breast density: How it affects performance indicators in screening programmes?
Posso M, Louro J, Sanchez M, Román M, Vidal C, Sala M, Baré M, Castells X on behalf of the BELE study group
Eur J Radiol. 2019;110:81-7
DOI: 10.1016/j.ejrad.2018.11.012

Article 2

Differences in breast cancer risk after benign breast disease by type of screening diagnosis
Louro J, Román M, Posso M, Comerma L, Vidal C, Saladie F, Alcántara R, Sanchez M, Quitana MJ, Del Riego J, Ferrer J, Peñalva L, Bargalló X, Prieto M, Sala M & Castells X
Breast. 2020;54:343-348
DOI: 10.1016/j.breast.2020.09.005

Article 3

Breast density, benign breast disease, and risk of breast cancer over time
Román M, Louro J, Posso M, Alcantara R, Penalva L, Sala M, del Riego J, Prieto M, Vidal C, Sánchez M, Bargalló X, Tusquets I & Castells X
European Radiology. 2021
DOI: 10.1007/s00330-020-07490-5

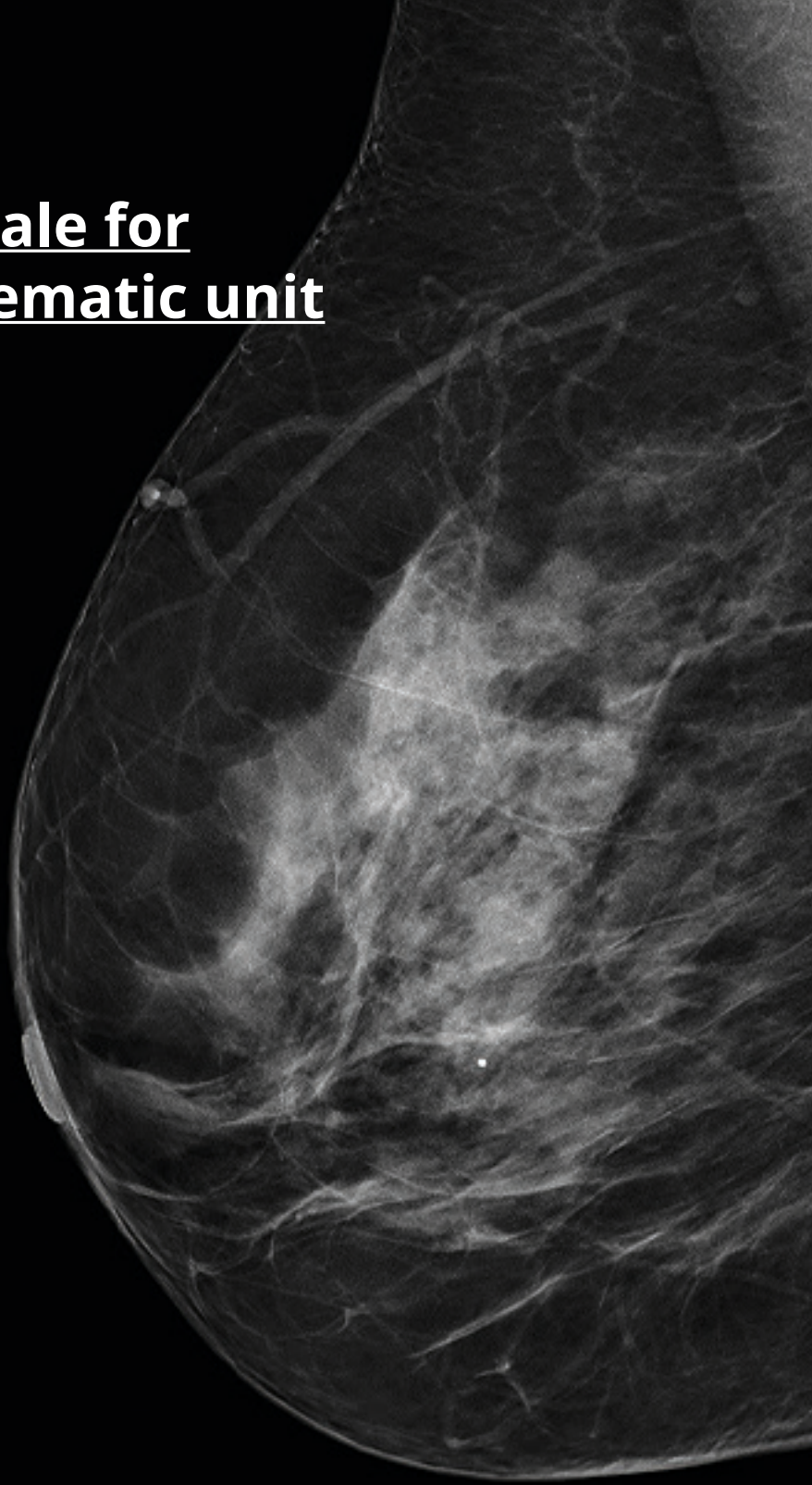
Article 4

A systematic review and quality assessment of individualized breast cancer risk prediction models
Louro J, Posso M, Hilton Boon M, Román M, Domingo L, Castells X & Sala M
Br J Cancer. 2019;121(1):76-85
DOI: 10.1038/s41416-019-0476-8

Article 5

Developing and validating an individualized breast cancer risk prediction model for women attending breast cancer screening
Louro J, Román M, Posso M, Vazquez I, Saladié F, Rodriguez-Arana A, Quintana MJ, Domingo L, Baré M, Marcos-Graguera R, Vernet-Tomas M, Sala M, Castells X on behalf of the BELE and IRIS Study Groups
PLoS One. 16(3): e0248930
DOI: 10.1371/journal.pone.0248930

Rationale for the thematic unit



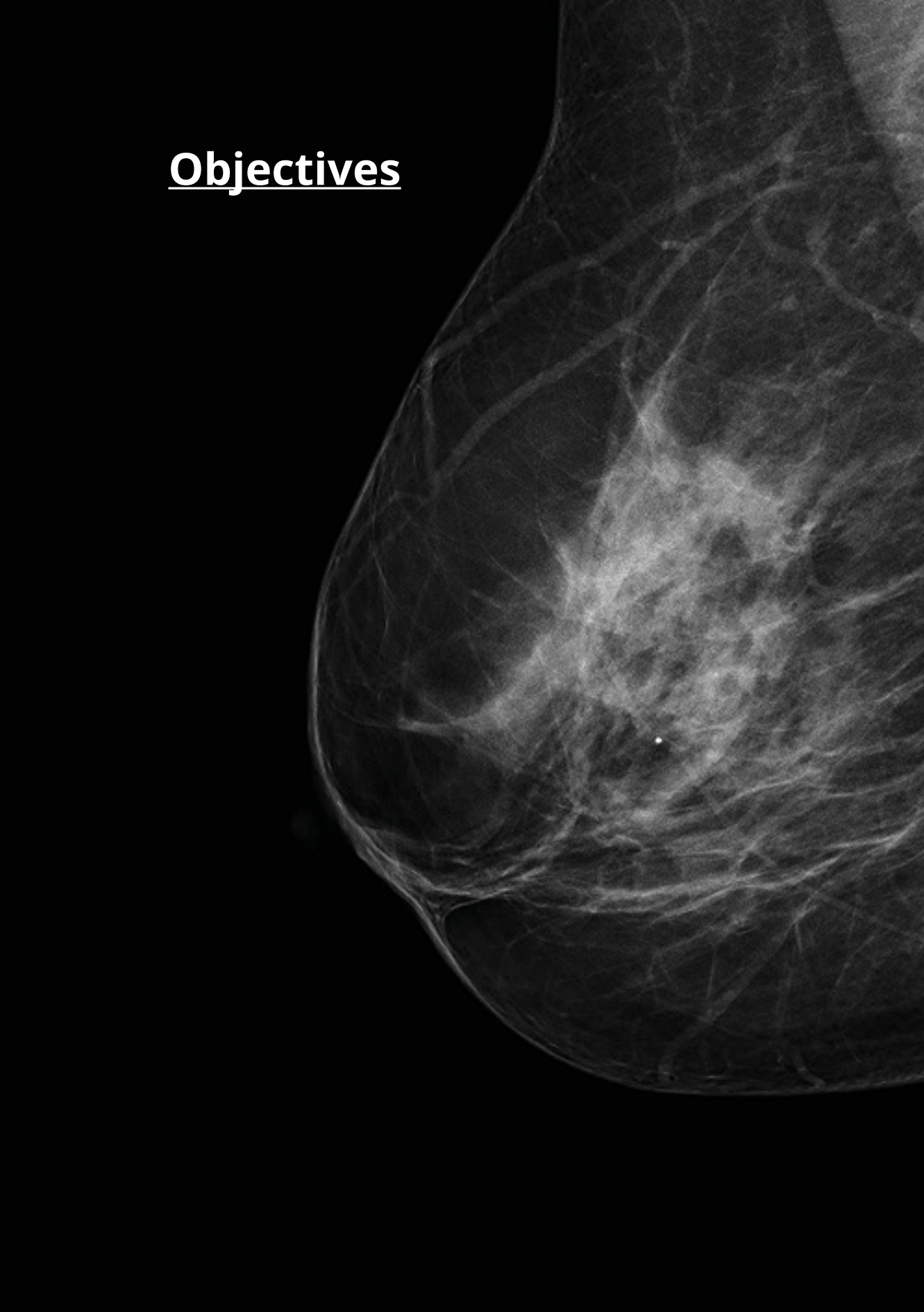
Breast cancer screening is one of the most widely evaluated population-based interventions. Despite the current debate about the balance between benefits and adverse effects of mammographic screening, health authorities at the national and European level continue to recommend population-based breast cancer screening for women aged 50 to 69 years. The current questioning of the effectiveness of screening highlights the need for its improvement, with new insights and novel data to evaluate possible alternatives in current early detection strategies. Breast cancer screening is offered to millions of women worldwide and therefore the impact of enhanced effectiveness of this practice could represent a substantial benefit for the whole population.

This thesis is framed within the current screening debate, with special emphasis on the evaluation of the different breast cancer risk factors, by means of large longitudinal screening cohorts, with the goal of developing an individualized breast cancer risk prediction model. The possibility of extending the information known to date about risk factors for breast cancer and the estimation of individual risk prediction models through the analysis of large population-based databases opens a window to a more accurate and reliable evaluation of mammographic screening. In particular, this thesis is based on 2 projects designed specifically to assess population-based screening and using longitudinal methodology that allow the analyses of the complicated information structures required for the evaluation of mammographic screening.

The longitudinal and complex structure of the database designed for the projects of this thesis has provided answers to research questions that were hitherto unfeasible. Furthermore, due to the potential of the data, it has been possible to extend their utility by answering other questions concerning breast cancer risk factors. As a complement to the work of this thesis and as part of the BEnign LEsion (BELE-2) and the IRIS (Individualized RISk) projects on which this thesis is based, other works have been carried out, presented as appendices in which the author has collaborated closely with the contribution of data and statistical analysis.

The results of this thesis will broaden overall knowledge of the different breast cancer risk factors and of how to estimate the breast cancer risk in women targeted for screening. Therefore, the results could help to make the shift from a one-size-fits all strategy to a risk-based personalized strategy, which could, in turn, improve the effectiveness of breast cancer screening.

Objectives



General objective

The general objective of the thesis is to deepen the analysis of population-based breast cancer screening. The aim of this thesis is to assess different breast cancer risk factors in order to develop and validate an individualized breast cancer risk prediction model.

Specific objectives

A better understanding of risk factors

- To analyze the impact of breast density on breast cancer screening, not only in terms of breast cancer risk but also in terms of masking negative results.

- To analyze the impact of benign breast disease in women participating in screening, specifically to assess differences in breast cancer risk across benign breast disease diagnosed at prevalent or incident screens.

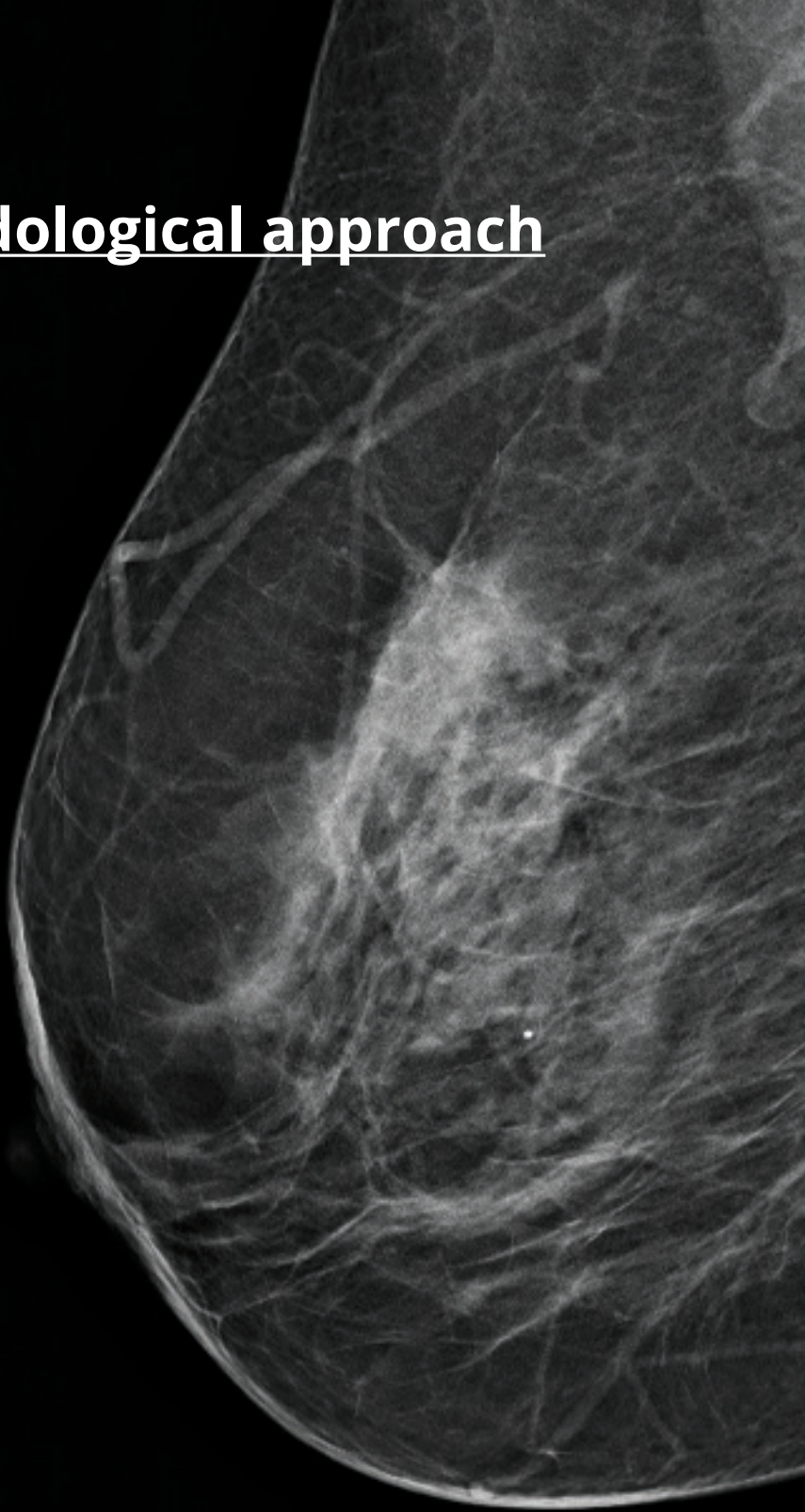
- To understand the combined effect of breast density and benign breast disease.

The breast cancer risk prediction model

- To carry out a systematic review of the existing breast cancer risk models in order to understand them and identify their weak points with a view to enhancing the design of a specific screening model.

- To develop and validate an individualized breast cancer risk prediction model.

Methodological approach



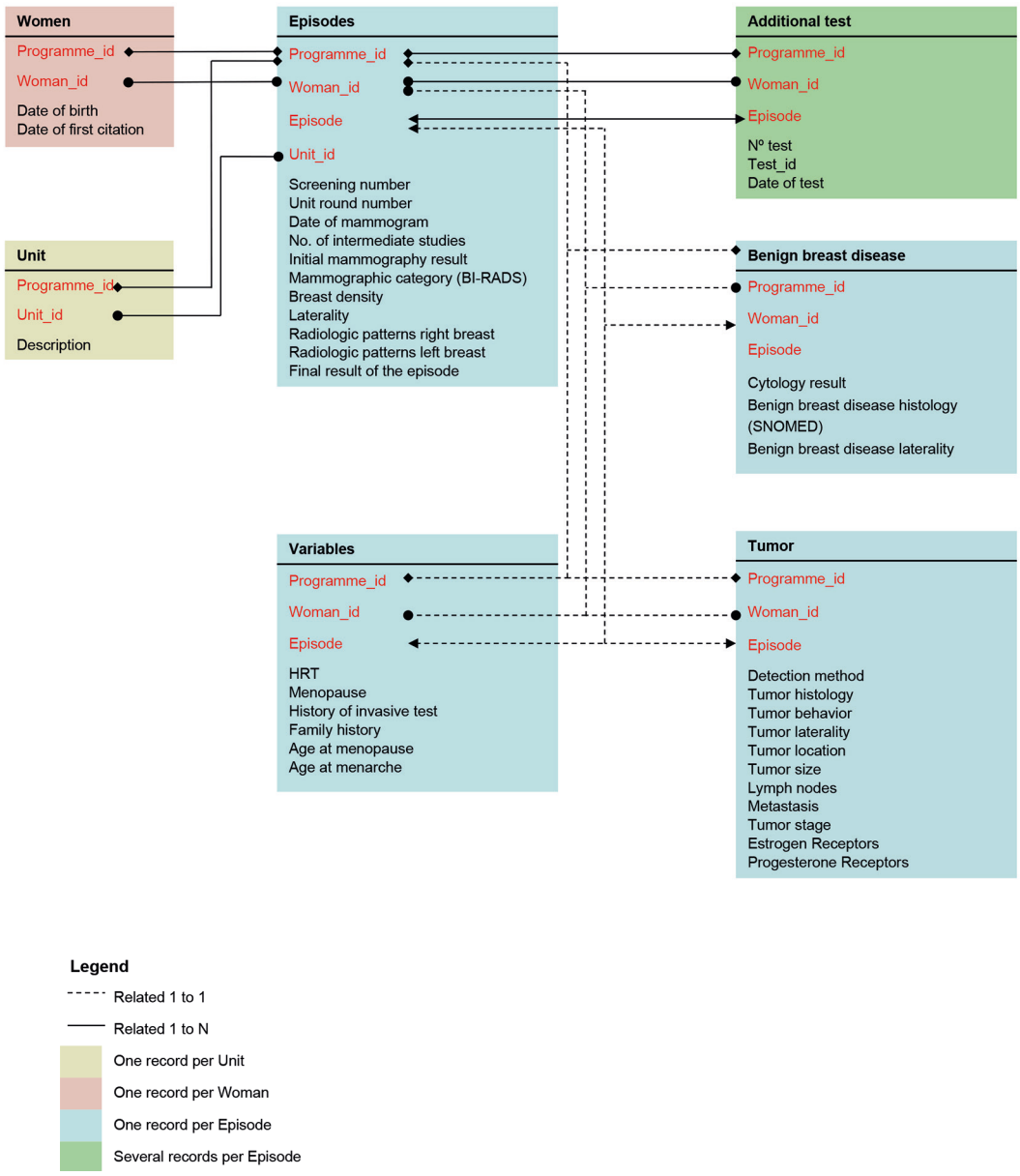


Figure 12. Structure of the database.

Source Figure 12 and 13: BEnign LEsion (BELE-2) and IRIS (Individualized RISk) projects.

Creation of the database

To construct the database used in this project, each of the 10 participating screening programs was asked to provide complete information on all women participating in the screening program at least once from the program launch until December 2015. These programs were Asturias, Barcelona- Àrea Metropolitana Sud, Cantabria, Costa de Ponent, Girona, Hospital Clínic i Provincial de Barcelona, Hospital de la Santa Creu i Sant Pau de Barcelona, Tarragona, Vallès Oriental and Vallès Occidental. All of them have independent and program-specific administrative structures to collect the information, so a detailed protocol was drawn up to define the variables and validate the information from them (see Appendix 3). The protocol of variables was developed and agreed upon with the help of the persons responsible for all the participating programs. Thus, ambiguity in the interpretation of the definitions was avoided, and the codification of the variables of interest was homogenized. In addition, information was sought on cancers up to December 2017, using hospital databases and cancer registries, to take into account possible interval cancers (cancers diagnosed after a negative mammogram and before the next call) in the population and to have information on the 2 years of follow-up after the last mammograms.

The information was collected in a multidimensional table structure with different levels of information (Figure 12). Parallel to data collection, a document with different validation rules was created to ensure data quality. A professional IT company was contracted to develop a virtual tool to apply the validation rules to the different program databases and return an error report. I resolved each of these errors with the different programs personally and individually until all the databases were validated. Once validated, the databases of the different programs were merged into a common database, which is that used for the various analyses.

Study population

The participating programs have structured information from at least 6 consecutive screening rounds, so that the information can be analyzed as a retrospective cohort (each woman must be identified by a unique code in each program participation). In addition, for each woman's participation, information was collected on both the woman's characteristics and the result of the mammogram, including additional scans. All women participating at least once in any of the programs are included. Women with a history of breast cancer or breast implant prior to first screening are excluded.

The project study period was from 1995 to December 2015 but the women were followed up until December 2017. The retrospective nature of the study involved the collection of information from a large volume of screening mammograms. The study contained information from 782,406 women who underwent 2,853,753 screening mammograms. In addition, information from 147,448 additional examinations was included to confirm or rule out malignancy and 18,573 cases of benign breast disease were diagnosed. Of these women, 12,102 were diagnosed with breast cancer during the screening process and 3,659 had an interval cancer. In our cohort, 78% of the participating women had at least 2 screening mammograms (n = 611,110), 61% had at least 3 (n = 483,079), and 22% had 6 or more (n = 174,068). The screening mammograms were performed in 35 different radiological units with an average of 81,535 mammograms per radiology unit.

Mammograms	2,853,753
Women screened	782,406
Additional explorations	147,448
Benign breast disease	18,573
Tumours detected in screening	12,102
Interval tumours	3,659

Figure 13. Characteristics of the BELE-2 / IRIS data base.

Statistical analysis

Model development

The main statistical analysis conducted for the different articles presented in this thesis was done using partly conditional Cox proportional hazard models. This type of model was chosen because it allowed the incorporation of the needs developed by the research team.

The model had to consider that:

- The response variable (breast cancer) is a dichotomous variable.
- Time is continuous and is equivalent to the time since a woman's first mammogram to the end of follow-up.
- It must consider repeated observations of the same woman throughout time.
- It must consider and update those variables that change over time (age, benign breast disease, etc.).

Partly conditional Cox proportional hazard models are an extension of the Cox proportional hazard model in which the hazard ratio of a given event over time is modeled (112). These models were developed in depth by Y. Zheng and P. J. Heagerty in their article "Partly Conditional Survival Models for Longitudinal Data" from 2005 (113). The main difference between these models and the Cox proportional hazards model is that the former takes into account repeated measurements of different women so that the different explanatory variables can be updated over time.

For example, a woman can start at time 0 without any diagnosis of benign breast disease but have a diagnosis at time 4. The model will take this information into account and the hazard ratio of having benign breast disease will be estimated considering that these women spent 4 units of time without benign breast disease.

The formulas of the classical (Equation 1) and the extension Cox proportional hazards model (Equation 2) are shown below.

Equation 1. Classical Cox proportional hazard model (112).

$$h(t|X) = h_0(t) \cdot e^{[\Sigma_{i=1}^{p} \beta_i X_i]}$$

Where:

$h_0(t)$ *is the baseline hazard.*

$X = (X_1, X_2, \dots, X_n)$ *are the predictor variables.*

$\beta = (\beta_1, \beta_2, \dots, \beta_n)$ *are the coefficients of the predictor variables.*

Equation 2. Partly Conditional extension of the Cox proportional hazard model (112).

$$h(t, X(t)) = h_0(t) \cdot e^{[\Sigma_{i=1}^{p_1} \beta_i X_i + \Sigma_{j=1}^{p_2} \delta_j X_j(t)]}$$

Where:

$h_0(t)$ *is the baseline hazard.*

p_1 *is the number of time-independent variables.*

p_2 *is the number of time-dependent variables.*

$X = (X_1, X_2, \dots, X_{p_1}, X_1(t), X_2(t), \dots, X_{p_2}(t))$ *are the predictor variables.*

$\beta = (\beta_1, \beta_2, \dots, \beta_{p_1})$ *are the coefficients of the time-independent variables.*

$\delta = (\delta_1, \delta_2, \dots, \delta_{p_2})$ *are the coefficients of the time-dependent variables.*

With this model we also can obtain estimators not only at 2 years, but also every 2 years (2, 4, 6, 8... up to 20 years) over the time a woman is screened.

In addition, for the analysis of breast density and screening performance indicators, generalized estimating equation (GEE) for repeated measures were used (114).

For the different model coefficient estimators in both Cox partly conditional and GEE models, robust standard errors were used to compute 95% confidence intervals using the robust Huber-White (sandwich) variance estimator (115) (Equation 3).

Equation 3. Robust Huber-White (sandwich) variance estimator formula (115).
Let i *index observations whose values are* y_i. *Let* $\theta \epsilon R^p$ *be a* $px1$ *parameter vector.*
Let $y \rightarrow f_i(y|\theta)$ *be a positive density. Let* Y_i *be independent with density* $f_i(\cdot\,|\theta)$.
Let $l(\theta) = \prod_{i=1}^{n} f_i(Y_i|\theta)$ *the likelihood function. The log likelihood function is therefore:*

$$L(\hat{\theta}) = \sum_{i=1}^{n} log\ f_i\,(Y_i|\theta)$$

The first and second partial derivatives of $\boldsymbol{L}$ *with respect to* θ *are given by:*

$$L'(\theta) = \sum_{i=1}^{n} g_i(Y_i|\theta)$$

$$L''(\theta) = \sum_{i=1}^{n} h_i\,(Y_i|\theta)$$

Where g_i *and* h_i *represent the second and first derivate of* $log\ f_i\,(Y_i|\theta)$ *with respect to* θ.
Then the robust Huber-White (sandwich) variance estimator $\hat{V}$ *is:*

$$\hat{V} = (-A)^{-1}B(-A)^{-1}$$

Where:

$$A = L''(\hat{\theta})$$

$$B = \sum_{i=1}^{n} g_i\,(Y_i|\hat{\theta}_0)^T g_i\,(Y_i|\hat{\theta}_0)$$

Notice that the square roots of the diagonal elements of $\hat{V}$ *are the robust standard errors, also called "Huber-White standard errors".*

Model validation

We assessed the internal validity of the model by means of its calibration and discriminatory accuracy. To perform internal validation, we split our cohort into 2 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 model (100) but we could also have performed cross validation or bootstrapping (116, 117).

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. The expected-to-observed (E/O) ratio assessed whether the number of women predicted to develop breast cancer by the model matched the actual number of cases of breast cancer diagnosed in the validation subcohort.

To account for censoring, the observed rate was estimated using the Kaplan-Meier estimator (118) (Equation 4).

Equation 4. Kaplan-Meier estimator of the survival function (118).

$$\hat{S}(t) = \prod_{i:\, t_i \leq t} (1 - \frac{d_i}{n_i})$$

Where:

t_i *is a time when at least one event happened.*

d_i *the number of events (in this case breast cancer diagnosis).*

n_i *the number of individuals known to have survived (in this case without a breast cancer diagnosis) up to time* t_i.

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. An E/O ratio of 1.0 indicates perfect calibration. We calculated the E/O ratio 95% CI using the formula of the standardized mortality ratio proposed by Breslow and Day (119) (Equation 5).

Equation 5. Confidence interval for E/O ratio adapted from Breslow and Day (119).

Let E/O_L *and* E/O_U *be the lower and upper value for the confidence interval of the expected-to-observed ratio* E/O.
Adapting the formula of the Breslow and Day standardized mortality ratio confidence intervals we define E/O_L *and* E/O_U *as:*

$$E/O_L = (E/O)[1 + \frac{1}{2N} Z^2_{\alpha/2}\{1 - (1 + 4N/Z^2_{\alpha/2})^{1/2}\}]$$

$$E/O_U = (E/O)[1 + \frac{1}{2N} Z^2_{\alpha/2}\{1 + (1 + 4N/Z^2_{\alpha/2})^{1/2}\}]$$

Where:
N *is the number of events (in this case breast cancer diagnosis).*
$Z_{\alpha/2}$ *denotes the* $100(1 - \frac{\alpha}{2})$ *percentile of the unit normal distribution.*

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 (120) (Equation 6).

Equation 6. Sensitivity, specificity and the AUC (120).

Let X *represent the predicted probability of developing the event before time* T *and let* D *be the event indicator. Let* f *be the probability density function of* X *for any cut-off point* $u \in (0,1)$.

We define:

$$Sensitivity = S(U) = P(D = 1) = \int_u^1 f(D = 1)dx$$

$$1 - Specificity = P(U) = P(D = 0) = \int_u^1 f(D = 0)dx$$

Hence, the AUC *corresponds to the area under the parametric plot of* $S(U)\ vs\ P(U)$ *and can be expressed as:*

$$AUC = \int_0^1 S(u)\frac{d}{du}P(u)du = P(X_i > X_j \mid D_i = 1, D_j = 0)$$

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 (121) (Equation 7).

Equation 7. Confidence interval for AUC adapted from Hanley and McNeil (121).

Let $AUROC_L$ *and* $AUROC_U$ *be the lower and upper value for the confidence interval of the AUC. Adapting the formula of the Hanley, we define* AUC_L *and* AUC_U *as:*

$$AUC_L = AUC - Z_{\alpha/2}\sqrt{\frac{AUC(1-AUC) + (N_1-1)(Q_1-AUC)^2 + (N_2-1)(Q_2-AUC^2)}{N_1N_2}})$$

$$AUC_U = AUC + Z_{\alpha/2}\sqrt{\frac{AUC(1-AUC) + (N_1-1)(Q_1-AUC)^2 + (N_2-1)(Q_2-AUC^2)}{N_1N_2}})$$

Where:

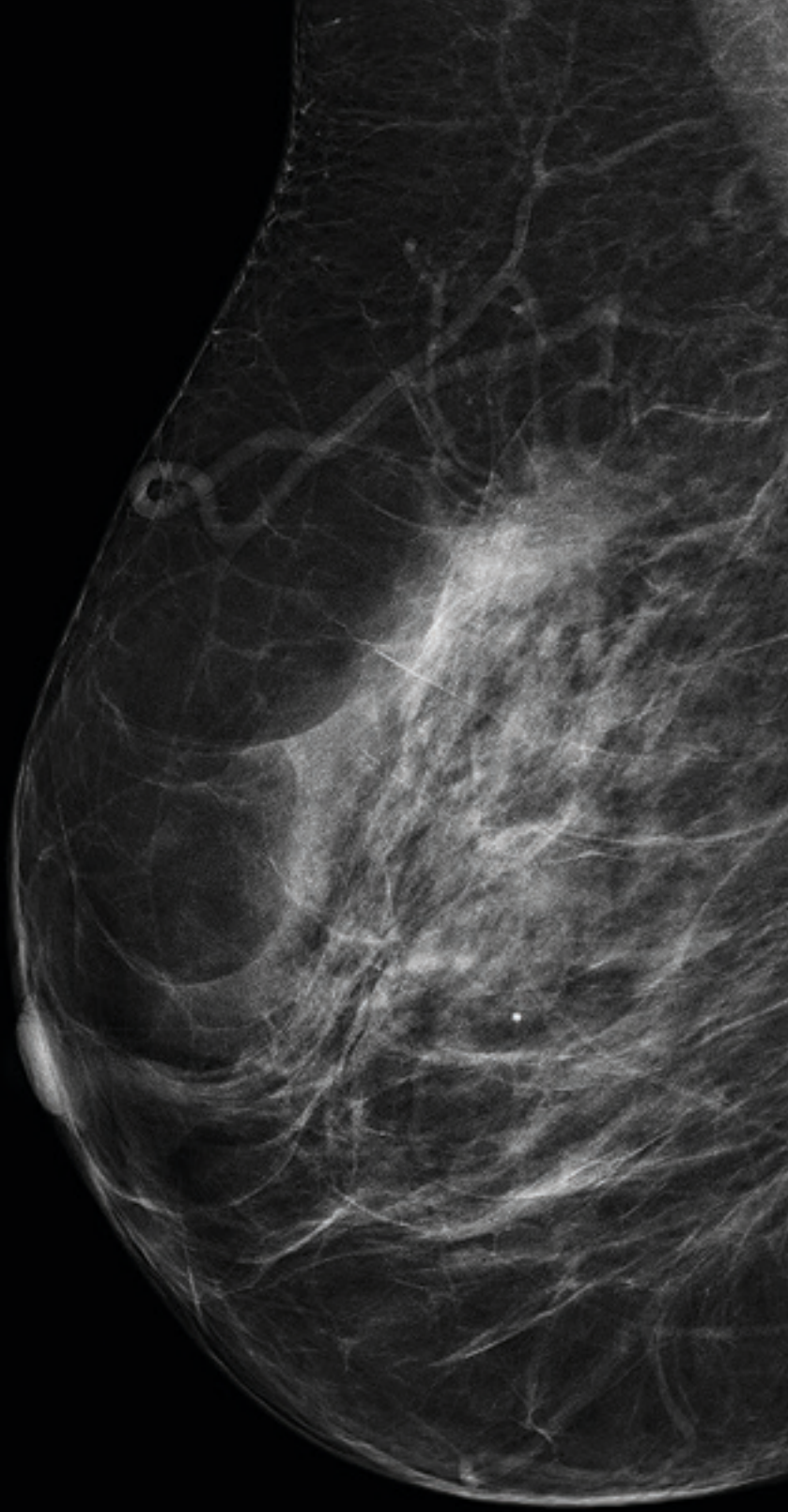
$$Q_1 = \frac{AUC}{2-AUC}$$

$$Q_2 = \frac{AUC}{1+AUC}$$

N_1 *the number of non-events (in this case without breast cancer diagnosis).*
N_2 *the number of events (in this case breast cancer diagnosis).*
$Z^2_{\alpha/2}$ *denotes the* $100\ (1-\frac{\alpha}{2})$ *percentile of the unit normal distribution.*

All statistical analyses of this thesis were performed using the statistical software SPSS version 23.0 (IBM, Armonk, NY, USA) and R version 3.5.0 (R Core Team, 2014).

Results



Article 1

Title: Mammographic breast density: How it affects performance indicators in screening programmes?

Authors: Posso M, Louro J, Sanchez M, Roman M, Vidal C, Sala M, Baré M, Castells X on behalf of the BELE study group

Journal: Eur J Radiol. 2019; 110:81-7

Impact Factor: 2.68 (Q2 Radiology, nuclear medicine & medical imaging)

DOI: 10.1016/j.ejrad.2018.11.012

Abstract:

Objectives: To investigate how breast density affects screening performance indicators in a digital mammography context.

Methods: We assessed the effect of breast density over the screen-detected and interval cancers rates, false-positives, specificity, sensitivity, recall rate, positive predictive value of recall (PPV-1), and PPV of invasive tests (PPV-2). Radiologists classified breast density using the BIRADS System. We used generalized estimating equations to account for within-woman correlation by means of the robust Huber-White variance estimator.

Results: We included 177,164 women aged 50–69 years who underwent 499,251 digital mammograms from 2004 to 2015 in Spain. According to the fibroglandular tissue percentage, 24.7% of mammograms were classified as BI-RADS 1 (<25% glandular), 54.7% as BI-RADS 2 (25–50% glandular), 14.0% as BI-RADS 3 (51–75% glandular) and 6.6% as BI-RADS 4 (>75% glandular). Overall, women with BI-RADS 3 had the highest screen-detected cancer rate (5.9 per 1000) and BI-RADS 4 the highest interval cancer rate (2.4 per 1000). Sensitivity decreased from

89.2% in women with BI-RADS 1 to 67.9% in BI-RADS 4. Both PPV-1 and PPV-2 decreased from 10.4% to 5.7% and from 49.8% to 32.4% in women with BI-RADS 1 and BI-RADS 4, respectively. Women aged 60–69 years with BI-RADS 4 had the lowest sensitivity (54.9%) and the highest interval cancer rate (3.8 per 1000).

Conclusions: Performance screening measures are negatively affected by breast density falling to a lower sensitivity and PPV, and higher interval cancer rate as breast density increases. Particularly women aged 60–69 years with >75% glandular breasts had the worst results and therefore may be candidates for screening using other technologies.

European Journal of Radiology 110 (2019) 81–87

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Mammographic breast density: How it affects performance indicators in screening programmes?

Margarita Posso[a,b], Javier Louro[a,b,c], Mar Sánchez[d], Marta Román[a,b], Carmen Vidal[e], María Sala[a,b], Marisa Baré[b,f], Xavier Castells[a,b,*], on behalf of the BELE study group

[a] Department of Epidemiology and Evaluation, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain
[b] Research Network on Health Services in Chronic Diseases (REDISSEC), Barcelona, Spain
[c] 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
[d] General Directorate of Public Health, Government of Cantabria, Santander, Spain
[e] Cancer Prevention and Monitoring Program, Catalan Institute of Oncology, Barcelona, Spain
[f] Clinical Epidemiology and Cancer Screening, Parc Taulí University Hospital, Sabadell, Spain

ARTICLE INFO

Keywords:
Breast neoplasms
Early detection of cancer
Mammography
Breast density

ABSTRACT

Objectives: To investigate how breast density affects screening performance indicators in a digital mammography context.
Methods: We assessed the effect of breast density over the screen-detected and interval cancers rates, false-positives, specificity, sensitivity, recall rate, positive predictive value of recall (PPV-1), and PPV of invasive tests (PPV-2). Radiologists classified breast density using the BIRADS System. We used generalized estimating equations to account for within-woman correlation by means of the robust Huber-White variance estimator.
Results: We included 177,164 women aged 50–69 years who underwent 499,251 digital mammograms from 2004 to 2015 in Spain. According to the fibroglandular tissue percentage, 24.7% of mammograms were classified as BI-RADS 1 (< 25% glandular), 54.7% as BI-RADS 2 (25–50% glandular), 14.0% as BI-RADS 3 (51–75% glandular) and 6.6% as BI-RADS 4 (> 75% glandular). Overall, women with BI-RADS 3 had the highest screen-detected cancer rate (5.9 per 1000) and BI-RADS 4 the highest interval cancer rate (2.4 per 1000). Sensitivity decreased from 89.2% in women with BI-RADS 1 to 67.9% in BI-RADS 4. Both PPV-1 and PPV-2 decreased from 10.4% to 5.7% and from 49.8% to 32.4% in women with BI-RADS 1 and BI-RADS 4, respectively. Women aged 60–69 years with BI-RADS 4 had the lowest sensitivity (54.9%) and the highest interval cancer rate (3.8 per 1000).
Conclusions: Performance screening measures are negatively affected by breast density falling to a lower sensitivity and PPV, and higher interval cancer rate as breast density increases. Particularly women aged 60–69 years with > 75% glandular breasts had the worst results and therefore may be candidates for screening using other technologies.

1. Introduction

Radiologists determine breast density based on the amount of radiopaque breast parenchyma that is visualized on the mammogram. Radiopaque areas correspond to regions in the breast that are rich in epithelial and stromal tissue while the non-dense, darker grey areas, correspond to regions that are predominantly fat [1].

The assessment of breast density in mammography screening has become relevant because it can limit the screening accuracy [2,3]. Several authors have reported that small breast cancers are likely to be easily diagnosed in a breast containing substantial fatty tissue. Conversely, it would be difficult to detect if the lesions are superimposed on dense tissue [4,5]. The role of breast density as a contributor to interval cancers has also been reported [3,6]. Thus, evaluating breast density as

Abbreviations: BI-RADS, American College of Radiology Breast Imaging Reporting and Data System; DCIS, ductal carcinoma in situ; GEE, generalised estimating equations; PPV-1, positive predictive value of recall; PPV-2, positive predictive value of invasive tests

* Corresponding author at: Department of Epidemiology and Evaluation, IMIM (Hospital del Mar Medical Research Institute), Passeig Maritím 25-29, 08003, Barcelona, Spain.
E-mail address: xcastells@parcdesalutmar.cat (X. Castells).

https://doi.org/10.1016/j.ejrad.2018.11.012
Received 27 June 2018; Received in revised form 9 November 2018; Accepted 12 November 2018

an influence on accuracy of detection is essential in screening programs.

Compared to film-screen, digital mammography can mitigate the breast density masking effect and improve cancer detection, especially in women with extremely dense breasts [7]. However, challenges when reading digital mammograms of women with dense breasts still exists, and whether this impact on the performance measures in screening needs further assessment in the light of modern digital technology.

Breast cancer screening in Spain started using digital mammography early in the 2000 decade, and spread rapidly afterwards [8]. Evaluating performance measures across breast dense categories provides valuable information for mammographic screening that can be used to guide clinical practice and screening policies. Our purpose was to assess the effect of breast density on the screening performance measures in a population-based program that uses digital mammography.

2. Materials and methods

2.1. Setting and study population

The Spanish, government funded, Breast Cancer Screening Program started in 1990 and became nationwide in 2006. All women aged 50–69 years biennially receive an invitation letter to participate in the program. The standard procedure for radiological performance is two-view mammography and double reading with consensus arbitration in case of disagreement. Mammograms were read by highly experienced radiologists who interpreted at least 1500 screening mammograms per year. Certified screening radiologists routinely evaluate mammograms. The BI-RADS® scale or equivalent is used to rate the probability of cancer. Women with positive mammographic findings, scored as 3; probably benign finding, 4; suspicious abnormality, 5; highly suspicious of malignancy, or 0; incomplete, are recalled for further assessments to confirm or rule out malignancy at the reference hospital of their screening geographic area.

We assessed information of digital screening exams performed in three Centres of the Spanish Breast Cancer Screening Program from July 2004 to December 2015 (Vallès Occidental, Barcelona- Àrea Metropolitana Sud, and Cantabria). From a total of 522,741 screening exams, we excluded 23,490 due to lack of information with regards to mammographic density. Hence, we included 499,251 digital exams performed on 177,164 women.

We included both screen-detected and interval cancers. Screen-detected cancers were diagnosed at routine screening. An interval breast cancer was defined as a breast carcinoma diagnosed after a negative screening test, or after a positive screening test where malignancy is finally ruled out, either before the next biennial invitation to screening, or within two years for women who had reached the upper age limit for screening. Invasive as well as in situ carcinomas (DCIS) were pathologically confirmed. Data on screening mammogram results, additional diagnostic tests, and pathological confirmation was obtained from the Breast Screening Centers database whereas we identified interval cancers by merging data from population-based cancer registries and hospital records. Each Ethics Committee at the participating institutions approved the study and informed consent was waived since we used anonymised retrospective data.

2.2. Breast density measurement

Breast density was determined per each mammogram by one or two radiologists using the Fourth Edition of the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS®) [9] and therefore all mammograms were categorized according to the percentage value of fibroglandular breast tissue as 1) < 25% glandular; 2) 25–50% glandular; 3) 51–75% glandular; or 4) > 75% glandular.

2.3. Accuracy measures

We assessed the cancer detection rate (screen-detected and interval cancers), false-positive rate, specificity, sensitivity, recall rate (frequency and type of additional tests), positive predictive value of recall (PPV-1), and PPV of invasive tests (PPV-2) according to breast density classification.

Breast cancer detection rates were defined as the number of cases per 1000 screening exams. False-positives were cases recalled for additional tests without an ultimate diagnosis of cancer. Sensitivity was defined as the number of screen-detected cancers divided by the number of screen-detected cancers plus interval cancers. Specificity was defined as the number of true-negative screening exams divided by the number of true-negatives tests plus false positives. PPV-1 was defined as the number of screen detected breast cancers divided by the number of recalls due to positive mammographic findings regardless of the additional test invasiveness. PPV-2 was defined as the number of screen detected breast cancers divided by the number of recalled exams including only invasive procedures (fine needle aspiration biopsy, core needle biopsy, open biopsy). The number of women needed to be recalled and to undergo an invasive procedure to detect one breast cancer was estimated by taking the inverse of PPV-1 (1/PPV-1) and PPV-2 (1/PPV-2), respectively.

2.4. Statistical analysis

The screening mammogram was the unit of analyses. Because women could have multiple screens during the study period, we used generalized estimating equations (GEE) to account for within-woman correlation in the performance indicators by means of the robust Huber-White (sandwich) variance estimator [10].

Screening accuracy measures were evaluated separately for the four breast density categories. Estimates of sensitivity and specificity, cancer detection rates, false-positive rates, PPV-1, and PPV-2 were stratified by type of screening (first or subsequent), and age at screening (50–59 or 60–69 years of age). Proportions across breast density categories were compared using the z-test for column proportions. The 95% confidence intervals (95% CIs) were calculated based on the standard errors obtained from the GEE models. P-values < 0.05 were considered statistically significant.

3. Results

The study included 499,251 digital mammograms from 177,164 women who underwent screening at age 50 to 69 years between 2004 and 2015. Among the mammograms analyzed, and according to its breast density, 123,292 (24.7%) were classified as BI-RADS 1, 272,964 (54.7%) as BI-RADS 2, 70,066 (14.0%) as BI-RADS 3, and 32,929 (6.6%) as BI-RADS 4. In terms of age, 280,312 (56.1%) mammograms were performed to women aged 50–59 years and 218,939 (43.9%) to woman aged 60–69 years. When classifying by type of screen, 103,308 (20.6%) were first screening examinations and 396.213 (79.4%) were subsequent screens. Overall, 2047 cancers were screen-detected and 550 were interval cancers.

The screen-detected invasive cancer rate increased from 2.23 per 1000 screening exams in the BI-RADS 1 group to 2.69 in the BI-RADS 2 group and to 3.95 in the BI-RADS 3 group. However, this rate decreased in the BI-RADS 4 group with a rate of 2.28. The rate of screen-detected DCIS increased with increasing density, ranging from 0.58 in women with BI-RADS 1 to 2.76 in those with BI-RADS 4. Regarding interval cancers, the rate of all malignancies (invasive cancer and DCIS) also increased with increasing density, from 0.34 to 2.40 per 1000 screening exams in BI-RADS 1 and BI-RADS 4, respectively. The false positive rate increased with increasing breast density for both additional imaging and invasive procedures. The overall proportion was 2.44%, 5.46%, 7.58% and 8.46% in women with BI-RADS 1, 2, 3 and 4, respectively

Table 1
Number and rate of screen-detected, interval cancer and false-positive results in mammographic screening according to breast density.

	BI-RADS 1 (< 25% glandular) (n = 123,292)	BI-RADS 2 (25-50% glandular) (n = 272,694)	BI-RADS 3 (50-75% glandular) (n = 70,066)	BI-RADS 4 (> 75% glandular) (n = 32,929)	Total (n = 499,251)
Screen detected cancers					
All malignant lesions, n (per 1000)	348 (2.82) [a,b]	1,116 (4.09) [a]	416 (5.94) [a]	167 (5.07) [b]	2,047 (4.10)
Invasive, n (per 1000)	275 (2.23) [a,b]	733 (2.69) [a]	277 (3.95) [b]	75 (2.28) [a,b]	1,360 (2.72)
DCIS, n (per 1000)	71 (0.58) [a,b]	379 (1.39) [a]	139 (1.98) [b]	91 (2.76) [a,b]	680 (1.36)
Unknown, n (per 1000)	2 (0.02)	4 (0.01)	0 (0.00)	1 (0.03)	7 (0.01)
Interval cancers					
All malignant lesions, n (per 1000)	42 (0.34) [c,d]	290 (1.06) [c,d]	139 (1.98) c	79 (2.40) [d]	550 (1.10)
Invasive, n (per 1000)	31 (0.25) [c]	239 (0.88) [c,d]	111 (1.58) [c,d]	73 (2.22) [d]	454 (0.91)
DCIS, n (per 1000)	6 (0.05)	31 (0.11)	13 (0.19)	2 (0.06)	52 (0.10)
Unknown, n (per 1000)	5 (0.04)	17 (0.06)	9 (0.13)	3 (0.09)	34 (0.07)
False positives					
All, n (per 1000)	3,012 (2.44) [e]	14,899 (5.46) [e]	5,314 (7.58) [e]	2,785 (8.46) [e]	26,010 (5.21)
Additional imaging, n (per 1000)	2,661 (2.16) [e]	13,422 (4.92) [e]	4,807 (6.86) [e]	2,436 (7.40) [e]	23,326 (4.67)
Invasive procedures, n (per 1000)	351 (0.28) [e]	1477 (0.54) [e]	507 (0.72) [e]	349 (1.06) [e]	2,684 (0.54)

DCIS: Ductal carcinoma in situ.
[a,b,c,d,e] Those values of the same row that share a same superscript are significantly different at $p < 0.05$ in a two-sided test for column proportions (z-test). Tests are adjusted using the Bonferroni correction for multiple comparison.

(Table 1).

Sensitivity decreased while increasing breast density ranging between a sensitivity of 89.2% in women with BI-RADS 1 and 67.9% in those with BI-RADS 4. This decrease was statistically significant except when comparing BI-RADS groups 2 versus 3 and BI-RADS 3 versus 4. Specificity significantly decreased with increasing breast density ranging from 97.5% to 91.5% in women with BI-RADS 1 and BI-RADS 4, respectively. Similarly to sensitivity and specificity, both PPV-1 (positive predictive value for all recall tests) and PPV-2 (positive predictive value for invasive procedures) decreased with increasing breast density but showing differences between the extreme groups. PPV-1 significantly decreased from 10.4% in women with BI-RADS 1, to 5.7% in those with BI-RADS 4. The decrease translates to 9.7 recalls needed for further workup to detect one breast cancer in women with BI-RADS 1, and 17.7 in women with BI-RADS 4. Regarding invasive procedures, 2.0 biopsies were required to detect one cancer in women with BI-RADS 1 whereas 3.1 biopsies were needed in those with BI-RADS 4 (Table 2).

Sensitivity decreased with increasing breast density in both women aged 50–59 and 60–69 years. A significant lower sensitivity in women aged 60–69 years was found in the extremely dense group compared with women aged 50–59 years (54.9% vs 73.1%) (Fig. 1). The lower sensitivity is explained by the high interval cancer rate amongst women with BI-RADS 4 aged 60 to 69 years (4 per 1000). Without considering the BI-RADS 4 group, the screen-detected cancer rate increased with increasing density in women aged 50 to 59 years as well as in those aged 60 to 69 years. In the younger age group it ranged from 2.1 to 6.0 and in the older age group from 3.5 to 5.8 per 1000 screening exams in women with BI-RADS 1 and BI-RADS 3, respectively (Fig. 1).

The analyses stratified by type of screening confirmed the lowest sensitivity in women with BI-RADS 4 (75.7% at first screen and 62.9% in subsequent screen). Sensitivity was significantly higher in women with BI-RADS 3 at first screen compared with subsequent screen (83.7% and 70.8%, respectively). Women with BI-RADS 4 also showed higher screen-detected cancer rate at first screen compared with subsequent screen (7.1 and 4.2, respectively)). We did not find differences when comparing interval cancer rates by type of screening in the different density groups (Fig. 2).

4. Discussion

Our results showed that the high breast density has a negative effect on the screening performance measures in a population-based program that uses digital mammography. We found that both sensitivity and positive predictive value were remarkably lower in women with BI-RADS 4. Compared to women with BI-RADS 1, the group with BI-RADS 4 had over a three-fold increased rate of interval cancer and false positives. Notably, women aged 60–69 years with BI-RADS 4 had the lowest sensitivity, which implied that one out of two breast cancers in

Table 2
Sensitivity, specificity, positive predictive value of recalls (PPV-1) and invasive procedures (PPV-2) in mammographic screening according to breast density with the 95% confidence intervals.

	BI-RADS 1 (< 25% glandular) % (95% CI)	BI-RADS 2 (25-50% glandular) % (95% CI)	BI-RADS 3 (50-75% glandular) % (95% CI)	BI-RADS 4 (> 75% glandular) % (95% CI)	Total % (95% CI)
Sensitivity	89.2 (85.7-91.9)	79.4 (77.2-81.4)	75.0 (71.2-78.4)	67.9 (61.8-73.4)	78.8 (77.2-80.4)
Specificity	97.5 (97.5-97.6)	94.5 (94.4-94.6)	92.4 (92.2-92.6)	91.5 (91.2-91.8)	94.8 (94.7-94.8)
PPV-1	10.4 (9.4-11.4)	7.0 (6.6-7.4)	7.3 (6.6-8.0)	5.7 (4.9-6.6)	7.3 (7.0-7.6)
1/PPV-1	9.7 (8.7-10.7)	14.4 (13.6-15.2)	13.8 (12.6-15.1)	17.7 (15.3-20.5)	13.7 (13.1-14.3)
PPV-2	49.8 (46.1-53.5)	43.0 (41.1-45.0)	45.1 (41.9-48.3)	32.4 (28.5-36.5)	43.3 (41.9-44.7)
1/PPV-2	2.0 (1.9-2.2)	2.3 (2.2-2.4)	2.2 (2.1-2.4)	3.1 (2.7-3.5)	2.3 (2.2-2.4)

CI: Confidence Intervals, PPV-1: Positive predictive value of recall, PPV-2: Positive predictive value of invasive tests, 1/PPV-1: Inverse of the positive predictive value of recall, 1/PPV-2: Inverse of the positive predictive value of invasive test.

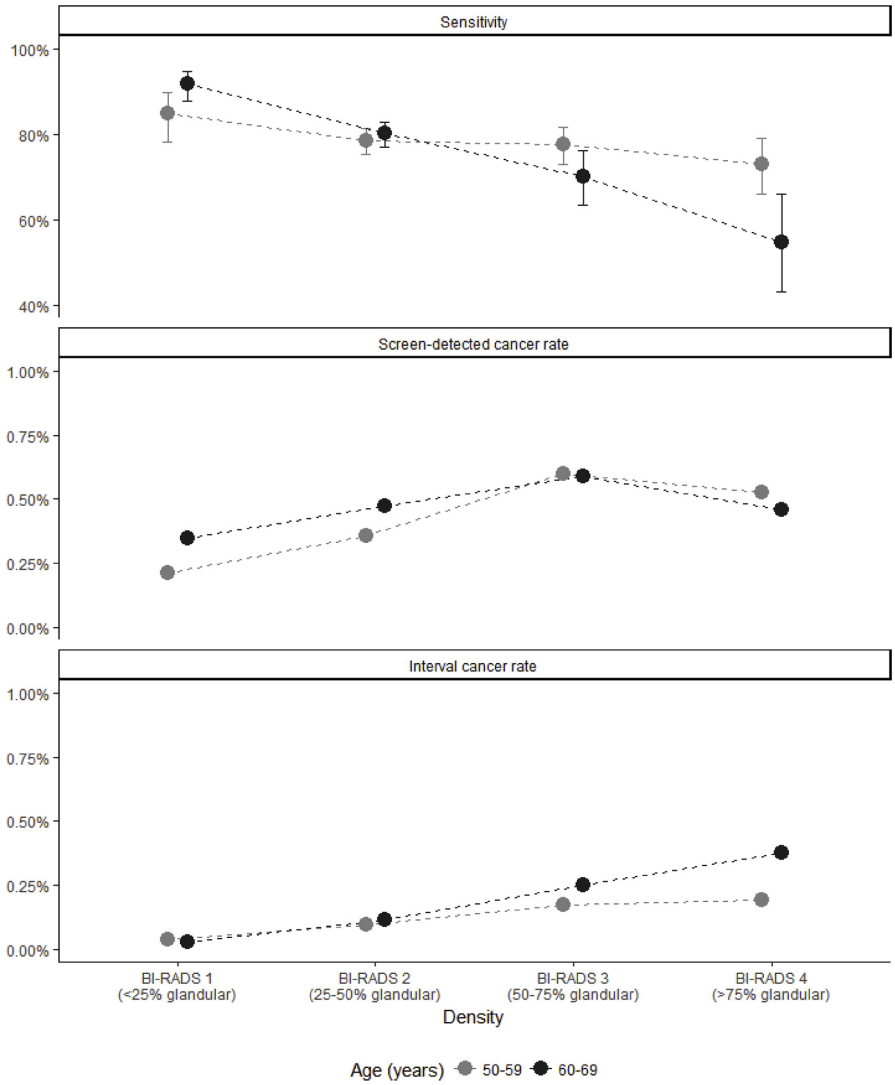


Fig. 1. Sensitivity, screen-detected cancer rate and interval cancer rate stratified by age group.

this group were diagnosed as an interval cancer.

Although sensitivity decreased with increasing breast density, the cancer detection rate shows an opposite trend. Two main factors can explain this finding. First, the masking effect of breast density reduces the likelihood of several tumours to be detected and therefore reduces the sensitivity. A plausible consequence is the higher interval cancer rate in women with dense breasts compared to fatty breasts, which is consistent with our results. Several studies have reported high interval cancer rates in women with BI-RADS 4, exceeding in some cases the amount of screen detected cancers [11,12]. Second, breast density acts as an independent risk factor for breast cancer [13,14] as it was observed in our study where both women with BI-RADS 3 and BI-RADS 4 had a higher rate of screen detected cancer than women with fatty breasts.

It also should be noticed that in our cohort the rate of DCIS constantly increased with increasing density whereas for invasive cancers this tendency was not conclusive. This can be explained by the fact that the observed density does not necessary imply histologic abnormality at the time of mammography screening. Several studies have demonstrated a strong relationship between mammographic density and histological precursors of breast cancer and also with DCIS [15,16]. Whether breast density is related or not to a biological phenotype promoting faster tumour growth or to a specific histological cancer type remains to be elucidated. However, our findings could be explained by the fact that in very dense breasts it is hard to spot small masses, asymmetries or distortions that are often the mammographic finding of invasive cancers, thereby leading to a lower incidence of screen-detected invasive cancers in women with BI-RADS 4. DCIS are on the other hand often associated with calcifications, and calcifications are almost as easy to spot in a dense breast as in a fatty breast. Thus, one

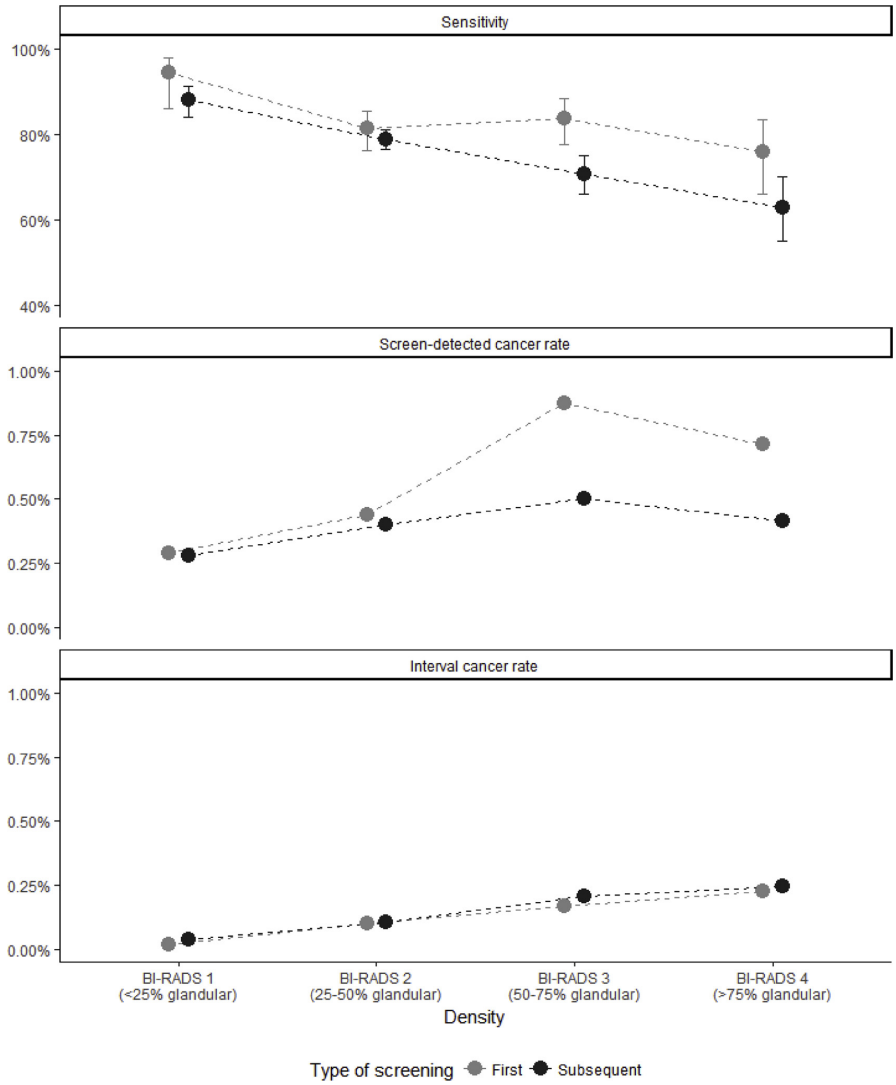


Fig. 2. Sensitivity, screen-detected cancer rate and interval cancer rate stratified by first and subsequent screening.

could expect the incidence of DCIS to be higher in women with BI-RADS 4 than in women with more fatty breasts because the extremely dense women have the highest risk of breast malignancy. This would also lead to a higher incidence of DCIS relative to invasive cancer in these women compared to women with less dense breasts. Overall, the increased risk can be even higher in women with dense breasts and therefore, as well as other factors [17], it should be considered when offering their follow-up strategies.

Previous studies conducted with screen-film mammography showed that sensitivity decreased from approximately 80% in women with BI-RADS 1 to 50% in women with BI-RADS 4 [18–20]. Digital mammography was expected to perform better in women with dense breasts, with several trials showing between 70% to 80% sensitivity in women with BI-RADS 3 and BI-RADS 4 in non-organized screening [7,21]. More recently, studies conducted in population-based European screening programs confirmed a sensitivity of about 70% in these groups [12,22], which is in agreement with our results. Overall, although the diagnostic accuracy of digital mammography outperforms conventional screen-film mammography, other techniques such as digital breast tomosynthesis, may improve screening performance measures among women with dense breasts [14].

Interestingly, we found that women aged 60 to 69 years with BI-RADS 4 showed the lowest sensitivity (54.9%) and the highest interval cancer rate (4 per 1000 screening exams). Breast density is expected to gradually decrease with increasing age after menopause [20,23]. However, there is a small proportion of women who remain at a high breast density ages 60 to 69 years. The explanation for why the sensitivity is lower in 60–69 years old women with BI-RADS 4 than in 50–59 year old women with BI-RADS 4 has not been established. However, it could be related to a higher incidence of cancer in older

women with dense breasts than in younger women with dense breasts. From our results, it seems that routine biennial screening in these group of women may not be as effective as in the average target population. These women may benefit from personalized screening strategies that combine new diagnostic tests.

In our study, women with dense breasts were more likely to be recalled for additional tests, including invasive procedures. Most additional tests were associated with an increased false positive rate and a decrease in the predictive value. In agreement with our findings, a previous study showed that women with dense breasts were more likely to undergo additional imaging tests [24]. However, they found that breast density was not significantly associated with biopsy and/or surgical consultation in women without additional imaging tests. The authors suggested that having imaging tests, especially ultrasound, was the factor associated with unnecessary biopsies in women with dense breasts. The distribution of diagnostic tests performed should be further evaluated, particularly in women aged 60–69 years and BI-RADS 4 because they showed low positive predictive values.

Almost 21% of mammograms in our cohort and up to 40% in other cohorts [13,25] were classified as BI-RADS 3 or BI-RADS 4, which represent a large proportion of screened women. Thus, to study breast density is helpful to better planning the screening process and resources needed, especially for women with dense breasts. Breast density is particularly relevant in the screening context, since it contributes more to the population risk than other much stronger but less common risk factors, such as BRCA mutations. In fact, some authors have proposed that breast density is the risk factor that increases far more the accuracy of a breast cancer risk prediction model [26]. Therefore, offering more accurate diagnostic tests to these women can positively affect the overall performance of screening programs by increasing sensitivity and/or decreasing interval cancers.

This study is based on a large cohort of screened women, involving more than 150,000 women followed for at least 10 years that allowed us to obtain robust conclusions. However, several limitations should be considered. First, variability between radiologists can affect the results since breast density measurements are inherently inaccurate depending on the subjective observation [25,27,28]. Despite this limitation, our results are consistent with those published by other European screening programs. Furthermore, highly trained radiologists performed breast density classification. Second, the BI-RADS edition that is referred to in this manuscript is the fourth edition published in 2003 [9] and differs from the current edition in how breast density is categorized [29], which focuses more on the qualitative description of breast density. We do not have data regarding the Fifth edition since it has been implemented since 2015. Although BI-RADS 1 to 4 is similar to BI-RADS a to d, it should be noticed that some women who, for example, had mostly fatty breast but with focal dense areas might now be classified as BI-RADS c, while they were previously classified as BI-RADS 2 [30]. Third, due to lack of information, we could not differentiate true interval cancers from false negatives and therefore we combined both of them in one category as interval cancers. This fact could lead us to a slight bias when estimating sensitivity.

5. Conclusions

Performance measures in screening mammography are negatively affected by breast density, falling to a lower sensitivity, positive predictive value and higher interval cancer rates. Although digital mammography is expected to have better results in women with dense breast, it seems that the performance improvements of this technique is less effective for screening women with BI-RADS 3 or BI-RADS 4. Women with dense breast may not obtain benefit from screening to the same extent as women with lower breast density. Particularly, women aged 60 to 69 years with BI-RADS 4 showed the lower sensitivity and higher interval cancer rate and therefore they may be more likely to have benefits from other screening technologies such as digital breast tomosynthesis.

Funding

This work was supported by Grants from Instituto de Salud Carlos III FEDER, [PI15/00098 and PI17/00047] and by the Network for Research into Healthcare in Chronic Diseases, REDISEC [RD12/0001/0015].

Role of the funding source

This funding source had no role in the design of this study and did not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

Acknowledgements

The authors acknowledge the Benign Lesion (BELE) Study Group, listed here in alphabetical order and grouped by institution: (a) IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain: Andrea Burón, Xavier Castells, Laia Domingo, Javier Louro, Margarita Posso, Ana Rodríguez-Arana, Marta Román, Maria Sala, Sònia Servitja, Mar Vernet; (b) Parc Taulí University Hospital, Sabadell, Spain: Marisa Baré; (c) Catalan Institute of Oncology, Cancer Prevention and Monitoring Program, Barcelona, Spain: Llucia Benito, Carmen Vidal; (d) Hospital de la Santa Creu i Sant Pau, Epidemiology Department, Barcelona, Spain: María Jesús Quintana, Judit Solà-Roca; (e) General Directorate of Public Health, Government of Cantabria, Santander, Spain: Mar Sánchez; (f) Principality of Astúrias Health Service, Spain: Miguel Prieto; (g) Fundació Lliga per a La Investigació i Prevenció Del Càncer, Universitat Rovira i Virgili, Tarragona, Spain: Jaume Galceran, Francina Saladié; (h) Hospital Santa Caterina, Girona, Spain: Joana Ferrer; (i) Catalan Cancer Plan, Department of Health, Barcelona, Spain: Josep Alfons Espinàs; (*j*) Private Foundation Asil Hospital, Granollers, Spain: Lupe Peñalva; and (k) Hospital Clinic, Preventive Medicine and Epidemiology Department, Barcelona, Spain: Isabel Torá-Rocamora, Xavier Bargalló. Javier Louro is a Ph.D. candidate at the Methodology of Biomedical Research and Public Health program, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain.

References

[1] N.F. Boyd, L.J. Martin, M. Bronskill, M.J. Yaffe, N. Duric, S. Minkin, Breast tissue composition and susceptibility to breast Cancer, JNCI J. Nat. Cancer Inst. 102 (2010) 1224–1237, https://doi.org/10.1093/jnci/djq239.
[2] D.S. A L Mousa, E.A. Ryan, C. Mello-Thoms, P.C. Brennan, What effect does mammographic breast density have on lesion detection in digital mammography? Clin. Radiol. 69 (2014) 333–341, https://doi.org/10.1016/j.crad.2013.11.014.
[3] K. Kerlikowske, W. Zhu, A.N.A. Tosteson, B.L. Sprague, J.A. Tice, C.D. Lehman, et al., Identifying women with dense breasts at high risk for interval Cancer, Ann. Intern. Med. 162 (2015), https://doi.org/10.7326/M14-1465 673–17.
[4] D.S.M. Buist, P.L. Porter, C. Lehman, S.H. Taplin, E. White, Factors contributing to mammography failure in women aged 40-49 years, JNCI J. Nat. Cancer Inst. 96 (2004) 1432–1440, https://doi.org/10.1093/jnci/djh269.
[5] J.J. Heine, P. Malhotra, Mammographic tissue, breast cancer risk, serial image analysis, and digital mammography, Acad. Radiol. 9 (2002) 298–316, https://doi.org/10.1016/S1076-6332(03)80373-2.
[6] S.H. Taplin, C.M. Rutter, C. Finder, M.T. Mandelson, F. Houn, E. White, Screening Mammography: clinical image quality and the risk of interval breast cancer, Am. J. Roentgenol. 178 (2002) 797–803, https://doi.org/10.2214/ajr.178.4.1780797.
[7] K. Kerlikowske, Comparative effectiveness of digital versus film-screen mammography in community practice in the United States, Ann. Intern. Med. 155 (2011), https://doi.org/10.7326/0003-4819-155-8-201110180-00005 493–34.
[8] M. Sala, M. Comas, F. Macià, J. Martinez, M. Casamitjana, X. Castells, Implementation of digital mammography in a population-based breast cancer screening program: effect of screening round on recall rate and cancer detection, Radiology 252 (2009) 31–39, https://doi.org/10.1148/radiol.2521080696.
[9] D'Orsi C, Mendelson E, Reston DI, Radiology VACO. Breast Imaging Reporting and Data System: ACR BI-RADS-Breast Imaging Atlas. 2003 n.d.
[10] P.J. Diggle, P. Heagerty, K. Liang, S.L. Zeger, Analysis of longitudinal data, Oxford Statistical Science Series, (2002).
[11] V.A. McCormack, Breast density and parenchymal patterns as markers of breast Cancer risk: a meta-analysis, Cancer Epidemiol. Biomarkers Prev. 15 (2006)

1159–1169, https://doi.org/10.1158/1055-9965.EPI-06-0034.
[12] L. Timmermans, L. Bleyen, K. Bacher, K. Van Herck, K. Lemmens, C. Van Ongeval, et al., Screen-detected versus interval cancers: effect of imaging modality and breast density in the Flemish Breast Cancer Screening Programme, Eur. Radiol. (2017) 1–10, https://doi.org/10.1007/s00330-017-4757-4.
[13] P.E. Freer, Mammographic breast density: impact on breast Cancer risk and implications for screening, RadioGraphics 35 (2015) 302–315, https://doi.org/10.1148/rg.352140106.
[14] C.I. Lee, L.E. Chen, J.G. Elmore, Risk-based breast Cancer screening: implications of breast density, Med. Clin. North Am. 101 (2017) 725–741, https://doi.org/10.1016/j.mcna.2017.03.005.
[15] N.F. Boyd, L.J. Martin, M.J. Yaffe, S. Minkin, Mammographic density and breast cancer risk: current understanding and future prospects, Breast Cancer Res. 13 (2011) 223, https://doi.org/10.1186/bcr2942.
[16] T.A. MacKenzie, L. Titus-Ernstoff, P.M. Vacek, B. Geller, J.E. Weiss, M.E. Goodrich, et al., Breast density in relation to risk of ductal carcinoma in situ of the breast in women undergoing screening mammography, Cancer Causes Control 18 (2007) 939–945, https://doi.org/10.1007/s10552-007-9035-3.
[17] A. Lafranconi, L. Pylkkänen, S. Deandrea, A. Bramesfeld, D. Lerda, L. Neamţiu, et al., Intensive follow-up for women with breast cancer: review of clinical, economic and patient's preference domains through evidence to decision framework, Health Qual. Life Outcomes (2017) 1–18, https://doi.org/10.1186/s12955-017-0779-5.
[18] P.A. Carney, D.L. Miglioretti, B.C. Yankaskas, K. Kerlikowske, R. Rosenberg, C.M. Rutter, et al., Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography, Ann. Intern. Med. 138 (2003) 168–175.
[19] M.T. Mandelson, N. Oestreicher, P.L. Porter, D. White, C.A. Finder, S.H. Taplin, et al., Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers, JNCI J. Nat. Cancer Inst. 92 (2000) 1081–1087.
[20] D. van der Waal, T.M. Ripping, A.L.M. Verbeek, M.J.M. Broeders, Breast cancer screening effect across breast density strata: a case-control study, Int. J. Cancer 140 (2016) 41–49, https://doi.org/10.1002/ijc.30430.
[21] E.D. Pisano, C. Gatsonis, E. Hendrick, M. Yaffe, J.K. Baum, S. Acharyya, et al., Diagnostic performance of digital versus film mammography for breast-cancer screening, N. Engl. J. Med. 353 (2005) 1773–1783, https://doi.org/10.1056/NEJMoa052911.
[22] S. Weigel, W. Heindel, J. Heidrich, H.W. Hense, O. Heidinger, Digital mammography screening: sensitivity of the programme dependent on breast density, Eur. Radiol. (2016) 1–8, https://doi.org/10.1007/s00330-016-4636-4.
[23] S.L. Heller, S. Hudson, L.S. Wilkinson, Breast density across a regional screening population: effects of age, ethnicity and deprivation, Br. J. Radiol. 88 (2015), https://doi.org/10.1259/bjr.20150242 20150242–0249.
[24] P.A. Carney, C.J. Kasales, A.N.A. Tosteson, J.E. Weiss, M.E. Goodrich, S.P. Poplack, et al., Likelihood of additional work-up among women undergoing routine screening mammography: the impact of age, breast density, and hormone therapy use, Prev. Med. 39 (2004) 48–55, https://doi.org/10.1016/j.ypmed.2004.02.025.
[25] B.L. Sprague, E.F. Conant, T. Onega, M.P. Garcia, E.F. Beaber, S.D. Herschorn, et al., Variation in mammographic breast density assessments among radiologists in clinical practice, Ann. Intern. Med. 165 (2016), https://doi.org/10.7326/M15-2934 457–11.
[26] J.A. Tice, D.L. Miglioretti, C.-S. Li, C.M. Vachon, C.C. Gard, K. Kerlikowske, Breast density and benign breast disease: risk assessment to identify women at high risk of breast cancer, J. Clin. Oncol. 33 (2015) 3137–3143, https://doi.org/10.1200/JCO.2015.60.8869.
[27] B.M. Keller, D.L. Nathan, S.C. Gavenonis, J. Chen, E.F. Conant, D. Kontos, Reader variability in breast density estimation from full-field digital mammograms, Acad. Radiol. 20 (2013) 560–568, https://doi.org/10.1016/j.acra.2013.01.003.
[28] A. Redondo, M. Comas, F. Macià, F. Ferrer, C. Murta-Nascimento, M.T. Maristany, et al., Inter- and intraradiologist variability in the BI-RADS assessment and breast density categories for screening mammograms, Br. J. Radiol. 85 (2012) 1465–1470, https://doi.org/10.1259/bjr/21256379.
[29] American College of Radiology, ACR BI-RADS Atlas: Breast Imaging Reporting and Data System, (2013).
[30] Irshad A, Leddy R, Lewis M, Cluver A, Ackerman S, Pavic D, et al. Changes in Breast Density Reporting Patterns of Radiologists After Publication of the 5th Edition BI-RADS Guidelines: A Single Institution Experience. American Journal of Roentgenology 2017; 209: 943–8. doi:https://doi.org/10.2214/AJR.16.17518.

Article 2

Title: Differences in breast cancer risk after benign breast disease by type of screening diagnosis

Authors: Louro J, Román M, Posso M, Comerma L, Vidal C, Saladié F, Alcántara R, Sanchez M, Quintana MJ, del Riego J, Ferrer J, Peñalva L, Bargalló X, Prieto M, Sala M & Castells X

Journal: Breast. 2020; 54:343-8

Impact Factor: 3.75 (Q1 Obstetrics & Gynecology)

DOI: 10.1016/j.breast.2020.09.005

Abstract:

Introduction: We aimed to assess differences in breast cancer risk across benign breast disease diagnosed at prevalent or incident screens.

Methods: We conducted a retrospective cohort study with data from 629,087 women participating in a long-standing population-based breast cancer screening program in Spain. Each benign breast disease was classified as non-proliferative, proliferative without atypia, or proliferative with atypia, and whether it was diagnosed in a prevalent or incident screen. We used partly conditional Cox hazard regression to estimate the adjusted hazard ratios of the risk of breast cancer.

Results: Compared with women without benign breast disease, the risk of breast cancer was significantly higher (p-value = 0.005) in women with benign breast disease diagnosed in an incident screen (aHR, 2.67; 95%CI: 2.24–3.19) than in those with benign breast disease diagnosed in a prevalent screen (aHR, 1.87; 95%CI: 1.57–2.24). The highest risk was found in women with a proliferative benign breast disease with atypia (aHR, 4.35; 95%CI: 2.09–9.08, and 3.35; 95%CI: 1.51–7.40 for

those diagnosed at incident and prevalent screens, respectively), while the lowest was found in women with non-proliferative benign breast disease (aHR, 2.39; 95%CI: 1.95–2.93, and 1.63; 95%CI: 1.32–2.02 for those diagnosed at incident and prevalent screens, respectively).

Conclusions: Our study showed that the risk of breast cancer conferred by a benign breast disease differed according to type of screen (prevalent or incident). To our knowledge, this is the first study to analyze the impact of the screening type on benign breast disease prognosis.

The Breast 54 (2020) 343–348

Contents lists available at ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Original article

Differences in breast cancer risk after benign breast disease by type of screening diagnosis

Javier Louro [a, b, c, d], Marta Román [a, b, c], Margarita Posso [a, b, c], Laura Comerma [e], Carmen Vidal [f], Francina Saladié [g], Rodrigo Alcantara [h], Mar Sanchez [i], M. Jesús Quintana [j, k], Javier del Riego [l], Joana Ferrer [m], Lupe Peñalva [n], Xavier Bargalló [o], Miguel Prieto [p], María Sala [a, b, c], Xavier Castells [a, b, c, *]

[a] IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain
[b] Research Network on Health Services in Chronic Diseases (REDISSEC), Barcelona, Spain
[c] Servei D'Epidemiologia I Avaluació, Hospital Del Mar, Barcelona, Spain
[d] 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
[e] Servei de Patologia, Hospital Del Mar, Barcelona, Spain
[f] Catalan Institute of Oncology (ICO), Barcelona, Spain
[g] Cancer Epidemiology and Prevention Service, Hospital Universitari Sant Joan de Reus, Institut D'Investigació Sanitària Pere Virgili, Universitat Rovira I Virgili, Reus, Spain
[h] Servei de Diagnòstic per La Imatge, Hospital Del Mar, Barcelona, Spain
[i] Direction General of Public Health, Cantabria, Spain
[j] Department of Clinical Epidemiology and Public Health, University Hospital de La Santa Creu I Sant Pau, IIB Sant Pau, Barcelona, Spain
[k] CIBER of Epidemiology and Public Health (CIBERESP), Spain
[l] Women's Imaging, Department of Radiology, UDIAT Centre Diagnòstic, Parc Taulí Hospital Universitari, Sabadell, Barcelona, Spain
[m] Radiology Department, Hospital de Santa Caterina, Salt, Girona, Spain
[n] Vallés Oriental Breast Cancer Early Detection Program, Spain
[o] Radiology Department, Hospital Clinic, Barcelona, Spain
[p] Breast Cancer Screening Program, Principality of Asturias, Spain

ARTICLE INFO

Article history:
Received 3 July 2020
Received in revised form
24 August 2020
Accepted 4 September 2020
Available online 3 October 2020

Keywords:
Breast neoplasms
Early cancer detection
Benign breast disease
Risk factors

ABSTRACT

Introduction: We aimed to assess differences in breast cancer risk across benign breast disease diagnosed at prevalent or incident screens.

Materials and methods: We conducted a retrospective cohort study with data from 629,087 women participating in a long-standing population-based breast cancer screening program in Spain. Each benign breast disease was classified as non-proliferative, proliferative without atypia, or proliferative with atypia, and whether it was diagnosed in a prevalent or incident screen. We used partly conditional Cox hazard regression to estimate the adjusted hazard ratios of the risk of breast cancer.

Results: Compared with women without benign breast disease, the risk of breast cancer was significantly higher (p-value = 0.005) in women with benign breast disease diagnosed in an incident screen (aHR, 2.67; 95%CI: 2.24–3.19) than in those with benign breast disease diagnosed in a prevalent screen (aHR, 1.87; 95%CI: 1.57–2.24). The highest risk was found in women with a proliferative benign breast disease with atypia (aHR, 4.35; 95%CI: 2.09–9.08, and 3.35; 95%CI: 1.51–7.40 for those diagnosed at incident and prevalent screens, respectively), while the lowest was found in women with non-proliferative benign breast disease (aHR, 2.39; 95%CI: 1.95–2.93, and 1.63; 95%CI: 1.32–2.02 for those diagnosed at incident and prevalent screens, respectively).

Conclusion: Our study showed that the risk of breast cancer conferred by a benign breast disease differed according to type of screen (prevalent or incident). To our knowledge, this is the first study to analyse the impact of the screening type on benign breast disease prognosis.

* Corresponding author. Hospital del Mar, Servei d'Epidemiologia I Avaluació. Passeig Marítim 25-29, Barcelona, 08003, Spain.
E-mail address: xcastells@parcdesalutmar.cat (X. Castells).

https://doi.org/10.1016/j.breast.2020.09.005

1. Introduction

Benign breast disease is associated with an increased risk of breast cancer both in the clinical setting [1,2] and in population-based screening [3]. Quantification of the increased risk according to the characteristics of each lesion is constantly under study. Benign breast disease lesions are most commonly classified as non-proliferative lesions, proliferative lesions without atypia, or proliferative lesions with atypia [4–6]. The risk of subsequent breast cancer is higher in proliferative lesions than in non-proliferative lesions, while the risk is highest in proliferative lesions with atypia [3,7].

Benign breast disease has been proposed as a key risk factor in several breast cancer risk prediction models [8–11]. These models are essential for the development of personalised screening strategies designed to improve the risk-benefit balance of breast cancer screening [12,13]. Therefore, it is important to fully understand differences in breast cancer risk in women diagnosed with benign breast disease.

In population-based breast cancer screening in Spain, women are invited to undergo a mammographic examination every 2 years from the age of 50–69 years. Mammographic examinations can therefore be classified as prevalent screens, ie, women's first participation in screening mammography, and incident screens, ie, all subsequent screening participations. The results of prevalent and incident screens differ in several screening outcomes, such as a higher detection rate and a higher recall rate at prevalent screens [14], resulting in higher sensitivity and lower specificity [15,16]. However, breast cancer diagnosed at prevalent screens has been shown to be less aggressive and to have slower growth [17,18]. Even so, no studies have evaluated the risk of breast cancer associated to bening breast disease according to type of examination (incident or prevalent). We hypothesised that benign breast disease diagnosed in a prevalent screen confer a lower risk of subsequent breast cancer than that diagnosed in incident screens, regardless of the benign breast disease subtype.

Using data from a long-standing population-based screening program in Spain, we aimed to assess differences in the risk of breast cancer after diagnosis of benign breast disease according to the screening type and histological subtype of benign breast disease.

2. Materials and methods

2.1. Study design and participants

The Spanish Breast Cancer Screening Program follows the recommendations of the European Guidelines for Quality Assurance in Breast Cancer Diagnosis [19]. At age 50 years, women are invited to undergo 2-dimensional digital screening mammography. Two projections (mediolateral-oblique and craniocaudal views) are interpreted according to the Breast Imaging Reporting and Data System (BI-RADS) [20] scale by trained breast radiologists. Women with abnormal mammographic findings are recalled for further assessments to confirm or rule out malignancy. Women without a breast cancer diagnosis are invited back for routine screening at 2 years.

We analysed data from seven centres in the Spanish Screening program, which routinely gathers information on benign breast disease diagnoses. The study population included 632,299 women who underwent at least one screening mammogram between 1994 and 2015 and who were followed up until December 2017. Due to the longitudinal nature of the study, women with breast cancer diagnosed in their first screen (n = 3,212) were excluded from the analyses as there was no time for follow-up, leaving left 629,087 women for the analysis.

2.2. Procedures

All women with screening mammograms scored BIRADS 0, 3, 4 or 5 were recalled for further assessments. If cancer could not be ruled out with non-invasive procedures when women were attending recall, core-needle or open biopsy was performed. All biopsies were examined and classified by hospital pathologists in each screening centre. All biopsies with a non-malignant classification were classified as benign breast disease. Following the criteria of Page et al. and Dupont et al. [4–6], we classified benign breast diseases as: non-proliferative, proliferative without atypia, and proliferative with atypia. Only women with asymptomatic benign breast disease diagnosed at screening were included in the study.

Both cancers detected at routine screening and interval cancers (those diagnosed within 24 months after a negative screening episode and before the next screening invitation) were included in the study regardless of whether they were invasive or in situ. Interval cancers were identified by merging population-based cancer registries and hospital-based cancer registries with data from screening participants. Benign breast disease identified at the same time as cancers were excluded from the benign breast disease group.

2.3. Analysis

We used the chi-squared test to compare proportions of different variables among those women without a benign breast disease diagnosis, those with a diagnosis in a prevalent screen and those with a diagnosis in an incident screen.

We calculated incident breast cancer rates using person-years at risk for women with and without a diagnosis of benign breast disease. Women without benign breast disease contributed person-years at risk from the date of the first screening mammogram until breast cancer diagnosis (screen-detected or interval cancer), benign breast disease diagnosis, or 2 years after the last mammographic examination, whichever came first. Women with benign breast disease contributed person-years at risk from the date of benign breast disease diagnosis until breast cancer diagnosis or 2 years after last screening examination, whichever occurred first.

We used a partly conditional Cox proportional hazards model to estimate the adjusted hazard ratios (aHR) and the 95% confidence intervals (95%CI) for the risk of breast cancer by screening type and benign breast disease subtype. These models are an extension of the Cox hazards model for repeated measures, which allowed us to update the changes in benign breast disease status during the study period. All analyses were adjusted for age and calendar year. We adjusted for age because women with benign breast disease diagnosed in a prevalent screen were expected to be younger than those diagnosed in incident screens. Adjustment by calendar year was included to capture possible differences in benign breast disease diagnosis techniques and classification during the study period. An interaction between screening type and benign breast disease subtype was tested; the interaction was found to be non-significant and was consequently not included in the final models (p for interaction = 0.83). Robust standard errors were used to estimate 95% confidence intervals. The proportional hazards assumption was assessed by plotting the log-minus-log of the survival function against log time for each predictive variable. The proportional hazards assumption was reasonable for all predictors.

We plotted the adjusted cumulative incidence curves by estimating the age- and calendar-year adjusted risk of cancer development of the average woman in each category with the partly

conditional Cox model. 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 3.5.0 (Development Core Team, 2014).

3. Results

We analysed information from 629,087 women who underwent 2,327,384 mammographic examinations between January 1994 and December 2015. During the study period, 9431 cases of breast cancer and 9184 cases of benign breast disease were diagnosed. We found no differences in the distribution of benign breast disease subtypes across incident and prevalent screens ($p = 0.48$) (Table 1). The proportion of breast cancer cases was significantly lower in women without benign breast disease than in those with benign breast disease (1.5% vs 2.7%, $p < 0.001$). Among women with a diagnosis of benign breast disease, the proportion of breast cancer cases was higher in women diagnosed in an incident screen than in those diagnosed in a prevalent screen (3.0% vs 2.4%, $p = 0.07$).

Women with benign breast disease had a higher risk of breast cancer than those without benign breast disease, regardless of the subtype of benign breast disease. The highest risk was found in women with proliferative benign breast disease with atypia (aHR, 3.82; 95%CI: 2.23–6.56), followed by those with proliferative benign breast disease without atypia (aHR, 3.19; 95%CI: 2.46–4.13) and those with non-proliferative benign breast disease (aHR, 1.95; 95%CI: 1.68–2.27) (Fig. 1). In addition, among women with benign breast disease, risk was higher in those diagnosed in an incident screen than in those diagnosed in a prevalent screen (aHR, 2.67; 95%CI: 2.24–3.19, and aHR, 1.87; 95%CI: 1.57–2.24, respectively).

The hazard ratios associated within each combination of screening type and benign breast disease subtype are shown in Fig. 2. Across benign breast disease subtypes, those diagnosed in an incident screen conferred a higher risk than those diagnosed in prevalent screens, although not statistically significant for al subtypes. Compared with women without a benign breast disease, the highest risk was found in those women with a proliferative benign breast disease with atypia (aHR, 4.35; 95%CI: 2.09–9.08, and 3.35; 95%CI: 1.51–7.40 for those diagnosed at incident and prevalent screens, respectively, p-value for comparison; $p = 0.634$), followed by women with proliferative benign breast disease without atypia (aHR, 3.83; 95%CI: 2.63–5.58, and 2.78; 95%CI: 1.95–3.96 for those diagnosed at incident and prevalent screens, respectively; p-value for comparison $p = 0.223$). The lowest was found in women with non-proliferative benign breast disease (aHR, 2.39; 95%CI: 1.95–2.93, and 1.63; 95%CI: 1.32–2.02 for those diagnosed at incident and prevalent screens, respectively, p-value for comparison; $p = 0.011$).

We examined the adjusted cumulative incident curves of breast cancer across the different classifications of benign breast disease and screening types. The probability of breast cancer diverged over time. The average 10-year breast cancer probability of women without a benign breast disease diagnosis was 1.9%. Among women with benign breast disease, the probability was higher in women diagnosed in an incident screen than in those diagnosed in a prevalent screen (average 10-year probability of breast cancer 4.9% vs 3.5%). Among benign breast disease subtypes, the highest probability of cancer was found in women with a proliferative benign breast disease with atypia followed by those with proliferative benign breast disease without atypia and those with a non-proliferative benign breast disease (average 10-year breast cancer probability 6.9%, 5.8% and 3.6%, respectively) (Fig. 3). The highest probability was found in women with proliferative benign breast disease diagnosed in an incident screen, with an average 10-year probability of breast cancer of 7.8% (Fig. 4).

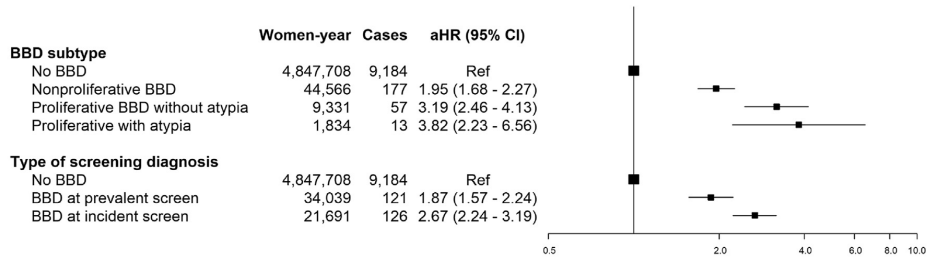
4. Discussion

In this study of more than 600,000 women with follow-up for more than 20 years, we found that a diagnosis of benign breast disease in an incident screen conferred a higher risk of subsequent breast cancer than diagnosis in a prevalent screen, regardless of the histological subtype. These findings highlight the importance of considering the screening type when benign breast disease was diagnosed in risk of breast cancer estimation. To our knowledge, this is the first study to include screening type in the assessment of the impact of benign breast disease on the risk of subsequent breast cancer.

Over the past decades, multiple studies have assessed the relationship between benign breast disease and the risk of breast cancer [2–6]. Particular efforts have been made to assess the association of this risk with the various benign breast disease subtypes [4–6]. As seen in previous reports, our study showed that women with benign breast disease had an increased risk of breast cancer [2]. Consistently, we found that the risk of breast cancer was highest among women with a proliferative benign breast disease with atypia, while proliferative benign breast disease without atypia conferred a higher risk than non-proliferative benign breast

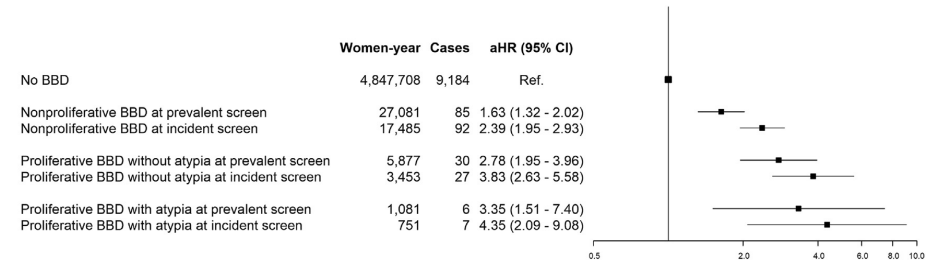
Table 1
Characteristics of the study population.

	No benign breast disease N = 619,864	Benign breast disease diagnosed in a prevalent screen N = 5,049	Benign breast disease diagnosed in an incident screen N = 4,174
Age at first screen			
50-54	360,020 (58.1%)	3,420 (67.7%)	2,687 (64.4%)
55-59	123,994 (20.0%)	758 (15.0%)	936 (22.4%)
60-64	100,956 (16.3%)	637 (12.6%)	499 (12.0%)
65-69	34,894 (5.6%)	234 (4.6%)	52 (1.2%)
Year of first screen			
<2005	299,173 (48.3%)	2,111 (41.8%)	2,856 (68.4%)
2005–2010	167,747 (27.1%)	1,220 (24.2%)	1,038 (24.9%)
>2010	152,944 (24.7%)	1,718 (34%)	280 (6.7%)
Type of benign breast disease			
No benign breast disease	619,864 (100%)	0 (0.0%)	0 (0.0%)
Non-proliferative	0 (0.0%)	3,948 (78.2%)	3,282 (78.6%)
Proliferative without atypia	0 (0.0%)	877 (17.4%)	728 (17.4%)
Proliferative with atypia	0 (0.0%)	224 (4.4%)	164 (3.9%)
Breast Cancer			
No	610,680 (98.5%)	4,928 (97.6%)	4,048 (97%)
Si	9,184 (1.5%)	121 (2.4%)	126 (3.0%)



The figure is based on two separate Cox regression models for subtype of benign breast disease and type of screening diagnosis.
All models were adjusted for age at screen and calendar year at screen.
BBD: Benign breast disease.

Fig. 1. Adjusted hazard ratios (aHR) of breast cancer incidence in women with benign breast disease compared with women with negative screening tests.



Adjusted for age at screen and calendar year at screen.
BBD: Benign breast disease.

Fig. 2. Adjusted hazard ratios (aHR) of breast cancer incidence in women with benign breast disease compared with women with negative screening tests testing the combined effect of type of benign breast disease, and round at benign breast disease diagnosis.

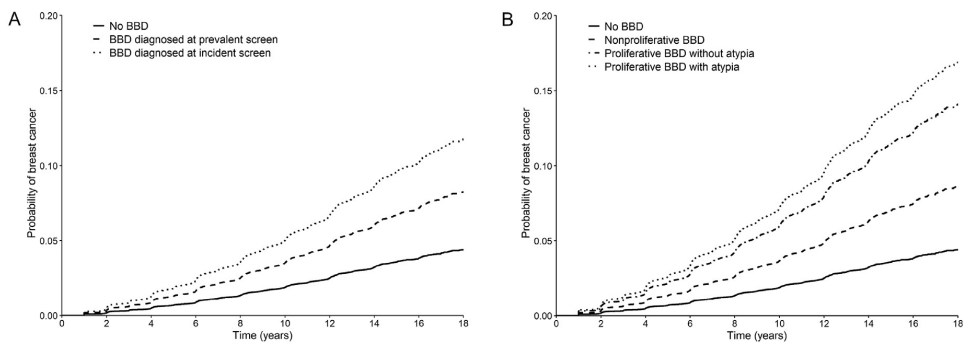


Fig. 3. Adjusted survival curves for breast cancer incidence based on Cox proportional hazards model for women with benign breast disease vs women with negative screening tests. Fig. 3 a. Solid line represents negative screening test group; dashed line represents benign breast disease diagnosed at prevalent round, dotted line benign breast disease diagnosed in incident round. Fig. 3 b. Solid line represents negative screening test group; dashed line represents nonproliferative benign breast disease, dashdotted line represent proliferative benign breast disease without atypia, dotted line represents proliferative benign breast disease with atypia.

disease [3]. Although the reduced sample size for some BBD subtypes led to non-significant differences in some subgroup comparisons, we found that the difference in risk across benign breast disease subtypes remained proportional within prevalent and incident screens and was systematically higher in those BBD diagnosed in incident screens. This finding is particularly relevant since 55% of benign breast diseases are diagnosed in prevalent screens. Screening type therefore provides key information for risk prediction in benign breast disease. Unless this information is included, the risk attributed to benign breast disease diagnosed in prevalent screens could be overestimated, and that for incident screens could be underestimated.

Previous studies performed in the last decade have assessed differences in breast cancer screening outcomes, such as cancer detection rates and false positive rates, by screening type [14–16]. Moreover, previous authors found differences in breast cancer characteristics depending on whether the cancers were diagnosed in a prevalent or incident screen [17,18], suggesting that latent

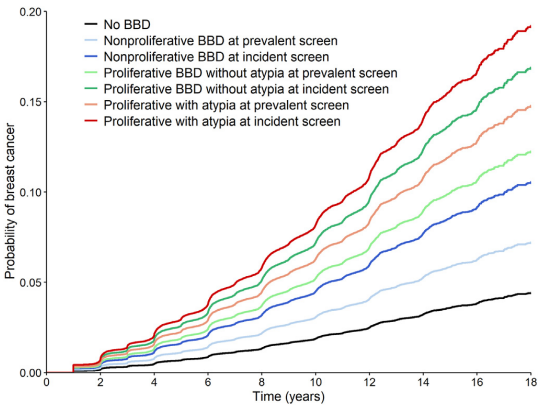


Fig. 4. Adjusted survival curves for breast cancer incidence based on Cox proportional hazards model for women with benign breast disease vs women with negative screening tests.

cancers diagnosed in prevalent screens have a slower growth pattern. The lower risk of breast cancer observed in women with benign breast disease diagnosed in prevalent screens might be partially explained by a slower growth pattern in prevalent benign breast disease. Age may play an important role in this effect, since women with benign breast disease diagnosed in incident screens are, on average, older than those diagnosed in a prevalent screen. However, to control for this potential confounding effect, we adjusted all analyses by age and calendar year.

The results of this study may have implications for clinical decisions on the follow-up of women with a diagnosis of benign breast disease, in which distinct follow-up strategies may be recommended depending on the benign breast disease subtype and screening type at diagnosis. The findings may also have implications for individualised risk prediction. The Breast Cancer Surveillance Consortium model [21] is being used in a large international randomised clinical trial to define the individual risk of the population targeted for breast cancer screening with a view to offering them personalised screening strategies [22]. Our findings reveal that screening type explained part of the risk associated with benign breast disease. Taking this variability into account could help to improve the discriminatory power of breast cancer risk prediction models, which is commonly moderate [11,23]. In addition, we found that women with proliferative benign breast disease diagnosed in an incident screen had a 4-fold higher risk of developing breast cancer than those without. This information, after further analysis with adjustment for other risk factors such as breast density [24], family history [25], or a risk score using information from single nucleoid polymorphisms [26], may be key to defining risk groups that could benefit from tailored screening strategies.

This study has some limitations. First, the number of cancers detected after proliferative benign breast disease with atypia was small because this subtype is uncommon, which limited our capacity to perform some subgroup analyses. Second, there is a possible bias produced by temporary changes both in the benign breast disease classification and in biopsy techniques (because fine-needle aspiration cytology has become practically obsolete). This bias is partially controlled as we adjusted our analysis by calendar year. Third, to be able to classify benign breast disease, we restricted our study to those benign breast diseases with available information on their histological subtype. Last, it was not possible to adjust the statistical analysis on other breast cancer risk factors such as breast density or familial history of breast cancer.

A strength of this study is that we analysed a large cohort of more than 600,000 women screened in a well-established population-based screening program with a 20-year follow-up. This is one of the largest cohorts analysing histopathologically confirmed benign breast disease, with nearly 10,000 diagnoses during follow-up.

In summary, our study shows that, regardless of the type of benign breast disease, women with benign breast disease diagnosed in an incident screen have a significantly higher subsequent risk of breast cancer than those with a benign breast disease diagnosed in a prevalent screen. To our knowledge, this is the first study to analyse this topic. It is important to consider this risk when developing risk-based personalised screening strategies.

Funding

This study was supported by grants from Instituto de Salud Carlos III FEDER (grant numbers: PI15/00098 and PI17/00047), and by the Research Network on Health Services in Chronic Diseases (RD12/0001/0015).

Ethics approval and consent to participate

The study was approved by the Clinic 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 register based study that used anonymised retrospective data and hence written consent was not required.

Declaration of competing interest

None declared.

Acknowledgments

The authors acknowledge the dedication and support of the entire Benign Lesion (BELE) and Individualised Risk (IRIS) Study Groups listed here in alphabetical order and grouped by institution: (a) IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain: Rodrigo Alcantara, Andrea Burón, Xavier Castells, Laia Domingo, Javier Louro, Margarita Posso, Ana Rodríguez-Arana, Marta Román, Maria Sala, Ignasi Tusquets, Ivonne Vazquez, Mar Vernet-Tomas; (b) Corporació Sanitària Parc Taulí, Sabadell, Spain: Marisa Baré, Javier del Riego; (c) Catalan Institute of Oncology, Barcelona, Spain: Llucia Benito, Carmen Vidal (d) Hospital Santa Caterina, Girona, Spain: Joana Ferrer; (e) Catalan Institute of Oncology, Girona, Spain: Rafael Marcos-Gragera; (f) Hospital de la Santa Creu i Sant Pau, Barcelona, Spain: Judit Solà-Roca, María Jesús Quintana; (g) General Directorate of Public Health, Government of Cantabria, Spain: Mar Sánchez; (h) Principality of Asturias Health Service, Spain: Miguel Prieto; (i) Fundació Lliga per a La Investigació i Prevenció Del Càncer, Tarragona, Spain: Francina Saladié, Jaume Galceran; (j) Hospital Clinic, Barcelona, Spain; Xavier Bargalló, Isabel Torá-Rocamora; (k) Vallés Oriental Breast Cancer Early Detection Program, Spain; Lupe Peñalva; (l) Catalonian Cancer Strategy, Barcelona, Spain: Josep Alfons Espinàs. Javier Louro is a Ph.D. candidate at the Methodology of Biomedical Research and Public Health program, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain.

References

[1] Carter CL, Corle DK, Micozzi MS, Schatzkin A, Taylor PR. A prospective study of the development of breast cancer in 16,692 women with benign breast disease. Am J Epidemiol 1988;128(3):467–77.

[2] 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.

[3] Castells X, Domingo L, Corominas JM, Tora-Rocamora I, Quintana MJ, Bare M, et al. Breast cancer risk after diagnosis by screening mammography of non-proliferative or proliferative benign breast disease: a study from a population-based screening program. Breast Canc Res Treat 2015;149(1):237–44.

[4] 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.

[5] Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. N Engl J Med 1985;312(3):146–51.

[6] Fitzgibbons PL, Henson DE, Hutter RV. Benign breast changes and the risk for subsequent breast cancer: an update of the 1985 consensus statement. Cancer Committee of the College of American Pathologists. Arch Pathol Lab Med 1998;122(12):1053–5.

[7] Dyrstad SW, Yan Y, Fowler AM, Colditz GA. Breast cancer risk associated with benign breast disease: systematic review and meta-analysis. Breast Canc Res Treat 2015;149(3):569–75.

[8] Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med 2004;23(7):1111–30.

[9] 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.

[10] Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. Ann Intern Med 2008;148(5):337–47.

[11] Louro J, Posso M, Hilton Boon M, Roman M, Domingo L, Castells X, et al. A systematic review and quality assessment of individualised breast cancer risk prediction models. Br J Canc 2019;121(1):76–85.

[12] 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.

[13] 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.

[14] Blanks RG, Given-Wilson R, Alison R, Jenkins J, Wallis MG. An analysis of 11.3 million screening tests examining the association between needle biopsy rates and cancer detection rates in the English NHS Breast Cancer Screening Programme. Clin Radiol 2019;74(5):384–9.

[15] 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.

[16] Hofvind S, Geller B, Vacek PM, Thoresen S, Skaane P. Using the European guidelines to evaluate the Norwegian breast cancer screening program. Eur J Epidemiol 2007;22(7):447–55.

[17] Anderson TJ, Lamb J, Alexander F, Lutz W, Chetty U, Forrest AP, et al. Comparative pathology of prevalent and incident cancers detected by breast screening. Edinburgh Breast Screening Project. Lancet 1986;1(8480):519–23.

[18] Paci E, Ponti A, Zappa M, Patriarca S, Falini P, Delmastro G, et al. Early diagnosis, not differential treatment, explains better survival in service screening. Eur J Canc 2005;41(17):2728–34.

[19] 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.

[20] Sickles EA, D'Orsi CJ, Bassett LW. ACR BI-RADS® mammography. In: ACR BI-RADS® atlas, breast imaging reporting and data System. Reston, VA: American College of Radiology; 2013.

[21] 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.

[22] MyPeBS. Randomized comparison of risk-stratified versus standard breast cancer screening in European women aged 40-70 (MyPeBS) [Internet]. 2017. Available from: www.brumammo.be/.../bmm-my-pebs-clinical-trial-protocol.pdf.

[23] McCarthy AM, Guan Z, Welch M, Griffin ME, Sippo DA, Deng Z, et al. Performance of breast cancer risk assessment models in a large mammography cohort. J Natl Cancer Inst 2019;112(5):489–97.

[24] Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007;356(3):227–36.

[25] Antoniou AC, Easton DF. Risk prediction models for familial breast cancer. Future Oncol 2006;2(2):257–74.

[26] 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.

Article 3

Title: Breast density, benign breast disease, and risk of breast cancer over time

Authors: Román M, Louro J, Posso M, Alcántara R, Peñalva L, Sala M, del Riego J, Prieto M, Vidal C, Sánchez M, Bargalló X, Tusquets I & Castells X

Journal: European Radiology. 2021. Epub ahead of print

Impact Factor: 4.10 (Q1 Radiology, nuclear medicine & medical imaging)

DOI: 10.1007/s00330-020-07490-5

Abstract:

Objectives: Assessing the combined effect of mammographic density and benign breast disease is of utmost importance to design personalized screening strategies.

Methods: We analyzed individual-level data from 294,943 women aged 50–69 years with at least one mammographic screening participation in any of 4 areas of the Spanish Breast Cancer Screening Program from 1995 to 2015 and followed up until 2017. We used partly conditional Cox models to assess the association between benign breast disease, breast density, and the risk of breast cancer.

Results: During a median follow-up of 8.0 years, 3,697 (1.25%) women had a breast cancer diagnosis and 5,941 (2.01%) had a benign breast disease. More than half of screened women had scattered fibroglandular density (55.0%). The risk of breast cancer independently increased with the presence of benign breast disease and with the increase in breast density (p for interaction = 0.84). Women with benign breast disease and extremely dense breasts had a threefold elevated risk of breast

cancer compared with those with scattered fibroglandular density and without benign breast disease (hazard ratio [HR] = 3.07; 95%CI = 2.01–4.68). Heterogeneous density and benign breast disease was associated with nearly a 2.5 elevated risk (HR = 2.48; 95%CI = 1.66–3.70). Those with extremely dense breast without a benign breast disease had a 2.27 increased risk (95%CI = 2.07–2.49).

Conclusions: Women with benign breast disease had an elevated risk for over 15 years independently of their breast density category. Women with benign breast disease and dense breasts are at high risk for future breast cancer.

European Radiology
https://doi.org/10.1007/s00330-020-07490-5

BREAST

Breast density, benign breast disease, and risk of breast cancer over time

Marta Román[1,2] · Javier Louro[1,2,3] · Margarita Posso[1,2] · Rodrigo Alcántara[4] · Lupe Peñalva[5] · Maria Sala[1,2] · Javier del Riego[6] · Miguel Prieto[7] · Carmen Vidal[8] · Mar Sánchez[9] · Xavier Bargalló[10] · Ignasi Tusquets[11] · Xavier Castells[1,2]

Received: 9 March 2020 / Revised: 6 October 2020 / Accepted: 9 November 2020

Abstract

Objectives Assessing the combined effect of mammographic density and benign breast disease is of utmost importance to design personalized screening strategies.

Methods We analyzed individual-level data from 294,943 women aged 50–69 years with at least one mammographic screening participation in any of four areas of the Spanish Breast Cancer Screening Program from 1995 to 2015, and followed up until 2017. We used partly conditional Cox models to assess the association between benign breast disease, breast density, and the risk of breast cancer.

Results During a median follow-up of 8.0 years, 3697 (1.25%) women had a breast cancer diagnosis and 5941 (2.01%) had a benign breast disease. More than half of screened women had scattered fibroglandular density (55.0%). The risk of breast cancer independently increased with the presence of benign breast disease and with the increase in breast density (p for interaction = 0.84). Women with benign breast disease and extremely dense breasts had a threefold elevated risk of breast cancer compared with those with scattered fibroglandular density and without benign breast disease (hazard ratio [HR] = 3.07; 95%CI = 2.01–4.68). Heterogeneous density and benign breast disease was associated with nearly a 2.5 elevated risk (HR = 2.48; 95%CI = 1.66–3.70). Those with extremely dense breast without a benign breast disease had a 2.27 increased risk (95%CI = 2.07–2.49).

Conclusions Women with benign breast disease had an elevated risk for over 15 years independently of their breast density category. Women with benign breast disease and dense breasts are at high risk for future breast cancer.

Key Points

- *Benign breast disease and breast density were independently associated with breast cancer.*
- *Women with benign breast disease had an elevated risk for up to 15 years independently of their mammographic density category.*

Keywords Breast neoplasms · Mass screening · Longitudinal studies · Mammographic density · Benign breast disease

✉ Xavier Castells
xcastells@parcdesalutmar.cat

[1] Department of Epidemiology and Evaluation, IMIM (Hospital del Mar Medical Research Institute), Passeig Marítim 25-29, 08003 Barcelona, Spain

[2] Research Network on Health Services in Chronic Diseases (REDISSEC), Barcelona, Spain

[3] European Higher Education Area, Department of Pediatrics, Obstetrics and Gynecology, Preventive Medicine and Public Health, Universitat Autónoma de Barcelona, Barcelona, Spain

[4] Department of Radiology, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain

[5] Breast Cancer Screening Technical Office, Private Foundation Asil Hospital, Granollers, Spain

[6] Department of Radiology, Parc Taulí University Hospital-UAB, Sabadell, Catalonia, Spain

[7] General Directorate of Public Health, Principality of Asturias Health Service, Asturias, Spain

[8] Cancer Prevention and Monitoring Program, Catalan Institute of Oncology, Barcelona, Spain

[9] General Directorate of Public Health, Government of Cantabria, Santander, Spain

[10] Department of Radiology, Hospital Clinic, Barcelona, Spain

[11] Department of Medical Oncology, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain

Published online: 06 January 2021

Abbreviations

BI-RADS	Breast Imaging Reporting and Data System
CI	Confidence interval
HR	Hazard ratio

Introduction

High breast density and benign breast disease are major risk factors for breast cancer. Women with a high mammographic density are associated with a two- to fourfold elevated risk compared with women with fatty breast tissue [1–4], while the presence of a benign breast disease doubles the risk of subsequent breast cancer [5]. Even though research has focused on each of these risk factors separately, little is known about the combined effect of these two factors, which has important implications to designing personalized strategies aimed at improving the effectiveness of breast cancer screening [6, 7].

A prior study with limited sample size found a significant interaction between mammographic density and different subtypes of benign breast disease [8]. This finding is in contrast to a more recent large study that found no combined effect of mammographic density and different subtypes of benign breast disease with the risk of breast cancer [9]. The study population in the abovementioned studies was limited to women with benign breast disease only, and they assessed the effects of different benign breast disease subtypes, but they did not assess the presence or absence of a benign breast disease. Women with a benign breast disease and high mammographic density could be at very high risk for breast cancer and they may benefit from more intense follow-up strategies or different screening modalities.

Our aim was to evaluate the joint association between benign breast disease and mammographic density with the risk of breast cancer in the context of population-based breast cancer screening by using individual-level data from the longstanding mammography program in Spain.

Methods

Setting and study population

Mammographic screening in Spain follows the recommendations of the European Guidelines [10] and is publicly funded. The program started in 1990 in a single setting and became nationwide in 2006. Population-based screening in Spain has been previously described in detail elsewhere [11–13]. Briefly, the program is organized into administrative screening settings responsible for the local application of screening in their area. Women aged 50 to 69 years are invited every 2 years for a two-view mammogram (craniocaudal, and mediolateral oblique). All screening mammograms are interpreted by 2 trained breast radiologists and classified according to the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) scale or equivalent [12]. Prior mammograms are available for comparison at subsequent screens. Women with abnormal findings on mammographic interpretation are recalled for further assessments, including additional imaging, ultrasound, and invasive procedures. If no malignancy is detected, women are referred back to screening at 2 years, while women diagnosed with breast cancer are referred for treatment. All breast biopsies are histopathologically confirmed by trained pathologists.

Data for this study were obtained from four centers of the Breast Cancer Screening Program that routinely gather information on mammographic density (Costa de Ponent, Vallés Oriental, Sabadell-Cerdanyola, and Cantabria). Data for the study comprised information about women screened between 1996 and 2015, with follow-up until December 2017 for interval cancer cases. The centers collect information on screening mammography examinations, recall, further assessments, and diagnoses performed in their defined catchment areas. Approval for data analyses was granted by the review boards of the institutions providing data. Informed consent was not required since we used anonymized retrospective data.

Our study population consisted of 295,837 women aged 50–69 years with at least one BI-RADS breast density examination at screening mammography between January 1996 and December 2015. We excluded 894 women diagnosed with breast cancer at first screen as they did not have any follow-up. This left 294,943 women for the analysis.

Definition of study variables

The results of breast biopsies were classified by trained pathologists at each screening center. Women were considered to have a benign breast disease if, after exclusion of malignancy in a biopsy, the pathology report specified a diagnosis of benign breast disease. If women had more than one biopsy, we ensured that none of them confirmed malignancy in order to establish the presence of a benign breast disease. Lobular carcinoma in situ on biopsy was considered a benign breast disease.

Breast cancer cases included all invasive cancers and ductal carcinoma in situ regardless of whether they were screen-detected cancers, or interval breast cancers (those diagnosed within 24 months after a negative screening exam and before the next screening invitation). Breast cancer cases were identified from the screening center databases, population-based cancer registries, the regional Minimum Data Set, and hospital-based cancer registries.

Breast density was determined by breast radiologists at the time of screening as part of routine mammogram interpretation, using the BI-RADS density categories [14]: almost

entirely fat (A), scattered fibroglandular density (B), heterogeneously dense (C), or extremely dense (D). If women had more than one mammographic density measure during the study period, we chose the breast density measured at the earliest screening examination. In case of discordance in breast density classification between radiologists, we chose the density classification assigned by the most expert radiologist. The BI-RADS B category was used as the reference group for mammographic density because it is the most common group and women with this category are considered to have an average risk of breast cancer.

Statistical analysis

Women were considered as the unit of analysis. For each woman, person-years at risk was calculated from the date of the first screen. For women with no history of benign breast disease, follow-up started at the date of the first screening participation and ended at the date of breast cancer diagnosis or 2 years after last mammography examination for follow-up of interval cancer cases, whichever came first. Women with a benign breast disease contributed person-years at risk to the "no benign breast disease group" from the date of the first screen until the date of diagnosis of first benign breast disease. They contributed person-years to the "benign breast disease group" from the date of diagnosis of the first benign breast disease. Similarly to women without a benign breast disease, follow-up for women with a benign breast disease ended at the date of breast cancer diagnosis or 2 years after last mammography examination for follow-up of interval cancer cases, whichever came first.

We compared the frequency distributions of various risk factors for women with and without breast cancer: presence or absence of a benign breast disease, BI-RADS density category at the earliest examination, and age at the first screen (in 5-year age groups). In addition, rates of breast cancer by the presence or absence of a previous benign breast disease and BI-RADS density category were calculated as the number of breast cancer cases divided by the number of women-years at risk in each group. Confidence intervals for rates were calculated using exact Poisson distribution.

We used a partly conditional Cox proportional hazards model to estimate the hazard ratios and 95% confidence intervals of the association between benign breast disease and mammographic density with the risk of breast cancer [15]. By using a partly conditional Cox model, we included all screening mammograms in each individual, incorporating benign breast diseases occurring after the first screening participation and accounting for within-woman correlation. We tested whether there was an interaction between previous benign breast disease and breast density with the risk of breast cancer.

To control for possible confounders such as changes in BI-RADS classification or technical changes over time, all models were adjusted for age at screen (continuous) and year of screen (continuous). The proportional hazards assumption was ascertained by plotting the log-minus-log of the survivor function against log time for each predictor variable. The tests found no evidence of violation of the proportional hazards assumption between the covariates and time.

We performed three sensitivity analyses. First, we limited the outcome to invasive breast cancer only. Women with DCIS were censored at the time of diagnosis and they were not included as breast cancer cases. Second, to test the effect of repeated breast density measures in the same person over the study period, we performed the analyses using the latest mammographic density measure for each woman instead of the earliest. Third, to test if women at higher risk had a higher proportion of interval cancer cases than those at lower risk, we performed an analysis of interval breast cancers as outcome. Women with screen-detected cancer were censored at the time of diagnosis and not included as breast cancer cases. All tests were two-sided with a 5% significance level. Statistical analyses were conducted in R 3.4.2 (R Foundation for Statistical Computing).

Results

Women in the study population had a median follow-up of 8.0 years. Out of the 294,943 women in the study population, 3697 (1.25%) had a diagnosis of breast cancer and 5941 (2.01%) had benign breast disease. The proportion of previous benign breast disease was higher among women with breast cancer than in those without breast cancer (2.6% and 2.0%, respectively) (Table 1). At baseline examination, BI-RADS scattered fibroglandular density (B) was the most common category for women with and without breast cancer (49.5 and 55.1%, respectively). Among women with breast cancer, there was a higher proportion of women with high mammographic density (BI-RADS density categories C or D), and a larger proportion of women aged 55 to 64 years.

The distribution of mammographic density categories by the presence or absence of benign breast disease is shown in Table 2. The most frequent mammographic density category was scattered fibroglandular for women with and without benign breast disease (50.6% and 55.1%, respectively). Compared with those without benign breast disease, women with benign breast disease had a higher proportion of extremely dense breasts (15.6% and 8.9%, in women with and without a benign breast disease, respectively), and a lower proportion of entirely fatty breasts (12.4% and 20.8%, respectively).

Table 3 shows the overall rates of breast cancer by the presence or absence of benign breast disease and by BI-RADS density categories. Among women with scattered fibroglandular breasts, the rate of breast cancer per 1000 women-years at risk was 2.43 (95%CI 1.68–3.17) and 1.46

Table 1 Baseline characteristics of women in the study population

	No breast cancer, $N = 291{,}246$	Breast cancer, $N = 3697$	p
Benign breast disease			
No	285 400 (98.0%)	3 602 (97.4%)	*
Yes	5 846 (2.0%)	95 (2.6%)	*
Breast density			
Almost entirely fat	60 343 (20.7%)	498 (13.5%)	*
Scattered fibroglandular	160 496 (55.1%)	1 831 (49.5%)	*
Heterogeneously dense	44 386 (15.2%)	724 (19.6%)	*
Extremely dense	26 021 (8.9%)	644 (17.4%)	*
Age at first screen			
50–54	187 829 (64.5%)	2 181 (59.0%)	*
55–59	56 149 (19.3%)	915 (24.7%)	*
60–64	35 455 (12.2%)	543 (14.7%)	*
65–69	11 813 (4.1%)	58 (1.6%)	*

*Different at $p < 0.05$ in a two-sided test of equality for column proportions (z-test). Tests are adjusted using the Bonferroni correction for multiple comparison. Using the Breast Imaging Reporting and Data System (BI-RADS) density categories: A = almost entirely fat; B = scattered fibroglandular densities; C = heterogeneously dense; D = extremely dense

(95%CI 1.40–1.53) for those with and without benign breast disease, respectively. For women with extremely dense breasts, the rates per 1000 women-years were 3.95 (95%CI 2.30–5.60) and 2.97 (95%CI 2.74–3.21) for those with and without benign breast disease, respectively.

The risks associated with each combination of mammographic density and benign breast disease are shown in Table 4. No significant interaction was found between previous benign breast disease and mammographic density (p for interaction = 0.84). The risk of breast cancer increased with increasing mammographic density and with the presence of benign breast disease. Compared with women with average breast density (scattered fibroglandular) and no benign breast disease, women with low breast density and no benign breast disease had a lower risk of future breast cancer (HR = 0.65; 95%CI = 0.59 to 0.72). In contrast, the risk for those with low breast density and a benign breast disease was similar to the average reference group (HR = 1.29; 95%CI = 0.64–2.58). Women with heterogeneous mammographic density had an elevated risk of breast cancer, both those with and without benign breast disease (HR = 2.48, 95%CI = 1.66–3.70; and HR = 1.58, 95%CI = 1.44–1.72, respectively). Similarly, women with extremely dense breasts were at the highest risk of future breast cancer, independently of whether they had benign breast disease or not (HR = 3.07; 95%CI = 2.01–4.68; and HR = 2.27, 95%CI = 2.07–2.49, for women with and without benign breast disease, respectively).

We examined the cumulative incidence curves of breast cancer for benign breast disease within breast density strata (Fig. 1). The figure depicts how the occurrence of breast cancer cases follows a staggered 2-year pattern given by the biennial screening participations of women in the programs. The risk of developing breast cancer diverged over time among women with and without benign breast disease, independently of the BI-RADS density strata. The probability of developing breast cancer 15 years after the index mammography examination for women without benign breast disease increased as mammographic density increased, ranging from 1.5% (95%CI = 1.3–1.7) for women with almost entirely fatty breasts to 4.8% (95%CI = 4.2–5.4) for those with extremely dense breasts. For women with a benign breast disease, the probability of breast cancer 15 years after the index

Table 2 Frequency and prevalence of breast density categories by presence or absence of a benign breast disease

	Bi-RADS breast density				
Benign breast disease	Almost entirely fat	Scattered fibroglandular	Heterogeneous	Extremely dense	Total
No BBD	60 106 (20.8%)	159 318 (55.1%)	43 842 (15.2%)	25.736 (8.9%)	289 002
BBD	735 (12.4%)	3 009 (50.6%)	1 268 (21.3%)	929 (15.6%)	5 941

BBD benign breast disease

Table 3 Overall rates of breast cancer by presence or absence of a benign breast disease

	Women-years at risk	Number of cases	Breast cancer rate (95%CI) per 1000 women-years (‰)
No benign breast disease			
Almost entirely fat	488,452	490	1.00 (0.91–1.09)
Scattered fibroglandular	1,222,576	1790	1.46 (1.40–1.53)
Heterogeneous	314,808	700	2.22 (2.06–2.39)
Extremely dense	209,134	622	2.97 (2.74–3.21)
Benign breast disease			
Almost entirely fat	4188	8	1.91 (0.59–3.23)
Scattered fibroglandular	16,901	41	2.43 (1.68–3.17)
Heterogeneous	7076	24	3.39 (2.03–4.75)
Extremely dense	5556	22	3.95 (2.30–5.60)

mammography ranged from 3.4% (95%CI = 1.1–5.7) for those with almost entirely fat breast to 7.0 (95%CI = 4.1–9.9) for extremely dense breast.

Sensitivity analyses excluding DCIS and including invasive breast cancer only as outcome were consistent with the full model and produced very similar estimates (Supplementary material, Table 1). In addition, the effect of choosing the latest rather than the earliest mammographic density measure had limited impact on the risk estimates (Supplementary material, Table 2). Lastly, the analyses of interval breast cancers as outcome had a small number of cases, which limited the statistical power. However, the analyses showed that the risk for interval breast cancer was higher in women with benign breast disease, and the risk increased as breast density increased (Supplementary material, Tables 3 and 4).

Discussion

In this analysis of data from more than 290,000 women with a follow-up of more than 15 years, we examined the combined contribution of benign breast disease and mammographic density to the risk of breast cancer. We found that the presence of benign breast disease and high mammographic density were independently associated with an increase in the risk of breast cancer. The risk diverged over the study period for women with and without benign breast disease across mammographic density categories. The highest risk for breast cancer was found in women with benign breast disease and extremely dense breasts. Few studies have evaluated the combined effect of benign breast disease and breast density on the risk of breast cancer at mammographic screening. We used long-term population-based data from women participating in breast cancer screening to assess how benign breast disease and breast density affect the risk of breast cancer.

Consistent with most studies, we found that mammographic density was associated with breast cancer risk [16–18]. We observed that higher BI-RADS breast density categories were associated with an increased risk of breast cancer. In addition, there is consistent evidence that women with a benign breast disease are at an increased risk of breast cancer [5, 19]. A prior large study assessed the combined effect of mammographic breast density and different subtypes of benign breast disease, in a population of women with the presence of benign breast

Table 4 Breast cancer risk by breast density and benign breast disease

	BI-RADS breast density, HR (95%CI)			
Benign breast disease	Almost entirely fat	Scattered fibroglandular	Heterogeneous	Extremely dense
No BBD	0.65 (0.59–0.72)	ref.	1.58 (1.44–1.72)	2.27 (2.07–2.49)
BBD	1.29 (0.64–2.58)	1.68 (1.24–2.29)	2.48 (1.66–3.70)	3.07 (2.01–4.68)

Hazard ratios relative to women with no benign breast disease and scattered fibroglandular density. Adjusted for age at first screen, and year of screen

p value for interaction between benign breast disease and breast density = 0.84

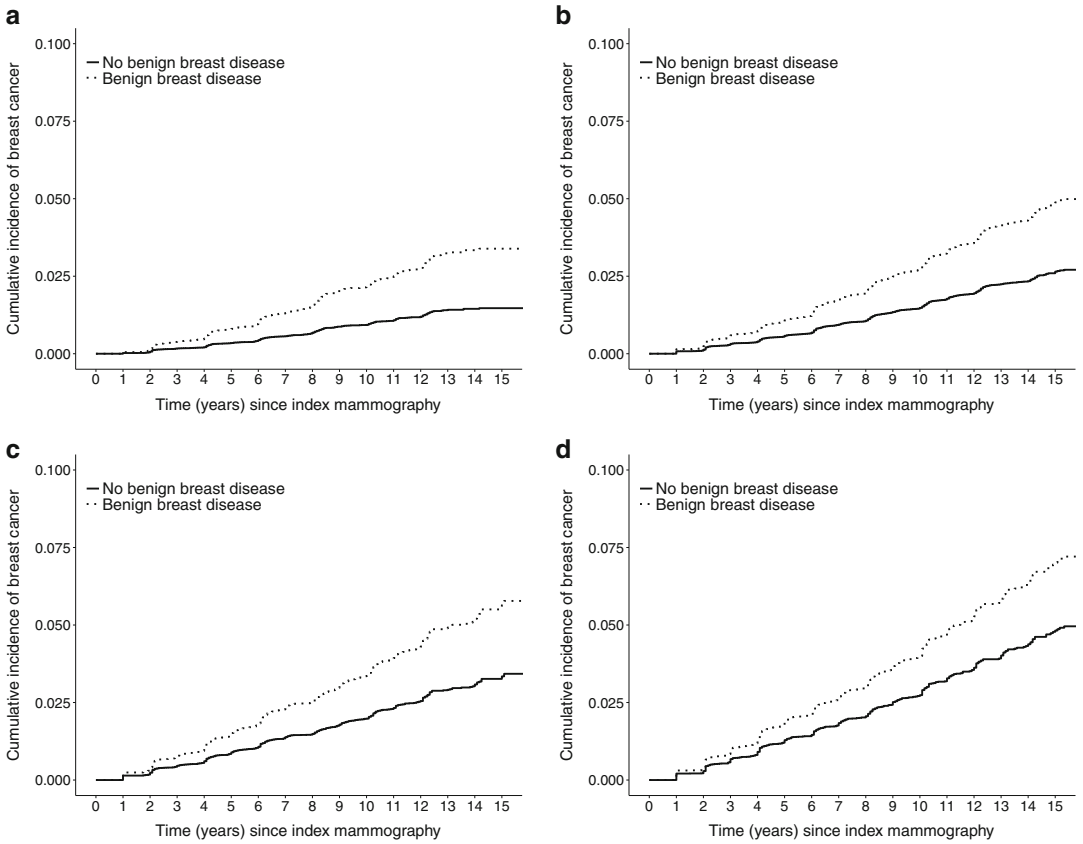


Fig. 1 Adjusted survival curves for each mammographic density strata based on Cox proportional hazards model for women with and without benign breast disease. Models are adjusted for age at the first screen and the year of screening. The solid line represents women without benign breast disease; the dashed line represents women with a benign breast disease diagnosis

disease [9]. That study also found that mammographic density and benign breast disease were independent risk factors for breast cancer. However, the study population in the abovementioned study consisted solely of women with a benign breast disease, limiting their ability to compare the effect of the presence or absence of a benign breast disease. Despite these differences in study population and referent group, the results of the two studies are consistent.

Previous benign breast disease and mammographic density are risk factors commonly included in most individualized risk prediction models for breast cancer, including the Breast Cancer Surveillance Consortium, and Tyrer-Cuzick models [20, 21]. The finding that both factors are independent risk factors for breast cancer consolidates their utility in risk prediction models. However, although risk prediction models for breast cancer have a high expected to observed ratio, their discrimination accuracy is moderate, with potential for improvement [22]. Women with benign breast disease diagnosed at mammography represent 2.0% of the study population of women screened. Among those, half of them had dense breast (heterogeneous or extremely dense). Although uncommon (1%), these women could benefit from more intense screened strategies or different screening techniques as ultrasound or MRI. On the other hand, 20% of the study population had almost entirely fat breast density and absence of a benign breast disease. These women could be candidates for de-intensified screening strategies with longer screening intervals (3 or 4 years). Results from randomized controlled trials like MyPeBS or Wisdom studies will provide evidence on the effectiveness of personalized strategies [23, 24].

We excluded from the analyses women with a screen-detected cancer diagnosed at prevalent screen. This was done because these women did not have any disease-free follow-up, and they could not contribute time at risk to the estimates. Because high breast density is associated with younger age, and there is a known masking effect of breast density, excluding prevalent cancers detected at mammography could entail that some cancers might have been missed in women with dense breast that would later come out as incident cancers. Thus, our associations may slightly overestimate the true strength of the associations between denser breast and the risk of subsequent breast cancer.

This study has several potential limitations. Breast density was reported by radiologists as part of routine screening practice, and their results are likely to be less precise than they would have been if an automated density measure had been used. The inter-rater agreement of the BI-RADS breast density classification is moderate in most studies [25, 26], even in our study population [27]. Despite the moderate inter-rater agreement, our results are consistent with those published by other studies using BI-RADS breast density [4, 28, 29]. Breast density was classified by highly trained radiologists with more than 1000 screening mammograms read per year. Another limitation of the study is that criteria for density classification in the BI-RADS Atlas have changed over time and women who might previously have been classified as BI-RADS density category B are now defined as BI-RADS C if there is any dense area that could mask a tumor [30]. Nevertheless, the BI-RADS classification has been shown to appropriately discriminate women at different risks for breast cancer, with a four-fold gradient in risk between BI-RADS categories A and D [31]. Also, there was a transition from film mammography to digital mammography during the study period. Screen-film mammography was the default technique at start-up. Full-field digital mammography was introduced in 2004 and gradually became widespread. In 2015, 59% of mammography examinations were performed with full-field digital mammography. The transition from film to digital mammography might have had an effect on mammographic density classification and the detection rate, although previous studies have shown that these outcomes were not affected by the introduction of digital mammography [32, 33]. To offset the effect of changes in BI-RADS density classification and the transition from screen-film to digital mammography, our analyses were adjusted by year of screen, but the effect of the adjustment on the risk estimates was minimal. Another limitation was that the number of breast cancer cases after diagnosis of a benign breast disease was small, particularly those associated with fatty breast and extremely dense breast, which limited our ability to perform analyses by subtype of benign breast disease, or other sub-group analyses. Analyses by subtype of benign breast disease would have been desirable as previous studies have shown that breast cancer risk varies by subtype [5, 9, 19]. Lastly, individual information on risk factors such as body mass index, hormone therapy use, and menopausal status were lacking because the screening centers did not collect these information. These factors are known to be associated with mammographic density and with breast cancer risk. Adjusting for these and other confounding factors would have been desirable and could have refined our estimates.

A major strength of this study is that the data were obtained from a well-established population-based screening program with an average participation rate of 67% of invited women, and a re-attendance rate of 91.2% [12]. We analyzed information obtained from over 15 years of follow-up, which guaranteed sufficient time to provide robust estimates. This is the first large longitudinal study to examine the effect of previous benign breast disease and mammographic density in population-based mammography screening.

Conclusion

In summary, we found that benign breast disease and high mammographic density independently predicted the risk of breast cancer. The risk of breast cancer was lowest in women whose breasts were almost entirely fat and highest in women with benign breast disease and extremely dense breasts. The risk remained elevated over 15 years. This information could be used when discussing the potential benefits and harms of personalized screening strategies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00330-020-07490-5.

Acknowledgements The authors acknowledge the dedication and support of the entire BELE (Benign Lesion) and IRIS (Individualized Risk) Study Groups, listed here in alphabetical order and grouped by institution: (a) IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain: Rodrigo Alcantara, Andrea Burón, Xavier Castells, Laura Comerma, Laia Domingo, Javier Louro, Margarita Posso, Marta Román, Maria Sala, Ignasi Tusquets, Ivonne Vazquez, Mar Vernet-Tomas; (b) Corporació Sanitària Parc Taulí, Sabadell, Spain: Marisa Baré, Javier del Riego; (c) Catalan Institute of Oncology, Barcelona, Spain: Llucia Benito, Carmen Vidal; (d) Hospital Santa Caterina, Girona, Spain: Joana Ferrer; (e) Catalan Institute of Oncology, Girona, Spain: Rafael Marcos-Gragera; (f) Hospital de la Santa Creu i Sant Pau, Barcelona, Spain: María Jesús Quintana, Judit Solà-Roca; (g) General Directorate of Public Health, Government of Cantabria, Spain: Mar Sánchez; (h) Principality of Astúrias Health Service, Spain: Miguel Prieto; (i) Fundació Lliga per a La Investigació i Prevenció Del Cáncer, Tarragona, Spain: Francina Saladié, Jaume Galceran; (j) Hospital Clinic, Barcelona, Spain; Xavier Bargalló, Isabel Torá-Rocamora; (k) Vallés Oriental Breast Cancer Early Detection Program, Spain; Lupe Peñalva; (l) Catalonian Cancer Strategy, Barcelona, Spain: Josep Alfons Espinàs.

Funding This work was supported by grants from Instituto de Salud Carlos III FEDER, (PI15/00098) and (PI17/00047), and from the Network for Research into Healthcare in Chronic Diseases, REDISECC (RD12/0001/0015).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Marta Román (mroman@parcdesalutmar.cat).

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise (Marta Román).

Informed consent Written informed consent was waived by the Institutional Review Board. Informed consent was not required since we used anonymized retrospective data

Ethical approval Data was obtained from the databases of the participating screening centres. The review boards of the institutions providing data (Costa de Ponent, Vallés Oriental, Sabadell-Cerdanyola, and Cantabria) granted approval for data analyses.

Study subjects or cohorts overlap Some study subjects or cohorts have been previously reported in:

Román M, Sala M, Baré M, Posso M, Vidal C, Louro J, Sánchez M, Peñalva L, Castells X, on behalf of the BELE study group. Changes in mammographic density over time and the risk of breast cancer: an observational cohort study. Breast. 2019; 46:108–115. https://doi.org/10.1016/j.breast.2019.04.007.

Posso M, Louro J, Sánchez M, Román M, Vidal C, Sala M, Baré M, Castells X; BELE study group. Mammographic breast density: How it affects performance indicators in screening programmes?. Eur J Radiol. 2019; 110:81–87. doi: 10.1016/j.ejrad.2018.11.012.

Román M, Hofvind S, von Euler-Chelpin M, Castells X. Long-term risk of screen-detected and interval breast cancer after false-positive results at mammography screening: Joint analysis of three national cohorts. Br J Cancer. 2019; 120(2):269–275. doi: 10.1038/s41416-018-0358-5.

Román M, Quintana MJ, Ferrer J, Sala M, Castells X, on behalf of the BELE study group. Cumulative risk of breast cancer screening outcomes according to the presence of previous benign breast disease and family history of breast cancer: supporting personalised screening. Br J Cancer 2017; 116(11):1480–1485. doi: 10.1038/bjc.2017.107.

Castells X, Torá-Rocamora I, Posso M, Román M, Vernet-Tomas M, Rodríguez-Arana A, Domingo L, Vidal C, Baré M, Ferrer J, Quintana MJ, Sánchez M, Natal C, Espinàs JA, Saladié F, Sala M. Risk of Breast cancer in Women with False-Positive Results according to the Mammographic Features. Radiology 2016; 280(2): 379–386. doi: 10.1148/radiol.2016151174.

Methodology
- retrospective
- observational
- multicenter study

References

1. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ (1998) Mammographic densities and breast cancer risk. Cancer Epidemiol Biomark Prev 7(12):1133–1144
2. Wolfe JN, Saftlas AF, Salane M (1987) Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study. AJR Am J Roentgenol 148(6): 1087–1092
3. Byrne C, Schairer C, Wolfe J et al (1995) Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst 87(21):1622–1629
4. McCormack VA, dos Santos Silva I (2006) Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomark Prev 15(6):1159–1169
5. Castells X, Domingo L, Corominas JM et al (2015) 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 149(1):237–244
6. Schousboe JT, Kerlikowske K, Loh A, Cummings SR (2011) Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. Ann Intern Med 155(1):10–20
7. Vilaprinyo E, Forne C, Carles M et al (2014) Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. PLoS One 9(2):e86858
8. Byrne C, Schairer C, Brinton LA et al (2001) Effects of mammographic density and benign breast disease on breast cancer risk (United States). Cancer Causes Control 12(2):103–110
9. Tice JA, O'Meara ES, Weaver DL, Vachon C, Ballard-Barbash R, Kerlikowske K (2013) Benign breast disease, mammographic breast density, and the risk of breast cancer. J Natl Cancer Inst 105(14):1043–1049
10. Perry N, Broeders M, de Wolf C, Trnberg S, Holland R, von Karsa L (2006) European guidelines for quality assurance in breast cancer screening and diagnosis. Available from: http://screening.iarc.fr/doc/ND7306954ENC_002.pdf
11. Castells X, Sala M, Ascunce N, Salas D, Zubizarreta R, Casamitjana M (2007) 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 Mèiques de Cataluña; 2006. Available from: http://aunets.isciii.es/ficherosproductos/70/AATRM%202006-01.pdf
12. Ascunce N, Salas D, Zubizarreta R, Almazan R, Ibanez J, Ederra M (2010) Cancer screening in Spain. Annonc Suppl 21(Suppl 3): iii43–iii51
13. Baré M, Bonfill X, Andreu X (2006) Relationship between the method of detection and prognostic factors for breast cancer in a community with a screening programme. J Med Screen 13(4):183–191
14. Sickles EA, D'Orsi CJ, Bassett LW (2013) ACR BI-RADS® Mammography. In: ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. American College of Radiology, Reston, VA
15. Zheng Y, Heagerty PJ (2005) Partly conditional survival models for longitudinal data. Biometrics. 61(2):379–391
16. Tice JA, Cummings SR, Ziv E, Kerlikowske K (2005) Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. Breast Cancer Res Treat 94(2):115–122
17. Vacek PM, Geller BM (2004) A prospective study of breast cancer risk using routine mammographic breast density measurements. Cancer Epidemiol Biomark Prev 13(5):715–722
18. Ziv E, Shepherd J, Smith-Bindman R, Kerlikowske K (2003) Mammographic breast density and family history of breast cancer. J Natl Cancer Inst 95(7):556–558
19. Hartmann LC, Sellers TA, Frost MH et al (2005) Benign breast disease and the risk of breast cancer. N Engl J Med 353(3):229–237
20. Tice JA, Miglioretti DL, Li C-S, Vachon CM, Gard CC, Kerlikowske K (2015) Breast density and benign breast disease: risk assessment to identify women at high risk of breast cancer. J Clin Oncol 33(28):3137–3143
21. Brentnall AR, Harkness EF, Astley SM et al (2015) Mammographic density adds accuracy to both the Tyrer-Cuzick

and Gail breast cancer risk models in a prospective UK screening cohort. Breast Cancer Res 17(1):147
22. Brinton JT, Hendrick RE, Ringham BM, Kriege M, Glueck DH (2019) Improving the diagnostic accuracy of a stratified screening strategy by identifying the optimal risk cutoff. Cancer Causes Control 30(10):1145–1155
23. MyPeBs. Randomized comparison of risk-stratified versus standard breast cancer screening in European women aged 40–70 (MyPeBS). 2017.
24. Esserman LJ, WISDOM Study and Athena Investigators (2017) The WISDOM Study: breaking the deadlock in the breast cancer screening debate. NPJ Breast Cancer 3:34
25. Ciatto S, Houssami N, Apruzzese A et al (2005) Categorizing breast mammographic density: intra- and interobserver reproducibility of BI-RADS density categories. Breast Edinb Scotl 14(4):269–275
26. Kerlikowske K, Grady D, Barclay J et al (1998) Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. J Natl Cancer Inst 90(23):1801–1809
27. Redondo A, Comas M, Macia F et al (2012) Inter- and intraradiologist variability in the BI-RADS assessment and breast density categories for screening mammograms. Br J Radiol 85(1019):1465–1470
28. Kerlikowske K, Ichikawa L, Miglioretti DL et al (2007) Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. J Natl Cancer Inst 99(5):386–395
29. Sartor H, Lang K, Rosso A, Borgquist S, Zackrisson S, Timberg P (2016) Measuring mammographic density: comparing a fully automated volumetric assessment versus European radiologists' qualitative classification. Eur Radiol 26(12):4354–4360
30. Irshad A, Leddy R, Lewis M et al (2017) Changes in breast density reporting patterns of radiologists after publication of the 5th Edition BI-RADS guidelines: a single institution experience. AJR Am J Roentgenol 209(4):943–948
31. Cummings SR, Tice JA, Bauer S et al (2009) Prevention of breast cancer in postmenopausal women: approaches to estimating and reducing risk. J Natl Cancer Inst 101(6):384–398
32. Sala M, Salas D, Belvis F et al (2011) Reduction in false-positive results after introduction of digital mammography: analysis from four population-based breast cancer screening programs in Spain. Radiology. 258(2):388–395
33. Harvey JA, Gard CC, Miglioretti DL et al (2013) Reported mammographic density: film-screen versus digital acquisition. Radiology. 266(3):752–758

Article 4

Title: A systematic review and quality assessment of individualized breast cancer risk prediction models

Authors: Louro J, Posso M, Hilton Boon M, Roman M, Domingo L, Castells X & Sala M

Journal: British Journal of Cancer; 121(1):76-85

Impact Factor: 5.79 (Q1 Oncology)

DOI: 10.1038/s41416-019-0476-8

Abstract:

Background: Individualized breast cancer risk prediction models may be key for planning risk-based screening approaches. Our aim was to conduct a systematic review and quality assessment of these models addressed to women in the general population.

Methods: We followed the Cochrane Collaboration methods searching in Medline, EMBASE and The Cochrane Library databases up to February 2018. We included studies reporting a model to estimate the individualized risk of breast cancer in women in the general population. Study quality was assessed by two independent reviewers. Results are narratively summarised.

Results: We included 24 studies out of the 2,976 citations initially retrieved. Twenty studies were based on four models, the Breast Cancer Risk Assessment Tool (BCRAT), the Breast Cancer Surveillance Consortium (BCSC), the Rosner & Colditz model, and the International Breast Cancer Intervention Study (IBIS), whereas four studies addressed other original models. Four of the studies included genetic information. The quality of the studies was moderate with some

limitations in the discriminative power and data inputs. A maximum AUC value of 0.71 was reported in the study conducted in a screening context.

Conclusion: Individualized risk prediction models are promising tools for implementing risk-based screening policies. However, it is a challenge to recommend any of them since they need further improvement in their quality and discriminatory capacity.

BJC British Journal of Cancer

www.nature.com/bjc

ARTICLE

Epidemiology

A systematic review and quality assessment of individualised breast cancer risk prediction models

Javier Louro[1,2,3], Margarita Posso[1,2], Michele Hilton Boon[4], Marta Román[1,2], Laia Domingo[1,2], Xavier Castells[1,2] and María Sala[1,2]

BACKGROUND: Individualised breast cancer risk prediction models may be key for planning risk-based screening approaches. Our aim was to conduct a systematic review and quality assessment of these models addressed to women in the general population.
METHODS: We followed the Cochrane Collaboration methods searching in Medline, EMBASE and The Cochrane Library databases up to February 2018. We included studies reporting a model to estimate the individualised risk of breast cancer in women in the general population. Study quality was assessed by two independent reviewers. Results are narratively summarised.
RESULTS: We included 24 studies out of the 2976 citations initially retrieved. Twenty studies were based on four models, the Breast Cancer Risk Assessment Tool (BCRAT), the Breast Cancer Surveillance Consortium (BCSC), the Rosner & Colditz model, and the International Breast Cancer Intervention Study (IBIS), whereas four studies addressed other original models. Four of the studies included genetic information. The quality of the studies was moderate with some limitations in the discriminative power and data inputs. A maximum AUROC value of 0.71 was reported in the study conducted in a screening context.
CONCLUSION: Individualised risk prediction models are promising tools for implementing risk-based screening policies. However, it is a challenge to recommend any of them since they need further improvement in their quality and discriminatory capacity.

British Journal of Cancer https://doi.org/10.1038/s41416-019-0476-8

BACKGROUND

Mammography screening has been associated with a reduction in breast cancer mortality and therefore organised breast cancer screening programmes using mammography have been well established worldwide.[1–4] Although there is not a single consensus, current screening programmes generally recommend biennial or triennial screening in Europe and annual or biennial screening in the US with variations in the recommended targeted age.[2–5] These recommendations usually consider age as the sole risk factor leading women to be invited for screening from age 40–50 until age 70–74, depending on the programmes.

The likelihood that a woman will benefit from screening mammography depends on her risk for developing clinically significant breast cancer in her lifetime. Taking individual risk factors beyond age into account should enable the classification of women into groups at varying risk of breast cancer. Personalised risk-based screening going beyond the current 'one-size fits all' recommendation may increase the effectiveness and benefit-harm balance of breast cancer screening. Individualised risk prediction models for breast cancer are a key element to develop risk-based screening approaches since they are designed to quantify the risk that can predict whether an individual woman would develop breast cancer in a defined period.[6]

A number of risk prediction models that include classical risk factors are commonly used in clinical contexts.[7] However, organised screening programmes do not use these models routinely. One reason for not including these models in screening context is the high uncertainty with regards to its applicability in screening settings. Also, the emergence of new risk prediction factors such as the expression of single nucleotide polymorphisms (SNPs) needs to be appropriately summarised before recommending one of the models into screening practice.

Like any other source of information, risk prediction models have limitations that should be evaluated before using them. A rigorous risk of bias assessment of the existing individualised risk models is needed to clarify the overall quality and applicability of each model. Therefore, the aim of this systematic review is to update the existing evidence, conduct a critical appraisal and risk of bias assessment and summarise the results of the individualised risk models which are used to estimate the risk of breast cancer in women in the general population.

METHODS

Data sources and searches

We performed a systematic review of the literature following the standard Cochrane Collaboration methods[8] and adhering to the PRISMA statement reporting recommendations.[9] A predetermined review protocol was registered (CRD42018089842) in the PROSPERO database (date of registration 1 March 2018). The Patient,

[1]Department of Epidemiology and Evaluation, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain; [2]Research Network on Health Services in Chronic Diseases (REDISSEC), Barcelona, Spain; [3]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 and [4]MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, Glasgow, UK
Correspondence: Margarita Posso (mposso@parcdesalutmar.cat)

Received: 10 January 2019 Accepted: 25 April 2019
Published online: 22 May 2019

Published by Springer Nature on behalf of Cancer Research UK

Intervention, Comparison, Outcomes (PICO) question of this systematic review is the following: Should individualised breast cancer risk prediction models vs. no risk prediction models be used to develop risk-based screening approaches for women in the general population?

We retrieved relevant literature by using a combination of controlled vocabulary and keyword search terms in the following databases: (i) Medline (accessed through PubMed); (ii) The Cochrane Library; and (iii) EMBASE (accessed through Ovid). Terms related to breast cancer recurrence were excluded in order to avoid retrieving citations out of the scope of this systematic review. We adapted the search algorithms to the requirements of each database and used validated filters to retrieve systematic reviews and primary studies as needed. We reviewed references of included studies that could potentially fulfil our eligibility criteria. The detailed search strategy is reported in Supplementary table 1.

We searched primary studies of individualised breast cancer risk models searching each database from its inception up to February 2018.

Study selection

Eligible studies were those published in English that reported a model to estimate the individualised risk of breast cancer in women in the general population. We included models that assessed more than one risk factor and reported the quantitative characteristics of the risk prediction model. If multiple publications were based on the same individualised risk model, the most extensive report of the model in terms of risk factors reported was chosen. We excluded external validation studies that replicated previous models without adding any additional information such as a new design for collecting the inputs data, modifications on the risk factors or the risk model method.

Articles identified from the search were loaded into EndNote X7.7.1 for Windows (2008, Version 12.0.4) and duplicates were removed.

Data extraction and quality assessment

One reviewer screened the search results based on title and abstract, and a second reviewer performed a quality check of the study screening by reviewing 20% of the references. Two reviewers independently confirmed eligibility based on the full text of the relevant articles. In case of disagreement between researchers, the inclusion of studies was determined by consensus. We reported the result of this process with a PRISMA flowchart (Fig. 1).

We used a predefined form to extract the following information from included studies: author, publication date, country, study design, the name of the model if available, sample characteristics, sample size, type of breast cancer, the method of analysis, and validation of the model. Data abstraction was conducted by one reviewer and checked by another.

Two reviewers carried out the assessment of the risk of bias independently and final quality assessment was based on consensus. We used the ISPOR-AMCP-NPC Questionnaire[10] to assess the relevance and credibility of each risk prediction study and the following sources of limitations: (i) internal and external validation; (ii) bias due to the study design for risk estimates; (iii) limitations in data inputs; (iv) appropriateness of the model analysis; (v) reporting bias; (vi) interpretation bias; and (vii) conflict of interest. The risk of bias for each domain was rated as low, high or unclear. For systematic reviews we used the AMSTAR 2 critical appraisal tool.[11]

Data synthesis and analysis

We evaluated the model validation by assessing both the discriminative power and the calibration accuracy estimated for the women in the general population. When available in the included publication, we extracted the area under the receiver operating characteristic curve (AUROC), the net reclassification index (NRI) and the expected observed (E/O) ratio. The NRI was not included in the tables because it was only reported in 2 out of 24 articles. The characteristics of the included models and the risk prediction outcomes reported preclude the possibility to pool data across studies. Therefore, a narrative synthesis has been conducted. Key study characteristics, validation and accuracy of individual risk models, and methodological quality are described in tables and summarised in a narrative manner. Results are presented according to the original model that they reported.

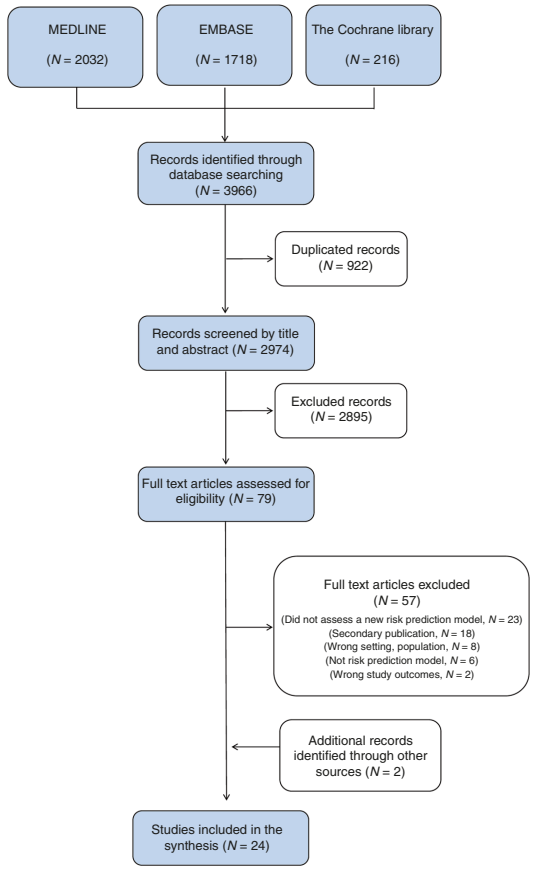


Fig. 1 PRISMA flowchart

RESULTS

Study inclusion

The database searches for primary studies retrieved 2974 citations, of which 79 were considered potentially relevant. These 79 studies were screened in full text. We found a systematic review of Anothaisintawee et al.,[7] which we used as a source of primary studies. In addition, two studies were included after a manual inspection of papers' references.[12,13] After the full text was checked, 24 studies[12–35] met the inclusion criteria and were considered in the evidence synthesis. Details about study inclusion with reasons for exclusion are described in the flowchart (Fig. 1), and a list of references to excluded studies is provided in Supplementary table 2.

Characteristics of the included studies

The included studies can be grouped according to the risk model that they reported, the Breast Cancer Risk Assessment Tool (BCRAT), the Breast Cancer Surveillance Consortium (BCSC), the Rosner & Colditz model, the International Breast Cancer Intervention Study (IBIS), and other original models. The study by Zhang et al.[13] is included in two of the groups (BCRAT and Rosner & Colditz models) because it provides information of both models and presents its results separately. A brief summary of the 24 included studies is presented in Table 1 and the extended characteristics in Supplementary table 3.

a. Breast Cancer Risk Assessment Tool 'BCRAT' model. This model was first published in the United States in 1989 assessing age, family history of breast cancer, age at first birth, menarche, and previous biopsies as risk factors for predicting individualised breast cancer risk.[22] After this first publication, eight studies were identified that were based on BRCAT model but modified the data collection design, assessed additional risk factors or changed the statistical method. In addition to the five risk factors proposed in 1989, other variables such as body mass index (BMI), weight, hormone replacement therapy (HRT), alcohol consumption, physical activity, diet, breast density, atypical hyperplasia, breast inflammatory disease, parity, a polygenic risk score or hormones information have been included in updated versions (Table 1).[13,14,16,17,20,23,25,26,30]
b. Breast Cancer Surveillance Consortium 'BCSC' model. One relevant variation of the BCRAT model opens the path to the emergence of the BCSC model first published by Tice et al. in 2008 in the United States.[31] In this study, Tice et al. used data from a cohort to create an individualised risk prediction model that combines age, family history, previous biopsies, breast density, and ethnicity. The BCSC model has been further evaluated by other authors[12,24,29,32] and it currently includes previous benign breast diseases and polygenetic risk score using SNPs as risk factors (Table 1).
c. Rosner & Colditz model. Parallel to the BCSC model, another model based on the 'Nurses' Health Study' cohort developed by Rosner & Colditz in 1996 was also developed in the United States. This model currently includes 11 risk factors: age, menarche, menopause, age at first birth, age at subsequent births, previous benign breast disease, HRT, family history, weight, BMI, alcohol consumption, and oestradiol levels.[18,19,27,28] In the same way as in the BCRAT, Zhang et al.[13] analysed this model adding breast density, a polygenic risk score and endogenous hormones as risk factors.
d. International Breast Cancer Intervention Study 'IBIS' model. The IBIS model[33] includes genetic information adding the BRCA genes and a hypothetical susceptibility gene.
e. Other models. Four studies reporting different models were also identified.[15,21,34,35] Apart from the above-mentioned risk factors, the models also assessed other variables such as abortion, breastfeeding, height, and previous mammography results. Particularly relevant is the Eriksson model[21] since it was the only one targeted to the screening population. In this study, the authors included risk factors that were available at mammography screening examination: age, BMI, HRT, family history, menopause, breast density, and presence of microcalcifications and/or masses in the screen-mammogram.

Discriminatory accuracy

Fifteen out of the 24 studies reported the discriminatory accuracy as the AUROC (Table 1 and Fig. 2).

a. BCRAT model. The first BCRAT model publication did not report the AUROC, however, later publications of this model reported a range that varied from 0.56 to 0.68. The three publications that included the original risk factors, age, family history of breast cancer, age at first birth, menarche, and previous biopsies, reported low AUROC values, 0.56 to 0.62.[14,20,23] Similarly, the AUROC reported by Boyle et al.[16] and Matsuno et al.[25] were 0.60 and 0.61, although these authors added BMI, HRT, alcohol, physical activity and diet, and ethnicity into the model. Zhang et al.[13] with the new variables reach an AUROC of 0.65 and Tice et al.[30] reported in 2005 a higher AUROC value of 0.68 which was obtained just adding breast density to the original five risk factors (Table 1). Zhang et al.[13] also reported the NRI to validate that his model improved the previous ones with a result of 8%.
b. BCSC model. The published value of the AUROC for the BCSC model was moderate, ranging from 0.64 to 0.69. Tice et al. included age, family history, previous biopsies, breast density reported by the Breast Imaging Reporting and Data System (BI-RADS), and ethnicity into the model in 2008 and obtained a value of 0.66 for the AUROC.[31] Instead of BI-RADS, Kerlikowske et al. assessed changes in breast density obtaining a similar result, 0.64.[24] Using previous benign breast disease, Tice et al. obtained a slightly higher AUROC value of 0.67 in 2015.[32] More recently, in 2015 and 2016, Vachon et al.[12] added to the model a polygenic risk score and Shieh et al.[29] a combination between a polygenic risk score and BMI reporting a value of 0.69 and 0.65 for the AUROC respectively (Table 1). Vachon et al.[12] also demonstrated the improvement of discriminatory accuracy estimating the NRI with a positive result of 11%.
c. Rosner & Colditz model. The discriminatory accuracy of this model varied from 0.61 to 0.68. The authors assessed age, family history, age at first birth, menarche, BMI, benign breast disease, menopause, HRT, age at subsequent births, alcohol, and weight. They obtained an AUROC of 0.64 and 0.61 for ER + /PR + and ER-/PR- tumours, respectively.[19] The addition of oestradiol levels to the model was tested by Rosner et al. who obtained a 0.65 AUROC value in 2008.[28] Finally the addition of a polygenic risk score, mammographic density and endogenous hormones by Zhang et al.[13] reached a 0.68 AUROC value (Table 1) and obtained an improvement of the discriminative accuracy also reflected in a NRI of a 9.5%.
d. IBIS model. The IBIS model original paper[33] does not include any validation and does not present the AUROC. Nevertheless, it has been externally validated showing an AUROC of 0.57 which increases to 0.61 when adding mammographic density.[36]
e. Other models. Overall, the AUROC values of these models were not higher than those shown by the above-mentioned models, varying from 0.62 to 0.64, although they included a large number of risk factors. However, the model reported by Eriksson et al.[21] did show an AUROC of 0.71 that was the highest AUROC value identified in this systematic review (Table 1). This model, in addition, is the only one that estimates a 2-year risk, while the rest of models estimate the risk at a longer time horizon. This could explain the difference in AUROC values since it becomes more difficult to predict risk as the time horizon increases.

Calibration accuracy

Nine out of the 24 studies reported the calibration accuracy as the E/O ratio (Table 1).

Table 1. Summary of included studies

Study ID	Targeted population	Risk factors (number of categories)	Discriminatory accuracy (AUROC)[a]	Calibration (E/O ratio)[a]
BCRAT model				
Banegas 2017	Hispanic, 25–79 years	Age (2), Menarche (3), Previous biopsies (2), Age at first birth (3), First degree breast cancer (2)	US-born: 0.56 Foreign-born: 0.62	US-born: 0.93[b] Foreign-born: 1.52[b]
Boyle 2004	Caucasian, 20–74 years	Age (n), Menarche (3), Age at first birth (3), First degree breast cancer (2), BMI (3), Alcohol (3), Physical activity (3), HRT (2), Diet beta-carotene/vitE (5), Diet fruits/vegetables (5)	0.6	1.03[b]
Chen 2006	Caucasian, 35–74 years	Age (2), Weight (6), Breast density (5), Menarche (3), Previous biopsies (3), Age at first birth (4), First degree breast cancer (3), Atypical hyperplasia (2)	None	None
Decarli 2006	Caucasian, 20–74 years	Age (2), Menarche (3), Previous biopsies (3), Age at first birth (4), First degree breast cancer (3)	0.59	0.96
Gail 1989	Caucasian, 20–79 years	Age (2), Menarche (3), Previous biopsies (3), Age at first birth (4), First degree breast cancer (3)	None	None
Gail 2007	African-American, 35–64 years	Age (2), Menarche (3), Previous biopsies (3), Age at first birth (4), First degree breast cancer (3)	0.56	0.93[b]
Matsuno 2011	Asian, 20–55 years	Age (2), Menarche (3), Previous biopsies (3), Age at first birth (4), First degree breast cancer (n), Ethnicity (6)	0.61	0.85[b]
Novotny 2006	Multiple ethnicities, 23–84 years	Age (2), Menarche (3), Previous biopsies (3), Age at first birth (4), First degree breast cancer (3), First degree family history of cancer (5), Parity (n), Breast inflammatory disease (2)	None	None
Tice 2005	Multiple ethnicities, older than 35 years	Age (2), Menarche (3), Previous biopsies (3), Age at first birth (4), First degree breast cancer (3), Breast density (4)	0.68	None
Zhang 2018	Caucasian 30–64 years	Age (2), Menarche (3), Previous biopsies (3), Age at first birth (4), First degree breast cancer (2), Polygenic Risk Score(n), Mammographic density (n), Estrone Sulphate (n), Testosterone (n), Prolactin (n)	0.65	None
Breast Cancer Surveillance Consortium 'BCSC' model				
Kerlikowske 2015	Multiple ethnicities, 35–74 years	Age (n), Ethnicity (6), First degree breast cancer (2), Previous biopsies (2), Changes in breast density (16)	0.64	5-years: 0.98 10-years: 0.95
Shieh 2016	Multiple ethnicities. Age was not specified	Age (n), Ethnicity (6), First degree breast cancer (2), Previous biopsies (2), Breast density (4), Polygenetic risk score (n), BMI (n)	0.65	None
Tice 2008	Multiple ethnicities, 35–84 years	Age (n), Ethnicity (4), First degree breast cancer (2), Previous biopsies (2), Breast density (BI-RADS)	0.66	1.03
Tice 2015	Multiple ethnicities, 35–74 years	Age (n), Ethnicity (4), First degree breast cancer (2), Breast density (4), Benign breast disease (6)	0.67	5-years: 1.04 10-years: 1.05
Vachon 2015	Multiple ethnicities. Age was not specified	Age (n), Ethnicity (6), First degree breast cancer (2), Previous biopsies (2), Breast density (4), Polygenetic risk score (*n*)	0.69	None
Rosner & Colditz model based on the 'Nurses' Health Study'				
Colditz 2000	Caucasian, 30–64 years	Age (*n*), Menarche (*n*), Age at first birth (*n*), Menopause (*n*), Age at subsequent births (*n*), Benign breast disease (2), HRT (*n*), First degree breast cancer (2), Weight (*n*), BMI (*n*), Alcohol (*n*)	None	None
Colditz 2004	Caucasian, 30–64 years	Age (*n*), Menarche (*n*), Age at first birth (*n*), Menopause (*n*), Age at subsequent births (*n*), Benign breast disease (2), HRT (*n*), First degree breast cancer (2), Weight (*n*), BMI (*n*), Alcohol (*n*)	ER+/PR+: 0.64 ER−/PR−: 0.61	None
Rosner 1996	Caucasian, 30–64 years	Age (*n*), Menarche (*n*), Age at first birth (*n*), Menopause (*n*), Age at subsequent births (*n*)	None	None
Rosner 2008	Caucasian, 30–64 years	Age (*n*), Menarche (*n*), Age at first birth (*n*), Menopause (*n*), Age at subsequent births (*n*), Benign breast disease (2), HRT (*n*), First degree breast cancer (2), Weight (*n*), BMI (*n*), Alcohol (*n*), Estradiol levels (*n*)	0.65	None
Zhang 2018	Caucasian 30–64 years	Age (*n*), Menarche (*n*), Age at first birth (*n*), Menopause (*n*), Age at subsequent births (*n*), Benign breast disease (2), HRT (*n*), First degree breast cancer (2), Weight (*n*), BMI (*n*), Alcohol (*n*), Polygenic Risk Score(*n*), Mammographic density (*n*), Estrone Sulphate (*n*), Testosterone (*n*), Prolactin (*n*)	0.68	None

Table 1 continued

Study ID	Targeted population	Risk factors (number of categories)	Discriminatory accuracy (AUROC)[a]	Calibration (E/O ratio)[a]
IBIS model				
Tyrer 2004	Multiple ethnicities. Age was not specified	Age (*n*), Gen phenotype (6), Family history (*n*, relationship, age), Menarche (*n*), Age at first birth (5), Menopause (*n*), Atypical Hyperplasia (2), Lobular carcinoma in situ (2), Height (3), BMI (5)	None	None
Other original models				
Barlow 2006	Multiple ethnicities, 35 to 84 years	Age (9), Age at first birth (4), Ethnicity (6), Menopause (2), First degree breast cancer (5), Previous biopsies (4), Breast density (5), HRT (3), BMI (5), Previous false positive or true negative screen result (2), Menopausal status (3)	Pre menopause: 0.63 Post menopause: 0.62	Pre menopause: 1.00 Post menopause: 1.01
Eriksson 2017	Caucasian, 40–74 years	Age (7), BMI (n), HRT (2), Breast cancer family history (2), Menopause (2), Breast density (4), Microcalcifications (5), Mases (n)	0.71	None
Ueda 2003	Asian women. Age was not specified.	Age (n), Menarche (3), Age at first birth (5), BMI (2), Breast cancer family history (2)	None	None
Tyrer 2004	Multiple ethnicities. Age was not specified	Age (*n*), Gen phenotype (6), Family history (*n*, relationship, age), Menarche (*n*), Age at first birth (5), Menopause (*n*), Atypical Hyperplasia (2), Lobular carcinoma in situ (2), Height (3), BMI (5)	None	None
Wang 2014	Asian, 35–70 years	Age (7), Menarche (2), Previous biopsies (2), Age at first birth (2), First degree breast cancer (2), Breastfeeding (2), Abortion (2)	0.64	None

[a]Discriminatory and calibration accuracy values represents the statistics published in the original articles for the general population Subgroup values are not reported here
[b]The original publication reported the Observed/Expected ratio. E/O ratios were calculated based on the original information. *AUROC* area under the receiver operating characteristic curve. *E/O* expected/observed, *BMI* body mass index, *HRT* hormone replace treatment, *ER* oestrogen receptor, *PR* progesterone receptor, *BI-RADS* Breast Imaging Reporting and Data System. (n): continuous variable in model

a. BCRAT model. Of the 10 studies derived of the BCRAT model, five reported the calibration accuracy. Banegas et al.[14] presented heterogeneous results depending on the provenance of the population, reporting an E/O ratio of 0.93 for US-born and 1.52 for foreign-born women. Although Matsuno et al.[25] added new variables to the original BCRAT model, the E/O ratio was 0.85, which was the lowest of the group, whereas the other studies published E/O ratios that varied from 0.93 to 1.03[16,20,23] (Table 1).
b. BCSC model. Tice et al. published in 2008 a value of 1.03 for the E/O ratio when looking at 5-year risk.[31] Using previous benign breast disease, they obtained a similar result in 2015, with an E/O ratio of 1.04 for 5-year risk and 1.05 for 10-year risk.[32] When Kerlikowske et al. assessed changes in breast density the ratio decreased obtaining a 0.98 for 5-year risk and 0.95 for 10-year risk.[24] The studies of Vachon et al. and Shieh et al. did not present validation regarding the calibration accuracy of the model (Table 1).
c. Rosner & Colditz model. Of the five studies based on the Rosner & Colditz model,[13,18,19,27,28] none of them reported calibration accuracy statistics of their models for the women in the general population.
d. IBIS model. The IBIS model original paper[33] does not report any calibration statistic. Nevertheless, other articles have validated it showing an E/O ratio of 1.67.[36]
e. Other models. The study Barlow et al.[15] was the only one that reported calibration accuracy and presented the closest E/O ratio to one of all the studies included in this review taking values of 1.00 and 1.01 for pre and post-menopausal status respectively (Table 1).

Quality assessment

The quality of the included studies was moderate due to some limitations in the discriminative power, study design, and data inputs. The studies did not show important limitations with regards to the validation, appropriateness of the model analysis, reporting or interpretation of the results (Fig. 3). A summary of the risk of bias assessment per each source of limitation is presented here and the detailed appraisal and judgements in Supplementary table 4.

Internal and external validation

Ten studies[14–17,20,23,25,26,30,31] validated their models by comparing the results with those published by Gail et al.,[22] three studies[24,29,32] compared with Tice et al.,[31] one[21] compared with both Gail et al.[22] and Tyrer et al.,[33] one[13] compared with both Gail et al.[22] and the results of a Rosner & Colditz model external validation[37] and three studies did not report the model validation in the primary articles.[19,22,34] Six studies assessed internal validation with a sample of the population that generated data for the model,[15,16,24,29,31,32] and four with an external population.[14,20,23,25] Despite not having reported the external validation in the primary articles, the Rosner & Colditz model[18,19,27,28] reported external validation in a subsequent article mentioned before.[37] Nine studies used the expected/observed event ratio to measure the calibration accuracy of the model.[14–16,20,23–25,29,31]

Bias due to the study design

Thirteen studies used a case-control design to obtain breast cancer risk estimates,[12–14,16,17,20–23,25,26,29,34] five studies used prospective cohorts,[15,18,19,27,28] and four models used retrospective cohorts.[24,30–32] The study of Wang et al.[35] and the study of Tyrer et al.[33] used risk estimates obtained from a systematic review of the literature.

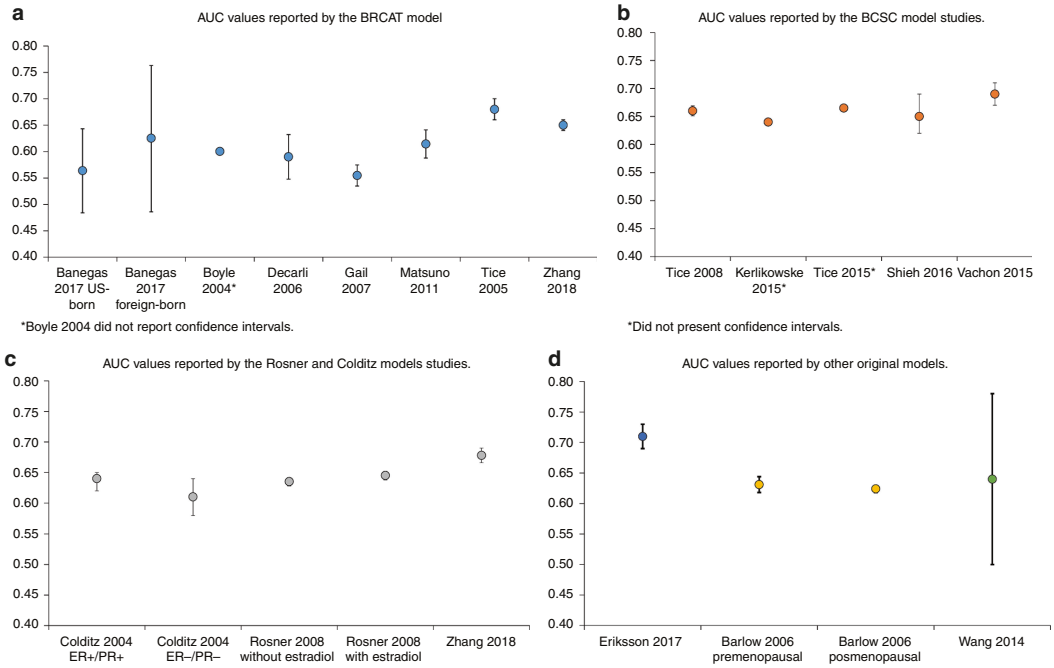


Fig. 2 Area under the ROC curve (AUROC) and Confidence Intervals reported by the included studies. **a** AUROC values reported by the BCRAT model studies. **b** AUROC values reported by the BCSC model studies. **c** AUROC values reported by the Rosner & Colditz model studies. **d** AUROC values reported by other original models

Limitations of data inputs

Sixteen studies obtained most of the input parameters from self-reported questionnaires.[13–20,22,23,25–28,30,34] The study of Matsuno et al.[25] also imputed ethnicity for women with missing data.

Appropriateness of the model analysis

Thirteen studies[12–17,20,22,23,25,26,29,34] used logistic regression to estimate the risk of having breast cancer according to the assessed risk factors, five used proportional hazard Cox models,[21,24,30–32] four used Poisson regression models,[18,19,27,28] and the other two studies used risk estimates obtained from a systematic review of the literature.[33,35]

Reporting bias

Twenty one studies reported all relevant and necessary information for the model creation.[12–23,25–29,31,33–35] Conversely, a critical lack of information was found in the other three studies.[24,30,32]

DISCUSSION

Summary of main results

This systematic review included 24 studies that aimed to estimate the individual risk of developing breast cancer in women in the general population. Twenty studies were based on four specific risk models (the BCRAT, the BCSC, the Rosner & Colditz and the IBIS model),[16–20,22–33] whereas four studies used other original models.[15,21,34,35] The most extensively used were the BCRAT, IBIS and the BCSC models. The number of risk factors included in the models ranged from five to 18. Other than age, which was the only risk factor present in all models, the BCRAT model also included family history, age at first birth, menarche, and previous biopsies. Breast density, benign breast disease, and polygenetic score were predominant in the BCSC model. Although during the last decade the models have shown improvements in their discriminatory accuracy, it remains at best moderate with a maximum AUROC value of 0.71 reported by Eriksson et al.[21] The calibration accuracy was very heterogeneous ranging from 0.85 to 1.52. Furthermore, the quality of the studies was not high due to limitations in the discriminative accuracy, study design, and data inputs.

Agreements and disagreements with other reviews

In this systematic review, we found that the number of individualised breast cancer risk prediction models has increased steadily over the past three decades. This finding is in agreement with the narrative overview published by Cintolo-Gonzalez et al. in 2017,[38] and it updates the results of a previous systematic review published by Anothaisintawee et al. in 2012.[7] In contrast to these reviews, however, our aim was to provide innovative information regarding the quality of the identified prediction models. Thus, we have identified and rigorously analysed the strengths and limitations of 24 individualised models in order to adjust our conclusions to the quality of the evidence.

We have identified two new trends with regards to the use and development of the models, which are the increased use of the BCSC model and the inclusion of common genetic variation in the prediction models. As compared to the information published in the review of Anothaisintawee et al.,[7] we found that in contrast to the BCRAT and Rosner & Colditz models that were the most frequently cited models up to 2010[7] the BCSC model has concentrated the attention of several authors during the last five years, although its discriminatory accuracy has not dramatically

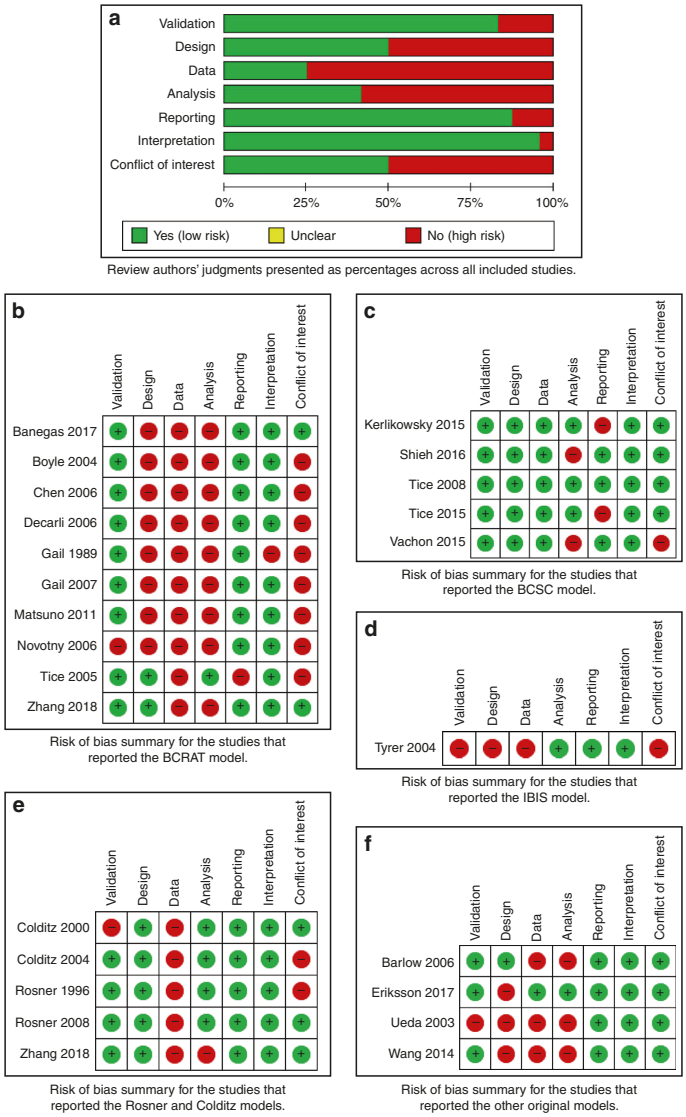


Fig. 3 Risk of bias summary: review authors' judgments about each risk of bias item for each included study. **a** Review authors' judgments presented as percentages across all included studies. **b** Risk of bias summary for the studies that reported the BCRAT model studies. **c** Risk of bias summary for the studies that reported the BCSC model studies. **d** Risk of bias summary for the studies that reported the Rosner & Colditz model studies. **e** Risk of bias summary for the studies that reported other original models

improved. Second, none of the models in the review of Anothaisintawee et al.[7] included genetic information as a risk factor. By contrast, we have identified four models including genetic information: the IBIS model[33] that includes genetic phenotype in their updated version, the BCSC model that includes a polygenetic score in both 2015[12] and 2016[29] publications, as well as the article by Zhang et al. that added a polygenic risk score to both the BCRAT and the Rosner & Colditz models.[13]

Most of the included studies reported the AUROC to determine the probability that a randomly chosen woman with disease would be correctly categorised as higher risk compared to a randomly chosen woman without disease. The discriminatory accuracy estimate does not express whether the model is more or less accurate in predicting the risk of specific individuals but measures the capacity of the model to determine which women are at higher/lower risk for developing breast cancer. Thus, both calibration accuracy and discriminatory accuracy should be assessed. Contrary to what is expected, we found that authors reported the E/O ratio only in less than half of the included studies. In addition to the AUROC value, the studies of Zhang et al.

and Vachon et al.[12,13] also reported an improvement in the net reclassification index (NRI) of the BCRAT, and Rosner & Colditz models, as well as in the BCSC model, respectively.

Overall, the information provided by the AUROC and the E/O ratio was consistent suggesting that the included models have moderate discriminatory accuracy and calibration accuracy when applied to the women in the general population. Nevertheless, it must be taken into account that despite the great importance of validation in terms of AUROC and E/O ratio, the presence of low values of AUROC or clearly different from 1 values of the E/O ratio does not mean that these models are useless. On the contrary, models are clinically useful even with moderate AUROC since they can reclassify individuals at the extremes of risk.[39] Thus, the verdict on risk models should not be based solely on these estimators. Instead, they need to be prospectively evaluated in clinical trials. In fact, there are currently two very large randomised trials assessing risk-based screening strategies. Both of them are using individualised models. Both the IBIS and the BCSC models are being tested in the European trial MyPeBS (My Personalised Breast Screening).[40] Also, the BCSC model is being tested in the US WISDOM trial (Women Informed to Screen Depending On Measures of risk).[41]

Applicability and completeness of evidence

The distribution of risk factors in such different populations may affect the applicability of the models to different contexts. The fact that different subtypes of breast cancer may have different genetic markers is widely accepted.[42] These differences, the nature of breast cancer itself and its low incidence may condition a low discriminatory accuracy of a model. In other words, in the general population, there is a low probability of having breast cancer (even in the highest risk group). This low probability may mean that the discriminatory power of a breast cancer risk model won't be as high as a risk model targeted to other common diseases such as cardiovascular events, for instance. Another potential limitation in the applicability in the screening context is the completeness and the number of included risk factors, which ranged from five to 18. Nevertheless, some potentially relevant risk factors such as genetic markers have been only included in few models. Recent studies[43,44] have shown that adding genetic information as a risk factor can increase the discriminative accuracy of the different models which opens the line for further evaluation. An evaluation that should first assess the calibration of these models in prospective cohort studies.

Overall, women are usually screened using mammography. Particularly in Europe, most programmes invite women for screening every 2 years.[2] The presence of some mammographic features in these screening mammograms may be related to the risk of developing breast cancer, as has been recently pointed out by some authors.[21,45] Only one of the 24 models identified in this systematic review included microcalcifications and masses found at mammography as risk factors in the model.[21] Time-changing variables such as radiological variables may not be as stable as personal history. However, in a screening context, this information is especially relevant because it is easily available from previous screening examinations.

Quality of the evidence

We found variability in the design of the studies that were used to obtain the cancer risk estimates. Notably, the study design used in the BCSC model was a cohort, which is a robust epidemiology design that allows developing and validating prediction models. Another frequently used design was the case-control study, nested or not. Contrary to the cohort study, time-changing variables may not be well obtained in case-control studies.

Regarding the external validation, the models showed some limitations given that few of them were further evaluated in different contexts. As far as we know, there are numerous scientific publications reporting external model validation in different settings and countries. These studies may help to understand the performance of a model in a specific context, but this issue was out of the scope of our review and, therefore, we have not included external validation studies. As an example of the relevance of these studies, we can inform that the BCRAT model has more than 50 articles informing the external validation of these models in different countries.[46] The Rosner-Colditz model has also been validated in several studies, one of the most complete validations being the one performed in 2013 by the authors themselves.[37] On the other hand, we found that although the Eriksson et al.[19] model reports the highest AUC (0.71), this model has not been externally validated, which increases the uncertainty about its applicability.

Also, there were limitations in data inputs, mostly due to the fact that in several models the information was provided by self-reported questionnaires that may affect the accuracy of the results. Finally, there is a limitation when comparing the AUROC or E/O ratio across the models given that there is great heterogeneity amongst them. The models were targeted to different populations, included different sets of risk factors, and often used different methodologies. We have taken into account all these variations and presented the results by model categories.

Potential biases in the review process

This systematic review was limited to studies published in English and did not involve an active search for grey literature, which is literature that is not formally published in sources such as books or journal articles. Therefore, some models may not have been identified. However, since we have conducted a comprehensive literature search in Medline, EMBASE and The Cochrane Library, we estimate that the loss of information due to the study selection criteria is low. Some key genetically oriented models, such as BOADICEA[47] and BRACAPRO[48] were not included in this review because they are aimed at high risk women and not useful for women in the general population in the screening context. Full-text screening and data abstraction process were performed by two researchers, which increase the quality of the review process. Moreover, as far as we know, this is the first review assessing the risk of bias of the identified risk prediction models.

CONCLUSIONS

The development of individualised breast cancer risk prediction models has increased over the last three decades, but the improvements in both the discriminatory power and calibration accuracy are still limited. Despite the time that has passed since the first model was published and a large number of available publications, only one model addressed to women attending a population-based screening programme[21] was identified. Currently, it is still a challenge to recommend any of the models as the standard for predicting individual risk in screening context. However, the models have been updated by adding new variables, such as common genetic variation or radiologic variables and have shown improvements in their quality as well as in their discriminative accuracy. These new variables need further evaluation to confirm its promising impact in the prediction capacity to propose personalised strategies for breast cancer screening.

ACKNOWLEDGEMENTS

The authors thank Ms. Lorea Galnares-Cordero for her contributions to the design of the search strategy and the initial retrieval of the citations. The authors also thank Ms. Julieta Politi and Mr. José María Montero-Moraga for their contribution in the screening process. Javier Louro is a Ph.D. candidate at the Methodology of Biomedical Research and Public Health program, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain.

AUTHOR CONTRIBUTIONS

J.L., M.P., M.S. and X.C. designed the study, and J.L. and M.P. wrote the manuscript. J. L., M.P. and M.H.B. performed the screening, data abstraction and quality assessment of included studies. M.S., L.D. and M.R. contributed to the analyses and interpreted the data. M.R., M.H.B. and X.C. collaborated in drafting the manuscript and revising it critically for important intellectual content. All authors read and approved the final manuscript.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41416-019-0476-8.

Competing interests: The authors declare no competing interests.

Ethics approval and consent to participate: All procedures performed in this study were in accordance with the ethical standards of the ethics committee of Parc de Salut Mar (CEIC Parc de Salut Mar) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Neither specific patient consent nor ethics committee's approval were required because we used published articles that were obtained from open access databases.

Data availability: The datasets analysed during the current study are publicly available from the corresponding author.

Consent for publication: Not applicable.

Funding: This work was partially supported by Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS) and by grants from Instituto de Salud Carlos III FEDER (grant numbers: PI15/00098 and PI17/00047). JL is core funded by the Research Network on Health Services in Chronic Diseases (RD12/0001/0015). MHB is core funded by the UK Medical Research Council (funding code: MC_UU_12017/15) and the Scottish Government Chief Scientist Office (funding code: SPHSU15). None of the funders participated in the design of the study, collection, analysis, or interpretation of data, or in writing the manuscript.

REFERENCES

1. The Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet. **380**, 1778–1786 (2012).
2. The European Commission Initiative on Breast Cancer (ECIBC). Recommendations from European Breast Guidelines. 2016. https://ecibc.jrc.ec.europa.eu/recommendations/.
3. U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann. Intern. Med.* **151**, 716–726 (2009). W-236.
4. Oeffinger, K. C., Fontham, E. T., Etzioni, R., Herzig, A., Michaelson, J. S., Shih, Y. C. et al. Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *JAMA* **314**, 1599–1614 (2015).
5. Mandelblatt, J. S., Stout, N. K., Schechter, C. B., van den Broek, J. J., Miglioretti, D. L., Krapcho, M. et al. Collaborative modeling of the benefits and harms associated with different U.S. Breast Cancer Screening Strategies. *Ann. Intern Med.* **164**, 215–225 (2016).
6. Steyerberg, E. W. *Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating.* (Springer Science, New York, 2009).
7. Anothaisintawee, T., Teerawattananon, Y., Wiratkapun, C., Kasamesup, V. & Thakkinstian, A. Risk prediction models of breast cancer: a systematic review of model performances. *Breast Cancer Res Treat.* **133**, 1–10 (2012).
8. Higgins, J. P. T., Green, S. (eds). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. 2011. http://handbook-5-1.cochrane.org/.
9. Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M. et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* **4**, 1 (2015).
10. Jaime Caro, J., Eddy, D. M., Kan, H., Kaltz, C., Patel, B., Eldessouki, R. et al. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health* **17**, 174–182 (2014).
11. Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J. et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* **358**, j4008 (2017).
12. Vachon, C. M., Pankratz, V. S., Scott, C. G., Haeberle, L., Ziv, E., Jensen, M. R., et al. The contributions of breast density and common genetic variation to breast cancer risk. *J. Natl. Cancer Inst.* **107**, dju397 (2015).
13. Zhang, X., Rice, M., Tworoger, S. S., Rosner, B. A., Eliassen, A. H., Tamimi, R. M. 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.* **15**, e1002644 (2018).
14. Banegas, M. P., John, E. M., Slattery, M. L., Gomez, S. L., Yu, M., LaCroix, A. Z., et al. Projecting Individualized Absolute Invasive Breast Cancer Risk in US Hispanic Women. *J. Natl. Cancer Inst.* **109**, djw215 (2017).
15. Barlow, W. E., White, E., Ballard-Barbash, R., Vacek, P. M., Titus-Ernstoff, L., Carney, P. A. et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. *J. Natl. Cancer Inst.* **98**, 1204–1214 (2006).
16. Boyle, P., Mezzetti, M., La Vecchia, C., Franceschi, S., Decarli, A. & Robertson, C. Contribution of three components to individual cancer risk predicting breast cancer risk in Italy. *Eur. J. Cancer Prev.* **13**, 183–191 (2004).
17. Chen, J., Pee, D., Ayyagari, R., Graubard, B., Schairer, C., Byrne, C. et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *J. Natl. Cancer Inst.* **98**, 1215–1226 (2006).
18. Colditz, G. A. & Rosner, B. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the Nurses' Health Study. *Am. J. Epidemiol.* **152**, 950–964 (2000).
19. Colditz, G. A., Rosner, B. A., Chen, W. Y., Holmes, M. D. & Hankinson, S. E. Risk factors for breast cancer according to estrogen and progesterone receptor status. *J. Natl. Cancer Inst.* **96**, 218–228 (2004).
20. Decarli, A., Calza, S., Masala, G., Specchia, C., Palli, D. & Gail, M. H. Gail model for prediction of absolute risk of invasive breast cancer: independent evaluation in the Florence-European Prospective Investigation Into Cancer and Nutrition cohort. *J. Natl. Cancer Inst.* **98**, 1686–1693 (2006).
21. 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.* **19**, 29 (2017).
22. Gail, M. H., Brinton, L. A., Byar, D. P., Corle, D. K., Green, S. B., Schairer, C. et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J. Natl. Cancer Inst.* **81**, 1879–1886 (1989).
23. Gail, M. H., Costantino, J. P., 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.* **99**, 1782–1792 (2007).
24. Kerlikowske, K., Gard, C. C., Sprague, B. L., Tice, J. A. & Miglioretti, D. L. One versus two breast density measures to predict 5 and 10-year breast cancer risk. *Cancer Epidemiol. Biomark. Prev.* **24**, 889–897 (2015).
25. Matsuno, R. K., Costantino, J. P., Ziegler, R. G., Anderson, G. L., Li, H., Pee, D. et al. Projecting individualized absolute invasive breast cancer risk in Asian and Pacific Islander American women. *J. Natl. Cancer Inst.* **103**, 951–961 (2011).
26. Novotny, J., Pecen, L., Petruzelka, L., Svobodnik, A., Dusek, L., Danes, J. et al. Breast cancer risk assessment in the Czech female population—an adjustment of the original Gail model. *Breast Cancer Res Treat.* **95**, 29–35 (2006).
27. Rosner, B. & Colditz, G. A. Nurses' health study: log-incidence mathematical model of breast cancer incidence. *J. Natl. Cancer Inst.* **88**, 359–364 (1996).
28. Rosner, B., Colditz, G. A., Iglehart, J. D. & Hankinson, S. E. 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.* **10**, R55 (2008).
29. Shieh, Y., Hu, D., Ma, L., Huntsman, S., Gard, C. C., Leung, J. W. et al. Breast cancer risk prediction using a clinical risk model and polygenic risk score. *Breast Cancer Res Treat.* **159**, 513–525 (2016).
30. Tice, J. A., Cummings, S. R., Ziv, E. & Kerlikowske, K. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. *Breast Cancer Res Treat.* **94**, 115–122 (2005).
31. Tice, J. A., Cummings, S. R., Smith-Bindman, R., Ichikawa, L., Barlow, W. E. & Kerlikowske, K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann. Intern Med.* **148**, 337–347 (2008).
32. Tice, J. A., Miglioretti, D. L., Li, C. S., Vachon, C. M., Gard, C. C. & Kerlikowske, K. Breast density and benign breast disease: risk assessment to identify women at high risk of breast cancer. *J. Clin. Oncol.* **33**, 3137–3143 (2015).
33. Tyrer, J., Duffy, S. W. & Cuzick, J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat. Med.* **23**, 1111–1130 (2004).
34. Ueda, K., Tsukuma, H., Tanaka, H., Ajiki, W. & Oshima, A. Estimation of individualized probabilities of developing breast cancer for Japanese women. *Breast Cancer* **10**, 54–62 (2003).
35. Wang, Y., Gao, Y., Battsend, M., Chen, K., Lu, W. & Wang, Y. Development of a risk assessment tool for projecting individualized probabilities of developing breast cancer for Chinese women. *Tumour Biol.* **35**, 10861–10869 (2014).

36. Brentnall, A. R., Harkness, E. F., Astley, S. M., Donnelly, L. S., Stavrinos, P., Sampson, S. et al. Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort. *Breast Cancer Res.* **17**, 147 (2015).
37. Rosner, B. A., Colditz, G. A., Hankinson, S. E., Sullivan-Halley, J., Lacey, J. V. Jr. & Bernstein, L. Validation of Rosner-Colditz breast cancer incidence model using an independent data set, the California Teachers Study. *Breast Cancer Res Treat.* **142**, 187–202 (2013).
38. Cintolo-Gonzalez, J. A., Braun, D., Blackford, A. L., Mazzola, E., Acar, A., Plichta, J. K. et al. Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications. *Breast Cancer Res Treat.* **164**, 263–284 (2017).
39. Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H. et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
40. MyPeBS. Randomized comparison of risk-stratified versus standard breast cancer screening in European women aged 40–70 (MyPeBS). 2017. http://www.brumammo.be/documents/docs/bmm-my-pebs-clinical-trial-protocol.pdf.
41. Esserman, L. J., Study, W. & Athena, I. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. *NPJ Breast Cancer* **3**, 34 (2017).
42. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA.* **98**, 10869–10874 (2001).
43. Garcia-Closas, M., Gunsoy, N. B., Chatterjee, N. Combined associations of genetic and environmental risk factors: implications for prevention of breast cancer. *J. Natl. Cancer Inst*. 2014;106.
44. Maas, P., Barrdahl, M., Joshi, A. D., Auer, P. L., Gaudet, M. M., Milne, R. L. et al. Breast cancer risk from modifiable and nonmodifiable risk factors among white women in the United States. *JAMA Oncol.* **2**, 1295–1302 (2016).
45. 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* **280**, 379–386 (2016).
46. Wang, X., Huang, Y., Li, L., Dai, H., Song, F. & Chen, K. Assessment of performance of the Gail model for predicting breast cancer risk: a systematic review and meta-analysis with trial sequential analysis. *Breast Cancer Res.* **20**, 18 (2018).
47. Antoniou, A. C., Pharoah, P. P., Smith, P. & Easton, D. F. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br. J. Cancer* **91**, 1580–1590 (2004).
48. Berry, D. A., Parmigiani, G., Sanchez, J., Schildkraut, J. & Winer, E. Probability of carrying a mutation of breast-ovarian cancer gene BRCA1 based on family history. *J. Natl. Cancer Inst.* **89**, 227–238 (1997).

Article 5

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

Authors: Louro J, Román M, Posso M, Vazquez I, Saladié F, Rodriguez-Arana A, Quintana MJ, Domingo L, Baré M, Marcos-Graguera R, Vernet-Tomas M, Sala M, Castells X on behalf of the BELE and IRIS Study Groups

Journal: PLoS One. 16(3): e0248930

Impact Factor: 2.74 (Q2 Multidisciplinary sciences)

DOI: 10.1371/journal.pone.0248930

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 on 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.

PLOS ONE

RESEARCH ARTICLE

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

OPEN ACCESS

Citation: Louro J, Román M, Posso M, Vázquez I, Saladié F, Rodriguez-Arana A, et al. (2021) Developing and validating an individualized breast cancer risk prediction model for women attending breast cancer screening. PLoS ONE 16(3): e0248930. https://doi.org/10.1371/journal.pone.0248930

Editor: Erin J. A. Bowles, Kaiser Permanente Washington, UNITED STATES

Received: September 23, 2020

Accepted: March 8, 2021

Published: March 23, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0248930

Data Availability Statement: Wwe have uploaded the database to the Harvard Dataverse online

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

repository. The data is accessible with DOI: https://doi.org/10.7910/DVN/3T7HCH.

Funding: This study was supported by grants from Instituto de Salud Carlos III FEDER [PI15/00098 and PI17/00047]; the Research Network on Health Services in Chronic Diseases [RD12/0001/0015]; and the Spanish Society of Epidemiology (SEE) [XV Alicia Llacer grant for the best research by a young researcher].

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: aHR, Adjusted hazard ratio; AUC, Area under the receiving operating characteristic curve; BI-RADS, Breast Imaging Reporting and Data System; BBD, Benign breast disease; DCIS, Ductal carcinoma in situ; E/O, Expected to observed; SNPs, Single Nucleotide Polymorphisms; 95%CI, 95% Confidence intervals.

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

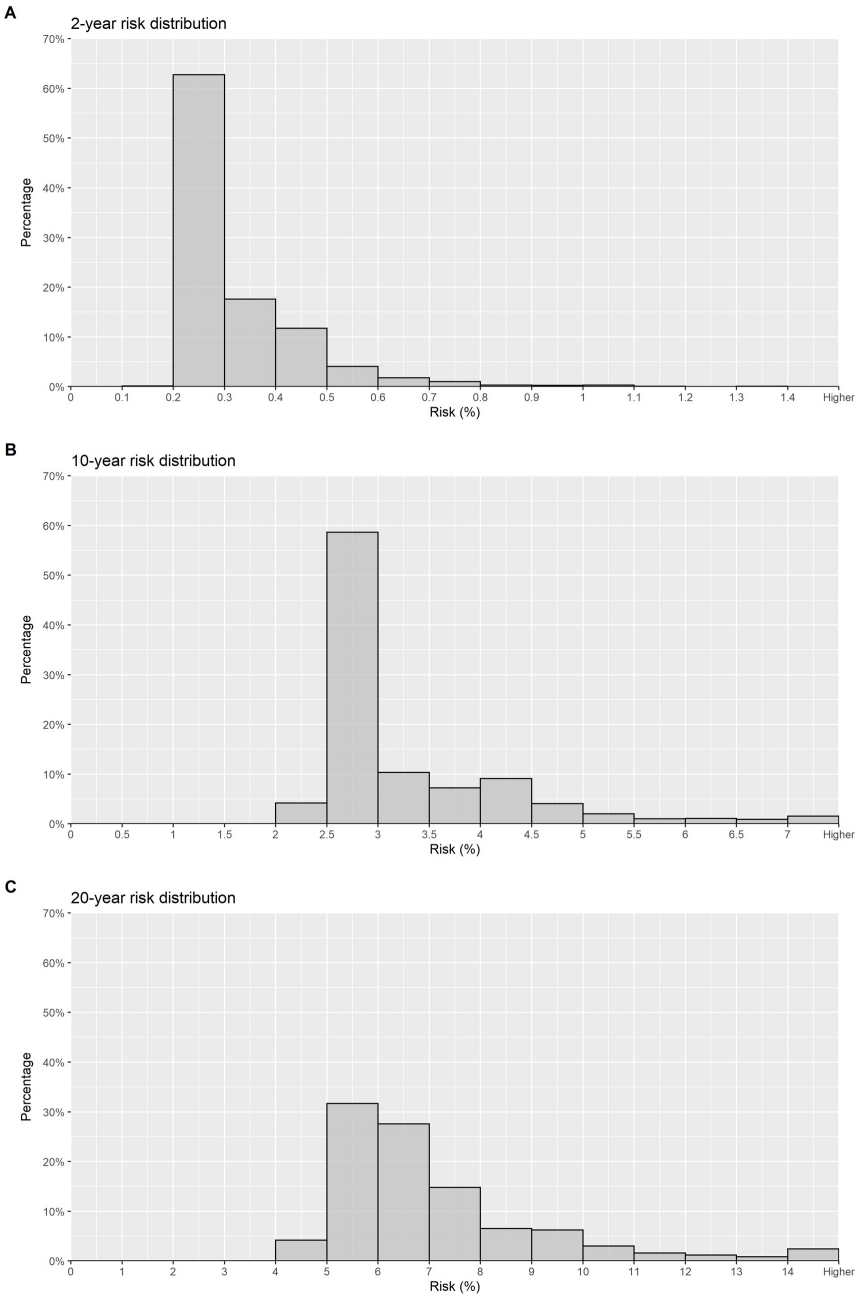


Fig 1. Distribution of the absolute cumulative risk estimates.

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

The authors acknowledge the dedication and support of the Benign Lesion (BELE) Study Group leaded by Xavier Castells (xcastells@parcdesalutmar.cat) and listed here in alphabetical order and grouped by institution: (a) IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain: Andrea Burón, Xavier Castells, Merce Comas, Jose Maria Corominas, Javier Louro, Ana Rodríguez-Arana, Marta Román, Maria Sala, Sonia Servitja, Mar Vernet-Tomas; (b) Corporació Sanitària Parc Taulí, Sabadell, Spain: Marisa Baré, Nuria Tora; (c) Catalan Institute of Oncology, Barcelona, Spain: Llucia Benito, Carmen Vidal (d) Hospital Santa Caterina, Girona, Spain: Joana Ferrer; (e) Catalan Institute of Oncology, Girona, Spain: Rafael Marcos-Gragera; (f) Hospital de la Santa Creu i Sant Pau, Barcelona, Spain: Judit Solà-Roca, María Jesús Quintana; (g) General Directorate of Public Health, Government of Cantabria, Spain: Mar Sánchez; (h) Principality of Astúrias Health Service, Spain: Miguel Prieto; (i) Fundació Lliga per a La Investigació i Prevenció Del Cáncer, Tarragona, Spain: Francina Saladié, Jaume Galceran; (j) Hospital Clinic, Barcelona, Spain; Xavier Bargalló, Isabel Torá-Rocamora; (k) Vallés Oriental Breast Cancer Early Detection Program, Spain; Lupe Peñalva; (l) Catalonian Cancer Strategy, Barcelona, Spain: Josep Alfons Espinàs.

The authors also acknowledge the dedication and support of the Individualized Risk (IRIS) Study Group leaded by Marta Román (mroman@parcdesalutmar.cat) and listed here in alphabetical order and grouped by institution: (a) IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain: Rodrigo Alcantara, Xavier Castells, Laia Domingo, Javier Louro, Margarita Posso, Maria Sala, Ignasi Tusquets, Ivonne Vazquez, Mar Vernet-Tomas; (b) Corporació Sanitària Parc Taulí, Sabadell, Spain: Marisa Baré, Javier del Riego; (c) Catalan Institute of Oncology, Barcelona, Spain: Llucia Benito, Carmen Vidal (d) Hospital Santa Caterina, Girona, Spain: Joana Ferrer; (e) Catalan Institute of Oncology, Girona, Spain: Rafael Marcos-Gragera; (f) Hospital de la Santa Creu i Sant Pau, Barcelona, Spain: Judit Solà-Roca, María Jesús Quintana; (g) General Directorate of Public Health, Government of Cantabria, Spain: Mar Sánchez; (h) Principality of Astúrias Health Service, Spain: Miguel Prieto; (i) Fundació Lliga per a La Investigació i Prevenció Del Cáncer, Tarragona, Spain: Francina Saladié, Jaume Galceran; (j) Hospital Clinic, Barcelona, Spain; Xavier Bargalló, Isabel Torá-Rocamora; (k) Vallés Oriental Breast Cancer Early Detection Program, Spain; Lupe Peñalva; (l) Catalonian Cancer Strategy, Barcelona, Spain: Josep Alfons Espinàs.

The authors also acknowledge the help of the (l) Biomedical Informatics Research Unit (GRIB) of the UPF; Alfons Gonzalez-Pauner, Ferran Sanz and (m) the Cardiovascular epidemiology and genetics group of the IMIM; Jaume Marrugat, Isaac Subirana, Joan Vila.

Javier Louro is a Ph.D. candidate at the Methodology of Biomedical Research and Public Health program, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

Author Contributions

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

Data curation: Javier Louro, Marta Román, Ivonne Vázquez, Ana Rodriguez-Arana.

Formal analysis: Javier Louro.

Funding acquisition: Marta Román, Xavier Castells.

Investigation: Javier Louro, Marta Román, Margarita Posso, Laia Domingo, Rafael Marcos-Gragera, María Vernet-Tomas, Xavier Castells.

Methodology: Javier Louro, Marta Román, Margarita Posso, Laia Domingo, Xavier Castells.

Project administration: Javier Louro, Xavier Castells.

Resources: Javier Louro, Francina Saladié, M. Jesús Quintana, Marisa Baré, Rafael Marcos-Gragera.

Supervision: Marta Román, María Vernet-Tomas, Maria Sala.

Validation: Maria Sala.

Writing – original draft: Javier Louro.

Writing – review & editing: Javier Louro, Marta Román, 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 Mediques 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

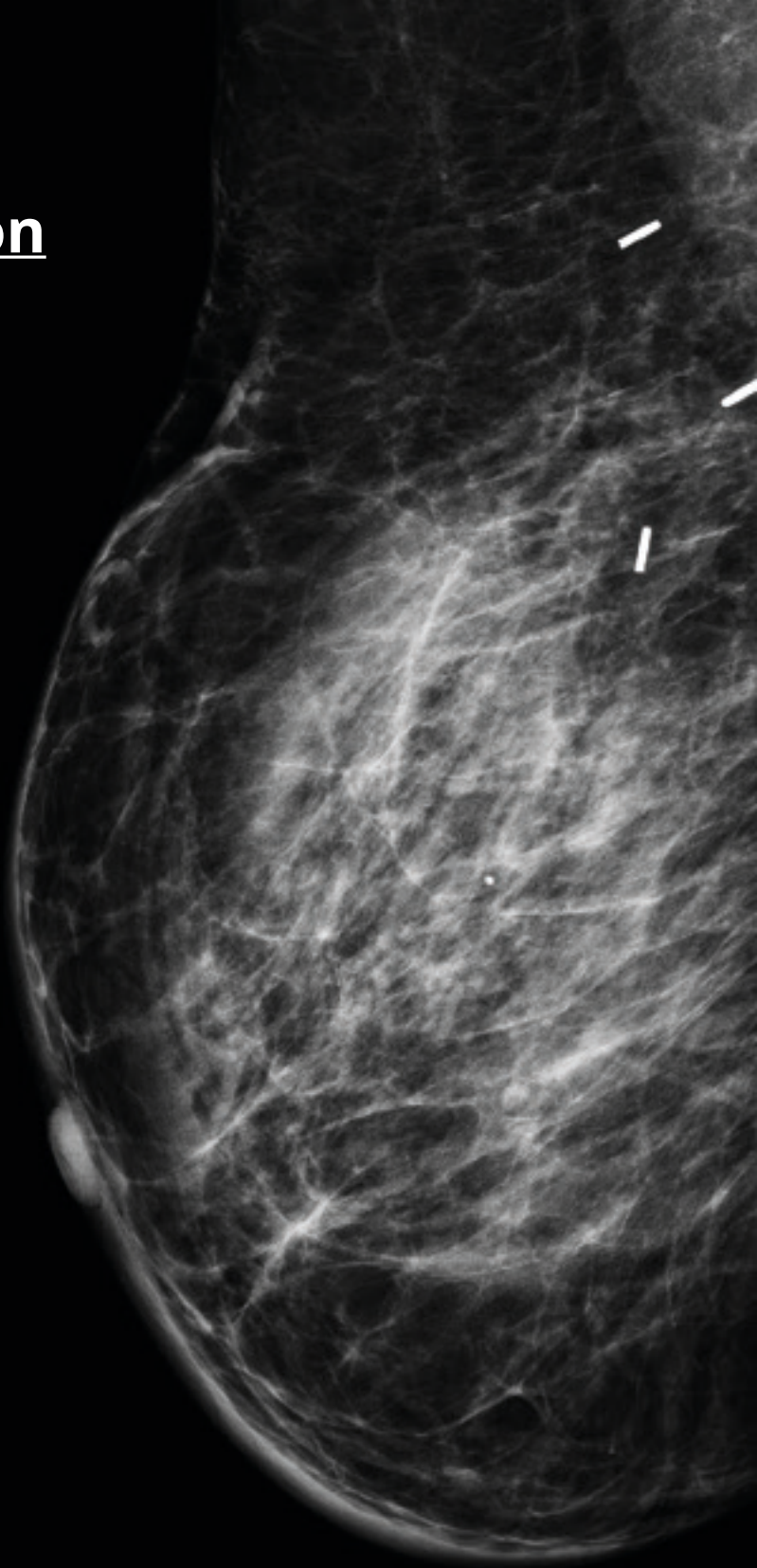
43. Posso M, Louro J, Sanchez M, Roman M, Vidal C, Sala M, et al. Mammographic breast density: How it affects performance indicators in screening programmes? Eur J Radiol. 2019; 110:81–7. https://doi.org/10.1016/j.ejrad.2018.11.012 PMID: 30599878

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

Discussion



Main findings

The results of the 5 articles support the main objective of this thesis, deepening understanding of the different breast cancer risk factors with large longitudinal screening cohorts and expanding knowledge of individual breast cancer risk prediction models.

In summary, the main findings of this research are:

Apropos of a better understanding of the risk factors

- Performance screening measures are negatively affected by increasing breast density. Sensitivity decreased from 89.2% in women with BI-RADS 1 to 67.9% in those with BI-RADS 4. The positive predictive value of both recall and invasive tests decreased from 10.4% to 5.7% and from 49.8% to 32.4% in women with BI-RADS 1 and BI-RADS 4, respectively.

- Compared with women without benign breast disease, the risk of breast cancer was significantly higher in women with benign breast disease diagnosed in an incident screen (aHR, 2.67; 95%CI: 2.24-3.19) than in those with benign breast disease diagnosed in a prevalent screen (aHR, 1.87; 95%CI: 1.57-2.24).

- The risk of breast cancer independently increased with the presence of benign breast disease and with the increase in breast density (p-value for interaction = 0.84) and remained elevated for over 15 years.

Apropos of the breast cancer risk prediction model

- Individualized risk prediction models are promising tools for implementing risk-based screening policies. However, it is challenging to recommend any of them yet, since they need further improvement in their quality and discriminatory capacity.

- We developed a risk prediction model based on family history, previous benign breast disease and previous mammographic features. All 3 risk factors were strongly associated with breast cancer risk, with the highest risk being found among women with a family history of breast cancer (aHR: 1.67), 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 on the time horizon selected.

Joint discussion of articles

The discussion is articulated in terms of the research questions central to this thesis. The specific discussion of each of the articles is presented in the articles themselves, where the limitations and strengths of each study are also described.

In the systematic review conducted as part of this research, we found that there was a large body of literature on breast cancer risk prediction models prior to this thesis (102-110). With the exception of one model (110), these models were not created specifically to predict the risk of women participating in breast cancer screening, but rather to support clinical decisions when attending a specific woman. These models have been analyzed individually, examining the variables used as well as performing a risk of bias assessment in order to grade the quality of the existing evidence. Although the models have improved during the last few decades, their discriminatory accuracy remains moderate with a maximum AUC value of 0.71 reported by the article focused precisely on screening. The calibration accuracy was widely heterogeneous ranging from 0.85 to 1.52. In contrast to the 2 previous systematic reviews about breast cancer

risk models (122, 123), the systematic review presented in this thesis provides innovative information regarding the quality of the identified prediction models. Furthermore, the quality of the studies was not high due to limitations in the discriminative accuracy, study design, and data inputs, hence, as discussed in the article, it is challenging to recommend application of any of the models in breast cancer screening.

The distribution of risk factors in such different populations may also affect the applicability of the models to different contexts. It is widely accepted that distinct subtypes of breast cancer may have different genetic markers (124). These differences, the nature of breast cancer itself and its low incidence may lead to a model having low discriminatory accuracy. In other words, in the general population, there is a "low" probability of having breast cancer, even in the highest risk group. Something similar happens with lung cancer, a heavy smoker (with an 80-pack-per-year history) is as much as 20-times more likely to develop lung cancer than a non-smoker, but the yearly incidence of lung cancer in people who have smoked heavily for 30 years, is just 2 to 3 per 1,000 (125). This low incidence limits the ability of cancer risk models to correctly discriminate who will develop the disease (poor discriminatory power) compared with other risk models targeting more incident diseases such as cardiovascular events (126), for instance.

Another potential limitation in the applicability of prediction models for breast cancer screening is the completeness and the number of included risk factors, which ranged in our review from five to 18. Nevertheless, some potentially relevant risk factors have only been included in a few models, indicating the need for further evaluation. There is an intense debate about which variables are most suitable for inclusion in a risk prediction model for breast cancer screening. There is probably no single answer to this question. The fewer variables used, the more feasible the model, but the more variables used, the greater the discriminatory power, as long as these variables truly modify breast cancer risk. In addition, the methodology of the model has to be appropriate. Therefore, for the proposed objective it is necessary to look for a pragmatic model, which has good discriminatory power, but which uses variables easily obtainable in a screening context. For this reason, to understand how to create a model in the most optimal way possible, we have to accurately analyze and assess the risk factors to be used. In the context of this thesis, we analyzed

characteristics of 2 variables that are key for risk prediction purposes, mammographic density, and benign breast disease, and we also included in our model other variables such as mammographic features or family history, in addition to age.

The results of the analysis of breast density clearly indicate that high breast density not only confers a higher biological risk but also produces masking on mammography, resulting in a lower sensitivity and positive predictive value. Almost 21% of mammograms in our cohort and up to 40% in other cohorts (127, 128) were classified as BI-RADS 3 or BI-RADS 4 density, representing a large proportion of screened women. This suggests that women with dense breasts could benefit from tests other than mammography and supports personalization. Other authors have already analyzed the use of ultrasound (78) or magnetic resonance imaging (79) in women with dense breasts as supplemental imaging with positive results. This is also in line with current clinical trials such as MyPeBS (My Personal Breast Screening) (82), in which women with dense breasts are offered magnetic resonance imaging complementary to mammography.

Another risk factor proposed for individual assessment of breast cancer risk is the presence of previous benign breast disease. The second article of this thesis studied this risk factor from a point of view that, to our knowledge, was used for the first time. We analyzed the differences in breast cancer risk depending on the type of screening, which reflects whether benign breast disease was diagnosed in a woman's first screening, called prevalent benign breast disease, or at any subsequent screening, called incident benign breast disease. We found that regardless of histological classification, women with an incident benign breast disease had a 42% higher risk than women with a prevalent one. This finding is particularly relevant since, in our population, 45% of benign lesions are diagnosed at incident screens. Screening type therefore provides key information for risk prediction since these differences in risk are not reflected in previous breast cancer risk models and may be used to improve the accuracy of predictions. Unless breast cancer risk prediction models include this information, the risk attributed to benign breast disease diagnosed in prevalent screens could be overestimated, and, likewise, the risk attributed to benign breast disease diagnosed in incident screens could be underestimated.

Moreover, the results of this study may have implications for clinical decisions on the follow-up of women with a diagnosis of benign breast disease. The recommended follow-up strategy may differ, depending on the benign breast disease subtype and screening type at diagnosis. In many screening programs, women diagnosed with a benign breast disease that confers a high breast cancer risk, such as atypical hyperplasia, are referred for close clinical follow-up. Therefore, a non-proliferative benign breast disease diagnosed in an incident screening may confer a higher risk than a proliferative benign breast disease diagnosed in a prevalent screen, and yet, to date, some of these women are screened less comprehensively. The evidence provided by this new classification opens the door to a more precise analysis of benign breast disease.

To determine how breast density and benign breast disease could be included in a breast cancer risk prediction model, the third article of this thesis analyses the interaction between them. To correctly model the risk, it is necessary to understand not only whether the variables actually define risk, but also whether they interact with each other. Otherwise, model could be biased and either under- or overestimate the risk of some of the variables. We found that the presence of benign breast disease and high mammographic density were independently associated with a higher risk of breast cancer. The risk diverged over the study period for women with and without benign breast disease across mammographic density categories. The finding that both factors are independent risk factors for breast cancer consolidates their utility in risk prediction models. Few studies have evaluated the combined effect of benign breast disease and breast density on the risk of breast cancer at mammographic screening. A prior large study (129) assessed the combined effect of mammographic breast density and different subtypes of benign breast disease. The study population in the abovementioned study consisted solely of women with a previous benign breast disease diagnosis, hampering comparison of the effect of the presence or absence of benign breast disease. Despite these differences in study population and reference group, the results of the 2 studies are consistent.

Finally, the fifth and last article of this thesis defines a breast cancer risk prediction model. In this model, we used benign breast disease, and 3 variables that have been analyzed as risk factors: age (83), family history (90) and suspicious findings at mammographic reading (97). We found

that a woman with a family history of breast cancer had a 67% higher risk than a woman without a family history. We also found that having a previous mammographic finding conferred a higher risk of breast cancer, the highest risk being in calcifications, with up to 2.52 times more risk than a woman with no previous mammographic finding. The model did not use all the evidence found in the articles of this thesis, but illustrated an approach to a specific risk model for breast cancer screening that allowed us to estimate the short- and long-term risk of breast cancer in women targeted for mammography screening. The internal validation of the model showed good calibration and modest discriminatory power, not far from the usual breast cancer risk prediction models, despite having a smaller number of variables, which makes it more affordable. 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. Nevertheless, this model is not yet able to be used for the proposed objective, the personalization of breast cancer screening, as it needs to be refined and updated by the addition of new variables.

From the articles presented in this thesis, it can be concluded that breast cancer risk prediction is an area of knowledge in continuous development because of its importance to screening programs. We have demonstrated that breast density and benign breast disease are variables that could be used for this purpose, as they may explain some of the variability in the breast cancer risk in women participating in screening. We have seen that family history and previous mammographic features also explain part of this variability. We have addressed the problem, but still without enough tools to make a model capable of meeting the proposed objective. Future lines of research should lead us to update the risk prediction model presented in this thesis with information on the type of screening at benign breast disease diagnosis, as we have found that it affects breast cancer risk regardless of histologic type. This finding may help to obtain a higher discriminatory power. In addition, the model should be updated with breast density, as dense breasts not only have a higher biological risk, but also a higher probability of masking on mammography.

There is still no established formula for the most optimal individual risk breast cancer prediction model for woman attending screening. A model for this purpose would have to include risk factors feasibly and easily obtainable at screening participation. Furthermore, the model would have to be developed with data from large screening cohorts, since the problem is longitudinal in nature. In addition, the model would be applied to a cohort that would be evaluated continuously. In other words, such a model would need to be not only valid for a baseline assessment, but could also re-evaluate the risk of each woman every 2 years, since risk can change over the years. The model would have to be validated both internally and in an external cohort, and would have to have good calibration and discrimination results to ensure that it could identify which women have a lower or higher risk of developing breast cancer.

Once such a model is developed, it will be feasible to calculate the individual risk of each woman attending screening. With this estimated risk, it would be possible to decide which screening test is best for each woman according to her individual characteristics. Moreover, it will be possible to leave behind biennial screening as we know it, and perhaps screen higher-risk women annually and lower-risk women every 3 years. This intervention is not free of controversy either. Like every population-based intervention, not all individuals will obtain the same benefits from personalized screening nor will they experience the same adverse effects. At an individual level, there will be women who would be harmed by the intervention, such as low-risk women who develop cancer shortly after having a mammogram and who may have a worse prognosis if it is detected at 3 years instead of 2 years. Nevertheless, these new strategies could be feasible in terms of resources and cost-effectiveness as long as we minimize harms to screened women and the benefits clearly outweigh the risks (74, 75). These new strategies should improve the balance of benefits and harms compared to average-risk women receiving biennial screening (76, 77). Thus, such a model is key to help decision-making on how to make the transition from the current "one size fits all" strategy to a more efficient personalized screening.

Limitations

The main limitations of this thesis are those characteristic of a retrospective cohort study. Although these studies provide information on a large number of women screened and over long periods of time, they are limited by the quality of the information available in the original data sources. Information in this thesis was drawn from the original databases of the screening programs participating in the BEnign LEsion (BELE-2) and the IRIS study projects. As previously mentioned, a detailed protocol for definitions and collection of variables was developed to ensure the homogeneity of the information collected. However, certain variables related to women's characteristics, such as family history and the presence of previous benign lesions, showed a significant volume of missing values, which in some cases was 40%. To assess the impact of these variables, a rigorous quality control was performed on the source and method of information, and various sensitivity analyses were conducted to assess the impact of this lack of information on the main study variables. However, the possible bias generated by the lack of these variables was controlled by simply restricting each of the analyses to the screening programs that had good information on these variables.

A specific limitation of the articles involving breast density is that variability among radiologists can affect the results since breast density measurements are inherently inaccurate depending on subjective observation. Nevertheless, our results are consistent with the previous literature, and in our cohort breast density was classified by highly trained radiologists.

Another limitation of the cohort used in this thesis is that the number of cancers detected after proliferative benign breast disease with atypia was small as it is an uncommon subtype, which limited our ability to perform some subgroup analyses. In addition, follow-up of women with diagnoses of this type (such as atypical hyperplasia) is shorter as they leave the screening by being sent to a more exhaustive clinical pathway.

In addition, the breast cancer risk prediction model described in the thesis was only internally validated and external validation of the results is needed to verify its predictive performance.

Lastly, the systematic review was limited to studies published in English and did not involve an active search for gray literature, which is literature that is not formally published in sources such as books or journal articles. Therefore, some models may not have been identified. However, since we conducted a comprehensive literature search in Medline, EMBASE and The Cochrane Library, we estimate that the loss of information due to the study selection criteria is low.

Strengths

As previously mentioned, this thesis is based on the BEnign LEsion (BELE-2) and the Individualized RISk (IRIS) projects. These projects allowed the availability of a joint database of different screening programs, which contains information on a large number of women, followed sequentially during their multiple participations in the screening. This cohort has individual-level information from more than 780,000 women participating in 10 large, well-established, population-based screening programs. As mentioned above, this is one of the largest databases with individualized information created to date for the assessment of breast cancer screening. In addition, information is available for a 20-year period, from the launch of the programs in 1995 until 2015. The intense work of drafting protocols, homogenization of criteria and data validation guarantees a high level of consistency in the information analyzed.

As discussed in the methods section, the methodology developed for the different articles in this thesis represents an advance in the evaluation of population-based screening from a longitudinal perspective and provides a consistent methodological approach that broadens the perspective on the assessment of invidualized risk.

All the papers in this thesis explore controversial aspects of population-based screening evaluation and provide unpublished answers that add to the knowledge of breast cancer screening practices, with particular emphasis on the evaluation of the different breast cancer risk factors and on individual risk prediction. Despite leaving some questions open, the articles in this thesis delve into a fundamental aspect of breast cancer screening— individual risk—on which all eyes are focused for future use in the development of personalized strategies. Other works carried out subsequently or simultaneously corroborate or extend the results of these works, and take them as a reference, lending credence to the findings presented.

To conduct this project, we had the direct participation of people with management responsibilities in the various participating programs. This allowed a greater translation of the research results into practice, and at the same time enables more relevant questions to be raised in the context of public health and community programs.

In addition to the analysis of the cohort, we conducted a comprehensive literature search of breast cancer risk prediction models following the standard Cochrane Collaboration methods and adhering to the PRISMA statement reporting recommendations, leading to a consistent systematic review. Full-text screening and data abstraction process were performed by 2 researchers, increasing the quality of the review process.

Future lines of research

This thesis forms part of the breast cancer screening research line of the Epidemiology and Evaluation Group of Hospital del Mar-IMIM. This group has been funded for many years by the CIBER of Epidemiology and Public Health and is currently funded by the Health Services and Chronic Diseases Research Network (REDISSEC), with numerous

projects and different initiatives to enhance knowledge of aspects related to mammographic screening and aiming to provide continuity to the issues presented.

Within this framework, the Individualized RISk (IRIS) study, within which this thesis is framed, is still active. The main objective of this project is to develop invidualized breast cancer risk prediction models for women participating in mammography screening based on their known risk factors. Specifically, the aim was to expand the BELE-2 cohort to collect information on breast density in those programs where it was not available, in order to include this variable in the model. To this end, we have worked jointly with the University of Valencia for the validation of an automated software capable of retrospectively reading the density of digital mammograms. This project has been significantly delayed due to the worldwide pandemic caused by the Coronavirus Disease 2019 (COVID-19) and the procurement process of breast density has not yet been completed. The model presented in this thesis will be updated with this variable as soon as this project is finished.

In addition, with the experience gained through the BELE-2 and IRIS projects, together with the collaborations initiated with other groups and the need to work in a broader framework, an international collaboration with the BreastScreen Norway group of the Cancer Registry of Norway has been formalized. We have been working for several months on a joint article that will provide continuity to this thesis. The objective of this collaborative work is to reach a better understanding of tumor growth in breast cancer cases diagnosed in women targeted for breast cancer screening. The article is well advanced and will soon be published. These results will serve to improve knowledge of the natural history of tumors diagnosed in women participating in screening and may be useful in developing more efficient screening strategies based on women's individual risk.

As a logical consequence of the worldwide study on breast cancer individual risk and breast cancer screening personalization, a clinical trial has been launched at the European level to compare a personalized risk-based screening strategy (based on the individual women's risk of developing breast cancer) to standard screening. This project, called MyPeBS, began in 2017 and it is still under recruitment. It is an

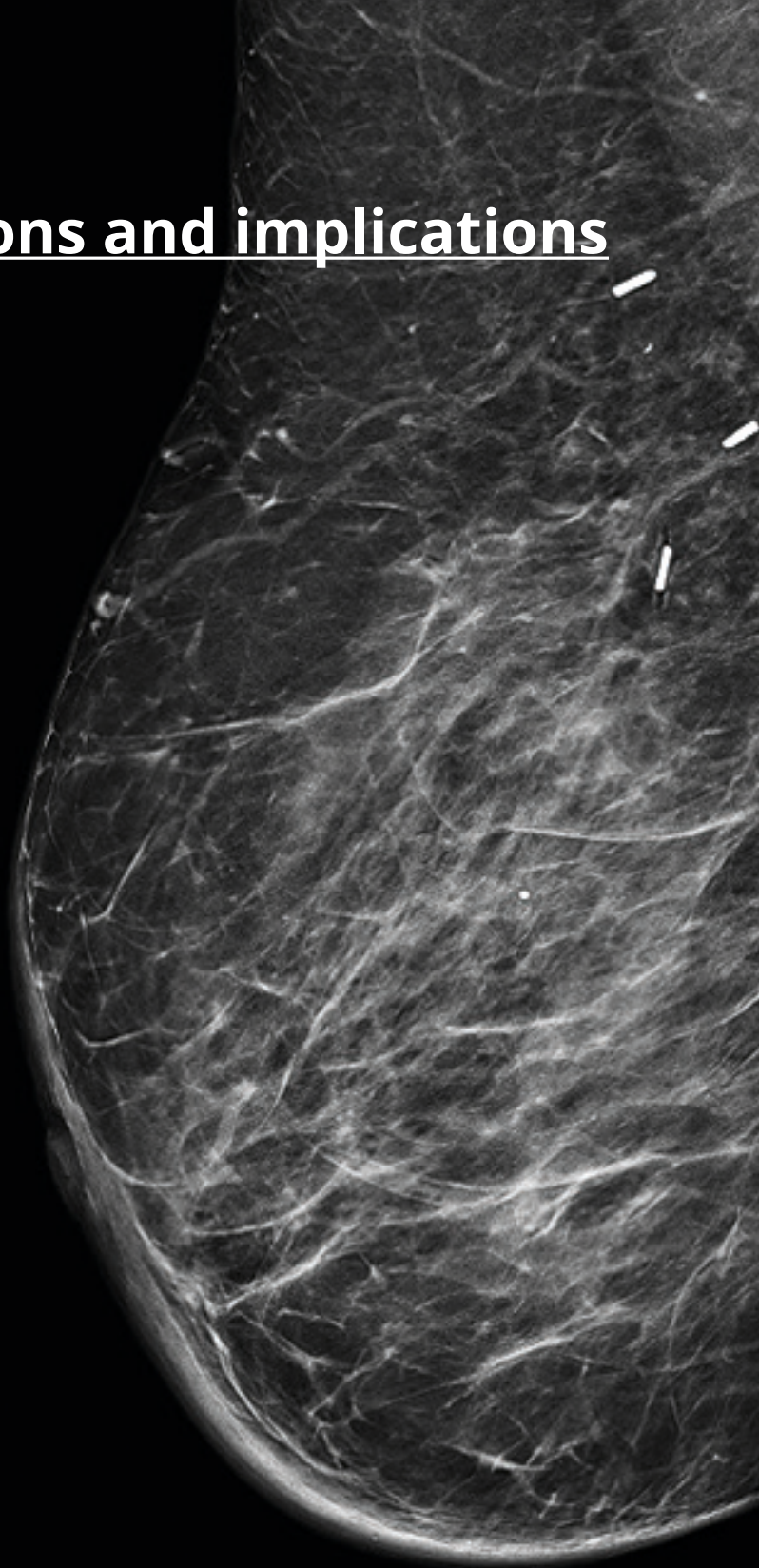
international project, funded by the European Union's Horizon 2020 research and involves 8 countries recruiting 85,000 women to participate in the clinical trial. It is expected that the results of this project will provide relevant information to eventually move worldwide from the current "one size fits all" breast cancer screening to fully personalized screening. This project is coordinated by a French group, Unicancer. Our group is part of the project as the only recruiting group in Spain and I personally will be the project manager of our recruiting center.

The new situation after the COVID-19 pandemic must also be considered for future decisions on breast cancer screening. Cancer screening programs have been considered as non-essential activities in most European countries and have therefore suffered a halt and/or delay in their activity. Currently, to reintroduce screening programs as soon as possible and at the same time reduce the risk of COVID-19 transmission, the programs have had to establish new safety guidelines. All of these safety measures reduce the mammographic capacity of the facility due to longer disinfection times and longer citation intervals to minimize contact among participants. These measures, coupled with the delay caused by periods of total disruption, present a challenge in the management of screening citations.

Our group has started a project whose main objective is to analyze the impact of the COVID-19 pandemic in the essential quality indicators of population-based mammographic breast cancer screening. We believe it is important to assess not only the participation and adherence rate, which may have been affected by the general fear of the pandemic after those first few months, but also the delay in diagnosis, which may lead to a higher number of breast cancers diagnosed at an advanced stage.

Hence, the possibility of distinguishing the most at-risk women in the population could also relieve the burden of mammograms in healthcare in a new reality where COVID-19 does not allow the same number of tests to be performed as before. As long as the number of mammograms performed remains lower than before the pandemic, it is important to prioritize the recall order based on breast cancer risk to try to minimize its effect on the various indicators of screening. Breast cancer risk prediction models may, therefore, be useful tools for post-pandemic reorganization of screening programs.

Conclusions and implications



Conclusions

- Performance screening measures are negatively affected by increasing breast density, which decreases sensitivity and positive predictive value.

- The risk of breast cancer conferred by benign breast disease is higher in women diagnosed in an incident screen than in those diagnosed in a prevalent screen, regardless of histological classification.

- Women with benign breast disease have an elevated risk for over 15 years independently of their breast density category.

- Individualized risk prediction models are promising tools for implementing risk-based screening policies. However, it is difficult to recommend a specific model since they all need further improvement in their quality and discriminatory capacity.

- We developed and validated a breast cancer risk prediction model able to estimate the short- and long-term breast cancer risk using information routinely reported at screening participation. Our model uses age, family history, mammographic findings, and benign breast disease.

- The model should be externally validated and updated with new variables. In the future, the model could help to guide individualized screening strategies aimed at improving the risk-benefit balance of mammography screening programs.

Recommendations and implications for public health

The results presented in this thesis improve the existing information to evaluate the balance between risk and benefits of mammographic screening. The assessment of the different risk factors, and the calculation of individual risk, is a substantial contribution to improve the effectiveness of population-based screening through personalization.

As discussed above, evidence supports personalization as the future of breast cancer screening and, to that end, it is necessary to perform breast cancer risk prediction models based on data from large screening cohorts and including risk factors easily obtainable at screening participation. The results of the various studies comprising this thesis are useful to better understand the different breast cancer risk factors and to know how to use them for individual risk assessment.

Our analysis showed that an elevated breast density not only confers a higher biological risk, but also produces a masking in mammograms, reducing their sensitivity and positive predictive value. This means that women with denser breasts do not benefit from mammography as much as women with fattier breasts and would benefit more from more accurate tests such as ultrasound or magnetic resonance imaging.

We have identified a new classification that allows us to better understand variability in breast cancer risk after a benign breast disease diagnosis, providing an innovative means of understanding these pathologies. Those benign breast disease diagnosed at a woman's first screening, called incident benign breast disease, have a higher risk than those diagnosed at subsequent screenings, regardless of their histological classification. This implies that we have identified another point at which certain women may benefit more from more comprehensive screening and the next step in screening could be to translate this finding into practice.

The joint analysis we have made of breast density and benign breast disease could be especially useful when defining in the future which variables should be part of the individualized risk models that will be used for the personalization of screening and how to include them in those models.

When we first approached a breast cancer risk prediction model, we found that women with a family history and previous mammographic findings also had a higher risk.

Hence, this thesis has led us to locate several variables that can be used to identify which women are at a higher risk of breast cancer and which are at lower risk and to deepen the knowledge of breast cancer risk prediction models. More exhaustive screening of high-risk women could reduce the number of interval cancers and less exhaustive screening of low-risk women could lead to a fewer number of false positives. The analysis presented in this thesis, therefore, provides new and useful tools to approach the next logical step, which is to update current "one-size-fits-all" breast cancer screening strategies to risk-based personalized breast cancer screening strategies.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021.

2. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, et al. Breast cancer statistics, 2019. CA Cancer J Clin. 2019;69(6):438-51.

3. Pollan M, Pastor-Barriuso R, Ardanaz E, Arguelles M, Martos C, Galceran J, et al. Recent changes in breast cancer incidence in Spain, 1980-2004. J Natl Cancer Inst. 2009;101(22):1584-91.

4. Ravdin PM, Cronin KA, Howlader N, Berg CD, Chlebowski RT, Feuer EJ, et al. The decrease in breast-cancer incidence in 2003 in the United States. N Engl J Med. 2007;356(16):1670-4.

5. Canfell K, Banks E, Moa AM, Beral V. Decrease in breast cancer incidence following a rapid fall in use of hormone replacement therapy in Australia. Med J Aust. 2008;188(11):641-4.

6. Glass AG, Lacey JV, Jr., Carreon JD, Hoover RN. Breast cancer incidence, 1980-2006: combined roles of menopausal hormone therapy, screening mammography, and estrogen receptor status. J Natl Cancer Inst. 2007;99(15):1152-61.

7. Katalinic A, Rawal R. Decline in breast cancer incidence after decrease in utilisation of hormone replacement therapy. Breast Cancer Res Treat. 2008;107(3):427-30.

8. Kerlikowske K, Miglioretti DL, Buist DS, Walker R, Carney PA, National Cancer Institute-Sponsored Breast Cancer Surveillance

C. Declines in invasive breast cancer and use of postmenopausal hormone therapy in a screening mammography population. J Natl Cancer Inst. 2007;99(17):1335-9.

9. Kliewer EV, Demers AA, Nugent ZJ. A decline in breast-cancer incidence. N Engl J Med. 2007;357(5):509-10; author reply 13.

10. Kumle M. Declining breast cancer incidence and decreased HRT use. Lancet. 2008;372(9639):608-10.

11. Allemani C, Minicozzi P, Berrino F, Bastiaannet E, Gavin A, Galceran J, et al. Predictions of survival up to 10 years after diagnosis for European women with breast cancer in 2000-2002. Int J Cancer. 2013;132(10):2404-12.

12. SEOM. Las cifras del cancer en españa. Available at: https://seom.org/seomcms/images/stories/recursos/Cifras_del_cancer_2020.pdf. 2020.

13. Pollan M, Michelena MJ, Ardanaz E, Izquierdo A, Sanchez-Perez MJ, Torrella A, et al. Breast cancer incidence in Spain before, during and after the implementation of screening programmes. Ann Oncol. 2010;21 Suppl 3:iii97-102.

14. Fidler HK, Boyes DA, Worth AJ. Cervical cancer detection in British Columbia. A progress report. J Obstet Gynaecol Br Commonw. 1968;75(4):392-404.

15. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease]. Bol Oficina Sanit Panam. 1968;65(4):281-393.

16. Wellings SR, Jensen HM. On the origin and progression of ductal carcinoma in the human breast. J Natl Cancer Inst. 1973;50(5):1111-8.

17. Wellings SR, Jensen HM, Marcum RG. An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. J Natl Cancer Inst. 1975;55(2):231-73.

18. Esserman LJ, Thompson IM, Reid B, Nelson P, Ransohoff DF, Welch HG, et al. Addressing overdiagnosis and overtreatment in cancer: a prescription for change. Lancet Oncol. 2014;15(6):e234-42.

19. Farabegoli F, Champeme MH, Bieche I, Santini D, Ceccarelli C, Derenzini M, et al. Genetic pathways in the evolution of breast ductal carcinoma in situ. J Pathol. 2002;196(3):280-6.

20. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Ten- to fourteen-year effect of screening on breast cancer mortality. J Natl Cancer Inst. 1982;69(2):349-55.

21. Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. BMJ. 1988;297(6654):943-8.

22. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. J Natl Cancer Inst. 2000;92(18):1490-9.

23. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L, et al. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet. 2006;368(9552):2053-60.

24. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet. 2002;359(9310):909-19.

25. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, et al. The Gothenburg Breast Screening Trial. Cancer. 2003;97(10):2387-96.

26. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L, et al. Screening for breast cancer: an update

for the U.S. Preventive Services Task Force. Ann Intern Med. 2009;151(10):727-37, W237-42.

27. Tonelli M, Connor Gorber S, Joffres M, Dickinson J, Singh H, Lewin G, et al. Recommendations on screening for breast cancer in average-risk women aged 40-74 years. CMAJ. 2011;183(17):1991-2001.

28. Independent UKPoBCS. The benefits and harms of breast cancer screening: an independent review. Lancet. 2012;380(9855):1778-86.

29. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev. 2013(6):CD001877.

30. Hofvind S, Ursin G, Tretli S, Sebuodegard S, Moller B. Breast cancer mortality in participants of the Norwegian Breast Cancer Screening Program. Cancer. 2013;119(17):3106-12.

31. Moss SM, Wale C, Smith R, Evans A, Cuckle H, Duffy SW. Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. Lancet Oncol. 2015;16(9):1123-32.

32. Myers ER, Moorman P, Gierisch JM, Havrilesky LJ, Grimm LJ, Ghate S, et al. Benefits and Harms of Breast Cancer Screening: A Systematic Review. JAMA. 2015;314(15):1615-34.

33. Iwamoto Y, Kaucher S, Lorenz E, Barnighausen T, Winkler V. Development of breast cancer mortality considering the implementation of mammography screening programs - a comparison of western European countries. BMC Public Health. 2019;19(1):823.

34. Dibden A, Offman J, Duffy SW, Gabe R. Worldwide Review and Meta-Analysis of Cohort Studies Measuring the Effect of Mammography Screening Programmes on Incidence-Based Breast Cancer Mortality. Cancers (Basel). 2020;12(4).

35. Zielonke N, Gini A, Jansen EEL, Anttila A, Segnan N, Ponti A, et al. Evidence for reducing cancer-specific mortality due to screening for breast cancer in Europe: A systematic review. Eur J Cancer. 2020;127:191-206.

36. Canelo-Aybar C, Ferreira DS, Ballesteros M, Posso M, Montero N, Sola I, et al. Benefits and harms of breast cancer mammography screening for women at average risk of breast cancer: A systematic review for the European Commission Initiative on Breast Cancer. J Med Screen. 2021:969141321993866.

37. Paci E, Broeders M, Hofvind S, Duffy SW, Euroscreen working g. The benefits and harms of breast cancer screening. Lancet. 2013;381(9869):800-1.

38. WHO Position Paper on Mammography Screening. WHO Guidelines Approved by the Guidelines Review Committee. Geneva 2014.

39. Paci E, Group EW. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. J Med Screen. 2012;19 Suppl 1:5-13.

40. Boyle P, Autier P, Bartelink H, Baselga J, Boffetta P, Burn J, et al. European Code Against Cancer and scientific justification: third version (2003). Ann Oncol. 2003;14(7):973-1005.

41. International Agency for Research on Cancer (IARC). Cancer Screening in the European Union – Report on the implementation of the Council Recommendation on cancer screening, European Commission, Brussels. Available at: https://ec.europa.eu/health/sites/health/files/major_chronic_diseases/docs/2017_cancerscreening_2ndreportimplementation_en.pdf. 2017.

42. European Commission Initiative on Breast Cancer Contributor G. European guidelines on breast cancer screening and diagnosis. Available at: https://healthcare-quality.jrc.ec.europa.eu/european-breast-cancer-guidelines. 2020.

43. Schunemann HJ, Lerda D, Quinn C, Follmann M, Alonso-Coello P, Rossi PG, et al. Breast Cancer Screening and Diagnosis: A Synopsis of the European Breast Guidelines. Ann Intern Med. 2020;172(1):46-56.

44. 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.

45. Ascunce N, Delfrade J, Salas D, Zubizarreta R, Ederra M, en nombre de Red de Programas de Cribado de C. [Breast cancer screening: characteristics and results of the Spanish programs]. Med Clin (Barc). 2013;141(1):13-23.

46. Red de Programas de Cribado de Cáncer. Programas de Cribado de Cáncer de Mama. Informe de evaluación 2017. Available at: www.cribadocancer.es. 2017.

47. van Schoor G, Moss SM, Otten JD, Donders R, Paap E, den Heeten GJ, et al. Increasingly strong reduction in breast cancer mortality due to screening. Br J Cancer. 2011;104(6):910-4.

48. Cronin KA, Feuer EJ, Clarke LD, Plevritis SK. Impact of adjuvant therapy and mammography on U.S. mortality from 1975 to 2000: comparison of mortality results from the cisnet breast cancer base case analysis. J Natl Cancer Inst Monogr. 2006(36):112-21.

49. Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. Lancet. 2001;358(9290):1340-2.

50. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? Lancet. 2000;355(9198):129-34.

51. de Koning HJ. Mammographic screening: evidence from randomised controlled trials. Ann Oncol. 2003;14(8):1185-9.

52. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database Syst Rev. 2006(4):CD001877.

53. Armstrong K, Moye E, Williams S, Berlin JA, Reynolds EE. Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. Ann Intern Med. 2007;146(7):516-26.

54. Morrison AS. Screening in chronic disease: New York: Oxford University Press; 1992.

55. Brett J, Austoker J. Women who are recalled for further investigation for breast screening: psychological consequences 3 years after recall and factors affecting re-attendance. J Public Health Med. 2001;23(4):292-300.

56. Brett J, Bankhead C, Henderson B, Watson E, Austoker J. The psychological impact of mammographic screening. A systematic review. Psychooncology. 2005;14(11):917-38.

57. Hofvind S, Ponti A, Patnick J, Ascunce N, Njor S, Broeders M, et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. J Med Screen. 2012;19 Suppl 1:57-66.

58. Carles M, Vilaprinyo E, Cots F, Gregori A, Pla R, Roman R, et al. Cost-effectiveness of early detection of breast cancer in Catalonia (Spain). BMC Cancer. 2011;11:192.

59. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. Ann Intern Med. 2007;146(7):502-10.

60. von Euler-Chelpin M, Risor LM, Thorsted BL, Vejborg I. Risk of breast cancer after false-positive test results in screening mammography. J Natl Cancer Inst. 2012;104(9):682-9.

61. Roman R, Sala M, De La Vega M, Natal C, Galceran J, Gonzalez-Roman I, et al. Effect of false-positives and women's characteristics on long-term adherence to breast cancer screening. Breast Cancer Res Treat. 2011;130(2):543-52.

62. Roman M, Castells X, Hofvind S, von Euler-Chelpin M. Risk of breast cancer after false-positive results in mammographic screening. Cancer Med. 2016;5(6):1298-306.

63. Roman M, Hofvind S, von Euler-Chelpin M, Castells X. Long-term risk of screen-detected and interval breast cancer after false-positive results at mammography screening: joint analysis of three national cohorts. Br J Cancer. 2019;120(2):269-75.

64. Roman R, Sala M, Salas D, Ascunce N, Zubizarreta R, Castells X. Effect of protocol-related variables and women's characteristics on the cumulative false-positive risk in breast cancer screening. Ann Oncol. 2012;23(1):104-11.

65. Puliti D, Miccinesi G, Zappa M, Manneschi G, Crocetti E, Paci E. Balancing harms and benefits of service mammography screening programs: a cohort study. Breast Cancer Res. 2012;14(1):R9.

66. Mandrik O, Zielonke N, Meheus F, Severens JLH, Guha N, Herrero Acosta R, et al. Systematic reviews as a 'lens of evidence': Determinants of benefits and harms of breast cancer screening. Int J Cancer. 2019;145(4):994-1006.

67. Blanch J, Sala M, Ibanez J, Domingo L, Fernandez B, Otegi A, et al. Impact of risk factors on different interval cancer subtypes in a population-based breast cancer screening programme. PLoS One. 2014;9(10):e110207.

68. Porter PL, El-Bastawissi AY, Mandelson MT, Lin MG, Khalid N, Watney EA, et al. Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst. 1999;91(23):2020-8.

69. Domingo L, Salas D, Zubizarreta R, Bare M, Sarriugarte G, Barata T, et al. Tumor phenotype and breast density in distinct categories of interval cancer: results of population-based mammography screening in Spain. Breast Cancer Res. 2014;16(1):R3.

70. Sala M, Domingo L, Louro J, Tora-Rocamora I, Bare M, Ferrer J, et al. Survival and Disease-Free Survival by Breast Density and Phenotype in Interval Breast Cancers. Cancer Epidemiol Biomarkers Prev. 2018;27(8):908-16.

71. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med. 2012;367(21):1998-2005.

72. Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. JAMA Intern Med. 2014;174(3):448-54.

73. Etzioni R, Xia J, Hubbard R, Weiss NS, Gulati R. A reality check for overdiagnosis estimates associated with breast cancer screening. J Natl Cancer Inst. 2014;106(12).

74. 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.

75. 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.

76. 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.

77. Roman M, Sala M, Domingo L, Posso M, Louro J, Castells X. Personalized breast cancer screening strategies: A systematic review and quality assessment. PLoS One. 2019;14(12):e0226352.

78. Yuan WH, Hsu HC, Chen YY, Wu CH. Supplemental breast cancer-screening ultrasonography in women with dense

breasts: a systematic review and meta-analysis. Br J Cancer. 2020;123(4):673-88.

79. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. N Engl J Med. 2019;381(22):2091-102.

80. van Ravesteyn NT, Schechter CB, Hampton JM, Alagoz O, van den Broek JJ, Kerlikowske K, et al. Trade-Offs Between Harms and Benefits of Different Breast Cancer Screening Intervals Among Low-Risk Women. J Natl Cancer Inst. 2021.

81. Esserman LJ, Study W, Athena I. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. NPJ Breast Cancer. 2017;3:34.

82. MyPeBS. Randomized Comparison Of Risk-Stratified versus Standard Breast Cancer Screening In European Women Aged 40-70 (MyPeBS) [Internet]. 2017. Available from: https://www.mypebs.eu/.

83. National Cancer Institute. Cancer Stat Facts: Female Breast Cancer. Available at: https://seer.cancer.gov/statfacts/html/breast.html. 2020.

84. Clavel-Chapelon F, Gerber M. Reproductive factors and breast cancer risk. Do they differ according to age at diagnosis? Breast Cancer Res Treat. 2002;72(2):107-15.

85. Colditz GA, Willett WC, Hunter DJ, Stampfer MJ, Manson JE, Hennekens CH, et al. Family history, age, and risk of breast cancer. Prospective data from the Nurses' Health Study. JAMA. 1993;270(3):338-43.

86. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer. 1997;71(5):800-9.

87. Brewer HR, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Family history and risk of breast cancer: an analysis accounting for family structure. Breast Cancer Res Treat. 2017;165(1):193-200.

88. Thompson D, Easton D. The genetic epidemiology of breast cancer genes. J Mammary Gland Biol Neoplasia. 2004;9(3):221-36.

89. Colditz GA, Kaphingst KA, Hankinson SE, Rosner B. Family history and risk of breast cancer: nurses' health study. Breast Cancer Res Treat. 2012;133(3):1097-104.

90. Collaborative Group on Hormonal Factors in Breast C. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet. 2001;358(9291):1389-99.

91. Carter CL, Corle DK, Micozzi MS, Schatzkin A, Taylor PR. A prospective study of the development of breast cancer in 16,692 women with benign breast disease. Am J Epidemiol. 1988;128(3):467-77.

92. 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.

93. Dyrstad SW, Yan Y, Fowler AM, Colditz GA. Breast cancer risk associated with benign breast disease: systematic review and meta-analysis. Breast Cancer Res Treat. 2015;149(3):569-75.

94. 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.

95. Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. N Engl J Med. 1985;312(3):146-51.

96. 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.

97. 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.

98. Sickles EDO, CJ.; Bassett, LW.; et al. ACR BI-RADS® Mammography. In: ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. Reston, VA, American College of Radiology. 2013.

99. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med. 2007;356(3):227-36.

100. 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.

101. Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. Nature. 1983;303(5920):767-70.

102. 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.

103. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. Ann Intern Med. 2008;148(5):337-47.

104. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med. 2004;23(7):1111-30.

105. Colditz GA, Rosner B. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the Nurses' Health Study. Am J Epidemiol. 2000;152(10):950-64.

106. Colditz GA, Rosner BA, Chen WY, Holmes MD, Hankinson SE. Risk factors for breast cancer according to estrogen and progesterone receptor status. J Natl Cancer Inst. 2004;96(3):218-28.

107. Rosner B, Colditz GA. Nurses' health study: log-incidence mathematical model of breast cancer incidence. J Natl Cancer Inst. 1996;88(6):359-64.

108. Rosner BA, Colditz GA, Hankinson SE, Sullivan-Halley J, Lacey JV, Jr., Bernstein L. Validation of Rosner-Colditz breast cancer incidence model using an independent data set, the California Teachers Study. Breast Cancer Res Treat. 2013;142(1):187-202.

109. 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.

110. 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.

111. Pashayan N, Antoniou AC, Ivanus U, Esserman LJ, Easton DF, French D, et al. Personalized early detection and prevention of breast cancer: ENVISION consensus statement. Nat Rev Clin Oncol. 2020;17(11):687-705.

112. Kleinbaum DG. Survival Analysis, a Self-Learning Text. Biometrical Journal. 1998;40(1):107-8.

113. Zheng Y, Heagerty PJ. Partly conditional survival models for longitudinal data. Biometrics. 2005;61(2):379-91.

114. Hardin JW. Generalized estimating equations. Hardin, Hardin. Boca Raton, Fla.: CRC Press; 2013.

115. Freedman DA. On the So-Called 'Huber Sandwich Estimator' and 'Robust Standard Errors'. The American Statistician. 2006;60:299–302.

116. 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):774-81.

117. Efron BT, R. Improvements on cross-validation: The .632+ bootstrap method. J Amer Statist Assoc. 1997;92(438):548-60.

118. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association. 1958;53(282):457-81.

119. Breslow NE, Day NE. Statistical methods in cancer research. IARC Workshop 25-27 May 1983. IARC Sci Publ. 1987:82.

120. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27(2):157-72; discussion 207-12.

121. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29-36.

122. Anothaisintawee T, Teerawattananon Y, Wiratkapun C, Kasamesup V, Thakkinstian A. Risk prediction models of breast cancer: a systematic review of model performances. Breast Cancer Res Treat. 2012;133(1):1-10.

123. Cintolo-Gonzalez JA, Braun D, Blackford AL, Mazzola E, Acar A, Plichta JK, et al. Breast cancer risk models: a comprehensive

overview of existing models, validation, and clinical applications. Breast Cancer Res Treat. 2017;164(2):263-84.

124. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869-74.

125. Fletcher RH, Fletcher SW, Fletcher GW. Clinical Epidemiology: The essentials. 5th edition: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2014.

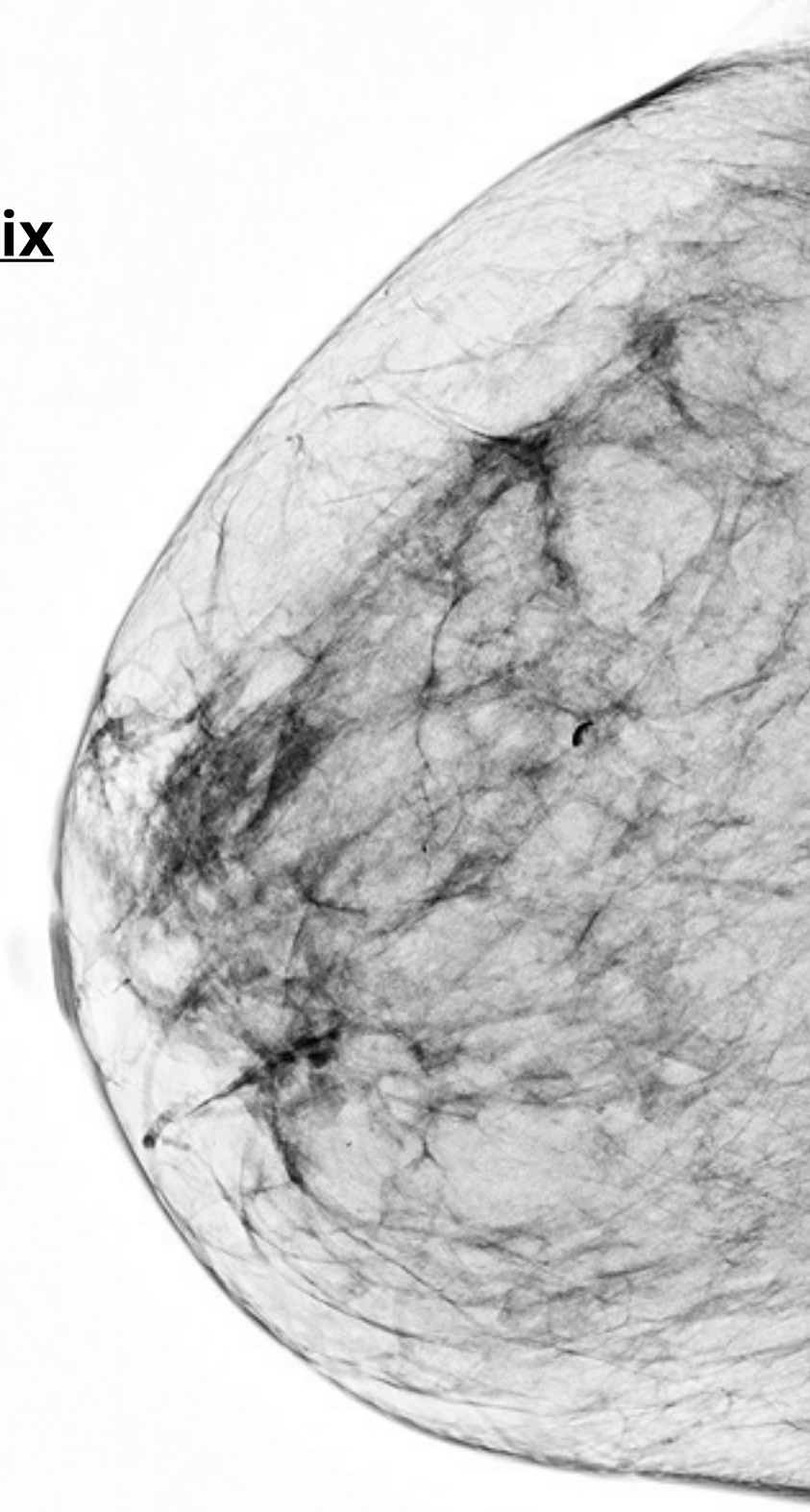
126. Marrugat J, D'Agostino R, Sullivan L, Elosua R, Wilson P, Ordovas J, et al. An adaptation of the Framingham coronary heart disease risk function to European Mediterranean areas. J Epidemiol Community Health. 2003;57(8):634-8.

127. Freer PE. Mammographic breast density: impact on breast cancer risk and implications for screening. Radiographics. 2015;35(2):302-15.

128. Sprague BL, Conant EF, Onega T, Garcia MP, Beaber EF, Herschorn SD, et al. Variation in Mammographic Breast Density Assessments Among Radiologists in Clinical Practice: A Multicenter Observational Study. Ann Intern Med. 2016;165(7):457-64.

129. Tice JA, O'Meara ES, Weaver DL, Vachon C, Ballard-Barbash R, Kerlikowske K. Benign breast disease, mammographic breast density, and the risk of breast cancer. J Natl Cancer Inst. 2013;105(14):1043-9.

Appendix



Appendix 1

Other articles produced with the database of the thesis

Article 6

Title: Risk of breast cancer two years after a benign biopsy depends on the mammographic feature prompting recall

Authors: Vernet-Tomás M, Louro J, Román M, Saladié F, Posso M, Prieto M, Vázquez I, Baré M, Peñalva L, Vidal C, Bargalló X, Sánchez M, Ferrer J, Espinàs JA, Quintana MJ, Rodriguez-Arana A & Castells X on behalf of the BELE Study Group

Journal: Maturitas. 2021 Feb;144:53-59

Impact Factor: 3.63 (Q1 Obstetrics & Gynecology)

DOI: 10.1016/j.maturitas.2020.10.024

Abstract:

Objective: We aimed to explore whether the type of mammographic feature prompting a false-positive recall (FPR) during mammography screening influences the risk and timing of breast cancer diagnosis, particularly if assessed with invasive procedures.

Study design: We included information on women screened and recalled for further assessment in Spain between 1994 and 2015, with follow-up until 2017, categorizing FPRs by the assessment (noninvasive or invasive) and mammographic feature prompting the recall.

Main outcome measures: Breast cancer rates in the first two years after FPR (first period) and after two years (second period).

Results: The study included 99,825 women with FPRs. In both periods, the breast cancer rate was higher in the invasive assessment group than in the noninvasive group (first period 12‰ vs 1.9‰, $p < 0.001$; second period 4.4‰ vs 3.1‰, $p < 0.001$). During the first period, the invasive

assessment group showed diverse breast cancer rates for each type of mammographic feature, with a higher rate for asymmetric density (31.9‰). When the second period was compared with the first, the breast cancer rate decreased in the invasive assessment group (from 12‰ to 4.4‰, $p < 0.001$) and increased in the noninvasive assessment group (from 1.9‰ to 3.1‰, $p < 0.001$).

Conclusion: In the context of mammography screening, the risk of breast cancer diagnosis during the first two years after FPR was particularly high for women undergoing invasive assessment; importantly, the risk was modified by type of mammographic feature prompting the recall. This information could help to individualize follow-up after exclusion of malignancy.

Article 7

Title: Changes in mammographic density over time and the risk of breast cancer: An observational cohort study

Authors: Román M, Sala M, Baré M, Posso M, Vidal C, Louro J, Sánchez M, Peñalva L & Castells X on behalf of the BELE Study Group

Journal: Breast. 2019 Aug;46:108-115

Impact Factor: 3.75 (Q1 Obstetrics & Gynecology)

DOI: 10.1016/j.breast.2019.04.007

Abstract:

Background: The effect of changes in mammographic density over time on the risk of breast cancer remains inconclusive.

Methods: We used information from four centres of the Breast Cancer Screening Program in Spain in the period 1996–2015. We analysed individual level data from 117,388 women first screened age 50–54, with at least two screening examinations. Breast density was determined using the BI-RADS classification (A to D in increasing order) at earliest and latest screening examination. Adjusted Poisson regression models were used to estimate the relative risk (RR) and 95% confidence intervals (95%CI) of the association between changes in mammographic density and breast cancer risk over time.

Results: During an average 5.8 years of follow-up, 1592 (1.36%) women had a breast cancer diagnosis. An increase in density category increased breast cancer risk, and a decrease in density decreased the risk, compared with women who remained in the same BI-RADS category. Women whose density category increased from B to C or B to D had a RR of 1.55 (95%CI = 1.24–1.94) and 2.32 (95%CI = 1.48–3.63),

respectively. The RR for women whose density increased from C to D was 1.51 (95%CI = 1.03–2.22). Changes in BI-RADS density were similarly associated with the risk for invasive cancer than for ductal carcinoma in situ.

Conclusions: Although a modest proportion of women changed BI-RADS density category, mammographic density changes modulated the risk of breast cancer and identified women at a differential risk. Using two longitudinal measures of BI-RADS density could help target women for risk-based screening strategies.

Appendix 2

Communications to congresses derived from the articles of this thesis

Louro J, Román M, Vidal C, Baré M, Sánchez M, Peñalva L,
Sala M & Castells X
Mammographic breast density in population-based screening programs in Spain
XXXV Congress of the Spanish Society of Epidemiology
Barcelona (Spain), 6-8 September 2017. Gac Sanit 2017; 31(Espec Congr):110

Louro J, Vidal C, Baré M, Sánchez M, Peñalva L & Sala M
Indicadores de proceso según la densidad mamaria en programas poblacionales de detección precoz de cáncer de mama en España
3er Congreso Español de la Mama
Madrid (Spain). 19-21 October 2017

Posso M, Louro J, Sala M, Román M, Domingo L & Castells X
Modelos de predicción de riesgo individual del cáncer de mama: Revisión sistemática
XXXVI Congress of the Spanish Society of Epidemiology
Lisbon (Portugal), 11-14 September 2018. Gac Sanit. 2018;32(Espec Congr): 120

Louro J, Román M, Quitana MJ, Saladié F, Prieto M, Bargalló X,
Posso M, Sala M , Castells X
An individualized breast cancer risk prediction model to personalize mammography screening
XXXVI Congress of the Spanish Society of Epidemiology
Lisbon (Portugal), 11-14 September 2018. Gac Sanit. 2018;32(Espec Congr): 311

Louro J, Castells X, Quintana MJ, Posso M, Sala M & Román M
An individualized breast cancer risk prediction model to personalize mammography screening
International Cancer Screening Conference
Rotterdam (The Netherlands), 3-5 June 2019

Posso M, Louro J, Román M, Domingo L, Castell X & Sala M
Individualized breast cancer risk prediction models in average-risk women: a systematic review and quality assessment
International Cancer Screening Conference
Rotterdam (The Netherlands), 3-5 June 2019

Louro J, Castells X, Alcántara R, Posso M, Prieto M, L. Peñalva, J. del Riego, C. Vidal & Román M
Mammographic density, benign breast disease, and the risk of breast cancer over time
XXXVII Congress of the Spanish Society of Epidemiology
Oviedo (España), 3-6 September 2019

Louro J, Román M, Posso M, Comerma L, Vidal C, Prieto M, Saladié F, Baré M, Sánchez M, Quintana MJ, Bargalló X, Ferrer J, Peñalva L, Sala M & Castells X
Differences in breast cancer risk after a benign breast disease according to the screening type
XII European Breast Cancer Conference.
Online due to Covid-19 pandemy, 2-3 October 2020
Abstract selected for press release

Appendix 3

Database protocol (in Spanish)

Contenido

Lesiones benignas de mama, densidad mamaria y asociación con el riesgo de cáncer en una cohorte de mujeres cribadas

Protocolo de las variables de estudio

Población de estudio

Todas las mujeres participantes en los programas de detección precoz de cáncer de mama que participan en el proyecto, con al menos una mamografía de cribado realizada desde el inicio de estos hasta el 31 de diciembre del 2015, ambos incluidos.

Los programas disponen de un identificador único por mujer que permite enlazar el historial de participaciones de una misma mujer a través de las distintas invitaciones para participar en el cribado mamográfico (como si de una cohorte se tratase). En estas condiciones, para cada mujer y para cada participación, se dispone de información sobre el resultado de la lectura mamográfica de cribado.

Se incluirá en la base de datos a las mujeres dentro de la edad diana de los programas de cribado, con al menos 1 cribado desde el inicio del programa hasta el 31 de diciembre del 2015 (al menos una vez participante), pudiendo tener su primer cribado (cribado inicial) en cualquier ronda del programa.

Se excluyen del estudio las mujeres no participantes en ninguna ronda dentro del periodo de estudio y aquellas con historia de cáncer mamario anterior al primer cribado.

Definición de cáncer de cribado

Se considerarán tumores detectados en el cribado todos aquellos tumores de mama primarios diagnosticados en mujeres participantes en alguno de los programas que forman parte del proyecto. Por definición estos tumores se detectan a lo largo del periodo de estudio a través de la realización de una mamografía de cribado, con sospecha de malignidad y diagnóstico final de cáncer.

Se incluirán todos los tumores malignos primarios de mama, tanto invasivos como in situ, de acuerdo con la Clasificación Internacional de Enfermedades para Oncología (CIE-O 3ª Edición).

Si una mujer presenta simultáneamente más de un tumor, se recogerá únicamente la información correspondiente al tumor que en el momento del diagnóstico presentara un estadio más avanzado (peor pronóstico), independientemente de que los tumores afectaran a una sola o ambas mamas.

Contexto

Con el fin de mejorar la comprensión de las variables que vamos a recoger para el estudio de los *"determinantes, evolución y características biológicas del cáncer de mama detectado en una cohorte de mujeres cribadas"* proponemos el siguiente

Ejemplo:
Imaginemos un programa de cribado iniciado en 1996, con cinco rondas, y una mujer nacida el 15/01/1948 y convocada a partir de la segunda ronda (fecha primera convocatoria 01/03/1998) y hasta la quinta. Supongamos que la mujer ha participado en la segunda (1998), cuarta (2002) y quinta (2004) ronda de una unidad radiológica del programa. Supongamos que su tercera mamografía de control en el programa (ronda 5 del programa) se realiza el 23/04/04. El programa contempla la realización de mamografía con única proyección y lectura doble con consenso. Tras la lectura por parte de los radiólogos no es posible descartar malignidad, ya que encuentran un patrón distorsionante con calcificaciones, con una clasificación de bi-rads 3 'probablemente benigno'. Además, se clasifica la densidad mamaria de la mujer como 3 'Densidad heterogénea' según la clasificación bi-rads. En fechas posteriores a la mamografía se le realizan consecutivamente las siguientes pruebas: ecografía, paaf y biopsia quirúrgica. Finalmente, no se diagnostica cáncer de mama, pero se le recomienda a la mujer la realización de una mamografía intermedia a los 6 meses. Pasados los 6 meses (15/11/04), se le realiza la mamografía y justo después de la mamografía (en esa misma fecha), se le realiza una ecografía y una biopsia escisional percutánea, con un resultado definitivo histológico de cáncer de mama.

Descripción de las variables

(ver categorías al final del documento)

Nota: Algunos programas aprovecharán la información proporcionada para el estudio BELE-1 y únicamente cargarán información parcial en este proyecto (seguimiento de las mujeres ya incluidas en la cohorte BELE-1, e información de las mujeres que se incorporan al cribado por primera vez en el periodo 2012–2015). Ya que todas las variables de este protocolo son comunes con el estudio BELE-1, se indica como nota a pie de página cuando sea preciso o deseable que las codificaciones asignadas coincidan entre la información del estudio BELE-1 y la nueva carga de datos..

1. Programas (1 registro por programa)

- ***Programa_id***[1]

Letras identificativas asignadas a cada uno de los programas participantes en el proyecto

  - *AS: Asturias*
  - *CT: Cantabria*
  - *GI: Girona*
  - *IC: ICO*
  - *PM: Parc Salut Mar*
  - *SB: Sabadell*
  - *SP: Sant Pau*
  - *TA: Tarragona*
  - *HC: Hospital clínic*
  - *GR: Vallès Oriental – Granollers*

Ejemplo:
'SP' (código asignado al programa del 'Hospital de Sant Pau')

- ***Nombre del Programa***

Descripción (etiqueta) del programa.

Ejemplo:
Sant PaU

[1] Para los programas participantes en el estudio BELE1 la clasificación asignada a la variable ***"Programa_id"*** será la misma.

2. Unidad Radiológica de Cribado (1-N registros por programa)

- ***URC_id[2]:***

Código de cada una de las Unidades de Cribado donde se realizan las exploraciones. Se entenderá como URCs, unidades organizativas independientes dentro de un mismo programa, sean estas fijas o móviles. Las URC se enumerarán de la siguiente forma: Letras identificativas del programa de cribado, seguidas de un número consecutivo. Se enumerarán las URC de forma sucesiva a partir del número 1, sin ningún orden específico En el caso de programas sin estructura de URC, bastará con poner las letras identificativas del programa de cribado seguidas del número 1 en esta variable.

Ejemplo:
Para la URC 1 de Asturias, tendremos el códico 'AS1'

- ***Descripción de la URC[3]***

Nombre y/o descripción (etiqueta) de la Unidad de Cribado donde se realizan las exploraciones. Los programas sin estructura de URC la etiquetarán con el nombre del propio programa. Aquellos programas que no identifiquen las URC, las etiquetaran con el nombre del programa seguido del número que le hayan asignado en la variable URC_id.

Ejemplo:
Para la URC 1 del ejemplo anterior (URC de Asturias), la descripción sería 'ASTURIAS1'

3. Mujeres (1 registro por mujer)

- ***Mujer_id[4]***

Número identificador, interno de los programas, de la mujer. Para un programa no puede haber dos mujeres con el mismo número identificador. Una mujer debe mantener durante todo el tiempo del estudio el mismo número, aunque cambie de unidad de exploración a lo largo del periodo.

---

[2] Para los programas participantes en el estudio BELE-1 la clasificación asignada a la variable ***"URC_id"*** será la misma que la asignada en el estudio BELE-1.

[3] Para los programas participantes en el estudio BELE-1 la ***"descripción de la URC"*** será la misma que la asignada en dicho estudio.

[4] Para los programas participantes en el estudio BELE-1 será **preciso** que este número identificador coincida con el identificador asignado en dicho estudio.

- ***Fecha de nacimiento***[5]

Fecha de nacimiento de la mujer. Valor 09/09/9999 si desconocido

Ejemplo:
15/01/1948

- ***Fecha 1ª citación en el programa***[6]

Fecha en la cual la mujer es invitada a participar por primera vez en la URC (independientemente de si participa o no, e independientemente si la fecha es anterior al 1 de enero del 2000). Valor 09/09/9999 si desconocido

Ejemplo:
01/03/1998

4. Episodios (1-N registros por mujer)

- ***Episodio***[7]

Indica el número ordinal de convocatoria (o invitación) de la mujer. El episodio incluye todo el proceso que transcurre desde la primera citación de una mujer para realizarse la mamografía rutinaria hasta el resultado final del episodio (cáncer, no cáncer, pendiente de exploraciones adicionales, pendiente mamografía intermedia, seguimiento incompleto, desconocido). Una mamografía intermedia no puede ser un episodio, aunque, dentro de un episodio, puede haber mamografías intermedias. **Solamente se van a registrar episodios de mujeres con mamografía realizada (participantes)**, de manera que los episodios de mujeres no participantes (cribado externo o no cribado) no generan registro en la base de datos. Por lo tanto, una mujer participante en una ronda de la URC concreta genera exactamente un episodio, y para esta misma mujer, la base de datos del proyecto actual contendrá tantos episodios como participaciones tenga la mujer desde el inicio del programa hasta el 31 de diciembre del 2011.

---

[5] Para los programas participantes en el estudio BELE-1 será **deseable** que esta fecha coincida con la fecha de nacimiento asignada en dicho estudio.

[6] Para los programas participantes en el estudio BELE-1 será **deseable** que esta fecha coincida con la fecha de 1ª citación en el programa asignada en dicho estudio.

[7] Para los programas participantes en el estudio BELE-1 será **preciso** que este identificador de episodios sea coherente con el identificador asignado en el BELE-1. Para las mujeres cribadas en el periodo 2012-2015 (nueva información a recoger), con participaciones anteriores a 2012, deberá de existir un orden lógico en la numeración de los episodios. Este identificador nos permitirá mantener la secuencia temporal entre los distintos estudios cuando se añada nueva información específica de este proyecto.

Se excluirán los episodios derivados de la petición de una mamografía por parte de una mujer por síntomas.

Ejemplo:
la base de datos de episodios contendría tres registros para esta mujer que serían rellenados con los números 1, 3 y 4 consecutivamente. El segundo episodio no formaría parte de la base de datos, ya que, a pesar de ser invitada, la mujer no participó en la 3ª ronda del programa y por lo tanto no generó registro.

- ***Número de cribado[8]***

Número ordinal que indica el número de exploraciones de cribado que una mujer lleva realizadas hasta el episodio objeto de estudio (incluyendo el mismo episodio). Este número siempre será menor o igual al número de episodio.

Ejemplo:
el campo número de cribados sería rellenado con los números 1, 2 y 3 para cada uno de los tres episodios que serían recogidos en la base de datos.

- ***Fecha mamografía de cribado***

Fecha de realización de la exploración de cribado (test de cribado)

Ejemplo:
23/04/04 (en el episodio codificado como 4)

- ***Número de estudios intermedios***

Número total de estudios con mamografía intermedia (controles avanzados) dentro del episodio objeto de estudio, realizados por indicación del programa independientemente de cuándo y dónde se realicen. Si dentro del episodio no se ha realizado ningún estudio intermedio, el valor de la variable será 0 (cero).

Ejemplo:
1 estudio intermedio (en el episodio codificado como 4)

- ***Resultado inicial de la mamografía de cribado***

Se indicará el **resultado inicial de la lectura mamográfica de cribado** para ese episodio, aunque posteriormente, y durante el mismo episodio, haya otros

[8] Para los programas participantes en el estudio BELE-1 será **preciso** que el número de cribado sea coherente con el valor asignado en el BELE-1. Para las mujeres cribadas en el periodo 2012-2015 (nueva información a recoger), con participaciones anteriores a 2012, deberá de existir un orden lógico en la numeración del cribado.

resultados y recomendaciones de otros test (mamografías intermedias, pruebas de imagen o exploraciones adicionales invasivas). En el caso de repetición técnica, se tomará como resultado inicial el que derive de la primera mamografía válida.

Se codificará de la siguiente manera:

- *1: Negativo.*
- *2: Exploraciones adicionales: mujeres a las que se les recomienda la realización de alguna prueba o exploración adicional para descartar malignidad independientemente de cuándo y dónde se realicen (estas pruebas pueden realizarse en la misma fecha de la exploración de cribado o en alguna fecha posterior previa reconvocatoria).*
- *3: Mamografía intermedia (Estudios intermedios): mujeres a las que a la vista del resultado de la mamografía de cribado se les recomienda la realización de una nueva mamografía antes de la secuencia que le correspondería de forma rutinaria (por ej. a los 3, 6 o 12 meses)*
- *99: Desconocido*

Ejemplo:
el campo resultado (inicial) de la mamografía sería rellenado con la etiqueta: exploraciones adicionales (en el episodio codificado como 4)

- ***Categoría mamografía de la prueba de cribado. (categórica)***

Corresponde a la clasificación BI-RADS de la lectura mamográfica del test de cribado, previo a la realización de exploraciones adicionales si fuesen necesarias. Se admite la categoría '0', en su caso. Se informará '99' en caso de BI-RADS desconocido.

- *0: Incierto*
- *1: Normal*
- *2: Benigna*
- *3: Probablemente Benigna*
- *4: Probablemente Maligna*
- *5: Maligna*
- *99: Desconocido*

Nota: *En el caso de los programas que recogen de manera independiente la clasificación Bi-Rads para cada mama, se informará aquel de los dos que indique una mayor sospecha de malignidad. La sospecha de malignidad de menor a mayor se establece del siguiente modo: 1 → 2 → 0 → 3 → 4 → 5*

Ejemplo:
el campo categoría mamográfica de la prueba de cribado sería rellenado con el valor 3: probablemente benigno (en el episodio codificado como 4)

- ***Densidad mamaria***

Se utilizará la clasificación cualitativa BI-RADS en 4 grupos para clasificar la mama según el nivel de tejido fibro-glandular. Se codificará el resultado global de ambas mamas. Se informará '99' en caso de densidad mamaria desconocida.

- *1: Completamente grasa (BI-RADS type 1)*
- *2: Densidad fibroglandular dispersa (BI-RADS type 2)*
- *3: Densidad heterogénea (BI-RADS type 3)*
- *4: Extremadamente densa (BI-RADS type 4)*
- *99: Desconocido*

Ejemplo:
el campo densidad mamaria sería rellenado con el valor 3: densidad heterogénea (en el episodio codificado como 4)

- ***Lateralidad (categórica)***

Indica en que mama se encuentra la sospecha radiológica de la lectura mamográfica. La categoría '*No pertinente*' se reserva para los casos con lectura mamográfica sin ningún hallazgo, y resultado inicial negativo. Si se han observado hallazgos en ambas mamas y se dispone de esta información se marcará la categoría '*Bilateral*'.

- *0: No pertinente*
- *1: Derecha*
- *2: Izquierda*
- *3: Bilateral*
- *99: Desconocido*

Ejemplo:
el campo lateralidad del episodio número 4 sería rellenado con el valor:
'2: Izquierda'

- ***Patrones radiológicos de la prueba de cribado (categórica)***

Hace referencia a los posibles patrones radiológicos observados durante la lectura mamográfica, sean estos malignos o benignos, e independientemente de si la mujer es reconvocada para hacerse exploraciones adicionales.
Para poder recoger información sobre múltiples patrones en los casos en que

se disponga de esta información, la descripción de la mamografía de la prueba de cribado se recoge en un bloque de 5 variables para cada mama. De esta manera se podrán informar diferentes combinaciones de patrones en cada mama.

Nota: *En el caso de que la* ***'lateralidad'*** *sea desconocida ('99: Desconocido'), pero se conozca el* ***'patrón radiológico'*** *de la lectura mamográfica, se rellenará esta información en la mama derecha únicamente. Esta codificación por defecto indicará que a pesar de que la* ***'lateralidad'*** *es desconocida, el* ***'patrón radiológico'*** *informado corresponde con alguna de las dos mamas, pero desconocemos exactamente a cuál.*

Mama derecha

- *Masa_D*

Imagen de Masa en mama derecha

  - *0: No*
  - *1: Sí*
  - *99: Desconocido*

- *Distorsión_D*

Imagen de Distorsión en mama derecha

  - *0: No*
  - *1: Sí*
  - *199: Desconocido*

- Calcificaciones_D

Imagen de Calcificaciones en mama derecha

  - *0: No*
  - *1: Sí*
  - *99: Desconocido*

- *Asimetría_D*

Imagen de Asimetría en mama derecha

  - *0: No*
  - *1: Sí*
  - *99: Desconocido*

- *Otros_D*

Observación en mama derecha de: Alteraciones de la piel / pezón; Prótesis; Cuerpos extraños; Cambios post-cirugía; Otros aspectos

- *0: No*
- *1: Sí*
- *99: Desconocido*

Ejemplo:
las variables 'masa_d', 'distorsión_d', 'calcificaciones_d', 'asimetría_d', y 'otros_d', serían rellenadas todas con el valor '0: No' en el episodio número 4

Mama izquierda

- *Masa_I*

Imagen de Masa en mama izquierda

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Distorsión_I*

Imagen de Distorsión en mama izquierda

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Calcificaciones_I*

Imagen de Calcificaciones en mama izquierda

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Asimetría_I*

Imagen de Asimetría en mama izquierda

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Otros_I*

Observación en mama izquierda de: Alteraciones de la piel / pezón; Prótesis; Cuerpos extraños; Cambios postcirugía; Otros aspectos

- *0: No*
- *1: Sí*

- *99: Desconocido*

Ejemplo:
las variabales 'distorsión_i ' y 'calcificaciones_i' serían rellenadas con el valor '1: Sí' en el episodio número 4, y las variables 'masa_i', 'asimetría_l', y 'otros_d' serían rellenadas con el valor '0: No' en ese mismo episodio número 4

- ***Resultado final del episodio***

El resultado final del episodio es el diagnóstico definitivo del mismo, tenga o no tenga exploraciones adicionales, tenga o no tenga mamografías intermedias, sea cual sea la casuística durante el episodio.

Valores posibles del resultado final del episodio y definición:

- *1: Cáncer: el diagnóstico definitivo (histológico) del episodio es de cáncer de mama.*
- *2: No cáncer: el resultado del episodio es de no cáncer de mama y se recomienda a la mujer un control rutinario.*
- *3: Pendiente estudios intermedios: mujeres con episodios incompletos. Como se incluirá a todas las mujeres con alguna mamografía realizada desde el inicio del programa hasta el 31 de diciembre de 2015 es posible que durante el último período haya mujeres que están todavía pendientes de la realización de un estudio intermedio y no se pueda establecer un resultado final del episodio.*
- *5: Seguimiento incompleto: mujeres a las que no se puede completar las exploraciones recomendadas en el episodio (pruebas adicionales, exploraciones intermedias, etc) por ejemplo por rechazo a las mismas por parte de la mujer, fallecimiento antes de completar proceso diagnóstico etc.*
- *99: Desconocido: corresponden a episodios de cribado pendientes de resolución, de los que se desconoce la información necesaria para asignarles un resultado.*

Ejemplo:
el campo resultado final del episodio número 4 sería rellenado con la etiqueta: '1: Cáncer'

5. Exploraciones adicionales (0-N registros por episodio)

- ***N prueba***

Es el contador del número de pruebas realizadas a una misma mujer dentro de cada episodio. Para cada prueba realizada a la mujer se le asignará el número de prueba que le corresponde.

- ***Prueba_id***

Se recogerán todas las pruebas realizadas en cada episodio, aunque sean del mismo tipo (Ej: dos PAAF). En el caso de una mujer-episodio con pruebas en ambas mamas en la misma fecha, se registrarán todas las pruebas realizadas. Se codificará de la siguiente manera:

- *1. Pruebas no Invasivas*
- *1.1. Otras proyecciones mamográficas*
- *1.2. Ecografía*
- *1.3. Resonancia magnética*
- *1.4. Otras pruebas no invasivas*
- *1.9. Prueba no invasiva desconocida*
- *2. Pruebas Invasivas*
- *2.1. PAAF*
- *2.2. Biopsia aguja gruesa*
- *2.3. Biopsia asistida por vacío*
- *2.4. Biopsia escisional percutánea*
- *2.5. Biopsia quirúrgica*
- *2.6. Otras pruebas invasivas*
- *2.9. Prueba invasiva desconocida*
- *99. Desconocido*

Ejemplo:
Dentro del episodio número 4 (tercer cribado de la mujer, ronda 5 del programa), en la base de datos de exploraciones adicionales constarían las siguientes 5 pruebas: ecografía, paaf, biopsia quirúrgica, ecografía, y biopsia escisional percutánea

- ***Fecha_prueba***

Fecha de realización de cada una de las Exploraciones adicionales realizadas. Se informará el valor 09/09/9999 si la fecha de la prueba es desconocida.

6. Variables de las mujeres (1 registro por episodio)

El objetivo es obtener información sobre el perfil de la mujer en cada episodio.

- ***THS (categórica)***

Se evaluará el uso del THS en el momento de hacerse la exploración de cribado. Se considerará que la mujer usaba THS si lo tomaba en el momento de la mamografía de cribado o en los 6 meses anteriores.

  - *1: Sí, en el momento del cribado o en los 6 meses anteriores.*
  - *2: No*
  - *99: Desconocido*

- ***Menopausia (categórica)***

Se recogerá el estado menopáusico de la mujer, es decir, si es posmenopáusica o en cambio pre- ó peri-menopáusica.

  - *1: Posmenopáusica*
  - *2: Premenopáusica o perimenopáusica*
  - *99: Desconocido*

- ***Antecedentes personales de prueba invasiva con resultado benigno o declaración por parte de la mujer de patología mamaria benigna***

Cuando se tenga conocimiento de que la mujer ha sufrido alguna prueba invasiva con resultado benigno, fuera del contexto del programa de cribado.

  - *1: Sí*
  - *2: No*
  - *99: Desconocido*

- ***Antecedentes familiares de cáncer de mama (categórica)***

Se considerará que una mujer tiene antecedentes familiares de cáncer de mama cuando tiene algún familiar de primer grado afectado (madre, hermanas o hijas) con cáncer de mama.

  - *1: Sí*
  - *2: No*
  - *99: Desconocido*

- ***Edad menopausia (numérica)***

Se informará la edad de la mujer en la menopausia cuando se recoja esta información. Se informará como '9: No pertinente' cuando la mujer sea pre-menopáusica.

- *9: No pertinente*
- *99: Desconocido*

- ***Edad de la menarquia (numérica)***

Se recogerá de manera numérica la edad en que la mujer declara haber tenido la menarquia.

- *99: Desconocido*

Nota: *Dado que la información de las variables de la Tabla 'Variables de las mujeres' se recoge de distinta forma según los programas, se dan las siguientes instrucciones de implementación:*

- *Si se recoge la información de la variable **para todos los episodios**, se informará esta variable en cada episodio según la codificación establecida en este protocolo. El valor 99 (desconocido) se utilizará en los casos en que no se disponga de información.*
- *Si se recoge la información de la variable en **un único campo que se sobrescribe** en las sucesivas visitas de la mujer (siendo por tanto los valores previos desconocidos), se informará el valor de esta variable en el último episodio de dicha mujer registrado en la base de datos. Los episodios anteriores de la mujer para esta variable se codificarán como 99 (desconocido).*
- *Si se recoge la información de la variable **únicamente en la primera participación de la mujer**, se informará el valor de esta variable en el primer episodio registrado de dicha mujer. Los episodios posteriores de esta mujer se codificarán como 99 (desconocido) para esta variable.*
- *Si la variable NO se recoge, todos los episodios se informarán como 99 desconocido).*

7. Lesiones benignas (1-N registros por mujer, máximo un registro por episodio)

Se recogerá en este apartado la información referente a las lesiones (benignas o con resultado negativo) identificadas durante el proceso de cribado, clasificadas mediante la realización de alguna prueba de carácter invasivo (PAAF, core biopsia, biopsia quirúrgica, etc).

1. Los casos registrados en esta tabla serán siempre episodios con resultado final **distinto** de "Cáncer". Los episodios con resultado final "Cáncer" generarán un registro en la tabla tumores, y no en esta.

2. Se generará un único registro para todo el conjunto de posibles pruebas invasivas que se realicen a la mujer en el episodio.
3. Si a una mujer se le realizan pruebas en ambas mamas (PAAF en la mama derecha, BAG en la mama izquierda), se recogerá únicamente la lesión de peor pronóstico.
4. En el caso de una mujer a la que se le realiza más de una prueba invasiva en el episodio (una PAAF y una biopsia quirúrgica, por ejemplo), se recogerá el resultado histológico de la prueba más fiable, que generalmente será la última prueba realizada (biopsia quirúrgica en este caso).

- ***Histología básica de las citologías (categórica)***

Nota: *Ya que en general las citologías (PAAF) no permiten una caracterización detallada de la histología, pero si una clasificación genérica del tipo de lesión, se propone esta primera clasificación genérica para no perder la información de las citologías (PAAF).*

- *0: Negativo células malignas*
- *1: Lesión no proliferativa*
- *2: Lesión proliferativa sin atipia*
- *3: Lesión proliferativa con atipia*
- *9: No pertinente*
- *99: Desconocido*

Nota: *Se excluye de la clasificación la categoría 'No concluyente'. En el caso de no tener constancia de la realización de más pruebas invasivas un resultado 'No concluyente' de la citología será codificado como '99: Desconocido'*
Nota: *Si en el episodio no se ha realizado ninguna citología (pero sí alguna biopsia), la "Histología básica de las citologías" se codificará como "9: No pertinente".*

- ***Histología de la lesión benigna***

Esta variable permitirá diferenciar el tipo histológico de las lesiones según la Clasificación Internacional de Enfermedades para Oncología CIE-O (3ª edición). Ver anexo 1.

- *99: Desconocido*

Nota: *Si la única prueba de carácter invasivo realizada en el episodio es una citología y se dispone de un resultado (sugerente) de la 'Histología de las lesiones benignas' para esta prueba, se informará esta variable, informándose como "99: Desconocido" cuando no se disponga de esta información.*
Nota: *Si en un mismo episodio se ha realizado alguna biopsia además de la*

citología, se informará la 'Histología de las lesiones benignas' con la información de la biopsia siempre que sea posible, ya que se considera esta prueba más fiable que la citología.

- ***Lateralidad (categórica)***

Indica en que mama se encuentra la lesión sospechosa que ha generado las exploraciones adicionales. Se informará como 'Desconocido' cuando no se disponga de información sobre la lateralidad de la lesión. Si se sabe que la mujer presentaba lesiones en ambas mamas se marcará la categoría 'Bilateral', a pesar de que únicamente se recogerán los resultados histológicos de la lesión de peor pronóstico.

- *1: Derecha*
- *2: Izquierda*
- *3: Bilateral*
- *99: Desconocido*

8. Variables referentes a los tumores

Se recogerá en este apartado la información referente a los tumores diagnosticados en el proceso de cribado. Tal y como se especifica en las definiciones iniciales de este mismo protocolo, si una mujer presenta simultáneamente más de un tumor, se recogerá únicamente la información correspondiente al tumor que en el momento del diagnóstico presentara un estadio más avanzado (peor pronóstico), independientemente de que los tumores afectaran a una sola o ambas mamas.

- ***Método de detección***

Se informará si el tumor ha sido detectado en el cribado mamográfico o como cáncer de intervalo.

- *1: Cribado*
- *2: Cáncer de Intervalo*
- *3: Fuera de cribado*

- ***Histología tumores***

Esta variable permitirá diferenciar el tipo histológico de los tumores según la Clasificación Internacional de Enfermedades para Oncología CIE-O (3ª edición) Ver anexo 2

- *99: Desconocido*

- ***Comportamiento tumoral***
    - *Se informará si es Ductal in situ o Invasivo*
    - *1: Ductal in situ*
    - *2: Invasivo*
    - *99: Desconocido*

- ***Lateralidad (categórica)***

Indica en que mama se encuentra el tumor diagnosticado. Se informará como 'Desconocido' cuando no se disponga de información sobre la lateralidad de la lesión. Si se sabe que la mujer presentaba tumores en ambas mamas se marcará la categoría 'Bilateral', a pesar de que únicamente se recogerá información del tumor de peor pronóstico en el momento del diagnóstico (estadio más avanzado).

- *1: Derecha*
- *2: Izquierda*
- *3: Bilateral*
- *99: Desconocido*

- ***Tamaño tumor (categórica)***

Descripción categórica del tamaño del tumor, de acuerdo con la clasificación TNM (ver anexo 3). El código 4 ('T1'), se utilizará cuando el tamaño del tumor sea ≤ 2 cm., pero no pueda determinarse si pertenece a los códigos 5 ('T1mic'), 6 ('T1a'), 7 ('T1b'), u 8 ('T1c'). Se informará como 'Desconocido' cuando no se disponga de información sobre el tamaño del tumor.

Nota: *Siempre que sea posible se recogerá el TNM patológico (pTNM), y en su defecto el TNM clínico cuando éste no esté disponible.*

- *1: Tx;*
- *2: T0;*
- *3: Tis;*
- *4: T1;*
- *5: T1mic;*
- *6. T1a;*
- *7. T1b;*
- *8: T1c;*
- *9: T2;*
- *10: T3;*
- *11: T4;*
- *12: T4a;*

- *13: T4b;*
- *14: T4c;*
- *15: T4d;*
- *99: Desconocido*

- ***Afectación ganglionar (categórica)***

Descripción de la afectación ganglionar, de acuerdo con la clasificación TNM (ver anexo 3). Se informará como 'Desconocido' cuando no se disponga de información sobre la afectación ganglionar del tumor.

- *1: Nx;*
- *2: N0;*
- *3: N1;*
- *4: N2;*
- *5: N2a;*
- *6: N2b;*
- *7: N3;*
- *8: N3a;*
- *9: N3b;*
- *10: N3c;*
- *99: Desconocido*

- ***Metástasis (categórica)***

Se informará de la posible presencia de metástasis, de acuerdo con la clasificación TNM (ver anexo 3). Se informará como 'Desconocido' cuando no se disponga de información sobre la presencia de metástasis.

- *1: Mx;*
- *2: M0;*
- *3: M1;*
- *99: Desconocido*

- ***Grado de diferenciación (categórica)***

Se obtiene al estudiar al microscopio las células tumorales. Por norma general los tumores in situ se clasificarán como 0. No aplicable. Se utilizan las siguientes categorías en la clasificación:

- *0: No aplicable*
- *1: Categoría I, las células más parecidas al tejido mamario normal (bien diferenciadas)*
- *2: Categoría II las intermedias*

- *3: Categoría III las menos parecidas al tejido normal y, por tanto, con peor pronóstico (mal diferenciadas)*
- *99. Desconocido*

Receptores hormonales

Los criterios para la interpretación de los análisis inmuno-histoquímicos de receptores de estrógenos y progesterona varía entre programas. Por norma general, se considera que un porcentaje ≥ al 10% de células teñidas es positivo. En caso de utilizar otro porcentaje como umbral, se especificará el umbral utilizado. Si se utiliza un "Score", se especificará el criterio utilizado. Si se utiliza la clasificación "Plus system" (-/+/++/+++), se considera la categoría '+' (o mayor) como positiva.

- ***Receptores de estrógeno (categórica)***

La presencia de los receptores determina el tratamiento específico del cáncer y, además, son factores pronósticos del cáncer. De acuerdo con los criterios estándares para la interpretación de los resultados de los análisis inmuno-histoquímicos, se clasificará como:

- *1: Negativo*
- *2: Positivo*
- *99: Desconocido*

- ***Receptores de progesterona (categórica)***

En función de la determinación inmunohistoquímica se clasificará en:

- *1: Negativo*
- *2: Positivo*
- *99: Desconocido*

Resumen de codificaciones y tablas

Programas (1 registro por programa)	**Codificación**
Programa_id	Código asignado a cada programa
Nombre del programa	Nombre (etiqueta) de cada programa
URC (1- N registros por programa)	**Codificación**
Programa_id	Código asignado a cada Unidad
URC_id	Radiológica de Cribado donde se realizan las exploraciones
Descripción URC	Descripción identificativa de la Unidad Radiológica de Cribado
Mujeres (1 registro por mujer)	**Codificación**
Programa_id	
Mujer_id	Número identificador interno del programa
Fecha de nacimiento	09/09/9999 Desconocido
Fecha 1a citación en el programa	09/09/9999 Desconocido
Episodios (1 registro por cada participación, 1-N registros por mujer)	**Codificación**
Programa_id	
Mujer_id	
URC_id	Código asignado a cada Unidad Radiológica de Cribado donde se realizan las exploraciones
Episodio	1,2,3,4,...
Número de cribado	1,2,3,4,...
Fecha mamografía de cribado	09/09/9999 Desconocido
Nº de estudios intermedios	0,1,2,3,4,...
Resultado inicial de la mamografía de cribado	1: Negativo 2: Exploraciones adicionales 3: Mamografía intermedia (Estudios intermedios) 99:Desconocido
Categoría mamografía de la prueba de cribado (Bi-Rads)	0: Incierto 1: Normal 2: Benigna 3: Probablemente Benigna 4: Probablemente Maligna 5: Maligna 99: Desconocido
Densidad mamaria	1: Completamente grasa 2: Densidad fibroglandular dispersa 3: Densidad heterogénea 4: Extremadamente densa 99: Desconocido

Lateralidad	0: No pertinente 1: Derecha 2: Izquierda 3: Bilateral 99: Desconocido
Masa_D (Patrones radiológicos mama derecha)	0: No 1: Sí 99: Desconocido
Distorsión_D (Patrones radiológicos mama derecha)	0: No 1: Sí 99: Desconocido
Calcificaciones_D (Patrones radiológicos mama derecha)	0: No 1: Sí 99: Desconocido
Asimetría_D (Patrones radiológicos mama derecha)	0: No 1: Sí 99: Desconocido
Otros_D (Patrones radiológicos mama derecha)	0: No 1: Sí 99: Desconocido
Masa_I (Patrones radiológicos mama izquierda)	0: No 1: Sí 99: Desconocido
Distorsión_I (Patrones radiológicos mama izquierda)	0: No 1: Sí 99: Desconocido
Calcificaciones_I (Patrones radiológicos mama izquierda)	0: No 1: Sí 99: Desconocido
Asimetría_I (Patrones radiológicos mama izquierda)	0: No 1: Sí 99: Desconocido
Otros_I (Patrones radiológicos mama izquierda)	0: No 1: Sí 99: Desconocido
Resultado final del episodio	1: Cáncer 2: No cáncer 3: Pendiente estudios intermedios 5: Seguimiento incompleto 99: Desconocido
Exploraciones adicionales (0-N registros por episodio)	**Codificación**
Programa_id	
Mujer_id	
Episodio	1,2,3,4,...
N_Prueba	1,2,3,4,...

Prueba_id	1.1. Otras proyecciones mamográficas 1.2. Ecografía 1.3. Resonancia magnética 1.4. Otras pruebas no invasivas 1.9. Prueba no invasiva desconocida 2.1. PAAF 2.2. Biopsia aguja gruesa 2.3. Biopsia asistida por vacío 2.4. Biopsia escisional percutánea 2.5. Biopsia quirúrgica 2.6. Otras pruebas invasivas 2.9. Prueba invasiva desconocida 99. Desconocida prueba
Fecha_prueba	09/09/9999 si desconocido
Variables de las mujeres (1 registro por episodio)	**Codificación**
Programa_id	
Mujer_id	
Episodio	1,2,3,4,...
THS	1: Sí, en el momento del cribado o en los 6 meses anteriores 2: No 99: Desconocido
Menopausia	1: Posmenopáusica 2: Premenopáusica o perimenopáusica 99: Desconocido
Antecedentes personales de prueba invasiva con resultado benigno	1: Sí 2: No 99: Desconocido
Antecedentes familiares de cáncer de mama	1: Sí 2: No 99: Desconocido
Edad menopausia	45, 46, 47, ... 9: No pertinente 99: Desconocido
Edad de la menarquia	10, 11, 12, ... 99: Desconocido
Lesiones benignas (0-N registros por episodio)	**Codificación**
Programa_id	
Mujer_id	
Episodio	1, 2, 3, 4, ...
Histología básica de las citologías	0: Negativo células malignas 1: Lesión no proliferativa 2: Lesión proliferativa sin atipia 3: Lesión proliferativa con atipia 9: No pertinente 99: Desconocido

Histología de la lesión benigna	Clasificación Internacional de Enfermedades para Oncología CIE-O (3ª edición) – VER ANEXO 1
Lateralidad	1: Derecha 2: Izquierda 3: Bilateral 99: Desconocido
Tumores (1 registro por tumor)	**Codificación**
Programa_id	
Mujer_id	
Episodio	1, 2, 3, 4, …
Método de detección	1: Cribado 2: Cáncer de Intervalo 3: Fuera de cribado
Histología tumores	Clasificación Internacional de Enfermedades para Oncología CIE-O (3ª edición) – VER ANEXO 2
Comportamiento tumoral	1: Ductal in situ 2: Invasivo 99: Desconocido
Lateralidad	1: Derecha 2: Izquierda 3: Bilateral 99: Desconocido
Tamaño tumor	1: Tx 2: T0 3: Tis 4: T1 5: T1mic 6. T1a 7. T1b 8: T1c 9: T2 10: T3 11: T4 12: T4a 13: T4b 14: T4c 15: T4d 99: Desconocido
Afectación ganglionar	1: Nx 2: N0 3: N1 4: N2 5: N2a 6: N2b 7: N3 8: N3a 9: N3b

	10: N3c 99: Desconocido
Metástasis	1: Mx 2: M0 3: M1 99: Desconocido
Grado de diferenciación	0: No aplicable 1: Categoría I 2: Categoría II 3: Categoría III 99. Desconocido
Receptores de estrógeno (Biomarcadores)	1: Negativo 2: Positivo 99: Desconocido
Receptores de progesterona (Biomarcadores)	1: Negativo 2: Positivo 99: Desconocido

Anexo 1

Categorías variables 'Histología lesiones benignas', según la Clasificación Internacional de Enfermedades para Oncología CIE-O (3ª edición):

Descripción	Código
Lesión benigna (sin especificar)	00001
Normal	00100
Normal citology	00120
Unsuitable sample	09000
Material unsuitable for diagnostics, no sample taken	09010
Capillary hemoangioma	09131
Malignant tumour tissue not found	09450
Traumatic lesion NOS	10000
Operation wound NOS	14020
Accessory structure NOS	22300
Ectopic mammary tissue	26030
Micro calcifications	30180
Ectasia NOS	32100
Galactocele	33220
Cyst	33400
Epidermoid cyst (atheroma)	33410
Haemorrhage	37000
Blue dome cyst	33710
Inflammation NOS	40000
Acute inflammation	41000
Abscess NOS	41740
Inflammation, chronic	43000
Plasma cell mastitis	43060
Granulomatous inflammation	44000
Reaction to foreign body	44140
Comedomastitis	46460
Fibrosis NOS	49000
Focal fibrosis	49001
Scar	49060
Necrosis SAI	54000
Necrosis in adipose tissue	54110
Calcareous deposit	55400
Atrophy NOS	58000
Lesión columnar con atipia	67020

Atipia citológica	69700
Atipia, sospechoso malignidad	69760
Involution	70800
Hypertrophy NOS	71000
Intraductal precancerous hyperplasia	71279
Hyperplasia	72000
Papillary hyperplasia	72050
Lobular hyperplasia	72100
Focal lobular hyperplasia	72101
Irregular lobular hyperplasia,uncertain benign/malignant	72105
Intraductal hyperplasia	72170
Ductal atypical hyperplasia	72175
Pseudoangiomatous stromal hyperplasia	72190
Lymphoid hyperplasia NOS	72200
Hiperplasia glandular	72420
Metaplasia	73000
Metaplasia, squamous	73220
Apocrine metaplasia	73310
Dysplasia	74000
Adenosis	74200
Sclerosing adenosis	74220
Adenosis, blunt duct	74240
Adenosis, florid	74260
Enfermedad fibroquística de la mama	74320
Fibrocystic disease, atypical	74325
Hamartoma	75500
Fibromatosis	76100
Proliferation phase	79120
Benign neoplasm	80000
Uncertain benign/malignant	80001
Tumor epitelial benigno	80100
Papilloma NOS	80500
Papillomatosis NOS	80600
Adenoma of the nippel	81400
Tubular adenoma NOS	82110
Atypical intraductal epithelial proliferation	85001
Intraductal papilloma	85030
Intracystic papilloma	85040
Papillomatosis, intraductal	85050
Subareolar, florid papillomatosis	85060
Atypical lobular hyperplasia	85201
Lobular carcinoma NOS in situ	85202

Fibroma	88110
Lipoma NOS	88500
Pleomorphic adenoma	89400
Fibroadenoma NOS	90100
Intracanalicular fibroadenoma	90110
Phyllodes tumour, benign	90200
Phylloid tumour NOS, uncertain benign/malignant	90201
Granular cell tumour NOS	95800
Malignant residue not found	99903

Anexo 2

Categorías variables 'Histología tumores', según la Clasificación Internacional de Enfermedades para Oncología CIE-O (3ª edición):

Descripción	Código
8000/3 Tumor maligno	80003
8001/3 Células tumorales malignas	80013
8002/3 Tumor maligno de células pequeñas	80023
8003/3 Tumor maligno de células gigante	80033
8004/3 Tumor maligno de celulas fusiformes	80043
8005/3 Tumor maligno, tipo células claras	80053
8010/2 Carcinoma in situ, SAI	80102
8010/3 Carcinoma SAI	80103
8010/6 Carcinoma, metastásico, SAI	80106
8011/3 Epitelioma maligno	80113
8012/3 Carcinoma de células grandes SAI	80123
8013/3 Carcinoma neuroendocrino de células grandes	80133
8014/3 Carcinoma de células grandes con fenotipo rabdoide	80143
8015/3 Carcinoma de células vidriosas	80153
8020/3 Carcinoma indiferenciado SAI	80203
8021/3 Carcinoma anaplásico SAI	80213
8022/3 Carcinoma pleomórfico	80223
8030/3 Carcinoma gigantocelular y fusocelular	80303
8031/3 Carcinoma de células gigantes	80313
8032/3 Carcinoma fusocelular	80323
8033/3 Carcinoma seudosarcomatoso	80333
8034/3 Carcinoma de células poligonales	80343
8035/3 Carcinoma de células gigantes con osteoclasto semejantes	80353
8041/3 Carcinoma de células pequeñas SAI	80413
8043/3 Carcinoma de células pequeñas tipo fusiforme	80433
8050/2 Carcinoma papilar in situ	80502
8050/3 Carcinoma papilar SAI	80503
8051/3 Papiloma verrugoso	80513
8052/2 Carcinoma papilar de células escamosas, no invasivo	80522
8052/3 Carcinoma papilar de células escamosas	80523
8070/2 Carcinoma in situ de células escamosas SAI	80702
8070/3 Carcinoma de células escamosas SAI	80703
8071/3 Carcinoma de células escamosas, tipo queratinizante SAI	80713
8072/3 Carcinoma de células escamosas, grandes, tipo no queratinizante	80723

8073/3 Carcinoma de células escamosas, pequeñas, tipo no queratinizante	80733
8074/3 Carcinoma de células escamosas, tipo fusocelular	80743
8075/3 Carcinoma de células escamosas, tipo adenoide	80753
8076/2 Carcinoma de células escamosas in situ con invasión dudosa del estroma	80762
8076/3 Carcinoma de células escamosas, microinvasor	80763
8078/3 Carcinoma de células escamosas con formación en formación de cuerno	80783
8140/2 Adenocarcinoma in situ SAI	81402
8140/3 Adenocarcinoma SAI	81403
8141/3 Adenocarcinoma escirroso (escirro)	81413
8143/3 Adenocarcinoma con diseminación superficial	81433
8147/3 Adenocarcinoma basocelular	81473
8190/3 Adenocarcinoma trabecular	81903
8200/3 Carcinoma adenoide-quístico	82003
8201/2 Carcinoma cribriforme, in situ	82012
8201/3 Carcinoma cribiforme	82013
8211/3 Adenocarcinoma tubular	82113
8230/2 Carcinoma ductal in situ, tipo sólido	82302
8230/3 Carcinoma sólido SAI	82303
8231/3 Carcinoma simple	82313
8240/3 Carcinoid tumor, SAI	82403
8246/3 Carcinoma neuroendocrino	82463
8251/3 Adenocarcinoma alveolar	82513
8255/3 Adenocarcinoma con subtipos mixtos	82553
8260/2 Adenocarcinoma papilar in situ, SAI	82602
8260/3 Adenocarcinoma papilar SAI	82603
8261/2 Adenocarcinoma in situ en adenoma velloso	82612
8261/3 Adenocarcinoma en adenoma velloso	82613
8310/3 Adenocarcinoma de células claras SAI	83103
8314/3 Carcinoma rico en lípidos	83143
8315/3 Carcinoma rico en glucógeno	83153
8320/3 Carcinoma de células granulares	83203
8323/3 Adenocarcinoma de células mixtas	83233
8401/2 Adenocarcinoma apocrino, in situ	84012
8401/3 Adenocarcinoma apocrino	84013
8440/3 Cistadenocarcinoma SAI	84403
8480/3 Adenocarcinoma mucinoso	84803
8481/3 Adenocarcinoma secretante de mucina	84813
8490/3 Carcinoma de células en anillo de sello	84903
8500/2 Carcinoma intracanalicular no infiltrante SAI	85002
8500/3 Carcinoma canalicular infiltrante	85003
8501/2 Comedocarcinoma no infiltrante	85012
8501/3 Comedocarcinoma SAI	85013

8502/3 Carcinoma juvenil de la mama	85023
8503/2 Adenocarcinoma papilar intracanalicular no infiltrante	85032
8503/3 Adenocarcinoma papilar intracanalicular, con invasión	85033
8504/2 Carcinoma intraquístico no infiltrante	85042
8504/3 Carcinoma intraquístico SAI	85043
8507/2 Carcinoma intraductal micropapilar	85072
8507/3 Carcinoma mixto de mama	85073
8508/3 Carcinoma hipersecretorio quístico	85083
8510/3 Carcinoma medular SAI	85103
8512/3 Carcinoma medular con estroma linfoide	85123
8513/3 Carcinoma medular atípico	85133
8514/3 Carcinoma ductal, típico desmoplástico	85143
8520/2 Carcinoma lobulillar in situ (D05.0)	85202
8520/3 Carcinoma lobulillar SAI (C50.-)	85203
8521/3 Carcinoma canalicular, infiltrante (C50.-)	85213
8522/2 Carcinoma intracanalicular y carcinoma lobulillar in situ (D05.7)	85222
8522/3 Carcinoma canalicular y lobulillar infiltrante (C50.-)	85223
8523/2 Carcinoma ductal infiltrante mixto con otros tipos de carcinoma, in situ	85232
8523/3 Carcinoma ductal infiltrante mixto con otros tipos de carcinoma	85233
8524/3 Carcinoma lobular infiltrante mixto con otros tipos de carcinoma	85243
8525/3 Adenocarcinoma polimorfo de grado bajo	85253
8530/3 Carcinoma inflamatorio (C50.-)	85303
8540/3 Enfermedad de Paget, mamaria (C50.-)	85403
8541/3 Enfermedad de Paget y carcinoma canalicular infiltrante de la mama (C50.-)	85413
8543/2 Enfermedad de Paget in situ y carcinoma intracanalicular de la mama (C50.-)	85432
8543/3 Enfermedad de Paget y carcinoma intracanalicular de la mama (C50.-)	85433
8550/3 Carcinoma de células acinosas	85503
8551/3 Cistadenocarcinoma de células acinosas	85513
8560/3 Carcinoma adenoescamoso	85603
8562/3 Carcinoma epitelial-mioepitelial	85623
8570/3 Adenocarcinoma con metaplasia escamosa	85703
8571/3 Adenocarcinoma con metaplasia ósea y cartilaginosa	85713
8572/3 Adenocarcinoma con metaplasia de células fusiformes	85723
8573/3 Adenocarcinoma con metaplasia apocrina	85733
8574/3 Adenocarcinoma con diferenciación neuroendocrina	85743
8575/3 Carcinoma metaplástico	85753
8800/3 Sarcoma SAI	88003
8801/3 Sarcoma fusocelular	88013
8802/3 Sarcoma de células gigantes	88023
8803/3 Sarcoma de células pequeñas	88033
8804/3 Sarcoma de células epitelioides	88043
8805/3 Sarcoma indiferenciado	88053

8806/3 Tumor desmoplástico de células redondas pequeñas	88063
8810/3 Fibrosarcoma SAI	88103
8811/3 Fibromixosarcoma	88113
8813/3 Fibrosarcoma fascial	88133
8814/3 Fibrosarcoma infantil	88143
8815/3 Tumor fibroso solitario, maligno	88153
8850/3 Liposarcoma SAI	88503
8851/3 Liposarcoma, bien diferenciado	88513
8852/3 Liposarcoma mixoide	88523
8853/3 Liposarcoma de células redondas	88533
8854/3 Liposarcoma pleomórfico	88543
8855/3 Liposarcoma mixto	88553
8857/3 Liposarcoma fibroplástico	88573
8858/3 Liposarcoma desdiferenciado	88583
8890/3 Leiomiosarcoma SAI	88903
8891/3 Leiomiosarcoma epitelioide	88913
8894/3 Angiomiosarcoma	88943
8895/3 Miosarcoma	88953
8896/3 Leiomiosarcoma mixoide	88963
8935/3 Sarcoma del estroma, SAI	89353
8980/3 Carcinosarcoma SAI	89803
8981/3 Carcinosarcoma, tipo embrionario	89813
8982/3 Mioepitelioma maligno	89823
8990/3 Mesenquimoma maligno	89903
8991/3 Sarcoma embrionario	89913
9020/3 Tumor filoide, maligno	90203
9120/3 Hemangiosarcoma	91203
9130/3 Hemangioendotelioma, maligno	91303
9133/3 Hemangioendotelioma epitelioide, maligno	91333
9580/3 Tumor de células granulares, maligno	95803
9581/3 Sarcoma alveolar de partes blandas	95813
9590/3 Linfoma maligno SAI	95903
9591/3 Linfoma maligno, no Hodgkin SAI	95913
9596/3 Limfoma Hodgkin y no Hodgkin compuesto	95963
9650/3 Enfermedad de Hodgkin SAI	96503
9651/3 Limfoma de Hodgkin, predominio linfocítico-histiocítico	96513
9652/3 Enfermedad de Hodgkin, celularidad mixta SAI	96523
9653/3 Enfermedad de Hodgkin con depleción linfocítica SAI	96533
9654/3 Enfermedad de Hodgkin con depleción linfocítica, tipo fibrosis difusa	96543
9655/3 Enfermedad de Hodgkin con depleción linfocítica, tipo reticular	96553
9659/3 Enfermedad de Hodgkin con predominio linfocítico, nodular	96593
9661/3 Granuloma de Hodgkin	96613

9662/3 Sarcoma de Hodgkin	96623
9663/3 Enfermedad de Hodgkin, tipo esclerosis nodular SAI	96633
9664/3 Enfermedad de Hodgkin, tipo esclerosis nodular, fase celular	96643
9665/3 Enfermedad de Hodgkin, tipo esclerosis nodular, con predominio linfocítico	96653
9667/3 Enfermedad de Hodgkin, tipo esclerosis nodular, con depleción linfocítica	96673
9670/3 Linfoma maligno, linfocítico de células pequeñas SAI	96703
9671/3 Linfoma maligno linfoplasmocítico	96713
9673/3 Linfoma maligno, linfocítico de diferenciación intermedia, difuso	96733
9675/3 Linfoma maligno mixto, de células pequeñas y grandes, difuso	96753
9680/3 Linfoma maligno, de células grandes, difuso SAI	96803
9684/3 Linfoma maligno, inmunoblástico SAI	96843
9687/3 Linfoma de Burkitt SAI	96873
9690/3 Linfoma maligno, folicular SAI	96903
9691/3 Linfoma maligno mixto, de células pequeñas hendidas y células grandes, folicular	96913
9695/3 Linfoma maligno, de células pequeñas hendidas, folicular	96953
9698/3 Linfoma maligno, de células grandes, folicular SAI	96983
9699/3 Linfoma marginal de la B-célula de la zona de Extranodal del tejido fino linfoide mucosa-asociado, SAI	96993
9701/3 Enfermedad de Sézary	97013
9702/3 Linfoma periférico de células T SAI	97023
9705/3 Linfoma periférico de células T, LAID (linfadenopatía angioinmunoblástica con disproteinemia	97053
9714/3 Linfoma de células grandes (Ki-1+)	97143
9719/3 Linfoma de la célula de Extranodal NK/T, tipo nasal	97193
9727/3 Precursor cell lymphoblastic lymphoma, NOS	97273
9728/3 Precursor B-cell lymphoblastic lymphoma	97283
9729/3 Precursor T-cell lymphoblastic lymphoma	97293
9731/3 Plasmocitoma SAI	97313
9734/3 Plasmacitoma, extramedular	97343
9740/3 Sarcoma de mastocitos	97403
9741/3 Mastocitosis maligna	97413
9750/3 Malignant histiocytosis	97503
9754/3 Langerhans cell histiocytosis, disseminated	97543
9755/3 Histiocystic sarcoma	97553
9756/3 Langerhans cell sarcoma	97563
9757/3 Interdigitating dendritic cell sarcoma	97573
9758/3 Follicular dendritic cell sarcoma	97583

Anexo 3

Clasificación Clínica TNM, según la definición de la UICC, 6ª Edición:

Tamaño del tumor primario

- *Tx: el tumor no se puede evaluar*
- *T0: no hay evidencia de tumor primario*
- *Tis: carcinoma in situ*
- *T1: este código se utilizará cuando el tamaño del tumor sea ≤ 2cm, pero no pueda determinarse si pertenece a las categorías '1mic', '1a', '1b' ó '1c'*
    - *T1mic: Microinvasion ≤0.1 cm de diámetro máximo. La microinvasion es la extensión de células cancerígenas a través de la membrana a los tejidos adyacentes con un foco no mayor de 0.1 cm. Si hay múltiples focos solo se utiliza el de mayor tamaño a efectos de clasificación de microinvasión (no utilizar la suma de todos los focos individuales). La presencia de múltiples focos de microinvasión debe registrarse, tal como se hace con los carcinomas múltiples invasivos*
    - *T1a: diámetro máximo >0.1 cm, pero ≤ 0.5 cm*
    - *T1b: diámetro máximo >0.5 cm, pero ≤ 1 cm*
    - *T1c; diámetro máximo >1 cm, pero ≤ 2cm*
- *T2; Tumor de diámetro máximo >2 cm, pero ≤ 5 cm*
- *T3; Tumor de diámetro máximo >5 cm*
- *T4; Tumor de cualquier tamaño con extensión directa a la pared del tórax o la piel solo como se describe en T4a a T4d. La pared torácica incluye las costillas, los músculos intercostales y el músculo serrato mayor, pero no los músculos pectorales*
    - *T4a; Extensión a pared torácica*
    - *T4b; Edema (incluyendo piel de naranja) o ulceración de la piel de la mama, o presencia de ganglios cutáneos satélites confinados en la misma mama*
    - *T4c; T4a y T4b conjuntamente*
    - *T4d; Carcinoma inflamatorio. El carcinoma inflamatorio se caracteriza por una induración cutánea difusa con un borde erisipeloide y, generalmente no se puede palpar ninguna masa subyacente. Si al realizar la clasificación anatomopatológica de un carcinoma inflamatorio clínico (T4d), la biopsia de la induración es negativa y no existe cáncer primario localizado que se pueda medir, la categoría T es pTx. Las categorías T1, T2, T3 pueden coexistir con la presencia de depresiones*

cutáneas, retracción de pezón o cualquier otra alteración cutánea exceptuando las descritas en las categorías T4b y T4d, sin que ello afecte la clasificación.

Ganglios linfáticos regionales

- *Nx; Los ganglios linfáticos regionales no se pueden valorar (extirpación previa)*
- *N0; No hay evidencia de metástasis ganglionares regionales*
- *N1; Metástasis móviles en ganglios axilares ipsilaterales*
- *N2; Metástasis en ganglios linfáticos axilares fijos o en ganglios de mama interna ipsilaterales clínicamente aparentes[Clínicamente aparentes: detectados por examen clínico o por estudios radiológicos -incluyendo linfoescintografía- o examen anatomopatológico visibles a simple vista] en ausencia de metástasis de ganglios linfáticos axilares*
  - *N2a; metástasis de ganglios linfáticos axilares fijados entre ellos o a otras estructuras*
  - *N2b; metástasis solo en ganglios linfáticos mamarios internos clínicamente aparentes1 y en ausencia de metástasis de ganglios linfáticos axilares clínicamente aparentes1*
- *N3; Metástasis en ganglios linfáticos infraclaviculares ipsilaterales con o sin afectación de ganglios axilares, o en ganglios linfáticos mamarios internos ipsilaterales clínicamente aparentes1 y en presencia de metástasis en ganglios linfáticos axilares clínicamente evidente, o metástasis de ganglios linfáticos supraclaviculares ipsilaterales con o sin afectación de ganglios linfáticos axilares o de mamaria interna.*
  - *N3a; metástasis solo en ganglios linfáticos mamarios internos clínicamente aparentes1 y en ausencia de metástasis de ganglios linfáticos axilares clínicamente aparentes1*
  - *N3b; metástasis en ganglios linfáticos mamarios internos y axilares ipsilaterales*
  - *N3c; metástasis de ganglios linfáticos supraclaviculares ipsilaterales*

Metástasis a distancia

- *Mx; Las metástasis a distancia no se pueden evaluar*
- *N0; No hay evidencia de metástasis a distancia*
- *M1; Metástasis a distancia*

Lesiones benignas de mama, densidad mamaria y asociación con el riesgo de cáncer en una cohorte de mujeres cribadas

Programa de control de calidad de la base de datos

Variables: tipo y rango de valores

Para cada tabla y campo (variable), se describe *'tipo'* y *'rango'* de valores. Los campos que, en términos de una base datos relacional, serán claves primarias aparecen en color rojo.

Se considera que un valor está fuera de rango, y por lo tanto es un valor imposible, cuando no está dentro del conjunto de valores predeterminados definidos para esa variable.

RESTRICCIÓN GENERAL: Los campos no pueden quedar vacíos, excepto que se indique expresamente en esta descripción.

1. Programas

- ***Programa_id***

Campo definido como texto de longitud 2. Para los programas participantes en el estudio BELE-1 la clasificación asignada a la variable "Programa_id" **será la misma**. Toma los siguientes valores:

Programa_id	***Nombre Programa***
AS	*Asturias*
CT	*Cantabria*
GI	*Girona*
IC	*ICO*
PM	*Parc Salut Mar*
SB	*Sabadell*
SP	*Sant Pau*
TA	*Tarragona*
HC	*Hospital Clinic*
GR	*Vallès Oriental – Granollers*

- ***Nombre del programa***

Campo definido como texto de longitud 30. Ver equivalencias en la definición de la anterior variable.

2. Unidad Radiológica de Cribado

- ***Programa_id***

Ya descrito

- ***URC_id***

Campo definido como texto de longitud 4. Se enumerarán de la siguiente forma: Letras identificativas del programa de cribado (dos dígitos), seguida de un número natural consecutivo, empezando por el 1, 2,..., etc. En el caso de programas sin estructura de URC, bastará con poner las letras identificativas del programa de cribado, seguidas del número 1 en esta variable. Para los programas participantes en el estudio BELE-1 la clasificación asignada a la variable "URC_id" **será la misma** que la asignada en dicho estudio.

Ejemplo: 'IC1'

- ***Descripción de la URC***

Campo definido como texto de longitud 50. Nombre y/o descripción (etiqueta) de la Unidad de Cribado donde se realizan las exploraciones. Los programas sin estructura de URC la etiquetarán con el nombre del propio programa. Aquellos programas que no identifiquen las URC con un nombre específico, las etiquetarán con el 'nombre del programa' seguido del mismo número que hayan asignado en el código URC_id en la variable anterior (nombre del programa1, nombre del programa2,..., nombre del programaN). Para los programas participantes en el estudio BELE-1 la clasificación asignada a la variable "Descripción de la URC" **será la misma** que la asignada en dicho estudio.

3. Mujeres

- ***Programa_id***

Ya descrito

- ***Mujer_id***

Campo definido como texto de longitud 15. Consiste en las dos letras identificativas del programa, más el número identificador interno del programa utilizado para cada mujer. Para los programas participantes en el estudio BELE-1 **será preciso** que este número identificador coincida con el identificador asignado en dicho estudio.

- ***Fecha de nacimiento***

Campo definido como texto de longitud 10, con formato 00/00/0000 correspondiente a día/mes/año (09/09/9999 si desconocido). No se añaden restricciones de valor para esta variable, pero se incluirán como valores poco probables los que estén fuera del rango comprendido entre 01/01/1919 y 31/12/1971, o bien 09/09/9999 si es desconocida. Para los programas participantes en el estudio BELE-1 **será deseable** que esta fecha coincida con la fecha asignada en dicho estudio.

- ***Fecha 1ª citación en el programa***

Campo definido como texto de longitud 10, con formato 00/00/0000 correspondiente a día/mes/año. No se añaden restricciones de valor para esta variable, pero se incluirán como valores poco probables los que estén fuera del rango comprendido entre 01/01/1989 y 31/12/2015, o bien 09/09/9999 si es desconocido. Para los programas participantes en el estudio BELE-1 **será deseable** que esta fecha coincida con la fecha asignada en dicho estudio.

4. Episodios

- ***Programa_id***

Ya descrito

- ***Mujer_id***

Ya descrito

- ***URC_id***

Ya descrito

- ***Episodio***

Campo definido como numérico. Se informará un número natural. La variable es ordinal, pero puede tener saltos, y no necesariamente comenzar en el 1. Para los programas participantes en el estudio BELE-1 **será preciso** que este identificador de episodios sea coherente con el identificador asignado en el BELE-1. Para las mujeres cribadas en el periodo 2012-2015 (ampliación de la cohorte de estudio), con participaciones anteriores a 2012, deberá de existir un orden lógico en la numeración de los episodios. Este identificador permitirá mantener la secuencia temporal entre los distintos estudios cuando se añada nueva información específica del proyecto.

- ***Número de cribado***

Campo definido como numérico. Se informará un número natural, comenzando con el 1. La variable es ordinal y **NO** puede tener saltos. Cuando se introduce un nuevo episodio con mamografía realizada, la variable "Número de cribado" incrementa necesariamente en una unidad el contador. Para los programas participantes en el estudio BELE-1 **será preciso** que el número de cribado sea coherente con el identificador asignado en el BELE-1. Para las mujeres cribadas en el periodo 2012 - 2015 (ampliación de la cohorte de estudio), con participaciones anteriores a 2012, deberá de existir un orden lógico en la numeración del cribado.

- ***Fecha mamografía de cribado***

Campo definido como texto de longitud 10, con formato 00/00/0000 correspondiente a día/mes/año. No se añaden restricciones de valor para esta variable, pero se incluirán como valores poco probables los que estén fuera del rango comprendido entre 01/01/1989 y 31/12/2015, o bien 09/09/9999 si es desconocida.

- ***Número de estudios intermedios***

Campo definido como numérico. Se informará un número natural entre 0 (cero) y 5 (cinco).

- ***Resultado inicial de la mamografía de cribado***

Campo definido como numérico. Ver equivalencias.

- *1: Negativo*
- *2: Exploraciones adicionales*
- *3: Mamografía intermedia*
- *99: Desconocido*

- ***Categoría mamografía de la prueba de cribado (Bi-Rads)***

Campo definido como numérico. Ver equivalencias.

- *0: Incierto*
- *1: Normal*
- *2: Benigna*
- *3: Probablemente Benigna*
- *4: Probablemente Maligna*
- *5: Maligna*
- *99: Desconocido*

- ***Densidad mamaria***

Campo definido como numérico. Ver equivalencias.

  - *1: Completamente grasa (BI-RADS type 1)*
  - *2: Densidad fibroglandular dispersa (BI-RADS type 2)*
  - *3: Densidad heterogénea (BI-RADS type 3)*
  - *4: Extremadamente densa (BI-RADS type 4)*
  - *99: Desconocido*

- ***Lateralidad***
  - *Campo definido como numérico. Ver equivalencias.*
  - *0: No pertinente*
  - *1: Derecha*
  - *2: Izquierda*
  - *3: Bilateral*
  - *99: Desconocido*

- ***Patrones radiológicos de la prueba de cribado (categórica)***

Esta información se recoge en un bloque de 5 variables para cada mama. Cada una de estas variables es numérica.

Mama derecha

- *Masa_D*

Campo definido como numérico. Ver equivalencias.

  - *0: No*
  - *1: Sí*
  - *99: Desconocido*

- *Distorsión_D*

Campo definido como numérico. Ver equivalencias.

  - *0: No*
  - *1: Sí*
  - *99: Desconocido*

- *Calcificaciones_D*

Campo definido como numérico. Ver equivalencias.

  - *0: No*
  - *1: Sí*
  - *99: Desconocido*

- *Asimetría_D*

Campo definido como numérico. Ver equivalencias.

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Otros_D*

Campo definido como numérico. Ver equivalencias.

- *0: No*
- *1: Sí*
- *99: Desconocido*

Mama izquierda

- *Masa_I*

Campo definido como numérico. Ver equivalencias.

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Distorsión_I*

Campo definido como numérico. Ver equivalencias.

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Calcificaciones_I*

Campo definido como numérico. Ver equivalencias.

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Asimetría_I*

Campo definido como numérico. Ver equivalencias.

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Otros_I*

Campo definido como numérico. Ver equivalencias.

- *0: No*
- *1: Sí*
- *99: Desconocido*

- *Resultado final del episodio*

Campo definido como numérico. Ver equivalencias

- *1: Cáncer*
- *2: No cáncer*
- *3: Pendiente estudios intermedios*
- *5: Seguimiento incompleto*
- *99: Desconocido*

La codificación *'3: Pendiente de estudios intermedios'* se reserva únicamente para aquellos episodios que por haberse realizado la mamografía de cribado durante el año 2015, puedan estar pendientes de la realización de estudios intermedios en la fecha de finalización de recogida de información (31/12/2015). Cualquier otro resultado final no conocido se informará como *'5: Seguimiento incompleto'*, o *'99: Desconocido'*.

5. Exploraciones adicionales (0-N registros por episodio)

- ***Programa_id***

Ya descrito

- ***Mujer_id***

Ya descrito

- ***Episodio***

Ya descrito

- ***N prueba***

Campo definido como numérico. Se informará un número natural, de manera consecutiva, comenzando con el 1.

- ***Prueba_id***
  - *Campo definido como numérico. Ver equivalencias:*

- *11: Otras proyecciones mamográficas*
- *12: Ecografía*
- *13: Resonancia magnética*
- *14: Otras pruebas no invasivas*
- *19: Prueba no invasiva desconocida*
- *21: PAAF*
- *22: Biopsia aguja gruesa*
- *23: Biopsia asistida por vacío*
- *24: Biopsia escisional percutánea*
- *25: Biopsia quirúrgica*
- *26: Otras pruebas invasivas*
- *29: Prueba invasiva desconocida*
- *99: Prueba desconocida*

- ***Fecha_prueba***

Campo definido como texto de longitud 10, con formato 00/00/0000 correspondiente a día/mes/año (09/09/9999 si desconocido). No se añaden restricciones de valor para esta variable, pero se incluirán como valores poco probables los que estén fuera del rango comprendido entre 01/01/1989 y 30/06/2016, o bien 09/09/9999 si es desconocida.

Nota: *La fecha de realización de pruebas abarca hasta el 30/06/2016 (6 meses más que la fecha máxima de realización de la mamografía de cribado). Este marco temporal de 6 meses se define para poder obtener información de las exploraciones complementarias realizadas para descartar o confirmar malignidad, de aquellas mamografías de cribado reconvocadas durante el último periodo de 2015.*

6. Variables de las mujeres (1 registro por episodio)

- ***Programa_id***

Ya descrito

- ***Mujer_id***

Ya descrito

- ***Episodio***

Ya descrito

- ***THS***

Campo definido como numérico. Ver equivalencias:

  - *1: Sí, en el momento de la mamografía o en los 6 meses previos*
  - *2: No*
  - *99: Desconocido*

- ***Menopausia***

Campo definido como numérico. Ver equivalencias.

  - *1: Posmenopáusica*
  - *2: Premenopáusica o perimenopáusica*
  - *99: Desconocido*

- ***Antecedentes personales de prueba invasiva con resultado benigno o declaración por parte de la mujer de patología mamaria benigna***

Campo definido como numérico. Ver equivalencias.

  - *1: Sí*
  - *2: No*
  - *99: Desconocido*

- ***Antecedentes familiares de cáncer de mama***

Campo definido numérico. Ver equivalencias.

  - *1: Sí;*
  - *2: No;*
  - *99: Desconocido*

- ***Edad menopausia***

Campo definido como numérico. Se informará un número natural. Se informará el valor '9: No pertinente' cuando la mujer sea pre-menopáusica. No se añaden restricciones de valor para esta variable, pero sí se incluirán valores poco probables.

  - *9: No pertinente*
  - *99: Desconocido*

- ***Edad de la menarquia***

Campo definido como numérico. Se informará un número natural. No se añaden restricciones de valor para esta variable, pero sí se incluirán valores poco probables.

  - *99: Desconocido*

7. Lesiones benignas (1-N registros por mujer, máximo un registro por episodio)

- ***Programa_id***

Ya descrito

- ***Mujer_id***

Ya descrito

- ***Episodio***

Ya descrito

- ***Histología de las citologías***

Campo definido como numérico. Ver equivalencias.

  - *0: Negativo células malignas*
  - *1: Lesión no proliferativa*
  - *2: Lesión proliferativa sin atipia*
  - *3: Lesión proliferativa con atipia*
  - *9: No pertinente*
  - *99: Desconocido*

- ***Histología de la lesión benigna***

Valor definido como texto de longitud 5 con el formato 00000. Los valores posibles son los del Código CIE-O (3ª Ed.). El valor 99 se empleará cuando la histología sea desconocida.

  - *99: Desconocido*

- ***Lateralidad***

Campo definido como numérico. Ver equivalencias.

  - *1: Derecha*
  - *2: Izquierda*
  - *3: Bilateral*
  - *99: Desconocido*

8. Variables referentes a los tumores

- ***Programa_id***

Ya descrito

- ***Mujer_id***

Ya descrito

- ***Episodio***

Ya descrito

- ***Método de detección:***

Campo definido como numérico. Ver equivalencias.

  - *1: Cribado*
  - *2: Cáncer de Intervalo*
  - *3: Fuera de cribado*

- ***Histología tumores***

Valor definido como texto de longitud 5 con el formato 00000. Los valores posibles son los del Código CIE-O (3ª Ed.). El valor 99 se empleará para designar que es desconocido.

  - *99: Desconocido*

- ***Comportamiento tumoral***

Campo definido como numérico. Ver equivalencias.

  - *1: Ductal in situ*
  - *2: Invasivo*
  - *99: Desconocido*

- ***Lateralidad***

Campo definido como numérico. Ver equivalencias:

  - *1: Derecha*
  - *2: Izquierda*
  - *3: Bilateral*
  - *99: Desconocido*

- ***Tamaño tumor***

Campo definido como numérico. Ver equivalencias.

  - *1: Tx*
  - *2: T0*
  - *3: Tis*
  - *4: T1*
  - *5: T1mic*
  - *6. T1a*

- *7. T1b*
- *8: T1c*
- *9: T2*
- *10: T3*
- *11: T4*
- *12: T4a*
- *13: T4b*
- *14: T4c*
- *15: T4d*
- *99: Desconocido*

- ***Afectación ganglionar***

Campo definido como numérico. Ver equivalencias.

- *1: Nx*
- *2: N0*
- *3: N1*
- *4: N2*
- *5: N2a*
- *6: N2b*
- *7: N3*
- *8: N3a*
- *9: N3b*
- *10: N3c*
- *99: Desconocido*

- ***Metástasis***

Campo definido como numérico. Ver equivalencias.

- *1: Mx*
- *2: M0*
- *3: M1*
- *99: Desconocido*

- ***Grado de diferenciación***

Campo definido como numérico. Ver equivalencias.

- *0: No aplicable*
- *1: Categoría I*
- *2: Categoría II*
- *3: Categoría III*
- *99. Desconocido*

- ***Receptores de estrógeno***

Campo definido como numérico. Ver equivalencias.

  - *1: Negativo*
  - *2: Positivo*
  - *99: Desconocido*

- ***Receptores de progesterona***

Campo definido como numérico. Ver equivalencias.

  - *1: Negativo*
  - *2: Positivo*
  - *99: Desconocido*

Rango de valores probables

- Las variables categóricas previamente definidas en la primera parte de este documento sólo podrán adoptar los valores indicados en su correspondiente descripción de valores posibles. Cualquier otro valor fuera de los definidos no será permitido e implicará la imposibilidad de cargar los datos.

- Para el resto de variables (no categóricas) se establecen las siguientes restricciones en los posibles valores que pueden tomar. Los valores fuera de estos rangos que se definen a continuación generarán una alerta, pero no incapacitarán la carga de los datos.

Tabla mujeres

1. La ***fecha de nacimiento*** deberá estar comprendida entre 01/01/1919 y 31/12/1971, o bien 09/09/9999 si es desconocida.

2. La ***fecha de 1ª citación en el programa*** deberá estar comprendida entre 01/01/1989 y 31/12/2015, o bien 09/09/9999 si es desconocida.

Tabla episodios

3. La ***fecha de mamografía de cribado*** deberá estar comprendida entre 01/01/1989 y 31/12/2015 ó bien 09/09/9999 si es desconocida.

4. El ***número de episodio*** deberá ser un número natural entre 1 y 15.

5. El ***número de cribado*** deberá ser un número natural entre 1 y 15.

6. El ***número de estudios intermedios*** deberá ser un número natural entre 0 y 5 (ambos inclusive).

Tabla exploraciones adicionales

7. La ***fecha de la prueba*** deberá estar comprendida entre 01/01/1989 y 30/06/2016, o bien 09/09/9999 si es desconocida.

8. El ***número de prueba*** deberá ser un número natural entre 1 y 20.

a variables de la mujer

9. La ***edad de la menopausia*** deberá ser un número natural comprendido entre 35 y 65, o bien 9 si la mujer es pre-menopaúsica, o 99 si es desconocido.

10. La ***edad de la menarquia*** deberá ser un número natural comprendido entre 8 y 20, o bien 99 si es desconocido.

Tabla lesiones benignas

11. La ***histología de la lesión benigna*** deberá tomar un valor según los posibles códigos de la clasificación SNOMED, Código CIE-O (3ª Ed.), o el valor 99 si es desconocido. Ver anexo 1.

Tabla de los tumores

12. La ***histología del tumor*** deberá tomar un valor según los posibles códigos de la clasificación SNOMED, Código CIE-O (3ª Ed.), o bien el valor 99 si es desconocido. Ver anexo 2.

Reglas lógicas de validación

Ningún campo podrá quedar informado como vacío. Todas las variables deberán ser informadas con algún valor. Se contemplan las categorías *desconocido* y/o *no pertinente* siempre que sea necesario para evitar los registros vacíos.

1. No pueden existir dos mujeres o más con el mismo código de ***"Mujer_id"*** en un mismo programa.

2. Para un mismo ***"Programa_id"*** y ***"Mujer_id"***, la variable ***"Episodio"*** no puede tener dos valores iguales (episodio duplicado).

3. La variable ***"Número de cribado"*** es igual o menor al número de "Episodio".

4. Las variables de ***Fecha*** ('fecha de nacimiento', 'fecha de 1ª citación', 'fecha de mamografía de cribado' y 'fecha prueba') tendrán formato 00/00/0000 en el orden día / mes / año y **NO** mes / día / año.

5. Si para una ***"Mujer_id"*** en un **episodio** concreto, se informa de la existencia de al menos un estudio intermedio, el ***"Resultado inicial de la mamografía de cribado"*** para ese mismo episodio debería ser diferente de 1 (Negativo).

6. Si para una ***"Mujer_id"*** en un **episodio** concreto se genera registro en la tabla **EXPLORACIONES ADICIONALES**, la variable ***"Resultado inicial de la mamografía de cribado"*** para ese mismo episodio debería ser diferente de 1 (Negativo).

7. Si el ***"Resultado final del episodio"*** para un episodio concreto es igual a 1 (Cáncer), el ***"Resultado inicial de la mamografía de cribado"*** deberá ser distinto de 1 (Negativo).

8. Si para una ***"Mujer_id"*** en un episodio concreto se genera registro en la tabla TUMORES, la variable ***"Resultado final del episodio"*** será codificada necesariamente como 1 (Cáncer).

9. Para un ***"Mujer_id"*** y ***"Episodio"*** concreto, si la variable ***"Resultado final del episodio"*** se codifica como 1 (Cáncer), no puede existir para esa misma mujer un episodio posterior.

10. Si para una ***"Mujer_id"*** en un episodio concreto la variable ***"Categoría mamográfica de la prueba de cribado (Bi-rads)"*** es 1 (normal), no debería de generarse registro en la tabla exploraciones.

11. Para una ***"Mujer_id"*** en un episodio concreto, si en la tabla **EPISODIOS** la variable **"Lateralidad"** se clasifica como 99 (Desconocida), los campos "Masa_I", "Distorsión_I", "Calcificaciones_I", "Asimetría_I", y "Otros_I" de la variable **"Patrones radiológicos de la prueba de cribado"** deberán clasificarse como 0 (No).

12. Para una ***"Mujer_id"*** en un episodio concreto, si en la tabla **EPISODIOS** la variable **"Lateralidad"** se clasifica como 1 (Derecha), los campos "Masa_I", "Distorsión_I", "Calcificaciones_I", "Asimetría_I", y "Otros_I" de la variable **"Patrones radiológicos de la prueba de cribado"** deberán clasificarse como 0 (No).

13. Para una ***"Mujer_id"*** en un episodio concreto, si en la tabla **EPISODIOS** la variable **"Lateralidad"** se clasifica como 2 (Izquierda), los campos "Masa_D", "Distorsión_D", "Calcificaciones_D", "Asimetría_D", y "Otros_D" de la variable **"Patrones radiológicos de la prueba de cribado"** deberán clasificarse como 0 (No).

14. Si para una ***"Mujer_id"*** en un episodio concreto se genera una ***"prueba_id"*** del tipo 2.1 (Citología) en la tabla **EXPLORACIONES ADICIONALES**, y **no genera registro en la tabla TUMORES**, debería generar un registro en la tabla **LESIONES BENIGNAS** (aunque sea desconocido) en ese mismo episodio.

15. Si para una ***"Mujer_id"*** en un episodio concreto se genera una ***"prueba_id"*** del tipo 2.2; 2.3; 2.4; o 2.5; (Biopsia) en la tabla **EXPLORACIONES ADICIONALES, y no genera registro en la tabla TUMORES**, debería generar un registro en la tabla **LESIONES BENIGNAS** (aunque sea desconocido) en ese mismo episodio.

16. Si para una ***"Mujer_id"*** en un episodio concreto se genera registro en la tabla **LESIONES BENIGNAS**, debería de generar alguna ***"prueba_id"*** del tipo 2.1; 2.2; 2.3; 2.4; 2.5; 2.6; o 2.9, en la tabla **EXPLORACIONES ADICIONALES**.

17. Si para una ***"Mujer_id"*** en un episodio concreto la variable ***"Histología básica de las citologías"*** es distinta de 9 (no pertinente), deberá de generar una ***"prueba_id"*** del tipo 2.1 (Citología) en la tabla **EXPLORACIONES ADICIONALES**.

18. Si para una ***"Mujer_id"*** en un episodio concreto la variable ***"Histología de la lesión benigna"*** es distinta de 99 (Desconocido), deberá de generar una ***"prueba_id"*** del tipo 2.1; 2.2; 2.3; 2.4; o 2.5; (Biopsia) en la tabla **EXPLORACIONES ADICIONALES**.

19. Si para una ***"Mujer_id"*** en un episodio concreto se genera registro en la tabla **TUMORES**, no deberá registro en la tabla **LESIONES BENIGNAS** en ese mismo episodio.

20. La variable ***"Menopausia"*** no puede ser codificada como 1 (posmenopáusica) en un episodio y en cualquier otro episodio posterior como 2 (pre o perimenopáusica).

21. La variable ***"Menopausia"*** si en un episodio se clasifica como 2 (premenopáusica), la variable ***"Edad de la menopausia"*** debería ser clasificada como 9 (No pertinente) en ese mismo episodio.

22. La variable ***"Menopausia"*** si en un episodio se clasifica como 1 (posmenopáusica), la variable ***"Edad de la menopausia"*** no debería ser clasificada como 9 (No pertinente) en ese mismo episodio.

23. La variable ***"Antecedentes personales de prueba invasiva con resultado benigno"*** si en un episodio se clasifica como 1 (Sí), en los subsiguientes no puede clasificarse como 2 (No). Debe seguir siendo clasificada como 1.

24. La variable ***"Antecedentes familiares de cáncer de mama"*** si en un episodio se clasifica como 1 (Sí), en los subsiguientes no puede clasificarse como 2 (No). Debe seguir siendo clasificada como 1.

25. La variable ***"Edad de la menopausia"*** si en un episodio se clasifica con un valor distinto de 9 (No pertinente), o 99 (Desconocido), la variable ***"Menopausia"*** debería ser distinta de 99 (Desconocido).

Valores poco probables

Los valores poco probables se definen para que la base de datos dé una señal de alerta (***warning***) indicando que podría haber una incoherencia.

1. Número de estudios intermedios ≥ 4.

2. N prueba ≥ 8 por episodio.

Individualized breast cancer risk prediction models applied to population-based screening mammography
PhD Thesis

Author
Javier Louro Aldamiz-Echevarría
jlouroae@gmail.com

University
UAB, Universitat Autònoma de Barcelona
Department of Pediatrics, Obstetrics & Gynecology,
and Preventative Medicine and Public Health

Carried out on
IMIM, Institut Hospital del Mar d'Investigacions Mèdiques

Graphic design
Helena Cepeda Martín

Printed at
Bookprint digital

1st Edition
Barcelona, April 2021