



**Universitat Autònoma de Barcelona (UAB)**

Facultat de Medicina

Doctorat en Pediatria, Obstetrícia i Ginecologia, Medicina Preventiva i Salut Pública

**REPTES EN L'AVALUACIÓ DEL CRIBRATGE POBLACIONAL**

**DEL CÀNCER DE MAMA:**

*L'efecte sobre la població no cribrada,*

*la influència de la via diagnòstica en les característiques dels tumors*

*i l'impacte de la prova d'imatge en els resultats*

**Tesi Doctoral**

Memòria presentada per Laia Domingo Torrell per optar al grau de Doctora per la Universitat Autònoma de Barcelona.

Signatura de la directora de tesi

Dra. Maria Sala Serra

Signatura de la doctoranda

Laia Domingo Torrell

Departament de Pediatria, Obstetrícia i Ginecologia, Medicina Preventiva i Salut Pública

**Universitat Autònoma de Barcelona, 2013**

Disseny de la coberta: Jordi Sales  
Impressió: Format Digital S.L.

Aquesta tesi doctoral s'ha realitzat al Servei d'Epidemiologia i Avaluació de l'Hospital del Mar-IMIM, a Barcelona, sota la direcció de la Dra. Maria Sala Serra. Es presenta com a compendi de publicacions.

Barcelona, Maig del 2013



## AGRAÏMENTS

Enmig de la voràgine d'aquests últims dies, costa trobar un moment per aturar-se, mirar enrere, i fer repàs de l'experiència que ha estat realitzar una Tesi Doctoral. Del primer que un es dona compte és que és impossible fer aquest camí sola. En el meu cas, l'experiència no pot ser més positiva i, a més a més, en tot moment he estat acompanyada de persones excepcionals. Crec que és el moment de donar les gràcies.

En primer lloc, vull agrair a la Maria Sala el fet d'haver-me obert les portes del Servei d'Epidemiologia i Avaluació, ara fa gairebé 5 anys. Des del primer moment, em va transmetre el seu suport i la seva confiança, animant-me a assumir responsabilitats i a prendre decisions. Al mateix temps que ha fomentat la meva implicació en projectes i publicacions, m'ha transmès els seus coneixements i la seva manera de treballar, donant-me així la seguretat imprescindible per ajudar-me a créixer professionalment i com a investigadora. Ha estat i continua sent un gust treballar al seu costat, encomanar-se del seu bon caràcter i la seva actitud resolutiva davant de qualsevol repte. Moltes gràcies Maria, al teu costat he après moltíssim.

A tots els companys i amics del Servei d'Epidemiologia i Avaluació. Un gran equip de persones, que m'han fet sentir com a casa. El bon ambient del Servei no es troba a tot arreu i part del mèrit el té en Xavier Castells, que és capaç d'identificar i potenciar el millor de cada persona. Així, aprofito per donar les gràcies al Xavier, pels grans moments de *brainstorming* i, òbviament, per la confiança que m'ha transmès en tot moment. Al Francesc Macià, per tenir sempre oberta la porta del seu despatx, disposat a resoldre qualsevol dubte lingüístic, referent al programa de cribratge de l'Hospital del Mar, o de qualsevol mena. Molt, molt especialment a tots aquells amb qui hem compartit viatges, sopars, cafès, tes i despatx, que més enllà de companys de feina, han esdevingut amics. Amb moltíssim carinyo, a l'Esther, pel seu suport diari, per haver-me cuidat tant i per la confiança i amistat fruit de tants moments compartits. A la Cristina i l'Andrea per l'alegria i la passió amb la que fan totes les coses i, evidentment, per estar sempre disposades a organitzar escapades de germanor. A la Mercè Comas i els seus comentaris suggerents i intel·ligents, que mai deixaran de sorprendre'm! A l'Anabel, per tot el temps que hem estat compartint el dia a dia al PRBB, espero que tinguis molta sort i que aviat puguem veure't doctorada. Al Jordi, per estar sempre de bon humor i disposat a donar un cop de mà en qualsevol cosa, sobretot si té a veure amb algun gràfic d'R... I a la Marta, per haver-me ensenyat que tot és possible, haver-me deixat compartir amb ella la seva gran experiència vital, i també, per totes les xerrades llargues i intenses que hem compartit. Un plaer treballar amb tots vosaltres.

Vull recordar-me també de totes les persones amb qui hem treballat conjuntament en diferents projectes. A tot l'equip de radiòlegs, oncòlegs i patòlegs de l'Hospital del Mar amb qui sempre hem pogut comptar per tirar endavant projectes i resoldre dubtes d'última hora. A tot l'equip del projecte INCA, pel temps i

l'esforç compartit, especialment a la Monste Rué i al grup de Lleida. I a la Lola i la Pepa de València, pel seu esforç i la seva constància en tirar endavant l'estudi INCA i també per haver contribuït enormement en fer d'aquest projecte un camp d'entrenament per afrontar qualsevol repte futur.

Viure durant uns mesos a Dinamarca ha estat una de les experiències més boniques que hauré viscut tot fent aquesta tesi. Els tres mesos a la Universitat de Copenhaguen em van servir per conèixer altres maneres de treballar, de relacionar-se, i també per conèixer-me millor a mi mateixa. Treballar al costat de l'Elsebeth Lynge i tot l'equip de Centre d'Epidemiologia i Cribratge, fou una experiència molt enriquidora. A més a més, tant dins com fora del despatx em vaig sentir com a casa, pels companys que vaig trobar i per les nombroses visites que vaig tenir. Grans moments en bicicleta, fred i pluja a Copenhaguen!

Tot i que la tesi pròpiament ha estat fruit de la feina i l'esforç dels últims 4 anys, molt probablement mai hauria arribat en aquest punt si anys enrere no hagués conegut a la Mercè i l'Olga i m'haguessin despertat l'interès per la recerca. Vull donar les gràcies molt, molt especialment a l'Olga, per la seva amistat, per la seva energia inesgotable, pels seus bons consells i per les moltes hores i projectes compartits.

I en aquest camí, com deia, he estat molt ben acompanyada... Gràcies a haver estudiat Farmàcia, a hores d'ara tinc unes amigues com la Lídia i la Núria. Amb elles, hem compartit moments de decisions importants, de quedar-se o de marxar, de dir que sí o dir que no. D'elles he après a no tenir por de les pròpies decisions, i a comprovar que l'amistat és resistent al temps i a l'espai! Lídia, sóc incapaç d'imaginar-me aquests últims anys sense les nostres nits a Gràcia, a Sants o per Skype! A l'Ester i a la Laura, per ser-hi sempre! Perquè hem sabut ajudar-nos en tot moment, i hem estat juntes fent ceràmica, acabant carreres i anant de congressos. Perquè setmana a setmana ens hem posat al dia, compartint i rient de les mil coses que han passat aquests últims anys. Juntes fem un gran equip! Espero seguir compartint aquest viatge amb vosaltres, no m'ho imagino d'una altra manera.

Al Jordi! Per la seva infinita paciència -24 hores al dia i 365 dies d'any. Per ensenyar-me a simplificar els problemes, pels bons consells, per saber sempre com fer-me somriure. I per l'oportunitat de tornar a viure al teu costat un moment important de la meua vida, com tants d'altres que ja hem compartit i per tots els que vindran. I a tots els Sant-Gisbert, per interessar-vos en tot moment i animar-me en aquest camí!

I finalment d'una vull agrair i vull dedicar aquest treball als meus pares, que se'ns dubte puc afirmar que seran les persones que els hi farà més il·lusió tenir aquesta tesi entre les mans. Per haver-me fet costat en totes les decisions i per haver-me ensenyat com d'important és esforçar-se per les coses, tenir una actitud positiva davant qualsevol situació i saber valorar els bons moments. Prometo ser-hi més a partir d'ara!

Moltes gràcies a tots!

## FINANÇAMENT

La realització d'aquesta tesi doctoral ha estat possible gràcies a diferents ajuts:

- Beca concedida pel CIBER d'Epidemiologia i Salut Pública (CIBERESP), en la convocatòria de l'any 2008 per accions específiques, pel projecte "*Características biológicas de los cánceres de intervalo: comparación con los cánceres de mama detectados en el marco del programa de detección precoz*" (AE08\_004).
- Projecte de recerca finançat pel Fondo de Investigación Sanitaria (FIS-ISCIII), en la convocatòria de l'any 2007, pel projecte "*Evaluación de la tasa de detección de cáncer y de resultados falsos positivos tras la implantación de la mamografía digital en los programas de detección precoz del cáncer de mama en España*" (PI07/90293).
- Beca concedida pel CIBER d'Epidemiologia i Salut Pública (CIBERESP) "*Ayuda para estancias breves en el extranjero con el fin de obtener la mención europea del doctorado*" en la convocatòria de l'any 2011, per realitzar una estada de 3 mesos al *Center for Epidemiology and Screening, Department of Public Health, University of Copenhagen*, Dinamarca.
- Ajut per a la finalització de tesis doctorals de la Fundació IMIM, en la convocatòria de l'any 2013.





## RESUM

**Introducció.** La detecció precoç del càncer de mama s'ha visualitzat com la millor estratègia per millorar el pronòstic de la malaltia i disminuir-ne la mortalitat a llarg termini. Seguint les recomanacions del Consell Europeu, molts països europeus han posat en marxa programes de caràcter poblacional que ofereixen majoritàriament mamografies biennals, a les dones d'entre 50-69 anys. Tanmateix, els resultats sobre l'efectivitat del cribratge en el context poblacional són controvertits, apuntant a la possibilitat de que l'impacte en la reducció de la mortalitat sigui menor del que generalment s'havia considerat i, en canvi, els efectes adversos majors. En vista de l'actual controvèrsia, aquesta tesi està centrada en tres qüestions que formen part del debat. Donat que l'estimació dels beneficis del cribratge es basen en la comparació entre població cribrada i no cribrada, s'han estudiat les taxes de detecció de càncer de mama en dones no cribrades, en una zona amb cribratge poblacional. Després, s'han explorat els càncers d'interval, que són indicadors a curt termini de la qualitat dels programes i dels quals, fins ara, pocs treballs han aportat dades referents al seu patró molecular i al seu pronòstic. Finalment, s'ha estudiat l'impacte de l'ús de la mamografia digital sobre les taxes de reconvoatòria, centrant-se en els patrons radiològics i les característiques dels tumors.

**Objectius.** La tesi té com a principal objectiu aprofundir en l'avaluació del cribratge poblacional del càncer de mama, concretament i) estudiant l'impacte de la introducció dels programes poblacionals de cribratge en la detecció del càncer de mama en el conjunt de la població diana; ii) avaluant l'associació entre la via diagnòstica (detecció en el cribratge o en l'interval entre mamografies) amb les característiques i el pronòstic dels tumors; i iii) estudiant l'impacte del tipus de prova d'imatge sobre els indicadors de qualitat del programa.

**Mètodes.** S'han realitzat quatre treballs a partir de l'estudi de tres cohorts de dones participants al cribratge a la ciutat de Barcelona, a Espanya, i a Dinamarca. Pel primer objectiu, s'estudiaren les taxes d'incidència del càncer de mama en la població diana de dues regions daneses, des de l'inici del primer programa poblacional, l'any 1991, fins l'any 2008. Es recolliren els diagnòstics de càncer de mama realitzats en el context del cribratge, entre intervals de cribratge i en dones no participants ( $n=3.290$ ,  $n=1.257$  i  $n=1.548$ , respectivament). Es calcularen les taxes en funció de la via diagnòstica i es compararen els principals indicadors de cribratge i les característiques dels tumors diagnosticats en les dues zones estudiades.

Pel segon objectiu, s'estudiaren els càncers d'interval en la cohort de dones cribrades al programa de l'Hospital del Mar de Barcelona entre els anys 1995-2008. Es dissenyà un estudi cas-control per comparar les característiques biològiques dels càncers d'interval ( $n=115$ ) i els càncers detectats en el cribratge

(n=115). Es procedí a la classificació radiològica dels càncers d'interval seguint les recomanacions europees per diferenciar entre veritables càncers d'interval, tumors falsos negatius, tumors ocults i de signes mínims. Es compararen les característiques clíniques al diagnòstic, el perfil molecular i el paper de la densitat mamària entre els grups d'estudi. Es realitzaren anàlisis bivariades i multivariades per estudiar les associacions entre les variables relacionades amb la dona i els tumors, entre grups d'estudi. A partir de la mateixa cohort s'avaluà la supervivència global i la supervivència lliure de malaltia en dones amb càncers detectats en el cribratge (n=98), dones amb veritables càncers d'interval (n=34) i s'incorporaren casos de dones amb tumors diagnosticats fora del cribratge (n=98). Es calculà la supervivència global (a partir de la mortalitat) i la supervivència lliure de malaltia (basada en les recaigudes) mitjançant corbes de Kaplan-Meier i posteriorment es realitzaren anàlisis de Cox ajustats per variables clínico-patològiques.

Pel tercer objectiu, s'avaluà l'impacte de la transició de la mamografia analògica a la digital a partir d'una cohort de 103.613 dones cribrades entre 1996 i 2007 en quatre programes poblacionals espanyols. Es compararen els patrons radiològics i els valors predictius positius (VPP=càncers diagnosticats, entre el total de dones reconvoques per proves complementàries) obtinguts en el període analògic i digital. També es compararen les característiques dels tumors diagnosticats amb ambdues tècniques.

Tant la gestió de les bases de dades com l'anàlisi estadística, es realitzaren amb el software SPSS (versió 12.0 i 18.0), i amb SAS (versió 9.2). Per la generació de gràfics s'utilitzà el paquet estadístic R (2.12.2). En tots els treballs s'assumeix un error alfa del 5%, considerant així, valors  $p < 0.05$  estadísticament significatius.

**Resultats.** i) La introducció del cribratge poblacional en el context danès suposà un pic de prevalença de diagnòstics de càncer de mama no només entre les dones participants en els programes (5.79 per 1000 persones-any a Copenhaguen i 4.47 a Fyn), sinó també entre les no participants (4.23 i 5.93, a Copenhaguen i Fyn, respectivament). A la zona de Fyn, la incidència de càncer de mama en població no participant fou inclús superior que en població participant. Els programes de cribratge d'ambdues regions presentaven diferències en els criteris d'elegibilitat de la població diana i en el patró de participació. ii) L'estudi dels càncers d'interval revelà diferències clíniques i moleculars en comparar-los amb els tumors detectats en el cribratge. A més, els càncers d'interval foren més freqüents entre dones amb elevada densitat mamària. Concretament, els veritables càncers d'interval presentaren una important associació amb el fenotip triple negatiu [OR= 8.85 (IC95%: 2.03–38.62)] i l'elevada densitat mamària [OR=3.08 (IC95%: 1.03–9.24)]. Les dones amb veritables càncers d'interval presentaren un major risc de recaigudes [HR=1.89; (IC95%: 0.67–5.31)] i de mortalitat [HR= 5.55 (IC95%: 1.61-19.15)], després d'ajustar per variables clínico-patològiques. iii) El VPP de la mamografia digital fou superior al de la mamografia analògica (7.0% vs. 5.5%, respectivament;  $p < 0.001$ ), especialment per les sospites per

masses i calcificacions. Es detectà una lleugera tendència creixent en la detecció de tumors *in situ* en el període digital.

**Conclusions.** i) Les dades procedents de dues zones daneses mostraren que la implementació dels programes poblacionals de detecció precoç de càncer de mama resultà en un pic de prevalença de diagnòstics de càncer de mama tant entre la població participant als programes com entre la no participant. Així mateix, mostraren com alguns indicadors de qualitat poden estar fortament influenciats pels criteris d'elegibilitat. ii) Els veritables càncers d'interval s'associaren amb un perfil molecular agressiu i amb l'elevada densitat mamària. Aquests tumors també s'associaren amb una menor supervivència global, després d'ajustar per variables clínic-patològiques. iii) La introducció de la mamografia digital en el context espanyol millorà el valor predictiu positiu de la prova d'imatge i apuntà a un augment de la detecció de tumors *in situ*.

**Continuïtat de la recerca.** L'experiència adquirida en l'estudi dels càncers d'interval fou clau per iniciar un projecte a nivell nacional per l'estudi de les taxes i els determinants dels càncers d'interval, amb informació de més de 600.000 dones cribrades i de més de 1.000 càncers d'interval classificats radiològicament (Estudi INCA). Aquest projecte també contemplava avaluar l'impacte de la introducció de la mamografia digital en el nombre de càncers d'interval i de tumors falsos negatius. Recentment, s'ha obtingut finançament per completar la informació amb dades de seguiment clínic. Per altra banda, en l'actual debat sobre el balanç entre riscos i beneficis del cribratge i de continus canvis tecnològics, hi ha una necessitat creixent de considerar noves estratègies de cribratge que, a igualtat de costos, permetin millorar els resultats. En aquest context, les dades procedents de grans cohorts poden resultar de gran utilitat per avaluar l'impacte de canvis en les actuals estratègies universals cap a d'altres més personalitzades, enfocades a maximitzar beneficis i disminuir efectes adversos en la població participant.



# ÍNDEX

AGRAÏMENTS .....	5
FINANÇAMENT .....	7
RESUM .....	9
ÍNDEX .....	13
ÍNDEX DE TAULES I FIGURES .....	15
Taules .....	15
Figures .....	15
PRESENTACIÓ .....	17
I. INTRODUCCIÓ .....	19
1. El càncer de mama, un problema de salut pública de primer nivell .....	22
1.1 Epidemiologia del càncer de mama .....	22
1.2 Factors de risc associats a la malaltia .....	23
1.3 Història natural del càncer de mama .....	25
1.4 L'heterogeneïtat biològica del càncer de mama i els seus perfils moleculars .....	26
1.5 Detecció i diagnòstic del càncer de mama .....	28
1.6 Factors pronòstic i factors predictius de la malaltia .....	29
1.7 Tractament del càncer de mama .....	30
2. Les pràctiques de cribatge poblacionals: posada en marxa, funcionament, efectes adversos i indicadors de qualitat .....	32
2.1 Evidència i desplegament dels programes de detecció precoç a Europa i Espanya .....	32
2.2 Efectes adversos de la detecció precoç .....	34
2.3 Indicadors recomanats per l'avaluació del cribatge poblacional .....	38
3. Reptes i dificultats en l'avaluació del cribatge poblacional del càncer de mama .....	41
3.1 El qüestionament del cribatge amb les actuals estratègies .....	41
3.2 Les dificultats per conèixer l'impacte real de la detecció precoç sobre la mortalitat: els biaixos associats al cribatge .....	42
3.3 Avaluar l'impacte del cribatge sobre la detecció del càncer: coexistència de població participant i no participant .....	45
3.4 Avaluar els efectes adversos del cribatge. El cas dels càncers d'interval .....	46
3.5 L'impacte de les noves tecnologies en els indicadors del cribatge .....	47
4. Presentació dels treballs que conformen la tesi .....	49

5. Justificació de la unitat temàtica .....	50
II. HIPÒTESIS I OBJECTIUS.....	51
1. Hipòtesis .....	53
2. Objectius .....	54
III. MÈTODES I RESULTATS .....	55
1. Article 1 .....	59
2. Article 2 .....	71
3. Article 3 .....	83
4. Article 4 .....	93
IV. DISCUSSIÓ.....	105
1. Resultats principals .....	107
2. Discussió conjunta dels articles .....	107
3. Limitacions .....	110
4. Fortaleses .....	111
5. Continuïtat i futures línies de recerca.....	112
V. CONCLUSIONS I IMPLICACIONS .....	115
1. Conclusions.....	117
2. Recomanacions i implicacions en Salut Pública .....	118
VI. ANNEX.....	121
PhD Dissertation - Extended Summary in English.....	123
VII. BIBLIOGRAFIA.....	137

## ÍNDEX DE TAULES I FIGURES

### Taules

<b>Taula 1.</b> Resum dels principals factors de risc associats amb el càncer de mama.....	24
<b>Taula 2.</b> Definició dels principals fenotips del càncer de mama, en funció de l'expressió de receptors d'èstrogen, de progesterona i el factor de creixement HER2. Es resumeixen les principals característiques relacionades amb el pronòstic i la resposta al tractament. ....	27
<b>Taula 3.</b> Recomanacions per la realització del cribratge mamogràfic a dones de risc intermedi per part de les principals agències internacionals.....	33
<b>Taula 4.</b> Resum dels treballs que conformen la tesi. S'indica el disseny, la població i el període d'estudi i els subjectes d'anàlisi. ....	58

### Figures

<b>Figura 1.</b> Taxes d'incidència i mortalitat en diferents països, estandarditzades per edat segons la població mundial. ....	23
<b>Figura 2.</b> Esquema dels diferents models d'evolució des de les cèl·lules normals fins a un tumor invasiu. ....	25
<b>Figura 3.</b> Esquema dels riscos i beneficis associats a la mamografia de cribratge. ....	38
<b>Figura 4.</b> Representació esquemàtica del biaix de durada. ....	43
<b>Figura 5.</b> Representació esquemàtica del biaix d'avanç diagnòstic.....	44





## PRESENTACIÓ

El càncer de mama és el càncer més freqüent en dones en països occidentals, i la primera causa de mort relacionada amb càncer entre les dones europees. El seu abast el converteix en un problema de salut pública de primer nivell, pel qual s'han dedicat nombrosos esforços per poder conèixer els determinants i promoure estratègies que en permetin millorar el pronòstic i així el seu impacte en la mortalitat. Els resultats de diferents assajos clínics realitzats a la dècada dels 70-80, van posar de manifest que la detecció precoç del càncer de mama mitjançant la mamografia, permetia la detecció en estadis precoços de la malaltia, fet que representava una millora en el pronòstic i una disminució de la mortalitat. Així, al llarg de les dècades dels 80-90 i fins a l'actualitat, la majoria de països europeus han posat en marxa programes amb diferents nivells d'organització (oportunistics o poblacionals), per fomentar la detecció precoç del càncer de mama. La gran majoria, ofereixen mamografies biennals a dones d'entre 50 i 69 anys.

Tanmateix, existeix certa controvèrsia pel que fa l'efectivitat del cribratge, especialment en el context poblacional. Se n'han posat en dubte tant els beneficis com els efectes adversos, argumentant que la reducció de mortalitat atribuïble a la detecció precoç seria inferior a la que generalment es suposa, i que ens canvi, els efectes adversos serien majors.

Aquesta tesi doctoral s'enmarca en la polèmica actual sobre el cribratge, aprofundint en tres qüestions que formen part del debat. Atès que l'estimació dels beneficis es basa en la comparació de població no cribrada i població cribrada, ens hem centrat en l'estudi de les taxes de detecció de càncer de mama en dones no participants al cribratge, en una zona on s'ofereix cribratge a nivell poblacional. Posteriorment, hem aprofundit en els càncers d'interval, indicadors de qualitat dels programes que a curt termini aporten informació sobre l'efectivitat del cribratge. Fins ara, relativament pocs treballs aporten dades sobre les característiques moleculars i evolució clínica de les diferents categories de càncer d'interval. En últim lloc, hem estudiat l'impacte de la transició de la mamografia analògica a la digital sobre les taxes de reconvoatòria dels programes de detecció precoç, centrant-nos en l'estudi dels patrons radiològics i les característiques del tumors.

Aquesta tesi, que es presenta com a compendi de publicacions, està composta per quatre treballs que pretenen visualitzar la complexitat de l'avaluació, al mateix temps que donen respostes concretes i inèdites sobre alguns dels problemes descrits.

Els diferents treballs s'han realitzat en el Servei d'Epidemiologia i Avaluació de l'Hospital del Mar-IMIM, sota la direcció de la Dra. Maria Sala. Els estudis centrats en els càncers d'interval s'han pogut dur a terme a partir de la informació de dones cribrades al programa de detecció precoç del càncer de mama de

l'Hospital del Mar i tractades al mateix hospital. Aquest estudi es va iniciar amb finançament del CIBERESP, i posteriorment es va ampliar amb recursos propis de l'hospital. El treball centrat en l'impacte de la introducció de la mamografia digital en el valor predictiu de la prova i en les característiques del tumors detectats amb ambdues tècniques forma part d'un projecte finançat pel FIS (PI07/90293) que, agrupant informació de quatre programes de cribratge espanyols, tenia com a objectiu principal avaluar l'impacte de la introducció d'aquesta nova tècnica en la detecció del càncer i l'especificitat de la prova. El treball centrat en l'evolució de les taxes de detecció de càncer dins i fora del cribratge s'ha realitzat amb dades de Dinamarca, i fou possible gràcies a una beca atorgada pel CIBERESP, per optar a la menció europea del doctorat. Amb aquesta ajuda vaig poder fer una estada de tres mesos al *Center of Epidemiology and Screening, Department of Public Health*, de la Universitat de Copenhaguen.

Personalment he estat directament implicada en el disseny i l'execució dels quatre treballs. Cal remarcar que en els treballs centrats en els càncers d'interval vaig participar activament en la fase de disseny, treball de camp, anàlisi de dades i redacció dels manuscrits. En els altres dos, la construcció de les bases de dades formava part de diferents projectes i la meva participació va ser en la concepció del treball, l'anàlisi estadística, la discussió de resultats i la redacció dels manuscrits.

# I. INTRODUCCIÓ

---



Frame i Carlson, l'any 1975 van apuntar quines eren les condicions necessàries que calia que es complissin per aplicar proves de detecció precoç per una malaltia en concret. En primer lloc, que la malaltia representés un problema de salut important, afectant de manera substancial en la qualitat i en l'esperança de vida. Segon, que es tractés d'una malaltia amb una etapa inicial asimptomàtica llarga i que se'n conegués la seva història natural. Tercer, que es disposés d'un tractament eficaç i acceptat per la població en cas de detectar la malaltia en els seus estadis inicials. Quart, que existís una prova de cribratge ràpida, segura, fàcil de realitzar, amb una elevada sensibilitat i especificitat, alt valor predictiu positiu i ben acceptada per metges i pacients. Cinquè, que es tractés d'una prova de cribratge amb una bona relació cost-efectivitat. I finalment, que la detecció precoç de la malaltia i el seu tractament en el període asimptomàtic disminuïssin la morbiditat i la mortalitat de manera global o cadascuna d'elles per separat<sup>1-4</sup>.

El Consell Europeu recomana la detecció precoç del càncer de mama, de cèrvix i de colon, tumors amb una incidència important a la població<sup>5</sup>. A Espanya s'ofereixen el cribratge del càncer de mama amb caràcter poblacional i el de cèrvix de manera oportunística. El cribratge poblacional de càncer colorrectal forma part de l'estratègia del Sistema Nacional de Salut i està en fase d'extensió en la majoria de comunitats autònomes<sup>6,7</sup>.

Degut a la inexistència d'una prevenció primària eficaç, la detecció precoç del càncer de mama s'ha visualitzat com la millor estratègia per millorar el pronòstic d'aquesta malaltia, esdevenint la malaltia neoplàsica amb una detecció precoç més estesa arreu del món, juntament amb una de les més estudiades. Cada any milions de dones són convidades als diferents programes i també cada any apareixen nous avenços, tant clínics, tecnològics com terapèutics. És imprescindible incorporar aquesta complexitat quan es vol profunditzar en el coneixement de la malaltia i en el paper que hi juga la detecció precoç. A més, per oferir unes pràctiques de detecció precoç de qualitat i optimitzar el cribratge, es fa necessari conèixer millor la història natural de la malaltia. Incorporar aquesta informació al cribratge, pot ajudar a millorar-ne l'efectivitat i a definir estratègies que en maximitzin els beneficis i en redueixin els efectes adversos.

La introducció s'estructura en tres apartats. El primer, repassa conceptes generals referents al càncer de mama. El segon, està dedicat a la detecció precoç de la malaltia, i aprofundeix en el funcionament dels

programes poblacionals de cribatge. Finalment, en el tercer apartat s'enumeren els principals reptes en l'avaluació del cribatge fent especial èmfasi en els aspectes en els quals està centrada la tesi.

## **1. El càncer de mama, un problema de salut pública de primer nivell**

El càncer de mama és la principal causa de mort per càncer en dones<sup>8</sup>, al mateix temps que és una de les neoplàsies en que s'ha millorat més el pronòstic al llarg de les últimes dècades. Si bé és cert que el risc es veu modificat per diferents aspectes reproductius<sup>9</sup>, cap d'ells resulta útil per prevenir la malaltia. Aquest fet, sumat a una història natural de la malaltia amb una fase preclínica llarga i a una elevada prevalença a nivell poblacional, han contribuït a fer de la prevenció secundària, és a dir, el diagnòstic precoç o cribatge, la millor estratègia per millorar-ne el pronòstic i, juntament amb la millora dels tractament, reduir-ne la mortalitat dràsticament.

A continuació es repassen els aspectes generals de la malaltia, des de l'epidemiologia, factors de risc, història natural i biologia, fins a la seva detecció, diagnòstic i tractament.

### **1.1 Epidemiologia del càncer de mama**

El càncer de mama representa, de lluny, el càncer més freqüent en dones a nivell mundial. S'estima que l'any 2008 es van diagnosticar 1.38 milions de nous casos, fet que representa el 23% del total de càncers diagnosticats, i es varen produir al voltant de 458,000 morts (el 14% de les morts per càncer)<sup>8</sup>.

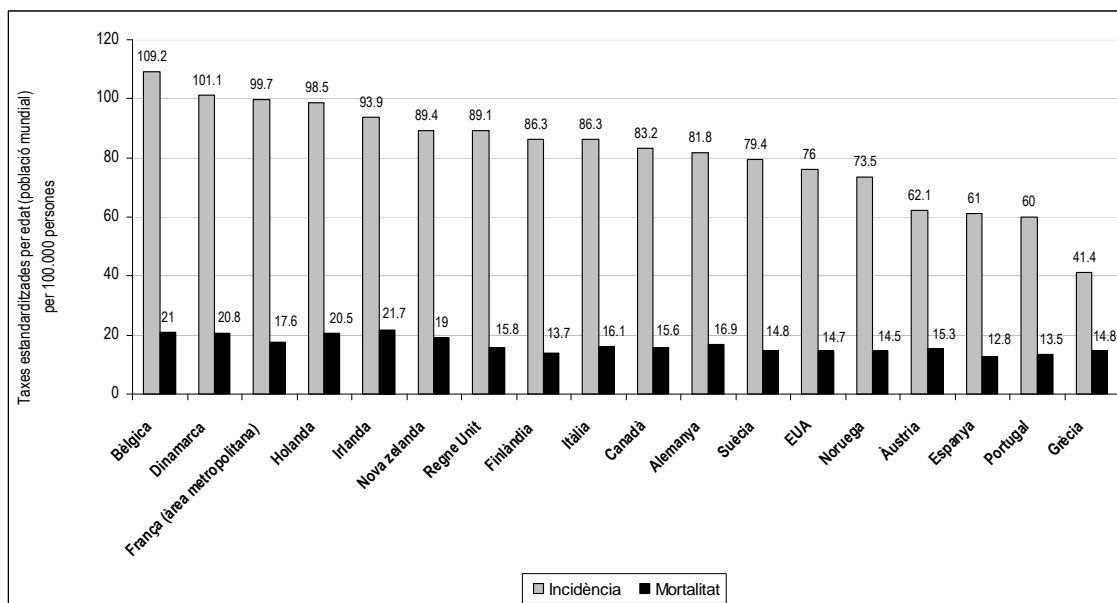
A nivell mundial, però, existeixen importants diferències al comparar les taxes entre països. Les més elevades es troben a Nord Amèrica i Europa, especialment a les zones de l'oest i nord d'Europa, i les més baixes en països asiàtics<sup>8,10,11</sup>

La incidència a nivell mundial ha experimentat una tendència creixent al llarg dels últims 40 anys. Aquest augment, especialment en les dècades dels 80 i 90, ha estat explicat per la combinació de diferents factors entre els quals destaquen l'envelliment de la població, canvis en aspectes reproductius i hormonal, així com la major disponibilitat de serveis de detecció precoç i la millora de les tècniques per la detecció i registre dels casos de càncer<sup>8,12,13</sup>. Tot i així, a partir de l'any 2000, en diferents països, entre ells Espanya, s'ha començat a experimentar un descens de la incidència<sup>14-19</sup>.

També al llarg de les dues últimes dècades, la mortalitat per càncer de mama ha disminuït substancialment, especialment en els països desenvolupats, convertint-se en una de les malalties oncològiques amb una major supervivència, situada actualment per sobre del 80% als 5 anys<sup>20</sup>. Tot i que les diferències entre països són considerables, entre 1989 i 2006 la mortalitat ha disminuït de mitjana un 19% en el context europeu<sup>21</sup>, de manera especialment marcada en dones menors de 50 anys (-37%) i en

el grup de 50-69 anys (-20%). Aquesta disminució s'atribueix principalment a la millora dels tractaments, a l'extensió de les pràctiques de detecció precoç i a canvis organitzatius, especialment a la introducció de les unitats funcionals i al treball multidisciplinari als hospitals<sup>22</sup>.

**Figura 1.** Taxes d'incidència i mortalitat en diferents països, estandarditzades per edat segons la població mundial.



Font de dades: GLOBOCAN 2008. <http://globocan.iarc.fr/>

A nivell espanyol, cada any es diagnostiquen al voltant de 20.000 nous casos i es produeixen 6.000 morts<sup>10</sup>. L'any 2008 la taxa d'incidència ajustada (utilitzant com a referència la població mundial) va ser de 61 casos nous per 100,000 dones i any, sent la mitjana de la Unió Europea de 77.1 casos nous per 100,000 dones i any. Espanya segueix sent un dels països europeus amb una menor incidència, tot i haver experimentat un gran augment al llarg dels últims anys. En la mateixa línia, la taxa de mortalitat també es troba per sota de la mitjana europea. A Espanya, la taxa de mortalitat ajustada l'any 2008 (utilitzant com a referència la població mundial) va ser de 12.9 morts per 100,000 dones i any, i la mitjana de la Unió Europea de 16.6 per 100,000 dones i any<sup>10</sup>.

## 1.2 Factors de risc associats a la malaltia

Entre el 70 i el 80% dels càncers de mama apareixen en dones sense factors de risc explicables. Tan sols entre un 5% i un 10% tenen un origen clarament genètic, relacionat amb els gens mutats BRCA1 i BRCA2<sup>23</sup>.

L'edat, com en la majoria dels càncers, és un factor de risc important de càncer de mama. En general, les taxes augmenten amb l'edat, especialment a partir dels 45-50 anys<sup>10</sup>. A partir dels 50 anys, aquest augment pronunciat de les taxes frena considerablement, fet que s'associa a l'arribada de la menopausa i que apunta a la rellevància dels factors reproductius i hormonals en l'etiologia de la malaltia.

A la Taula 1 queden resumits els principals factors de risc relacionats amb el càncer de mama, juntament amb la direcció i la representació de la magnitud de l'efecte, de manera aproximada.

**Taula 1.** Resum dels principals factors de risc associats amb el càncer de mama.

Factor de risc	Direcció de l'efecte*
Mutació dels gens BRCA1 o BRCA2	↑↑
Història familiar de càncer de mama en familiars de primer grau	↑↑
Lesions benignes	↑↑
Elevada densitat mamària	↑↑
Edat del primer naixement >30 anys vs. <20 anys	↑↑
Menopausa >54 anys vs. <45 anys	↑↑
Elevats nivells endògens d'estrogen	↑↑
Tractament hormonal substitutiu	↑
Exposició a radiacions	↑↑
Menarquia <12 anys vs. >14 anys	↑
Consum d'alcohol (≥1 beguda al dia)	↑
Elevat índex de massa corporal (post menopausa)	↑
Elevat índex de massa corporal (pre menopausa)	↓
Tamoxifè	↓
Estils de vida sedentaris	↑
Raça blanca	↑

Adaptat a partir d'Adami et al. Textbook of Cancer Epidemiology, 2002 i Warner. Clinical Practice. Breast cancer screening. N Engl J Med 2011; 365(11):1025-32.<sup>9,24</sup>

\*Les fletxes indiquen de manera simplificada la direcció i la magnitud de l'efecte: ↑ indica que el factor augmenta el risc de manera moderada. ↑↑ indica que el factor augmenta el risc de manera important. ↓ indica que el factor disminueix el risc de càncer de mama.

Entre els diferents factors de risc destaca la densitat mamària<sup>25</sup>, per la importància que ha anat guanyant en els darrers anys. Diferents hipòtesis i mecanismes biològics s'han proposat per explicar la relació entre l'elevada densitat mamària i el risc de càncer de mama. Per una banda, s'atribueix a l'elevada densitat mamària un efecte emmascarador en la mamografia, dificultant la detecció de tumors petits, i d'histologies especialment lobulars que es presenten sense microcalcificacions<sup>26</sup>. Per altre banda, també s'ha observat que l'associació que es manté fins al cap de deu anys, fet que implica que diferents mecanismes biològics podrien estar implicats en aquesta associació, més enllà a l'efecte emmascarador<sup>9</sup>. Entre ells, s'ha descrit una associació amb els nivells d'estrogen i amb una major proporció de factors de creixement de tipus insulínics (IGF) en el teixit fibrós de la mama, que estimularia el desenvolupament de lesions benignes existents<sup>27</sup>. Actualment, hi ha nombrosos estudis enfocats a elucidar el paper de la densitat mamària en el risc de desenvolupar càncer, molts dels quals fan èmfasi al potencial que pot tenir incorporar aquesta variable en els models de predicció de risc de càncer i en la personalització del cribratge<sup>28</sup>. De fet, recentment, en alguns estats dels Estats Units s'ha aprovat una llei que obliga als professionals a

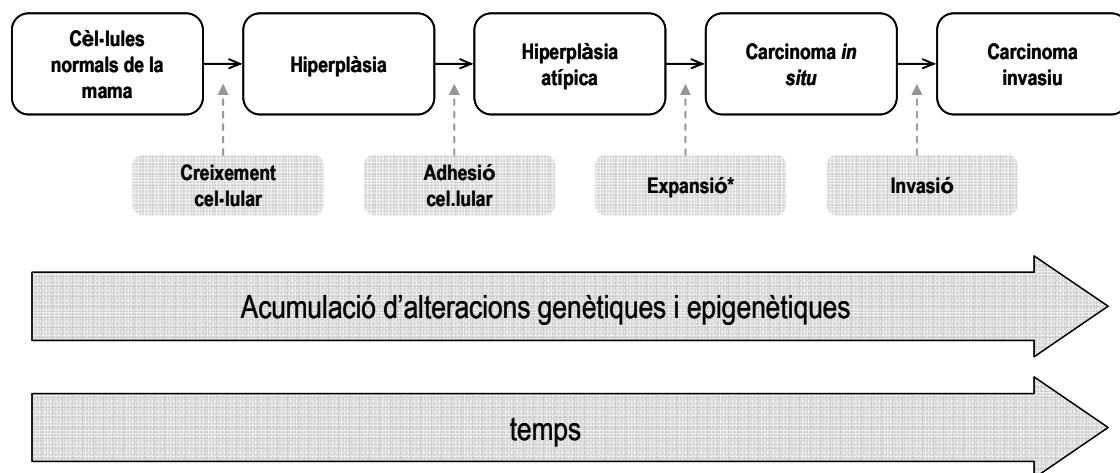


informar a les dones que tenen mames denses i advertir-les que aquest fet els hi augmenta la probabilitat de desenvolupar un càncer ([www.areyoudense.com](http://www.areyoudense.com)).

### 1.3 Història natural del càncer de mama

El càncer de mama és una malaltia heterogènia i complexa, amb diferents característiques morfològiques, biològiques, pronòstiques i amb diferent resposta al tractament. Respecte al seu origen i desenvolupament, existeixen diferents hipòtesis. La més coneguda, el model de Wellings Jensen, postula que el càncer de mama invasiu es desenvolupa a partir d'una lesió benigne de la mama, a partir de l'acumulació de múltiples alteracions genètiques<sup>29,30</sup>. Aquest model presenta una transició progressiva des de l'epiteli normal fins a la presència d'un carcinoma invasiu, passant per diferents fases: lesions amb hiperplàsia no atípica, hiperplàsia atípica, carcinoma *in situ* i carcinoma invasiu<sup>31-33</sup>. Alteracions genètiques i epigenètiques detectades en la lesió benigne i identificades també en els tumors *in situ* i invasius corroboren aquesta via, a més de l'augment del risc de desenvolupar càncer pels subtipus més avançats<sup>31,34-36</sup>. Tot i així, tan sols un percentatge de lesions benignes acaba evolucionant a tumor invasiu, passant per les formes no invasives o *in situ*<sup>37</sup>.

**Figura 2.** Esquema dels diferents models d'evolució des de les cèl·lules normals fins a un tumor invasiu.



Adaptat a partir d' Allred et al. Ductal carcinoma *in situ* and the emergence of diversity during breast cancer evolution. Clin Cancer Res 2008;14:370-378<sup>35</sup>

El model lineal es basa en les hipòtesis de Wellings i Jensen que, de forma resumida, proposen una progressió aparentment contínua des de les cèl·lules normals fins a desenvolupar un tumor invasiu, on al llarg de llargs períodes temporals s'acumularien diferents alteracions. Per altre banda, diferents teories proposen que aquesta progressió no sempre es compleix, i en aquest cas atribueixen a les alteracions genètiques i epigenètiques un rol més important, que en alguns casos implicarien la progressió de cèl·lules 'normals' a formes *in situ* i després invasives, sense la necessitat de transicionar al llarg de llargs períodes de temps per totes les formes premalignes.

Per altra banda, existeix certa evidència que no totes les lesions segueixen aquest model lineal. Farabegoli et al.<sup>38</sup> apunten que els carcinomes *in situ* són precursors possibles dels tumors invasius, però no obligats. Postulen que alguns subtipus es desenvolupen a partir de l'evolució de cèl·lules progenitores, i que posteriorment aquestes serien afectades per diferents mutacions. Les dues hipòtesis són fruit de múltiples treballs i observacions, i donada l'elevada complexitat i heterogeneïtat de la malaltia, no semblen ser incompatibles.

Majoritàriament, les formes *in situ* es diagnostiquen en dones asimptomàtiques i prop del 80-90% a partir de microcalcificacions detectades en la mamografia<sup>39</sup>. Aquestes microcalcificacions acostumen a alinear-se al llarg dels conductes mamaris, a ser granulars i amb densitats irregulars<sup>39</sup>. De totes maneres, algunes formes benignes de patologia mamària, com l'hiperplàsia ductal, el carcinoma lobular *in situ* o l'hiperplàsia lobular atípica poden manifestar-se radiològicament de manera semblant<sup>40</sup>.

L'interès respecte els carcinomes *in situ* ha crescut en paral·lel a la introducció de la detecció precoç i a la millora de les tècniques diagnòstiques. S'estima que a mitjans del segle XX representaven al voltant de l'1-2% de nous diagnòstics, mentre que actualment suposen el 20-30%<sup>37,41</sup>. Pel fet que no tots els carcinomes *in situ* evolucionen a formes invasives, l'augment de la seva detecció amb la introducció dels programes poblacionals, preocupa pel possible sobrediagnòstic que podrien representar.

#### **1.4 L'heterogeneïtat biològica del càncer de mama i els seus perfils moleculars**

Els tumors de mama, més enllà de la diversitat histològica, mostren importants diferències a nivell genètic i molecular que contribueixen en augmentar la seva l'heterogeneïtat i complexitat. A l'any 2000, Perou & Sorlie<sup>42</sup>, van proposar una classificació taxonòmica basada en firmes d'expressió genètica que va permetre diferenciar els càncers en quatre subtipus moleculars: *Luminal-like*, *normal-like*, *basal-like* i HER2-positiu. Aquesta classificació es basa en l'origen de la cèl·lula precursora del càncer a partir de les diferents entitats que conformen l'epiteli mamari (cèl·lules luminals o glàndules secretores, cèl·lules basals o mioepitelials i cèl·lules progenitores o *stem cells*). Aquests tres tipus cel·lulars tenen diferents funcions, morfologia i característiques dins la glàndula mamària i per definició expressen immunofenotips diferents.

Les tècniques d'immunohistoquímica han permès traslladar els coneixements sobre la heterogeneïtat genètica del càncer de mama a la pràctica clínica. Aquestes tècniques es basen en reaccions immunoenzimàtiques que permeten visualitzar certes proteïnes presents a l'interior o a la superfície de les cèl·lules mitjançant el microscopi tradicional, de manera relativament senzilla i poc costosa. Principalment es determina la presència de receptors d'estrogen (RE), receptors de progesterona (RP) i del factor de creixement HER2. La combinació d'aquests biomarcadors representa una bona aproximació als diferents

perfils d'expressió genètica mitjançant la classificació en fenotips moleculars, utilitzats en el dia a dia de l'atenció oncològica ja que són útils per valorar la possibilitat d'administrar tractaments neoadjuvants<sup>43</sup>.

La següent taula resumeix els principals fenotips, basats en la detecció per tècniques d'immunohistoquímica de biomarcadors amb valor predictiu per càncer de mama<sup>43</sup>.

**Taula 2.** Definició dels principals fenotips del càncer de mama, en funció de l'expressió de receptors d'estrogen, de progesterona i el factor de creixement HER2. Es resumeixen les principals característiques relacionades amb el pronòstic i la resposta al tractament.

Fenotip	Expressió biomarcadors	Característiques principals	Resposta a tractament
<b>Luminal A</b>	RE+, RP+, HER2-	Bon pronòstic, baixa expressió de gens de proliferació.	Elevada taxa de resposta a hormonoteràpia.  No es beneficien de la quimioteràpia.
<b>Luminal B</b>	RE+, RP-, HER2- RE+, RP+/-, HER2+*	Moderada/baixa expressió de RE i major expressió de gens de proliferació	Es beneficien de l'hormonoteràpia i la quimioteràpia.
<b>HER2</b>	RE-, RP-, HER2+	Alta expressió de gens relacionats amb l'amplificació d'HER2. Pitjor pronòstic que els subtipus Luminals	Bona taxa de resposta al tractament amb l'anticòs monoclonal <i>trastuzumab</i> (Herceptin®).  Elevada taxa de resposta a quimioteràpia que inclogui antraciclina i/o taxans.
<b>Basal-like triple negatiu</b>	RE-, RP-, HER2-	Expressió de gens característics de les cèl·lules epitelials basals. S'associen a la disfunció del BRCA i EGFR, amb elevada taxa de proliferació i mal pronòstic	No respon ni a hormonoteràpia ni teràpia amb anticòssos.  Bona resposta a la quimioteràpia.

\*Basat en la classificació de consens establerta l'any 2011<sup>43</sup>. Goldhirsch et al. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St Gallen International. Ann Oncol 2011; 22: 1736–1747

Mitjançant les tècniques immunohistoquímiques és possible detectar en les cèl·lules tumorals l'expressió de proteïnes o biomarcadors, que actualment no tenen aplicació terapèutica o valor predictiu, però sí valor pronòstic. És el cas de la presència de la proteïna anòmala p53 o el factor Ki67, dels quals se'n coneix el mecanisme cel·lular, i a partir del qual se'n deriva la seva interpretació a nivell de pronòstic. La proteïna p53 és un supressor tumoral que juga un paper important en l'apoptosi, en l'estabilitat genòmica i en la inhibició d'angiogènesis. En alguns casos, la proteïna es veu alterada i deixa de ser funcional, perdent la capacitat supressora tumoral i afavorint la progressió del tumor. El factor Ki67 és un antigen present a la fase de divisió cel·lular, però absent en la resta del cicle cel·lular, fet que el converteix en un bon marcador de proliferació tumoral. Actualment, existeixen nombrosos estudis centrats en la identificació de

biomarcadors cel·lulars, la majoria d'ells amb l'objectiu a llarg termini de trobar noves dianes terapèutiques o identificar marcadors precoços de la malaltia<sup>44</sup>.

### 1.5 Detecció i diagnòstic del càncer de mama

Generalment, el diagnòstic d'un cas de càncer de mama és fruit de la combinació de proves d'imatge i proves addicionals, que permeten confirmar la malignitat de la sospita radiològica i al mateix temps oferir informació histològica i patològica per afinar el tractament.

Les proves d'imatge més utilitzades són la mamografia, l'ecografia i la ressonància magnètica, les quals normalment es realitzen en combinació amb una exploració física que inclogui palpació. D'aquests procediments el més important, el més específic i el més utilitzat és, sens dubte, la mamografia.

La mamografia és una prova d'imatge basada en rajos X que es realitza sobre la mama i que permet reconèixer les lesions en una fase inicial, utilitzant una baixa dosi de radiació i amb un cost acceptable. A partir de la imatge mamogràfica, el radiòleg és capaç d'identificar certs signes radiològics que seran indicadors d'una probable lesió, tant benigne com maligne. Els signes que més sovint s'associen amb patologia mamària són masses, distorsions, microcalcificacions i asimetries. Tot i així, la mamografia no és una prova diagnòstica, i per tant, davant una sospita de lesió maligne resulta imprescindible realitzar proves complementàries per establir el diagnòstic definitiu. Tradicionalment, s'ha utilitzat i encara s'utilitza la mamografia analògica o convencional, amb sistemes de pantalla-pel·lícula. Tanmateix, des de principis de la dècada del 2000 s'ha anat introduint el mamògraf digital, per les avantatges en la manipulació, definició i emmagatzematge de les imatges<sup>45,46</sup>.

L'ecografia mamària és un mètode auxiliar que en algunes circumstàncies pot arribar a esdevenir la principal tècnica de diagnòstic per la imatge. La seva principal avantatge recau en la distinció entre la naturalesa sòlida o cística de les lesions nodulars identificades per mamografia. També resulta útil en l'estudi de les mames amb un component glandular important, o mames denses<sup>47,48</sup>. Les altres tècniques, com la ressonància magnètica, tenen importància en casos concrets com en el cribatge de pacients d'alt risc, però no s'utilitzen de rutina<sup>49,50</sup>. Tanmateix, el seu ventall d'indicacions cada cop és més ampli i són especialment utilitzades en el seguiment de les lesions postquirúrgiques així com en la valoració de l'extensió de les lesions que es realitza abans de recomanar o no el tractament conservador<sup>50,51</sup>.

Per arribar al diagnòstic de certesa, però, cal recórrer a tècniques invasives que permetin accedir a una mostra amb material biològic. Entre les més utilitzades destaca la biòpsia per punció-aspiració amb agulla fina, o PAAF (en anglès, *fine needle aspiration cytology*, FNAC), les core-biòpsies o les biòpsies quirúrgiques. La primera es tracta d'una citologia a partir de la qual s'obté una mostra de teixit útil per determinar la presència de cèl·lules malignes, però sense conservar-ne la histologia; és possible realitzar-la de manera ambulatoria fet que suposa una avantatge en termes assistencials. Les core-biòpsies són

biòpsies guiades per proves d'imatge, que es realitzen amb pistoles automàtiques i que permeten obtenir material biològic a partir del qual s'estableix el diagnòstic definitiu. Finalment, la biòpsia quirúrgica també permet obtenir una mostra de teixit biològic per arribar al diagnòstic definitiu, tractant-se en aquest cas d'una intervenció més agressiva.

### 1.6 Factors pronòstic i factors predictius de la malaltia

Amb la mostra obtinguda a partir de les proves invasives, es realitza l'examen patològic per establir el diagnòstic final de càncer o descartar-lo. Davant d'un diagnòstic de càncer, es determinen diferents paràmetres amb valor pronòstic i predictiu que aporten informació útil per la presa de decisions sobre el maneig i el tractament de la malaltia. Els factors pronòstic són paràmetres determinats en el moment del diagnòstic o després de la cirurgia que s'associen amb l'evolució clínica del tumor en absència de tractament. En canvi, els factors predictius, aporten informació sobre el grau de resposta a un tractament específic per la malaltia. En el diagnòstic del càncer de mama, els paràmetres que es determinen de rutina són:

La **histologia**, a partir de l'aparença del teixit tumoral. La majoria dels casos deriven de l'epiteli de les cèl·lules que formen els ductes o els lòbuls mamaris; aquests càncers es classifiquen com a carcinomes ductals o lobulars, respectivament, i representen més del 90% de diagnòstics. Per altre banda, en funció de si existeix o no invasió del teixit que envolta les cèl·lules cancerígenes, els carcinomes es classificaran en formes invasives o *in situ*. El sistema internacional de classificació de malaltia (International Classification of Diseases for Oncology, ICD-O) es revisa periòdicament i és el que s'utilitza per la classificació de la localització i la morfologia dels tumors, entre ells de les mama.

El **grau histològic** dóna informació a partir de la comparació de l'aparença de les cèl·lules cancerígenes amb les cèl·lules del teixit normal. Les cèl·lules normals de la mama són cèl·lules diferenciades, amb funcions concretes. Els tumors poc diferenciats o d'alt grau, són tumors més avançats, fet que s'associa amb un pitjor pronòstic. S'acostuma a utilitzar la classificació segons Scarff-Bloom-Richardson modificada<sup>52</sup>, que agrupa els tumors en casos molt diferenciats (I), moderadament diferenciats (II) o poc diferenciats (III).

L'**afectació dels ganglis limfàtics** continua sent el factor pronòstic més important de la malaltia. Des de la dècada del 2000 s'ha generalitzat l'ús de la tècnica del "gangli sentinella", per determinar l'estat ganglionar, fet que ha suposat un gran avanç especialment en els casos poc avançats<sup>53</sup>. El gangli sentinella és el primer nòdul limfàtic on les cèl·lules tumorals arribarien en cas de disseminació, i es determina la seva afectació mitjançant un radioisòtop. En cas que el gangli sentinella sigui positiu, es procedeix al buidament ganglionar de la regió axil·lar; quan és negatiu s'assumeix que la resta de ganglis tampoc estan afectats i no es fa cap intervenció. La posada en pràctica d'aquesta tècnica ha suposat un

gran avanç i ha permès evitar el buidament ganglionar en un percentatge molt important de càncers de mama, contribuint a la millora de la qualitat de vida per les pacients, la disminució de la morbiditat per limfoedema i la reducció de costos pel sistema sanitari<sup>54</sup>.

**L'estadi** és també una classificació internacional que es basa en la mida del tumor (T), l'afectació limfàtica (N) i les metàstasis (M) en el moment del diagnòstic. La classificació TNM es pot dur a terme en base a l'estudi clínic o patològic de la peça tumoral, i segueix els criteris de la Unió Internacional pel Control del Càncer (Union for International Cancer Control, UICC). Una mida gran, afectació ganglionar i presència de metàstasis tindran estadis majors així com pitjor pronòstic. Els principals estadis són:

Estadi 0: formes pre-invasives o *in situ*.

Estadi 1-3: el tumor es troba localitzat a la mama o als ganglis limfàtics regionals

Estadi 4: el càncer s'ha estès a diferent òrgans.

**L'estatus de certs receptors** a la superfície de les cèl·lules tumorals és un altre paràmetre que es determina rutinàriament des de que hi ha teràpies que actuen a aquest nivell. Generalment es determina l'expressió de receptors d'estrògens, de progesterona i HER2.

## 1.7 Tractament del càncer de mama

El tractament del càncer de mama ha canviat considerablement al llarg de les últimes dècades i de fet es considera que la millora en les opcions terapèutiques així com els canvis en els models organitzatius per abordar-la, són factors que han contribuït d'una manera significativa a la millora del pronòstic de la malaltia i a la reducció de la mortalitat<sup>22</sup>.

Les principals estratègies terapèutiques són la cirurgia, la radioteràpia, la quimioteràpia, la hormonoteràpia i la teràpia biològica o amb anticossos. En funció de l'estadi en el qual s'ha diagnosticat el càncer i la determinació dels altres factors pronòstic i predictius es decideix l'estratègia terapèutica, que pot implicar utilitzar únicament algun dels esmentats tractaments o, normalment, estratègies combinades.

La cirurgia per l'excisió del tumor i del teixit sa que l'envolta, acostuma a ser el primer tractament. El tipus de cirurgia, radical o conservadora, s'elegeix en funció del tipus i l'abast del càncer. La cirurgia radical o mastectomia es realitza en casos avançats de la malaltia i generalment s'acompanya de la reconstrucció de la mama. Tanmateix, cada cop és més freqüent la cirurgia conservadora en combinació amb radioteràpia o quimioteràpia postoperatòria, la qual cosa limita les recurrències locals<sup>9</sup>. L'abast de la cirurgia dependrà del tipus i la grandària del tumor, juntament amb la quantitat de teixit que ha de ser extret i la mida de les mames.

La radioteràpia està indicada principalment en els casos on s'ha realitzat cirurgia conservadora i, mitjançant dosis controlades de radiació, té com a principal objectiu eliminar les cèl·lules cancerígenes que puguin quedar al teixit mamari. S'ha vist que millora el risc de recaiguda local, i en alguns casos també la supervivència global<sup>55</sup>.

La quimioteràpia implica l'administració de fàrmacs citotòxics, per eliminar les cèl·lules tumorals, generalment després de la cirurgia. En aquest cas es denomina teràpia adjuvant. En alguns casos, però, es realitzen sessions de quimioteràpia abans de la cirurgia, per intentar disminuir la mida del tumor. Aquesta pràctica s'anomena teràpia neoadjuvant i s'acostuma a realitzar en casos de tumors de grans dimensions. S'utilitzen diferents fàrmacs, l'elecció dels quals es fa en funció de les característiques i el grau de disseminació del càncer. La quimioteràpia és la teràpia d'elecció en la majoria de casos diagnosticats en estadis 2-4 i resulta especialment efectiva pels tumors ER-. Ciclofosfamida i doxorubicina, ciclofosfamida, doxorubicina i docetaxel, o ciclofosfamida, metotrexat i fluorouracil, són algunes de les combinacions de fàrmacs més utilitzades. Totes elles, tenen associats importants efectes adversos, com la pèrdua temporal de cabell i trastorns digestius.

La teràpia hormonal només és efectiva pels tumors RE+ i/o RP+. Amb el tractament hormonal s'espera disminuir els nivells d'hormones circulants, i així interferir en el creixement tumoral. L'elecció del tractament també es fa en funció de les característiques dels tumors, alhora que es té en compte l'edat i l'estat hormonal de les dones. Tot i que en alguns casos s'administra abans de la cirurgia per disminuir les dimensions del tumor i facilitar la seva excisió, generalment s'administra com a teràpia adjuvant en el postoperatori. El tamoxifè impedeix la unió de l'estrogen al seu receptor de membrana, inhibint el seu efecte sobre el creixement tumoral. Els inhibidors de l'aromatasa (anastrozol, exemestan i letrozol) actuen bloquejant l'aromatasa, un enzim implicat en la producció d'estrògens després de la menopausa. L'excisió o supressió ovàrica només es fa en casos de dones premenopàusiques (on es produeixen aquestes hormones abans de la menopausa).

Els tumors que expressen el factor de creixement HER2, són susceptibles de beneficiar-se de la teràpia amb *trastuzumab* (Herceptin®). Aquest fàrmac és un anticòs monoclonal, que es va començar a comercialitzar a partir de l'any 2004 i que ha demostrat millorar la supervivència lliure de malaltia dels casos HER2+ i diagnosticats en estadis 1-3 fins a valors pròxims al 87%<sup>56</sup>. Aquesta dada és especialment rellevant ja que els tumors HER2+ representen al voltant del 25-30% dels càncers de mama. En absència de tractament, la presència d'aquest factor s'associa a mal pronòstic i a l'augment del risc de recaigudes, ja que es tracta d'un factor de creixement que estimula la divisió cel·lular. Tot i així, es tracta d'un medicament que requereix administració parenteral i amb importants efectes adversos.

## **2. Les pràctiques de cribratge poblacionals: posada en marxa, funcionament, efectes adversos i indicadors de qualitat**

La detecció precoç del càncer de mama mitjançant mamografia té com a principal objectiu reduir la mortalitat global per aquesta malaltia a partir de la detecció dels tumors en fases poc avançades. Per assolir aquest objectiu de perspectiva poblacional, és un requisit indispensable aconseguir un àmplia participació, fet que implica que un gran nombre de dones sanes se sotmetin periòdicament a exàmens mamogràfics. Des d'una òptica poblacional s'espera que la implementació del cribratge es tradueixi a mitjà o llarg termini amb una disminució de la mortalitat de la malaltia en la població, encara que a nivell individual no totes les dones participants es beneficiïn directament de la participació als programes organitzats. Tot i les diferents estimacions, es calcula que per cada 1000 dones cribrades biennalment durant 20 anys, entre els 50-51 i els 68-69 anys i seguides fins als 79 anys (el benefici de cribrar una dona de 69 anys trigará uns anys a aparèixer) es diagnosticaran 71 càncers de mama, s'evitaran entre 7 i 9 morts, es sobrediagnosticaran 4 tumors i 200 dones experimentaran un resultat fals positiu<sup>57</sup>. Tanmateix, el recent informe Marmot, dut a terme per un panell independent on s'han avaluat els riscos i beneficis del cribratge, conclou que cada mort per càncer evitada, haurà suposat el sobrediagnòstic de 3 tumors<sup>58</sup>.

Els següents punts resumeixen aspectes clau de les pràctiques poblacionals de cribratge del càncer de mama. Es repassa l'evidència que en justificà el seu desplegament en base dels beneficis del cribratge i l'abast actual en el context europeu. També es descriuen els efectes adversos i els principals indicadors recomanats per l'avaluació dels programes i finalment es repassen les controvèrsies que actualment envolten les pràctiques de cribratge.

### **2.1 Evidència i desplegament dels programes de detecció precoç a Europa i Espanya**

La implementació de programes poblacionals de cribratge arreu d'Europa es va fer en base als resultats de diversos assajos aleatoritzats i controlats realitzats en diferents contextos: Nova York (EUA)<sup>59</sup>, Canadà<sup>60</sup>, Malmö (Suècia)<sup>61</sup>, Estocolm (Suècia)<sup>62</sup>, Göteborg (Suècia)<sup>63</sup>, Edimburg (Regne Unit)<sup>64</sup>. Malgrat que alguns han estat qüestionats metodològicament (bàsicament pels criteris d'aleatorització) en global, els resultats dels assajos demostraren una reducció de la mortalitat en dones participants d'entre 50-69 anys, variable entre un 15% i 35%, segons l'edat, els anys de seguiment, el nombre de dones participants i la periodicitat de la mamografia<sup>7,65,66</sup>.

Al llarg de la dècada de 1990, la majoria de societats científiques començaren a recomanar el cribratge mamogràfic per la detecció precoç del càncer de mama, i nombrosos països posaren progressivament en marxa programes d'abast poblacional<sup>67</sup>. Actualment, el Consell Europeu recomana el cribratge



poblacional del càncer de mama a les dones d'entre 50-69 anys, i que els programes segueixin els estàndards establerts per les "Guies Europees de Garantia de Qualitat en el Cribatge de Càncer de Mama"<sup>5,68,69</sup>.

L'any 2007, 26 dels 27 estats membres de la Unió Europea tenien programes per la prevenció de càncer de mama, la majoria dels quals d'abast poblacional (Alemanya, Àustria, Bèlgica, Dinamarca, Eslovènia, Espanya, Estònia, Finlàndia, França, Hongria, Irlanda, Itàlia, Holanda, Luxemburg, Malta, Polònia, Portugal, República Txeca, Regne Unit, Romania i Xipre). En tan sols 11 països, però, el desplegament del cribatge era complet a nivell de tot el país. Tot plegat suposa que a la Unió Europea prop de 59 milions de dones es troben en l'edat diana segons les recomanacions i d'aquestes, el 41% viuen en països amb cobertura total, i el 44% en països amb cobertura parcial<sup>70</sup>.

També val a dir, que tot i l'existència de les guies, que tenen l'objectiu d'uniformitzar les pràctiques de cribatge, existeixen diferències organitzatives. El major consens es troba en la periodicitat del cribatge que és de 2 anys a tots els països, excepte al Regne Unit i a Malta que és de 3 anys. Respecte l'edat de la població diana, si bé la majoria cobreixen el rang recomanat (12 països, 50-69 anys), existeixen alguns casos on el rang d'edat és menor (50-59 anys, 50-65 anys) o inclús superior (45-69 anys, 40-70 anys). La majoria de diferències entre programes, però, recauen en els criteris d'elegibilitat i en el protocol de les pràctiques de cribatge.

**Taula 3.** Recomanacions per la realització del cribatge mamogràfic a dones de risc intermedi per part de les principals agències internacionals

<b>Organització</b>	<b>Any publicació guies</b>	<b>Recomanacions sobre grups d'edat i periodicitat</b>
European Comission	2006	Edat 50-69 anys, cada 2 anys
IARC	2002	Edat 50-69 anys, cada 2 anys
US Preventive Task Force	2009	Edat 50-74 anys, cada 2 anys; Edat 40-49 i ≥75 anys, individualitzar la decisió (recomanen cada 2 anys, en cas de realitzar-se)
American Cancer Society	2010	Edat ≥40 anys, cada any
National Cancer Institute	2010	Edat ≥40 anys, cada 1-2 anys
Canadian Task Force on Preventive Health Care	1998-2001	Edat 50-69 anys, cada 1-2 anys; Edat 40-49 anys individualitzar la decisió (recomanen cada 1 o 2 anys, en cas de realitzar-se)
National Health Service United Kingdom	2011	Edat 47-73 anys, cada 3 anys

Recopilació personal

## L'extensió del cribratge a Espanya i Catalunya

A Espanya, l'any 1990 es va posar en marxa a Navarra el primer programa de detecció precoç de càncer de mama. Des de llavors i gradualment s'han anat iniciant a la resta de Comunitats Autònomes fins arribar l'any 2006 a la cobertura total de la població diana. A Catalunya, el primer programa fou el de l'Hospital del Mar, a Barcelona, que s'inicià l'any 1995 i que actualment convoca cada dos anys a més de 34.000 dones.

L'organització dels diferents programes depèn de les comunitats autònomes i es coordinen a través de la *Red de Programas de Cribado* (<http://www.programascancerdemama.org>)<sup>71</sup>. Els responsables dels programes es reuneixen anualment, i posen en comú els resultats d'indicadors relacionats amb l'organització, els recursos i altres elements destinats a garantir la seva qualitat a partir de l'avaluació específica i conjunta dels programes.

Tots els programes utilitzen la mamografia biennal com a prova de cribratge. El grup d'edat de la població diana és de 50 a 64 anys, però la majoria de programes han optat per ampliar-lo fins al 69 anys; en 6 comunitats autònomes també s'inclou el grup de 45 a 49 anys. No obstant, continuen existint algunes diferències importants, especialment en el mètode de lectura, simple o doble<sup>7</sup>. Pel que fa al tipus de mamografia, des de l'any 2000, alguns programes han optat per substituir els mamògrafs analògics convencionals per mamògrafs digitals. Actualment, en 9 comunitats autònomes s'ha introduït total o parcialment la mamografia digital, i es preveu que aquesta tècnica es vagi estenent a tots els programes<sup>71</sup>.

## 2.2 Efectes adversos de la detecció precoç

La posada en marxa dels programes de detecció precoç porta implícit que la població participant, és a dir, població 'sana', pugui experimentar certs efectes adversos. És ben sabut que no totes les dones es beneficien per igual del fet de participar al cribratge. A la majoria de dones participants, no se'ls hi diagnosticarà cap tumor al llarg de la seva història de cribratge, però és probable que pateixin algun dels efectes adversos que se'n deriven.

Existeixen cinc efectes adversos associats al cribratge mamogràfic, de diferent abast, freqüència i conseqüències<sup>72</sup>: 1) la radiació associada a la mamografia; 2) el dolor causat per la prova; 3) les conseqüències clíniques, psicològiques i sobre el comportament envers el programa després d'experimentar un resultat fals positiu; 4) la falsa sensació de seguretat i el retràs diagnòstic associat als resultats falsos negatius; i, 5) el risc de sobrediagnòstic i sobretractament de lesions que en absència de cribratge mai haurien estat diagnosticades.

L'efecte de **la radiació** i **el dolor** que comporta el fet de realitzar la mamografia són els efectes adversos més lleus, tot i ser al mateix temps els més freqüents. En el cas de la radiació, no s'han realitzat treballs

que específicament avaluïn el risc de càncer associat a la radiació de la mamografia. Sí que existeixen alguns resultats provinents d'estudis observacionals, els quals descriuen un possible augment del risc, proporcional a la dosi de radiació i al nombre d'exposicions<sup>72</sup>. Tot i així, és important subratllar que fora de del cribratge poblacional organitzat les dones sovint es realitzen més exàmens mamogràfics en l'àmbit de la sanitat privada, ja sigui abans o després de l'edat recomanada o bé entre els intervals de cribratge. Aquest fet, òbviament augmentarà la dosi de radiació rebuda al llarg de la vida.

Els treballs centrats en el dolor associat a la mamografia, troben resultats variats, que van des del 28% al 77% de dones que reporten haver sofert dolor durant la realització de la mamografia<sup>72</sup>. Aquesta variabilitat de resultats en part s'atribueix a la subjectivitat de la mesura, però també a les diferències que s'han detectat en funció de la fase del cicle menstrual, o a l'ansietat. És important remarcar que els treballs que estudien l'efecte que pot tenir el dolor en el retorn al cribratge, mostren que poques dones deixaran de realitzar-se mamografies per aquest motiu<sup>72,73</sup>.

Els resultats **falsos positius** apareixen quan davant d'una sospita mamogràfica es recomanen proves addicionals (altres proves d'imatge no invasives, o proves que impliquin algun procediment invasiu) i el resultat d'aquestes acaba descartant el diagnòstic d'un càncer. Les Guies Europees recomanen que el percentatge de dones reconvoques per exploracions addicionals no superi el 7% en cribratges inicials i del 5% en cribratges successius<sup>69</sup>. S'estima que el risc acumulat de presentar com a mínim un resultat fals positiu al llarg de la història de cribratge en el context europeu, és del 19.7%<sup>74</sup>. Concretament a Espanya, el risc acumulat de fals positiu és del 20.4% (95%IC: 20.0-20.8), i varia entre el 51.4% i el 7.5% en funció dels perfils de risc de les dones<sup>75</sup>. Els resultats falsos positius s'associen amb estrès psicològic i angoixa davant un possible diagnòstic de càncer<sup>76</sup>. Per altre banda, en el context espanyol s'ha comprovat que les dones que han experimentat un resultat fals positiu són menys adherents al programa<sup>77,78</sup>, tot i que en d'altres contextos l'efecte sobre la participació en següents invitacions no es veu modificat<sup>79</sup>, o inclús augmenta<sup>80</sup>. El caràcter poblacional del cribratge fa que, en valors absoluts, el nombre de falsos positius sigui l'efecte advers amb un major impacte tant sobre el volum de dones a les quals pot arribar a afectar com als recursos i la càrrega assistencial que suposen.

Per altre banda, l'existència de casos **falsos negatius** disminueix el nombre de càncers detectats precoçment i per tant, afecta l'efectivitat del programa de cribratge i el seu potencial en la reducció de la mortalitat basat en la detecció precoç. A més, un resultat negatiu en el cribratge pot produir una falsa sensació de seguretat a les dones, que les faci estar menys alerta a possibles signes de la malaltia.

Per determinar els casos falsos negatius, és necessari prèviament conèixer els càncers d'interval. Segons les Guies Europees, els càncers d'interval són els càncers de mama que apareixen després d'un episodi de cribratge negatiu, que pot incloure o no valoracions addicionals, i abans de la següent invitació al programa<sup>69</sup>. Aquesta definició queda subjecte a la periodicitat de la prova, a més d'estar influenciada per la disponibilitat de les dades de diagnòstics de càncer realitzats fora del circuit del cribratge. Així, es tracta

d'una definició amb múltiples limitacions, però tanmateix molt útil per l'obtenció d'indicadors que a curt termini aporten informació sobre l'efectivitat del cribratge.

En el context europeu, els càncers d'interval suposen al voltant del 20-30% de càncers diagnosticats en dones participants en el cribratge<sup>61</sup>. Les guies utilitzen el valor de la incidència proporcional de càncer d'interval (proporció de càncers d'interval respecte la taxa d'incidència basal de càncer de mama) per establir els criteris de qualitat i els valors de referència. Els nivells desitjables s'han de situar per sota del 50% durant el primer any i per sota del 30% al llarg del segon any<sup>69</sup>. Per altra banda, recomanen que els falsos negatius no representin més del 20% del total de càncers d'interval.

Per la identificació dels càncers d'interval és imprescindible conèixer els casos incidents de càncer de mama en la població diana. Aquest procés requereix el creuament de les dades del programa (dones participants) amb registres que permetin identificar els càncers de mama diagnosticats fora del circuit de cribratge, tant en sanitat pública com privada. La font de dades d'elecció són els registres de càncers poblacionals, però en aquelles zones on no n'hi hagi cal recórrer a altres fonts, com els registres de tumors hospitalaris, el conjunt mínim de dades hospitalàries (CMBD-H) o registres dels propis programes.

Un cop identificats, el següent repte és classificar-los radiològicament a partir de la revisió de la mamografia de cribratge immediatament anterior al diagnòstic i la pròpia mamografia del diagnòstic. Aquest procés de revisió, l'han de realitzar radiòlegs amb experiència i seguint una metodologia consensuada. Finalment, els càncers d'interval es poden classificar radiològicament en cinc categories<sup>69</sup>:

- Veritables càncers d'interval: tumors que no eren visibles en l'anterior mamografia de cribratge, i en canvi ho són en la mamografia diagnòstica.
- Tumors ocults: són tumors invisibles tant en la mamografia de cribratge com en la diagnòstica.
- Signes mínims: aquells tumors que tot i ser visibles a la mamografia diagnòstica, tan sols presentaven signes mínims de sospita en la mamografia de cribratge.
- Falsos negatius: tumors que eren clarament visibles en la mamografia de cribratge. Es poden subclassificar en: 1) errors d'interpretació radiològica, o 2) errors tècnics, en cas que una mamografia de qualitat deficient per mala posició o mala qualitat tècnica pugui haver contribuït a la no identificació del càncer.
- Inclassificables: en aquells casos on no es disposi de la mamografia de cribratge i/o la mamografia del diagnòstic i que per tant, no es pugui classificar en cap de les categories anteriors.

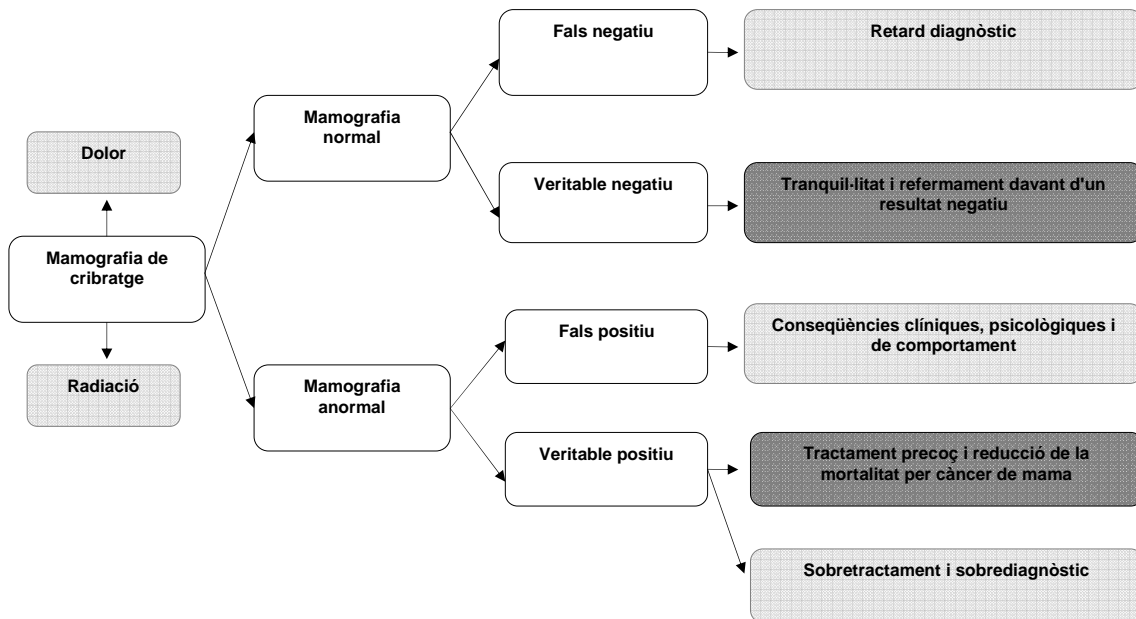
Els diferents subtipus, però, conformen un grup heterogeni, tant per la seva pròpia definició com des d'un punt de vista biològic. Mentre els falsos negatius són majoritàriament deguts a errors d'interpretació, els veritables càncers d'interval -que representen prop de la meitat dels càncers d'interval- semblen ser

tumors amb característiques biològiques que els fan menys susceptibles a ser detectats amb mamografies biennals. Tanmateix, tot i l'interès global d'aconseguir disminuir el nombre de càncers d'interval, es fa necessari conèixer l'heterogeneïtat dels subtipus i factors de risc específics per definir estratègies de millora.

Finalment, el **sobrediagnòstic** es defineix com la detecció de tumors en el cribratge, que sense l'existència d'aquest mai haguessin donat simptomatologia al llarg de la vida de la dona<sup>82</sup>. Alguns casos s'atribueixen al diagnòstic de tumors que haguessin restat asimptomàtics al llarg de la vida de la dona; altres casos esdevenen sobrediagnòstic tot i ser tumors detectats precoçment, pel fet de que la dona es mor d'altres causes abans de que els tumors haguessin desenvolupat símptomes en absència de cribratge. Per tant, en ambdues situacions les dones sofreixen els efectes adversos del cribratge i del tractament de la malaltia, sense oportunitat de beneficiar-se'n. Per l'impacte emocional i del propi tractament, es considera l'efecte advers més greu atribuït a la detecció precoç. Tot i així, com que avui en dia no és possible saber quins casos són els sobrediagnosticats, aquest efecte es converteix en quelcom teòric i que inclús pot portar a situacions paradoxals, tal i com apunten Raffle & Gray<sup>83</sup> en descriure'l com la "paradoxa de la popularitat". Contra major sigui l'efecte del sobrediagnòstic, més persones creuran que han estat beneficiades per les pràctiques de cribratge.

El nombre de casos sobrediagnosticats tan sols es podrà estimar de manera 'teòrica' a nivell poblacional a partir de les dades recollides durant els anys que s'ha realitzat el cribratge. Existeixen un gran nombre d'aproximacions matemàtiques, sense que hi hagi un únic mètode pel seu càlcul. Això implica que els resultats variïn enormement, i que siguin susceptibles a interpretacions esbiaixades. Les estimacions van des de l'1%<sup>84</sup> al 50%<sup>85,86</sup>. Un recent treball de revisió que estudia els resultats publicats en el context europeu, considerant només aquells que tenen en compte la incidència basal de càncer de mama esperada en absència de cribratge, juntament amb el control del biaix de durada (veure pàg. 40), situa el sobrediagnòstic en els programes d'Europa entre l'1 i el 10%<sup>82</sup>. També recentment, l'informe Marmot estima que el 19% de tumors diagnosticats al cribratge són sobrediagnòstics<sup>58</sup>.

**Figura 3.** Esquema dels riscos i beneficis associats a la mamografia de cribratge.



Adaptat d'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. Review<sup>72</sup>.  
 Les caselles amb el fons gris clar, indiquen els efectes adversos que s'atribueixen al cribratge i les de color gris més fosc, els beneficis.

### 2.3 Indicadors recomanats per l'avaluació del cribratge poblacional

La complexa organització dels programes de cribratge de càncer de mama, tant des d'un punt de vista administratiu, com multidisciplinari i assistencial, requereix una contínua avaluació que garanteixi que es compleixen els estàndards de qualitat recomanats a nivell internacional. L'èxit del cribratge de càncer de mama depend tant dels propis resultats dels programes i el seu impacte en la salut de la població, com de la seva organització, implementació, execució i acceptabilitat.

A nivell organitzatiu resulta fonamental disposar de fonts d'informació de qualitat a l'hora d'identificar la població diana i utilitzar sistemes eficaços d'invitació. Per altre banda, també és imprescindible disposar de serveis de diagnòstic i de tractament amb capacitat per donar serveix a la població coberta pel programa. Finalment, les guies insisteixen en la importància de fomentar la màxima participació entre la població diana<sup>69</sup>.

L'objectiu de reduir la morbiditat i la mortalitat per càncer de mama, s'assoleix a llarg termini, i tan sols serà possible de corroborar anys o inclús dècades després de la implementació del cribratge. No obstant, les guies europees defineixen diferents indicadors de qualitat que a curt termini ofereixen informació necessària per garantir l'efectivitat dels programes. Així mateix, en la majoria de casos n'estableixen valors de referència, els quals permeten als programes avaluar diferents aspectes i conèixer els punts on cal centrar esforços per millorar-ne la qualitat.

Els principals indicadors predictius de la capacitat del programa per produir un impacte en la població en termes de reducció de mortalitat són:

- Participació global: Percentatge de dones a les quals se'ls realitza l'exploració del cribratge, d'entre el total de dones convidades.
- Cobertura: Percentatge de la població diana que es realitza mamografies periòdiques de cribratge, contemplant també les mamografies realitzades fora del programa.
- Adherència: Percentatge de dones que són explorades en una ronda de cribratge d'entre el total que participaren en l'anterior i tornen a ser convidades.

Com a indicadors referents a la patologia detectada i les seves característiques, destaquen:

- Taxa de detecció: Proporció de dones entre les quals es confirma mitjançant l'estudi anatomopatològic un càncer de mama, respecte el total de dones que s'han realitzat una mamografia de cribratge.
- Característiques dels tumors detectats: per exemple, el percentatge de tumors intraductals d'entre el total de càncers detectats, el percentatge de càncers invasius amb una mida superior als 10mm, o el percentatge de tumors invasius sense invasió ganglionar d'entre el total de càncers detectats.

Com a indicadors del procés i la qualitat del cribratge, en destaquen:

- La sensibilitat: la capacitat de detectar el tumor quan aquest hi és present, o la ràtio de càncers de mama correctament identificats en el cribratge (veritables positius) sobre el total de càncers identificats en el cribratge i no identificats en el cribratge (veritables positius i càncers d'interval).

La sensibilitat dels programes es calcula a partir del total de càncers d'interval. En canvi, quan es vol conèixer la sensibilitat de la prova, s'utilitza el nombre de falsos negatius. En els valors de referència de les guies europees no s'especifica la sensibilitat tot i que es reconeixen els càncers d'interval com un indicador vàlid de la sensibilitat del cribratge<sup>69</sup>. La majoria de treballs que estudien la sensibilitat del cribratge, reporten valors entre el 65 i 85%<sup>81</sup>.

- L'especificitat: és la capacitat del test de donar un valor negatiu quan no hi ha cap tumor, o la ràtio d'individus que, en efecte, no tenen la malaltia (veritables negatius), d'entre el total que han estat identificats com a negatius en el procés de cribratge (veritables negatius i falsos positius).

Per poder tenir un estimador absolutament correcte de l'especificitat, caldria seguir a tots aquells casos amb resultat negatiu per descartar els possibles errors i provar la seva negativitat. S'utilitza, però, el nombre de falsos positius, que es calcula a partir dels casos de dones reconvidades (per existir alguna sospita en la mamografia de cribratge) on es descarta el diagnòstic de càncer. Així, pel càlcul de l'especificitat, s'acostumen a utilitzar les dades directes del procés de cribratge (resultats negatius i

resultats falsos positius). En general, els valors de l'especificat dels programes es situen entre el 94%-97%<sup>87</sup>.

- Valor predictiu positiu: és la probabilitat que un individu presenti la malaltia quan el resultat de la prova és positiu (veritables positius), és a dir, el nombre de càncers detectats sobre el total de dones amb mamografies positives i amb valoracions addicionals (veritables positius i falsos positius).

Per l'avaluació del cribratge resulta de gran importància trobar el balanç entre sensibilitat i especificitat, per tal de poder assolir una ràtio acceptable de casos veritables positius i falsos positius. Les guies europees recomanen que per aconseguir un augment dels beneficis (reducció de falsos negatius) i una disminució de la càrrega social i psicològica (reducció de falsos positius) cal posar en pràctica un rigorós control de qualitat, la capaciació sistemàtica del personal de salut i el seguiment de les dones, juntament amb l'avaluació anual de resultats del cribratge.



### 3. Reptes i dificultats en l'avaluació del cribratge poblacional del càncer de mama

Per estimar l'efectivitat del cribratge de manera fiable, és necessari realitzar estudis en el marc dels assajos clínics aleatoritzats i, posteriorment, en el context poblacional. S'han realitzat nombrosos treballs, avaluacions, revisions sistemàtiques i metanàlisis i, en general, tots ells mostren grans controvèrsies respecte a l'efectivitat del cribratge. Aquestes controvèrsies, en part s'expliquen per les diferents estratègies utilitzades en els estudis i a la gran diversitat de polítiques de cribratge existents en els diferents països<sup>88-91</sup>.

A aquesta situació, cal sumar-hi certs reptes metodològics dels estudis realitzats en el marc del cribratge mamogràfic. En primer lloc, degut a que el càncer de mama és una malaltia de progressió lenta, es necessiten llargs períodes temporals per verificar la disminució de la mortalitat atribuïble a la intervenció. En segon lloc, cal incorporar la informació dels efectes adversos derivats del cribratge. I finalment, cal tenir present que el context on es desenvolupa el cribratge és canviant i resulta imprescindible disposar d'una evidència científica actualitzada. Concretament en el cas concret del càncer de mama, els canvis tècnics i en el tractament de la malaltia al llarg dels últims anys han estat de gran rellevància.

Els següents apartats desenvolupen aquests punts, fent èmfasi en els aspectes més controvertits del cribratge. Al seu temps, aquestes són algunes de les qüestions en les quals s'emmarquen els treballs que conformen la tesi.

#### 3.1 El qüestionament del cribratge amb les actuals estratègies

La publicació l'any 2000 a la revista *The Lancet* d'un metanàlisi que posava en dubte la reducció de la mortalitat atribuïda al cribratge en base als resultats dels assajos clínics, és considerat per molts el punt d'inflexió a partir del qual s'ha posat sobre la taula la idoneïtat del cribratge del càncer de mama amb les actuals estratègies. En aquest metanàlisi, realitzat per el *Nordic Cochrane Centre* i firmat per dos autors danesos<sup>92</sup>, es revisaven les dades de set assajos clínics, cinc dels quals els autors consideraven metodològicament inacceptables, bàsicament per una inadequada aleatorització. Amb les dades procedents dels dos assajos més robustos (Canadà i Malmö), conclouen que no es podia demostrar la reducció de la mortalitat per càncer de mama atribuïda al cribratge. Peter Gøtzsche, en el mencionat treball i en següents publicacions, posa en qüestió una possible sobreestimació dels beneficis del cribratge i una subestimació dels efectes adversos (especialment el sobrediagnòstic) i critica la informació que es fa arribar a les dones, massa esbiaixada segons ell, a favor del cribratge<sup>93</sup>.

Els capítols més recents d'aquest intens debat es troben en la publicació d'un monogràfic de la revista *Journal Medical Screening*<sup>94,95</sup>, juntament amb l'Informe Marmot comentat anteriorment<sup>58</sup>. Ambdós treballs, conclouen, amb matisos, que amb l'evidència actual el cribratge de càncer de mama continua

sent recomanable i que els beneficis del cribratge justifiquen els efectes adversos que s'hi associen. En el primer cas es revisen sistemàticament les dades de múltiples estudis observacionals realitzats en el context europeu que avaluen tant els beneficis dels cribratge com els efectes adversos. Descriuen un 25%-31% de reducció de la mortalitat, un 17% de risc de presentar un resultat fals positiu i un 6.5% de sobrediagnòstic<sup>57</sup>. El monogràfic, però, no inclou cap treball referent a la sensibilitat o als càncers d'interval en l'exhaustiva revisió de la situació del cribratge en el context europeu. A l'infome Marmot s'apunta que la reducció de mortalitat atribuïble al cribratge és del 20% (IC95%: 11-27%) i el percentatge de casos sobrediagnosticats del 19% (quan s'expressa com a proporció de casos diagnosticats en el grup participant al cribratge durant el període actiu del programa), mentre que subratllen la necessitat de revisar la informació que es fa arribar a les dones per tal que sigui transparent i permeti una presa de decisions informada<sup>58</sup>.

Últimament el debat sembla que comença a focalitzar l'interès en la personalització del cribratge<sup>28</sup>. Amb aquest enfocament es pretenen proposar estratègies alternatives a les actuals "one size fits all", encaminades a maximitzar els beneficis (intensificar el cribratge o utilitzar diferents tecnologies entre els grups de major risc) i reduir els efectes adversos. Tot i la voluntat d'avançar en aquest aspecte, encara queden interrogants i qüestions per resoldre per tal de poder definir grups específics de dones que es puguin beneficiar d'un canvi en les actuals estratègies de detecció precoç del càncer de mama.

### **3.2 Les dificultats per conèixer l'impacte real de la detecció precoç sobre la mortalitat: els biaixos associats al cribratge**

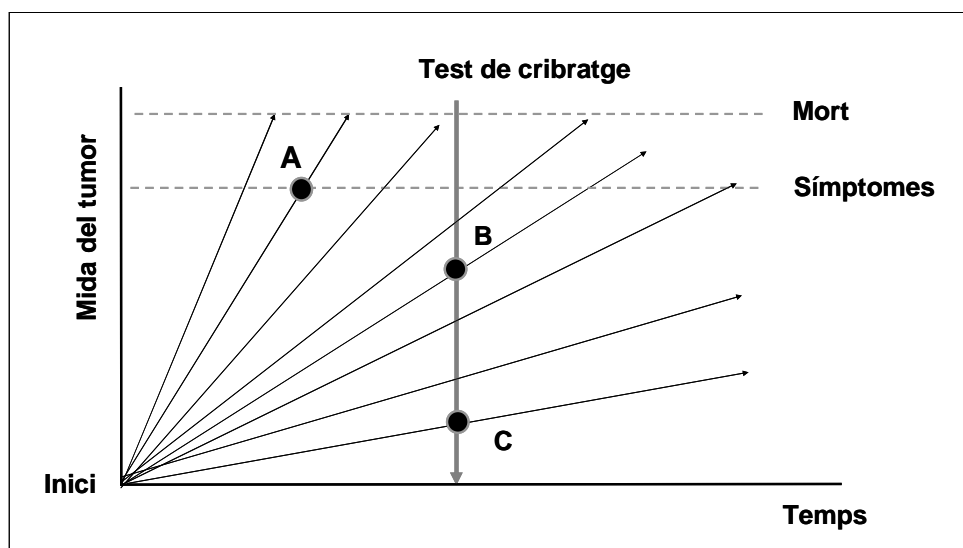
Per l'estudi de l'impacte del cribratge en termes de reducció de la mortalitat, és necessària la comparació dels resultats de mortalitat específica per càncer de mama o mortalitat global d'un grup de dones cribrades amb un grup de dones que no hagin estat intervingudes. Per obtenir estimadors fiables cal assegurar la possibilitat de comparar els grups, a més de tenir en compte els possibles biaixos que puguin afectar els resultats. A continuació s'apunten els principals biaixos propis del cribratge, el biaix de selecció, el biaix de durada i el biaix d'avanç diagnòstic, i les estratègies proposades per controlar-los<sup>83</sup>.

**Biaix de selecció o biaix del participant sa** (*Self-selection bias* o *healthy screenee effect*). La població participant als programes de detecció precoç acostuma a tenir hàbits més saludables i a estar en millor estat de salut que la no participant<sup>83</sup>. Aquest aspecte contribueix en presentar un millor pronòstic de la malaltia, fet podria suposar una sobreestimació dels resultats del cribratge, ja que la població que en fa ús compta amb un percentatge més elevat d'individus sans<sup>96</sup>. L'únic disseny capaç d'evitar el biaix de selecció són els assajos aleatoris controlats. Actualment, la possibilitat de dur a terme aquests estudis és pràcticament impossible, ja que la pràctica extensiva de la detecció precoç no podria garantir la no

contaminació dels grups control. Quan s'utilitzen dissenys observacionals, s'acostumen a utilitzar estratègies basades en les que s'utilitzen en assajos clínics i que permeten corregir el fet que les dones no participants al cribatge tenen un major risc de morir per càncer de mama, no atribuïble a la detecció precoç, que les que hi participen<sup>97</sup>. Recentment, Autier i Boniol<sup>98</sup> han demostrat com aquest ajust en dissenys cas-control no és vàlid quan altres factors a part del cribatge poden influenciar sobre la mortalitat per càncer de mama. Tanmateix, aquest biaix no afecta els resultats quan es comparen grups de dones cribrades, com en l'anàlisi dels càncers d'interval<sup>99</sup>.

**Biaix de durada** (*Length time bias*). Es tracta d'un fenomen pel qual els tumors de creixement lent, menys agressius, tenen un període asimptomàtic més llarg i per tant tenen una major probabilitat de ser detectats en el cribatge. Aquest fet els hi confereix un avantatge artificial en el temps de supervivència. El cas extrem del biaix de durada és el sobrediagnòstic, és dir, aquells casos de creixement molt lent o no progressius, que sense l'existència del cribatge mai haguessin arribat a donar simptomatologia. De manera contrària, els tumors de creixement més ràpid presenten una menor probabilitat de ser detectats en les proves de cribatge i es troben sobrerrepresentats entre els tumors que apareixen per simptomatologia, en l'interval entre mamografies. Aquest fet implica que les característiques dels tumors detectats en el cribatge sigui millor que les dels tumors detectats clínicament (fora del cribatge) i que les comparacions entre ambdós grups no es pugui realitzar de forma directa<sup>99</sup>.

**Figura 4.** Representació esquemàtica del biaix de durada.



Adaptat de *Clinical Epidemiology. The essentials, 4th edition*. Robert H Fletcher y Suzanne W Fletcher, Lippincott Williams & Wilkins 2005<sup>100</sup>.

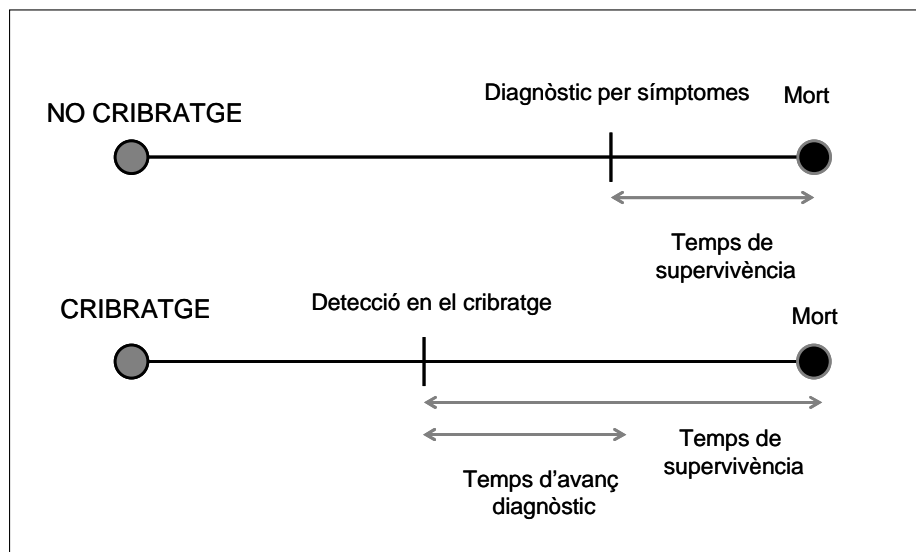
(A) Tumors de creixement ràpid, els quals presenten una menor probabilitat de ser detectats en les proves de cribatge. (B) Tumors de creixement més lent, i amb una fase preclínica llarga, els quals tenen una major probabilitat de ser detectats en les proves de cribatge, abans de que donin simptomatologia. (C) Tumors de creixement molt lent, o inclús no progressius, que en

absència de cribratge mai haguessin arribat a donar simptomatologia. Es tracta de tumors sobrediagnosticats pel cribratge, el cas extrem del biaix de durada.

El biaix de durada es basa en l'heterogeneïtat biològica del càncer de mama. Per estimar l'impacte del cribratge sobre la mortalitat en estudis observacionals, Duffy et al.<sup>99</sup> proposen una correcció basada en anàlisis de sensibilitat dels diferents grups de tumors en funció de la seva agressivitat. El paràmetre resultant es té en compte en l'ajust quan s'estima l'efecte del cribratge en la mortalitat. En els treballs centrats en la supervivència en funció de la via diagnòstica, es proposa la introducció en els models de riscos proporcionals, informació de caràcter biològic, que permeti controlar l'heterogeneïtat biològica que està relacionada amb el pronòstic del càncer de mama<sup>101</sup>. Generalment aquesta aproximació es fa introduint informació d'expressió de biomarcadors, com els RE, l'HER2, o el Ki67, en els models matemàtics<sup>102</sup>.

**Biaix d'avanç diagnòstic (Lead time bias).** L'efecte d'aquest biaix s'explica pel fet de que el temps de supervivència en els casos detectats en el cribratge aparentment és més llarg, però no perquè sigui reflex d'un millor pronòstic de la malaltia sinó perquè aquesta ha estat detectada precoçment. És a dir, el temps de supervivència s'ha començat a comptabilitzar abans. Aquest fet pot comportar que independentment que dos individus acabessin morint al mateix moment, el cas detectat abans -mitjançant les proves de cribratge, quan encara era asimptomàtic- presentaria un període de supervivència més llarg.

**Figura 5.** Representació esquemàtica del biaix d'avanç diagnòstic



Considerant la supervivència des del moment del diagnòstic, el fet d'avançar el diagnòstic significa que s'allarga el període de supervivència.

En aquest cas, són el temps que el tumor passa en fase preclínica (*sojourn time*) i la periodicitat de la prova de cribratge, els factors que condicionen el temps pel qual s'avança el diagnòstic. Des del punt de vista matemàtic s'han proposat múltiples aproximacions<sup>99,103,104</sup>, basades en l'estimació del període addicional de seguiment atribuïble a la detecció precoç. L'obtenció d'aquest paràmetre s'aconsegueix a partir de la distribució del *sojourn time* i la taxa de transició a estadi clínic<sup>99</sup>. En els treballs amb una perspectiva més clínica, s'opta per la introducció de variables relacionades amb el creixement del tumor (mida, estadi) com a aproximació al moment del diagnòstic<sup>101,105,106</sup>.

A la pràctica, és imprescindible tenir en compte aquests tres biaixos a l'hora d'interpretar els resultats on s'analitza tant la reducció de mortalitat atribuïda a la detecció precoç com la supervivència de les dones en funció de la via de diagnòstica.

### **3.3 Avaluar l'impacte del cribratge sobre la detecció del càncer: coexistència de població participant i no participant**

Com s'ha apuntat en apartats anteriors, l'avaluació del principal benefici del cribratge, és a dir, la disminució de la mortalitat, tant sols es pot realitzar anys després de la implementació dels programes. La introducció d'aquest concepte temporal comporta certes dificultats, entre les quals destaca conèixer la població participant al llarg del temps i considerar les variables canviants (tractaments, canvis organitzatius) que puguin condicionar el pronòstic del càncer.

Pot resultar relativament senzill conèixer la població que participa als programes i la que no a partir del creuament de registres amb informació censal, per identificar la població diana, i les bases de dades dels programes. En cap cas, però, es pot considerar que la població no participant no estigui exposada a la detecció precoç. En la majoria de països, el cribratge poblacional coexisteix amb cribratge oportunístic o fora del circuit dels programes, que pot tenir un abast variable, inclús entre diferents zones en un mateix país<sup>107-109</sup>. El cribratge oportunístic s'atribueix majoritàriament a l'àmbit de la medicina privada i resulta per tant, difícil de comptabilitzar.

L'existència de cribratge oportunístic abans de la introducció dels programes de cribratge, però quan ja existia certa conscienciació entre la població dels beneficis de la detecció precoç, pot suposar que els estimadors de l'incidència basal del càncer de mama no reflecteixin una situació d'absència de cribratge<sup>69</sup>. Per altra banda, en algunes zones no és estrany que certs professionals recomanin realitzar mamografies entre els intervals de cribratge o fora de l'edat diana a totes les dones, o a grups específics. Resulta important conèixer l'abast del cribratge oportunístic en una àrea determinada mitjançant, per exemple, les enquestes de salut, ja que tant els valors basals de la incidència de càncer de mama, com les característiques dels tumors detectats fora del programa poden variar. A les guies europees es recomana conèixer la "cobertura del programa de cribratge", és dir la població que es beneficia de la detecció precoç tant si és mitjançant un programa poblacional o la medicina privada. Encara que es tracti de dades

aproximades, seran rellevants per la correcta interpretació dels resultats en comparar població participant i no participant en una zona concreta<sup>69</sup>.

En molts països, disposar d'informació a nivell poblacional presenta múltiples limitacions, tant per l'obtenció de les dades, com per la seva qualitat i les dificultats en el procés de creuament entre registres. En aquest aspecte, els països escandinaus són els que tenen uns sistemes d'informació més complets, amb registres poblacionals de tumors, de mortalitat, de patologies, etc<sup>110,111</sup>. A més a més, el procés de creuament es simplifica enormement ja que cada individu disposa d'un codi d'identificació personal compartit per tots els registres. Tot plegat explica que la gran majoria de treballs que avaluen l'impacte del cribratge en la mortalitat per càncer de mama, o en la detecció de la malaltia s'hagin realitzat en aquests països<sup>22,91,112,113</sup>.

### **3.4 Avaluar els efectes adversos del cribratge. El cas dels càncers d'interval**

Els efectes adversos més estudiats han estat els falsos positius i el sobrediagnòstic<sup>74,82</sup>. Respecte l'estudi dels càncers d'interval, hi ha hagut un interès creixent que ha focalitzat part de la recerca en aquest grup de tumors. Bàsicament els treballs s'han centrat en determinar el nombre de càncers d'interval per poder estimar la sensibilitat del programa i en la seva caracterització des d'un punt de vista descriptiu<sup>81,114-116</sup>. L'heterogeneïtat, però, caracteritza aquests càncers i en dificulta el seu estudi. En primer lloc la pròpia definició de càncer d'interval és subjecte a la periodicitat del cribratge i pot comportar mala classificació. Tot i l'obvietat d'aquest punt, resulta de gran rellevància ja que suposarà que casos que en un programa amb mamografies biennals són considerats càncers d'interval, en un altre, amb intervals anuals, per exemple, no ho serien. Un segon aspecte clau és la seva classificació en subtipus a partir de la classificació radiològica de les mamografies de cribratge i diagnòstiques. Tot i existir recomanacions sobre la metodologia i les categories radiològiques<sup>69,117</sup>, la gran majoria de treballs presenten classificacions modificades, realitzades amb l'objectiu de facilitar el procés de revisió o bé de presentar grups amb un major nombre de casos. Aquest fet també pot diluir certes diferències i matisos que hi puguin haver entre categories.

Els càncers d'interval agrupen principalment tumors falsos negatius, és a dir, existents en l'anterior mamografia, i veritables càncers d'interval, tumors indetectables en l'anterior cribratge. Només aquest fet, ja apunta que l'etiologia dels càncers d'interval pot ser variada i cal profunditzar en aquestes diferències per tal disposar una descripció acurada d'aquests tumors. Tant si el seu estudi s'enfoca des de l'òptica del programa, com a indicador de qualitat, com si exploren els seus determinants, la diferenciació dels subtipus serà clau. Per la millora del programa amb les actuals estratègies, la categoria més rellevant seran els falsos negatius, ja que una disminució del seu nombre pot conduir a la millora dels indicadors del programa (sensibilitat) i a llarg termini, de l'efectivitat. En canvi, en el nou paradigma de replantejament del cribratge, poder definir un grup de dones més susceptibles a desenvolupar un càncer d'interval pot

ajudar a la definició d'estratègies personalitzades. L'edat o l'elevada densitat mamària són alguns dels factors que s'han associat amb un major risc de càncer d'interval<sup>118,119</sup>. En cas que es disposés d'una definició més acurada de perfils de risc de dones més susceptibles a desenvolupar un càncer d'interval, seria possible valorar la introducció de tecnologies específiques per grups de dones amb un major risc<sup>120</sup>. Per exemple, la ressonància magnètica o les ecografies són tècniques que s'ha demostrat que augmenten la sensibilitat en mames denses, a costa, però, d'una disminució de l'especificitat<sup>121-123</sup>.

Des del punt de vista clínic, els càncers d'interval constitueixen un grup de tumors amb un pitjor pronòstic que els càncers detectats en el cribratge<sup>120,124</sup>. Aquest pitjor pronòstic en general s'explica pel retard diagnòstic, ja que els càncers detectats en el cribratge es diagnostiquen en fase preclínica, mentre que els càncer d'interval ho fan després d'haver donat simptomatologia. Més enllà d'aquest retard diagnòstic diferents autors han descrit diferències a nivell molecular, associades amb una pitjor agressivitat entre els càncers d'interval. En aquest grup, el percentatge de tumors que expressen receptors hormonals (RE, RP) resulta menor, mentre que el percentatge de tumors amb una taxa de proliferació cel·lular elevada (sobrexpressió del factor de creixement Ki67) i poc diferenciats seria major<sup>125</sup>. Tot i així, les diferències moleculars entre els subtipus de càncer d'interval han estat poc explorades, com tampoc ho han estat les seves implicacions en el pronòstic dels càncers.

### **3.5 L'impacte de les noves tecnologies en els indicadors del cribratge**

A finals de la dècada de 1990 i principis del 2000, es dugueren a terme els primers assajos aleatoris controlats per avaluar l'efectivitat de la mamografia digital en comparació a la convencional com a prova de cribratge del càncer de mama<sup>126,127</sup>. Després dels resultats que mostraven una igual capacitat diagnòstica, apuntant a una major efectivitat en dones joves i amb mames denses<sup>128</sup>, la tècnica digital es va anar estenent entre els programes de cribratge. A més, aspectes relacionats amb l'emmagatzematge de les imatges, el seu processat i qualitat van ser altres factors determinants en la seva implementació<sup>45</sup>. Actualment, la transició d'analògic a digital en el nostre context forma part de les polítiques sanitàries i, en els pròxims anys, el revelat analògic quedarà en desús.

Al llarg de la dècada del 2000 molts programes han canviat progressivament els mamògrafs analògics pels digitals, fet que requereix conèixer l'impacte d'aquesta transició en la pràctica diària, més enllà dels resultats dels assajos. Per poder realitzar aquestes avaluacions és necessari que transcorri un cert temps. A partir de l'any 2005 es començaren a publicar els primers resultats en el context poblacional. Com en d'altres aspectes relatius al cribratge, tot i existir consens en que la tècnica és com a mínim tan bona com l'anterior, s'ha generat certa controvèrsia a la vista de certs resultats contradictoris<sup>127,129</sup>. Respecte la taxa de detecció en diferents contextos, entre ells l'espanyol, no s'observen diferències significatives entre ambdues tècniques<sup>130,131</sup>. En canvi, en altres països s'ha descrit un augment de la taxa de detecció<sup>132</sup>. Igualment, respecte els falsos positius els resultats també són variats<sup>130-132</sup>. La gran

majoria de treballs, però, coincideixen en detectar un major percentatge de tumors *in situ*<sup>127,131,132</sup> amb la tècnica digital. Aquest augment, però, encara no s'ha pogut confirmar que estigui sent degut a un *avanç de l'avanç diagnòstic*, o bé a un possible sobrediagnòstic de tumors. Aquest punt implica conèixer més a fons la naturalesa dels casos diagnosticats i conèixer des del punt de vista radiològic quines són les diferències en les sospites reportades pels radiòlegs i de quina manera afecta al valor predictiu de la prova.

Per completar l'avaluació de la transició analògic-digital també serà important conèixer el nombre de càncers d'interval. Fins ara, pocs treballs n'han estudiat l'impacte, tot i que en un recent estudi realitzat a Noruega, Hoff et al.<sup>133</sup> no es descriuen diferències estadísticament significatives respecte el percentatge de falsos negatius en ambdós escenaris.



#### 4. Presentació dels treballs que conformen la tesi

Aquesta tesi doctoral està composta per 4 treballs publicats en format d'articles científics en revistes internacionals indexades en les àrees de 'Salut pública, mediamambiental i ocupacional', 'Oncologia' i 'Radiologia, medicina nuclear i imatge mèdica'.

**Article 1.** Seventeen-years overview of breast cancer inside and outside screening in Denmark.

Laia Domingo, Katja K. Jacobsen, My von Euler-Chelpin, Ilse Vejborg, Walter Schwartz, Maria Sala, Elsebeth Lynge.

Acta Oncol. 2013 Jan;52(1):48-56.

**Article 2.** Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain.

Laia Domingo, Maria Sala, Sònia Servitja, Josep Maria Corominas, Francesc Ferrer, Juan Martínez, Francesc Macià, Maria Jesús Quintana, Joan Albanell, Xavier Castells.

Cancer Causes Control 2010 Aug;21(8):1155-64.

**Article 3.** Aggressiveness features and outcomes of true interval cancers: comparison between screen-detected and symptom-detected cancers.

Laia Domingo, Jordi Blanch, Sònia Servitja, Josep Maria Corominas, Cristiane Murta-Nascimento, Antonio Rueda, Maximino Redondo, Xavier Castells, Maria Sala.

Eur J Cancer Prev 2013 Jan;22(1):21-8.

**Article 4.** Differences in radiological patterns, tumour characteristics and diagnostic precision between digital mammography and screen-film mammography in four breast cancer screening programmes in Spain.

Laia Domingo, Anabel Romero, Francesc Belvis, Mar Sánchez, Joana Ferrer, Dolores Salas, Josefa Ibáñez, Alfonso Vega, Francesc Ferrer, Maria Soledad Laso, Francesc Macià, Xavier Castells, Maria Sala.

Eur Radiol. 2011 Sep;21(9):2020-8.

## 5. Justificació de la unitat temàtica

El cribratge de càncer de mama és una de les intervencions poblacionals més avaluades, i actualment més qüestionada. Efectivament, resten nombroses preguntes per resoldre, la majoria de les quals fan referència al balanç entre riscos i beneficis. No obstant, les autoritats sanitàries el continuen recomanant<sup>68,87</sup>. El qüestionament actual del benefici del cribratge fa necessari avançar en la seva avaluació, amb nous punts de vista i noves dades, per disposar dels elements necessaris per proposar canvis en les actuals estratègies de diagnòstic precoç. La personalització del cribratge en funció de la presència de certs factors de risc de la malaltia és una via, però cal més informació per poder-la plantejar i en un futur, implementar. Aquesta tesi s'emmarca en l'actual controvèrsia del cribratge, aprofundint en algunes qüestions que formen part del debat.

El càlcul dels beneficis del cribratge a nivell poblacional en termes de detecció dels tumors en estadis precoços i disminució de la mortalitat per càncer de mama, està basat en la comparació dels resultats a llarg termini de cohorts de dones cribrades i dones no cribrades. Tanmateix, s'assumeix que les dones no participants no es veuen afectades per la detecció precoç. Disposar de més informació en aquesta qüestió pot ser útil en la interpretació dels resultats derivats d'aquest tipus de treballs.

La quantificació dels efectes adversos ha guanyat importància al llarg dels últims anys. Els falsos positius i el sobrediagnòstic han estat els més estudiats, mentre que els càncers d'interval han quedat relegats a un segon pla. Per l'impacte que pot tenir una reducció del seu nombre en l'augment de l'efectivitat del cribratge, últimament existeix un interès creixent per conèixer millor els seus determinants, aprofundir en la seva biologia i en la seva evolució clínica.

Finalment, la implementació de la mamografia digital és una realitat en molts països i sembla que amb els anys desplaçarà completament el revelat analògic. Caldrà doncs, avaluar l'impacte d'aquest canvi en termes de riscos i beneficis per les dones cribrades i tenir en compte aquestes dades en l'actualització de guies i recomanacions.

La tesi s'ha dut a terme en el Servei d'Epidemiologia i Avaluació de l'Hospital del Mar-IMIM, i s'emmarca en la línia de recerca centrada en l'avaluació de les pràctiques de cribratge poblacionals <sup>75,77,78,134-138</sup>.

## **II. HIPÒTESIS I OBJECTIUS**

---



## 1. Hipòtesis

### Hipòtesi general

El cribratge del càncer de mama és una activitat d'abast poblacional en la majoria de països europeus. El seu disseny i organització, i també les característiques de la prova de detecció precoç, influeixen tant en els beneficis (detecció de tumors en estadis precoços i reducció de la mortalitat per càncer de mama), com en els efectes adversos (falsos positius, sobrediagnòstic i càncers d'interval).

### Hipòtesis específiques

Sobre l'impacte de la introducció del cribratge en la detecció de càncer de mama:

1. La posada en marxa d'un programa de detecció precoç afecta en el comportament de les taxes de detecció de càncer de mama en el conjunt de la població diana.
2. Les diferències organitzatives entre programes de cribratge poden influenciar les taxes de detecció dins dels programes de cribratge i les taxes de càncers d'interval.

Sobre l'impacte de la via diagnòstica en les característiques dels tumors i en el seu pronòstic:

3. Les característiques dels càncers d'interval (estadi clínic, grau histològic, marcadors biològics) són diferents a les dels tumors detectats en el cribratge.
4. Les dones amb càncer d'interval, i especialment amb veritable càncer d'interval, presenten un període lliure de malaltia més curt que les dones amb tumors diagnosticats en el cribratge.

Sobre l'impacte de les característiques de la prova de cribratge en els indicadors de qualitat:

5. La precisió diagnòstica de la mamografia digital és superior a la de la mamografia analògica.
6. Els canvis en la precisió diagnòstica entre tècniques es poden atribuir a diferències en els valors predictius positius de les diferents sospites radiològiques.
7. Els tumors diagnosticats amb la tècnica digital es detecten en estadis menys avançats.

## 2. Objectius

L'**objectiu general** de la tesi és aprofundir en l'avaluació del cribratge poblacional del càncer de mama específicament en els següents aspectes:

- L'impacte de la introducció del cribratge en la detecció de càncer de mama
- L'impacte de la via diagnòstica en les característiques dels tumors i en el seu pronòstic
- L'impacte de les característiques de la prova de cribratge en els indicadors de qualitat.

**Els objectius específics**, treballats en diferents estudis, són els següents:

Sobre l'impacte de la introducció del cribratge en la detecció de càncer de mama:

1. Descriure les taxes de detecció del càncer de mama, en el conjunt de la població diana i diferenciant segons la via diagnòstica (detecció en el cribratge, en d'interval entre cribratges o fora del circuit de cribratge).
2. Aprofundir en l'efecte de les diferències organitzatives dels programes poblacionals en les taxes de detecció dins i fora del programa.

Sobre l'impacte de la via diagnòstica en les característiques dels tumors i en el seu pronòstic:

3. Comparar les característiques biològiques dels càncers d'interval amb els tumors detectats en les proves de cribratge.
4. Estimar l'impacte de la via diagnòstica i altres variables (característiques biològiques, biomarcadors) sobre la supervivència de les dones amb càncer de mama.

Sobre l'impacte de les característiques de la prova de cribratge en els indicadors de qualitat:

5. Comparar els valors predictius positius de la mamografia de cribratge abans i després de la implementació de la mamografia digital.
6. Comparar els valors predictius positius de les diferents sospites radiològiques (patrons radiològics) en dones reconvoques a partir de mamografies analògiques i mamografies digitals.
7. Comparar les característiques dels tumors diagnosticats amb mamografia analògica i mamografia digital.

### **III. MÈTODES I RESULTATS**

---





La tesi està dividida en quatre treballs, publicats en revistes científiques com a articles de recerca original. Tots s'han realitzat en el context del cribratge poblacional, però cada un d'ells analitzen diferents poblacions.

Per facilitar la comprensió dels conjunt de la tesi, a continuació es presenta una taula resum on s'enumeren els objectius específics de la tesi i les referències dels articles originals que hi donen resposta, juntament amb el disseny i les principals característiques de les poblacions d'estudi.

Posteriorment, es presenten els quatre articles que conformen la tesi.

**Taula 4.** Resum dels treballs que conformen la tesi. S'indica el disseny, el període d'estudi i la població d'estudi.

Objectius	Articles	Disseny d'estudi	Període d'estudi	Població d'estudi
<b>L'impacte de la introducció del cribratge en la detecció de càncer de mama</b>				
1. Taxes de detecció dins i fora del cribratge	Domingo et al. Acta Oncol. 2013 Jan;52(1):48-56.	Cohort retrospectiva	Abril 1991 – Abril 2008 Seguiment fins al Gener del 2011, per la identificació dels càncers d'interval	Població diana dels programes de la regió de Copenhagen i Fyn, Dinamarca. Edat: 50-69 anys 3.290 càncers detectats en el cribratge 1.257 càncers d'interval 1.548 càncers diagnosticats en dones no participants
2. Impacte de les diferències organitzatives sobre les taxes de detecció				
<b>L'impacte de la via diagnòstica en les característiques dels tumors i en el seu pronòstic</b>				
3. Característiques biològiques dels càncers d'interval vs. càncers de cribratge	Domingo et al. Cancer Causes Control 2010 Aug;21(8):1155-64.	Cas-control aniuat en una cohort	Novembre 1995- Setembre 2008	Dones participants al programa de cribratge de càncer de mama de l'Hospital del Mar, Barcelona Edat: 50-69 anys 115 càncers detectats en el cribratge 115 càncers d'interval
4. Pronòstic dels tumors en funció de la via diagnòstica	Domingo et al. Eur J Cancer Prev 2013 Jan;22(1):21-8.	Estudi de seguiment	Novembre 1995 – Setembre 2008 Seguiment fins al Desembre de 2010	Dones diagnosticades i tractades de càncer de mama a l'Hospital del Mar, Barcelona Edat: 50-69 anys 97 càncers detectats en el cribratge 34 veritables càncers d'interval 97 càncers diagnosticats en dones no participants
<b>L'impacte de les característiques de la prova de cribratge en els indicadors de qualitat.</b>				
5. VPP de la mamografia analògica i la mamografia digital				
6. VPP segons els patrons radiològics en mamografia analògica i digital	Domingo et al. Eur Radiol. 2011 Sep;21(9):2020-8.	Cohort retrospectiva	Gener 1996- Desembre 2007	Dones cribrades en 4 programes poblacionals de càncer de mama espanyols: Cantàbria, València, Barcelona i Girona. Edat: 45/50-69 anys 171.191 mamografies analògiques 91.647 mamografies digitals
7. Comparació de les característiques dels tumors detectats en mamografia analògica i digital				

## 1. Article 1

**Títol:** Seventeen-years overview of breast cancer inside and outside screening in Denmark. Seventeen-years overview of breast cancer inside and outside screening in Denmark

**Autors:** Domingo L, Jacobsen KK, Euler-Chelpin M, Vejborg I, Schwartz W, Sala M, Lynge E.

**Revista:** Acta Oncol. 2013 Jan;52(1):48-56.

**Factor d'impacte** (2011): 3.330 (Q2 Oncologia)

**DOI:** 10.3109/0284186X.2012.698750

### Resum:

**BACKGROUND:** Long-term data on breast cancer detection in mammography screening programs are warranted to better understand the mechanisms by which screening changes the breast cancer pattern in the population. We aimed to analyze 17 years of breast cancer detection rates inside and outside screening in two Danish regions, emphasizing the influence of organizational differences of screening programs on the outcomes.

**MATERIAL AND METHODS:** We used data from two long-standing population-based mammography screening programs, Copenhagen and Fyn, in Denmark. Both programs offered biennial screening to women aged 50-69 years. We identified targeted, eligible, invited and participating women. We calculated screening detection and interval cancer rates for participants, and breast cancer incidence in non-screened women (= targeted women excluding participants) by biennial invitation rounds. Tumor characteristics were tabulated for each of the three groups of cancers.

**RESULTS:** Start of screening resulted in a prevalence peak in participants, followed by a decrease to a fairly stable detection rate in subsequent invitation rounds. A similar pattern was found for breast cancer incidence in non-screened women. In Fyn, non-screened women even had a higher rate than screening participants during the first three invitation rounds. The interval cancer rate was lower in Copenhagen than in Fyn, with an increase over time in Copenhagen, but not in Fyn. Screen-detected cancers showed tumor features related with a better prognosis than tumors detected otherwise, as more than 80% were smaller than 20 mm and estrogen receptor positive.

**CONCLUSION:** Data from two long-standing population-based screening programs in Denmark illustrated that even if background breast cancer incidence and organization were rather similar, performance indicators of screening could be strongly influenced by inclusion criteria and participation rates. Detection rates should be interpreted with caution as they may be biased by selection into the screening population.



## ORIGINAL ARTICLE

**Seventeen-years overview of breast cancer inside and outside screening in Denmark**LAIA DOMINGO<sup>1,2,3</sup>, KATJA KEMP JACOBSEN<sup>1</sup>, MY VON EULER-CHELPIN<sup>1</sup>, ILSE VEJBORG<sup>4</sup>, WALTER SCHWARTZ<sup>5</sup>, MARIA SALA<sup>2,3</sup> & ELSEBETH LYNGE<sup>1</sup>

<sup>1</sup>Department of Public Health, University of Copenhagen, Denmark, <sup>2</sup>Epidemiology and Evaluation Department, Hospital del Mar-IMIM, Barcelona, Spain, CIBER Epidemiología y Salud Pública, <sup>3</sup>EHEA Doctoral Program in Public Health, Department of Pediatrics, Obstetrics and Gynecology, Preventive Medicine and Public Health, Universitat Autònoma de Barcelona (UAB), Bellaterra, Barcelona, Spain, <sup>4</sup>Center of Diagnostic Imaging, Copenhagen University Hospital, Rigshospitalet, Denmark, and <sup>5</sup>Mammography Screening Program, Odense University Hospital, Denmark

**Abstract**

**Background.** Long-term data on breast cancer detection in mammography screening programs are warranted to better understand the mechanisms by which screening changes the breast cancer pattern in the population. We aimed to analyze 17 years of breast cancer detection rates inside and outside screening in two Danish regions, emphasizing the influence of organizational differences of screening programs on the outcomes. **Material and methods.** We used data from two long-standing population-based mammography screening programs, Copenhagen and Fyn, in Denmark. Both programs offered biennial screening to women aged 50–69 years. We identified targeted, eligible, invited and participating women. We calculated screening detection and interval cancer rates for participants, and breast cancer incidence in non-screened women (= targeted women excluding participants) by biennial invitation rounds. Tumor characteristics were tabulated for each of the three groups of cancers. **Results.** Start of screening resulted in a prevalence peak in participants, followed by a decrease to a fairly stable detection rate in subsequent invitation rounds. A similar pattern was found for breast cancer incidence in non-screened women. In Fyn, non-screened women even had a higher rate than screening participants during the first three invitation rounds. The interval cancer rate was lower in Copenhagen than in Fyn, with an increase over time in Copenhagen, but not in Fyn. Screen-detected cancers showed tumor features related with a better prognosis than tumors detected otherwise, as more than 80% were smaller than 20 mm and estrogen receptor positive. **Conclusion.** Data from two long-standing population-based screening programs in Denmark illustrated that even if background breast cancer incidence and organization were rather similar, performance indicators of screening could be strongly influenced by inclusion criteria and participation rates. Detection rates should be interpreted with caution as they may be biased by selection into the screening population.

Breast cancer is the most common cancer among women worldwide, representing a major public health issue [1]. Over the last three decades its epidemiology has varied considerable: while the mortality has been declining in high-income countries [2], the breast cancer incidence has been increasing until the early 2000s, where after a downturn was reported in many developed countries [3,4]. Several factors have likely contributed to the observed trends, including a significant drop in the hormone therapy use, widespread use of effective systemic treatment modalities,

and mammographic screening programs having past the phase of prevalence peaks [2–4].

The goal of screening mammography is to reduce breast cancer mortality by means of early detection of tumors. Achieving the expected reduction in mortality will depend on the quality of screening procedures and participation. Women who regularly participate in screening may profit from it by detection of cancer at an earlier stage, which are associated with a better prognosis than those detected outside screening [5]. Screening might though also have

negative side effects as false positive tests and over-diagnosis. The benefit of attending a screening program is reduced for women diagnosed with interval cancer (i.e. cancers diagnosed after a negative screening result and before the next invitation to screening). These tumors are detected clinically, thus with a delay in time of diagnosis and worse prognosis features. The interval cancer rate has been recognized as a valid indicator of screening quality and the sensitivity of the screening procedure [6].

Denmark is one of the countries with the highest breast cancer incidence in the world [1]. Mammography screening was introduced in 1991 in the municipality of Copenhagen, and in the county of Fyn in 1993. Nationwide roll out of screening took place in 2007–2010. Although the background cancer incidence does not differ greatly across Danish regions [7], some organizational differences exist between the above mentioned screening programs in criteria for definition of eligible women and in invitation procedures. The existence of the Danish unique identification number allows the link of screening data with population-based registers and the identification of all diagnosis of breast cancer, both detected at screening, as interval cancers, and outside screening. The chance to have accurate information from screened and non-screened women from two regions with long-term screening performance allowed to look for effects of organizational differences in areas with fairly similar background incidence rates. In addition, with nearly 20 years of follow-up, to our knowledge, this is the longest studied scenario showing trends overtime of screen-detected breast cancers, interval cancers and breast cancers diagnosed outside screening. Differences reported both for background cancer incidences and screening protocols as well as for screening performance indicators in European countries [8–11], call for nationwide evaluations of screening tradeoffs.

The aim of the current work was to determine the rates of screen-detected, interval cancers and cancers diagnosed outside screening over 17 years in Denmark (Copenhagen and Fyn), emphasizing the influence of organizational differences in the screening programs to changes of breast cancer detection patterns in the population.

## Material and methods

### *Setting and study population*

In this study we used data from the first eight invitations rounds of the Copenhagen organized screening program (April 4, 1991 to April 14, 2008) and from the first six invitation rounds of the organized screening program in Fyn (November 1, 1993 to December

31, 2005). Both programs offered, by a personal letter, biennial screening to women aged 50 to 69. Screening took place in central clinics, in Fyn supplemented by a mobile van. At first screen two-view mammography was used, at subsequent screens one- or two-view depending on breast density, and from 2004 onwards all women were offered two-view mammography. All mammograms were independently double read by specially trained radiologists. During the study period some organizational and technological changes occurred. The most important being the introductions of high frequency ultrasound devices and stereotactic guided biopsies, of two-view mammography for all screens, and of digital mammography [12].

We distinguished between the *target* population, all women within the catchment area, target age and time period; the *eligible* population, target population excluding women ineligible due to breast cancer related reasons; the *invited* population, eligible population excluding those not invited for other reasons; and the *participating* population, invited women who participated.

*Copenhagen, Target:* The Copenhagen program targeted women aged 50–69 at the start of each invitation round and residing in the catchment area at any time during the invitation round. This means that the Copenhagen program targeted fixed birth cohorts in each invitation round. *Copenhagen, Eligible:* From its implementation, the program aimed at including information from the Danish Breast Cancer Cooperative Group (DBCG) to avoid inviting women operated for breast cancer within the past 18 months. As this register was not fully updated at the time, women operated within the past 18 months were not systematically excluded from invitation. As the program progressed, information was collected on women with breast cancer surgery within the last 18 months, bilateral mastectomy or breast prostheses where mammography was technically not possible. Once identified, these women were exempted from invitation. *Copenhagen, Invited:* Furthermore, women could notify the program if they did not want to be invited. From Copenhagen, invitation data were available.

For the Fyn program some assumptions had to be made as the program did not systematically keep files on targeted and/or invited women. *Fyn, Target:* The Fyn program targeted women aged 50–69 and residing in the catchment area at the date of invitation. This means that the Fyn program targeted a dynamic population, and this had to be taken into account in identifying the target population. *Fyn, Eligible:* The Fyn county had its own IT-system, which enabled exclusion from invitation of women with breast cancer during the last five years, and from

1997 also women with excisions from in situ or non-specified neoplasms. *Fyn, Invited:* As in Copenhagen, women could notify the program if they did not want to be invited. Furthermore, non-participants in the previous round were not invited unless they contacted the program. Invited women with a mammography within the last six months could cancel their screening participation, as they were controlled elsewhere in the health care system.

#### *Data collection*

The Danish Civil Registration System (CRS) was used to identify the target population. The CRS is based on the Danish unique identification number. This register is updated every day and includes demographic information, including migration and death, for all persons ever residing in Denmark since 1968. Each of the two programs kept Mammography Registers on the outcome of screening. We furthermore used data from the Danish Cancer Register (DCR), and the DBCG.

Individual screening data were retrieved from the Mammography Registers and linked by use of the unique identification number with the target populations. This linkage allowed us to distinguish between participants and other targeted women. The Mammography Registers included information on date of screening mammogram, outcome of screening and (where relevant) date of screen-detected cancer diagnosis, including carcinoma in situ (CIS) and invasive cancers. Interval cancers were defined as breast cancers diagnosed within two years of a negative screen or before next screen, whichever came first. Interval cancers and cancers among other targeted women were retrieved by merging the Mammography Registers with the DCR, for the identification of invasive cancers and with the DBCG, for the identification of CIS.

Tumor-related information was retrieved from the DCR and DBCG. From DCR, we obtained data on morphology (International Classification of Diseases -10 Classification), and from DBCG we obtained data on tumor size, lymph node involvement, and estrogen receptor (ER) status.

#### *Statistical analysis*

For targeted women in each screening round, six outcomes were possible: 1) participation in screening without breast cancer diagnosis; 2) participation with screen-detected breast cancer; 3) participation with interval breast cancers; 4) false positive rate (screen-positive women – screen-detected cases); 5) non-screened women without breast cancer diagnosis and; 6) non-screened women with breast cancer. To take the varying length of invitation rounds

into account, we calculated incidence rates for each invitation round. Detection rate in participants was calculated as screen-detected breast cancers divided by person years among participants. Interval cancer rate in participants was calculated as interval cancers divided by person years at risk in up to two years among participants declared negative at screening. Incidence rate in non-screened women was calculated as diagnosed breast cancers in non-screened women divided by the person years accumulated by these women. The women's screening round was defined as the time (first, second, etc) that she participated in the screening program, regardless of the number of the program's invitation round. Data were processed using SAS (version 9.2), SPSS (version 18.0) and R statistical software programs (version 2.12.2).

#### *Ethics*

Use of screening data and tumor-related information was approved by the Danish Data Inspection Agency (2008-41-2191).

#### **Results**

The number of targeted, eligible, invited and participating women presented clear differences in both studied screening programs (Table I). In Copenhagen, the proportion of targeted women identified as non-eligible was small, the proportion of eligible women who had declined invitation was larger, and so was the proportion of invited women who did not participate. By contrast, in Fyn more targeted women were identified as non-eligible, a large proportion of women were not invited because they had declined invitation or because they had not participated previously, and consequently more of the invited women participated in Fyn than was the case in Copenhagen. For these reasons, we present breast cancer rates for participants and non-screened women (including non-eligible, non-invited, and non-participants), respectively.

From 1991, with the implementation of a breast cancer population-based screening program in Copenhagen, it had accounted 214,088 screening participations and 139,461 targeting of non-screened women. It should be noted that a given woman can have been targeted in more than one invitation round. From 1993 onwards, the Fyn program has accounted 272,634 screening participations and 90,692 targeting of non-screened women (Table II).

The target populations of Copenhagen and Fyn had fairly similar breast cancer incidence rates (Table II). Copenhagen had a clear prevalence peak in screen-detected cancers 5.79 per 1000 person years or 11.88

Table I. Number of target population, eligible population, invited women and participants by invitation round in Copenhagen and Fyn screening programs, Denmark.

Dates for invitation round	Length of invitation round (years)	Target population	Eligible population	Eligible/target population (%)	Invited women	Invited/eligible (%)	Participant women	Participants/invited (%)	Participants/target (%)
<b>Copenhagen screening program</b>									
<b>Invitation rounds</b>									
1 4 April 1991–23 April 1993	2.05	45 151	43 092	95.44	43 092	100	30 388	70.52	67.30
2 26 April 1993–30 May 1995	2.01	42 069	41 159	97.84	38 766	94.19	26 109	67.35	62.06
3 1 June 1995–24 March 1997	1.82	41 121	40 037	97.36	36 653	91.55	25 153	68.62	61.17
4 25 March 1997–19 April 1999	2.07	41 686	40 304	96.68	36 864	91.46	25 427	68.98	61.00
5 20 April 1999–31 March 2001	1.95	42 340	40 865	96.52	37 502	91.77	25 059	66.82	59.19
6 1 April 2001–31 May 2003	2.17	43 420	41 815	96.30	33 677	80.54	25 271	75.04	58.20
7 1 June 2003–31 December 2005	2.58	45 051	42 330	93.96	35 340	83.49	26 205	74.15	58.17
8 1 January 2006–14 April 2008	2.28	52 710	50 798	96.37	41 576	81.85	30 476	73.30	57.82
<b>Fyn screening program</b>									
<b>Invitation rounds</b>									
1 1 November 1993–31 December 1995	2.16	56 112	49 111	87.52	49 111	100	41 519	84.54	73.99
2 1 January 1996–31 December 1997	2.00	58 009	52 870	91.14	48 057	90.90	44 117	91.80	76.05
3 1 January 1998–31 December 1999	2.00	59 697	54 577	91.42	48 712	89.25	44 892	92.16	75.20
4 1 January 2000–31 December 2001	2.00	61 247	55 968	91.38	48 839	87.26	45 817	93.81	74.81
5 1 January 2002–31 December 2003	2.00	63 049	57 685	91.49	50 134	86.91	47 458	94.66	75.27
6 1 January 2004–31 December 2005	2.00	65 212	60 233	92.36	53 485	88.80	48 831	91.30	74.88

per 1000 participants. Hereafter the rate decreased, and it remained fairly stable across the subsequent invitation rounds (Figure 1). The breast cancer incidence in non-screened women showed a pattern similar to that of screened women though at a lower level. From the second invitation round onwards, the incidence rate in non-screened women increased steadily. From invitation round four and onwards, the rates for participating and non-screened women overlapped. The interval cancer rate started at 0.97 per 1000 person years, and this rate increase slightly over time being 1.47 per 1000 person years in the eighth invitation round.

The start of screening in Fyn resulted also in a prevalence peak among participating women at 4.47 per 1000 person years or 9.66 per 1000 participants, where after the incidence decreased and remained fairly stable over time (Table II). A similar pattern was seen for non-screened women. However, in Fyn at the start of screening the incidence in non-screened women was higher than the incidence in participants. From invitation round four and onwards, the two sets of rates overlapped. The interval cancer rate in Fyn started at 1.08 per 1000 person years, increasing to a higher level in the next three invitations round where after it decreased again.

In Copenhagen, interval cancers constitute 25% [555/(1622 + 555)] of breast cancers in participants. Of these, 168 (30%) were detected within 0–11 months of the previous negative screen and the remaining 70% within 12–23 months. In Fyn, interval cancers accounted for 30% [702/(1668 + 702)] of breast cancers in participants, with 31% detected within the first 11 months and 69% within 12–23 months of the negative screen. The division of the breast cancer rate among screened women into screen-detected and interval cancer can be used as an indicator of sensitivity. In the first invitation round of the Copenhagen program this proportion was 86% = 5.79/(5.79 + 0.97). In the subsequent invitation rounds this proportion varied between 66% and 74% with no particular trend. In the first invitation round of the Fyn program the proportion was 81% = 4.47/(4.47 + 1.08), and in subsequent invitation rounds the proportion varied between 61% and 72%, also with no particular trend.

Copenhagen had a high false positive rate of 5.6% during the first invitation round (Table II). This rate decreased over subsequent invitation rounds to stabilize at 1.4–1.5% from the sixth round onwards. Fyn had a much follow false positive rate starting out at 1.7% during the first invitation round, and stabilizing at 0.8% during the last two invitation rounds.

Detection data for participants by women's screening round showed a peak at first screening followed by a dramatic decrease and a steady increase



Table II. Number of breast cancers (invasive and carcinoma in situ) detected at screening in participants, diagnosed as interval cancers in participants or diagnosed in non-screened women. Proportion per 1000 women, and rate per 1000 person years. Copenhagen and Fyn screening programs, Denmark.

	Screened women					Non-screened women				Total per 1000 py (4)			
	Participants	Screen-detected cancers, of which (CIS)	per 1000 py (1)	Interval cancers, of which (CIS)	per 1000 py (2)	False positive rate, %	Non-screened women	Diagnosed cancers, of which (CIS)	per 1000 py (3)				
<b>Copenhagen screening program</b>													
Invitation round													
1	30 388	361 (44)	11.88	5.79	58 (2)	1.93	0.97	5.6%	14 763	128 (8)	8.67	4.23	5.91
2	26 109	164 (17)	6.28	3.13	65 (6)	2.51	1.25	4.0%	15 960	62 (0)	3.95	1.96	3.45
3	25 153	156 (18)	6.20	3.41	59 (3)	2.36	1.18	2.5%	15 968	70 (3)	4.38	2.41	3.81
4	25 427	147 (18)	5.78	2.79	73 (1)	2.89	1.44	2.4%	16 260	108 (4)	6.64	3.21	3.80
5	25 059	145 (22)	5.79	2.97	66 (3)	2.65	1.32	1.8%	17 281	94 (6)	5.44	2.79	3.69
6	25 271	180 (42)	7.12	3.28	62 (1)	2.47	1.24	1.5%	18 149	109 (4)	6.01	2.77	3.73
7	26 205	227 (40)	8.66	3.36	83 (2)	3.20	1.60	1.4%	18 846	163 (5)	8.65	3.35	4.07
8	30 476	242 (47)	7.94	3.48	89 (2)	2.94	1.47	1.4%	22 234	162 (5)	7.29	3.20	4.10
Total (Rounds 1-8)	214 088	1622 (248)	7.58	3.57	555 (20)	2.61	1.31	2.6%	139 461	896 (35)	6.43	3.02	4.09
<b>Fyn screening program</b>													
Invitation round													
1	41 519	401 (59)	9.66	4.47	89 (4)	2.16	1.08	1.7%	14 593	187 (11)	12.81	5.93	5.58
2	44 117	236 (35)	5.35	2.67	124 (6)	2.83	1.41	1.1%	13 892	89 (7)	6.41	3.20	3.87
3	44 892	216 (21)	4.81	2.41	140 (4)	3.13	1.57	1.1%	14 805	90 (8)	6.08	3.04	3.74
4	45 817	273 (35)	5.96	2.98	128 (4)	2.81	1.41	1.0%	15 430	90 (1)	5.83	2.92	4.01
5	47 458	257 (19)	5.42	2.71	112 (3)	2.37	1.19	0.8%	15 591	94 (7)	6.03	3.01	3.67
6	48 831	285 (31)	5.84	2.92	109 (4)	2.25	1.12	0.8%	16 381	101 (5)	6.17	3.08	3.80
Total (Rounds 1-6)	272 634	1668 (200)	6.12	3.02	702 (25)	2.59	1.30	1.1%	90 692	651 (39)	7.18	3.54	4.10
Total	486 722	3290 (448)	6.76	3.27	1257 (45)	2.60	1.30	1.8%	230 153	1548 (74)	6.73	3.22	4.10

CIS: Carcinoma in situ.

(1) Person years at risk to develop a screen-detected cancer were estimated as number of participants multiplied by length of invitation round.

(2) Person years at risk to develop an interval cancer were estimated as number of participants, minus participants with screen-detected cancers, multiplied by 2.

(3) Person years at risk to develop a cancer outside screening were estimated as number of non-screened women multiplied by length of invitation round.

(4) For simplicity, for each invitation round based on the total of screen-detected cancers, interval cancers, and cancers in non-screened women, although part of the interval cancers were diagnosed during the next invitation round.

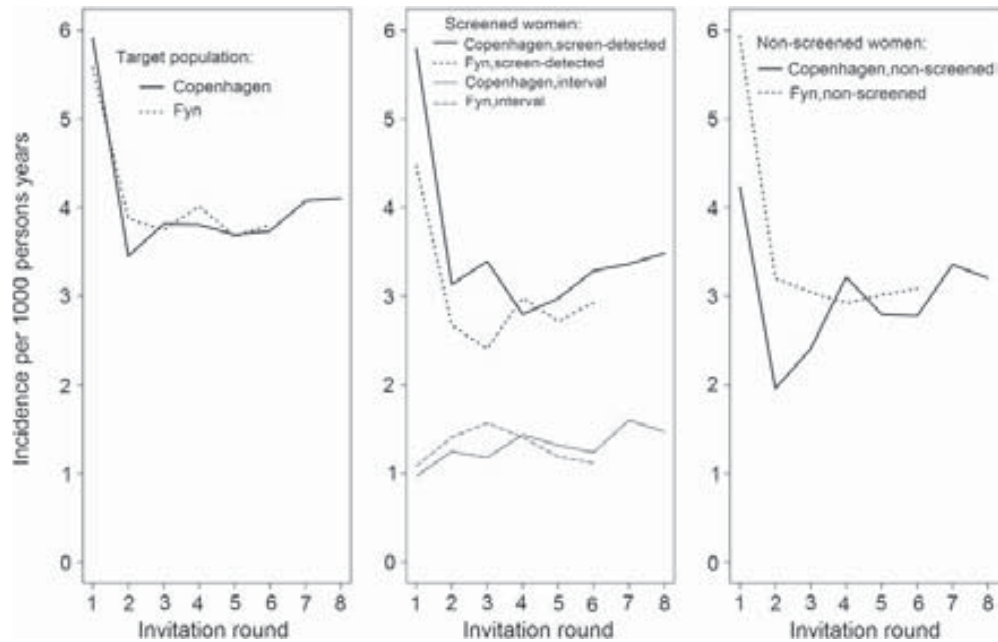


Figure 1. Incidence rates of breast cancer in target population in Copenhagen (1991–2008) and Fyn (1993–2005). Incidence rates of screen-detected breast cancer and interval cancer in participants, and of breast cancers in non-screened women in Copenhagen and Fyn, Denmark.

as the population aged by increasing screen number. The interval cancer rate was stable across women's screening round (Table III).

In situ cases constituted 15% of the screen-detected cases in Copenhagen, and 12% in Fyn. Screen-detected cancers showed tumor features related with a better prognosis, as more than 80% were smaller than 20 mm and expressed ER (Table IV). The Copenhagen profile of screen-detected cancers was better than the one from Fyn, as 41% in Copenhagen vs. 35% in Fyn were less than/equal to 10 mm. Among interval cancers the proportion of large invasive cancer (> 30 mm) was more than twice as high as in screen-detected cancers but still lower than in non-screened women. In Copenhagen nearly 4% of cases in non-screened women were CIS, while the percentage in Fyn was 6%. In Copenhagen, 78% of interval cancers and 76% of cancers in non-participants were ER positive, these percentages being 76% and 77%, respectively, in Fyn.

## Discussion

As expected due to the lead time, the start of organized screening in Denmark resulted in a prevalence peak of screen-detected breast cancers. It was highly surprising, however, that the start of screening resulted also in a "prevalence peak" among non-screened women. In Fyn it was furthermore so that the non-screened women had a higher breast cancer incidence than the screening participants during the first three invitation rounds. Screen-detected tumors

in Copenhagen presented a better profile than those detected in Fyn, as 41% in Copenhagen vs. 35% in Fyn were less than/equal to 10 mm, and 15% vs. 12%, were respectively in situ carcinomas.

The strengths of this study were the long study period allowing an overview over nearly 20 years, and the completeness and accuracy of the data. The current work was based on data from organized population-based screening and on a data from a population-based cancer register. This allowed accurate identification of breast cancer cases detected both inside and outside of screening, and the use of unique personal identification numbers ensured complete linkage. All screening data were obtained from computerized records and came from two screening programs in Denmark whose performance indicators reached the European standards [12,14]. Within the Danish health care system women with symptoms of breast cancer can be referred for mammography by their general practitioner. This system was in place already before the organized screening programs started. The use of opportunistic screening mammography has been very limited in Denmark [13].

The study had some limitations as well. First, we could not divide interval cancers into those deriving from false negative screens and those developed after screening, this information would provide better insight into the screening process. Nonetheless, it has been reported from other screening settings that false negative cases account for about one third of interval cancers or less [15–17]. Secondly, misclassification

Table III. Screen-detected breast cancers and interval cancers. Proportion per 1000 participants by women's screening round in Copenhagen and in Fyn screening programs, Denmark.

Participants		Screen-detected cancers	%	Interval cancers	%
Copenhagen screening program					
Women's screening round					
1st	69 647	646	9.3	150	2.2
2nd	22 597	149	6.6	58	2.6
3rd	41 320	206	5.0	111	2.7
4th	30 466	199	6.5	86	2.8
5th	22 045	165	7.5	58	2.7
6th	15 119	137	9.1	47	3.1
7th	9053	83	9.2	32	3.6
8th	3841	37	9.6	13	3.4
Fyn screening program					
Women's screening round					
1st	75 626	599	7.9	182	2.4
2nd	61 955	270	4.4	176	2.9
3rd	50 364	245	4.9	147	2.9
4th	38 735	231	6.0	86	2.2
5th	28 554	188	6.6	77	2.7
6th	17 400	135	7.8	34	2.0

of the detection mode cannot be excluded. Some interval cancers could be classified as screen-detected if symptomatic women waited for the screening visit instead of making an immediate appointment with a physician. This misclassification would attenuate differences on tumor characteristics between screen-detected and interval cancers.

At the start of the Copenhagen program, only women treated for breast cancer within the past 18 months were considered non-eligible, and 95% of the target population was invited. Only 71% of the invited women participated. However, up to now none of the studies from the Scandinavian countries have indicated that non-participants in screening should be particularly burdened by known breast cancer risk factors [7,18]. The "prevalence peak" in non-screened women was therefore unlikely to be explained by a selection bias. A possible explanation for the "prevalence peak" might be increased awareness of breast cancer along with the start of the screening program. Women with symptoms might then have sought help before they were invited to the program. The clear deficit in incidence among non-screened women in the second invitation round speaks in favor of the hypothesis.

The situation at the start of the screening program in Fyn was different. Here, 12% of the targeted women were identified as non-eligible. In addition to women already diagnosed with breast cancer within the last five years, the non-eligible women in Fyn from 1997 onwards included also those with a prior excision of in situ or non-specified neoplasms, and women with a recent mammogram could cancel participation. As women with a previous benign breast

lesions are at an increased risk of breast cancer, the non-eligible women thus represented a high risk group. Furthermore, the participation rate among invited women in Fyn was 85%, considerably higher than in Copenhagen. As Fyn is a mixed urban-rural area, and Copenhagen is the capital this difference reflects the well-known urban-rural gradient in screening participation [19]. The high ineligibility rate in Fyn combined with the high participation rate in invited women meant that non-screened women in Fyn formed a selected group at higher risk of breast cancer than the screening participants. Our data thus illustrated that in addition to differences in screening performance, the organizational structure as invitation procedures and the screening setting as participation rate may contribute to variations in breast cancer occurrence both inside and outside of screening.

Throughout the study period, the proportion of ineligible women was higher in Fyn than in Copenhagen. This means that the participants in Fyn formed a breast cancer wise more healthy population than the participants in Copenhagen. Such a selection bias may explain why the cancer detection rate was lower in Fyn than in Copenhagen, particularly in the first three invitation rounds. There might though also be true performance differences between the programs. The Fyn program started later than the Copenhagen program and deliberately aimed at a lower recall rate. As showed previously, this policy resulted in a lower cumulative risk of false positive tests in Fyn than in Copenhagen, being 9% and 16%, respectively [20]. The interval cancer rates were though higher in Fyn than in Copenhagen during the first three invitation rounds. While the specificity seem to be higher in Fyn than in Copenhagen indicated by the lower false positive rates, the sensitivity seem to have been lower indicated by the interval cancer rate. These differences were in particular seen during the first three invitation rounds.

These very different observations from otherwise similar mammography screening programs in a small homogeneous country suggest caution in comparison of screening outcomes from different programs. In the European Guidelines [6] an acceptable detection rate at initial screen is set to two times the population background incidence rate, and to 1.5 for subsequent screens. Our analysis has shown that if certain groups of high risk women are exempted from the programs, then detection rates will not be comparable across programs even when the population background incidence rate is controlled for as suggested in the European Guidelines.

Interval cancer rates were slightly higher than those reported in Sweden [9], The Netherlands [21], Finland [22], Italy [11] and Norway [23], but similar

Table IV. Comparison of tumor characteristics between Copenhagen and Fyn regions (Denmark), of screen-detected cancers, interval cancers and cancers detected in non-screened women.

	Screen-detected cancers		Interval cancers		Cancer in non-participants	
	Copenhagen n = 1621 (%)	Fyn n = 1667 (%)	Copenhagen n = 555 (%)	Fyn n = 702 (%)	Copenhagen n = 897 (%)	Fyn n = 651 (%)
Tumor type (1)						
In situ	248 (15.3)	200 (12.0)	20 (3.6)	25 (3.6)	35 (3.9)	39 (6.0)
Invasive	1372 (84.7)	1467 (88.0)	535 (96.4)	677 (96.4)	861 (96.1)	612 (94.0)
Unknown (2)	1	0	0	0	1	0
Tumor's morphology (3)						
Ductal	1045 (77.5)	1110 (77.6)	377 (70.5)	519 (76.7)	632 (73.4)	442 (72.2)
Lobular	186 (13.8)	187 (13.1)	99 (18.5)	72 (10.6)	112 (13.0)	64 (10.5)
Others	118 (8.8)	134 (9.4)	59 (11.3)	86 (12.7)	117 (13.6)	106 (17.3)
Unknown (2)	23	36	0	0	0	0
Tumor's size (3)						
0–10 mm	536 (41.1)	493 (34.6)	103 (21.7)	93 (14.9)	103 (15.1)	66 (13.1)
11–20 mm	604 (46.3)	659 (46.3)	220 (46.3)	263 (42.1)	263 (38.5)	212 (42.0)
21–30 mm	121 (9.3)	188 (13.2)	96 (20.2)	174 (27.9)	186 (27.2)	128 (25.4)
> 30 mm	43 (3.3)	84 (5.9)	56 (11.8)	94 (15.1)	131 (19.2)	99 (19.6)
Unknown(2)	68	43	60	53	178	107
Lymph node involvement (3)						
Negative	970 (74.3)	966 (68.0)	280 (58.7)	321 (51.4)	332 (48.5)	242 (47.5)
Positive	335 (25.7)	454 (32.0)	197 (41.3)	303 (48.6)	353 (51.5)	268 (52.6)
Unknown (2)	67	47	58	53	176	102
ER status (3)						
Negative	169 (13.0)	163 (11.5)	101 (22.4)	152 (24.5)	165 (23.6)	114 (22.9)
Positive	1131 (87.0)	1251 (88.5)	370 (77.6)	469 (75.5)	535 (76.4)	385 (77.1)
Unknown (2)	72	53	58	56	161	113

(1) Include in situ and invasive cancers.

(2) Tumors with missing information were excluded for calculation of percentage.

(3) Include only invasive cancers.

than those found in UK for the 0–24 months after screen [8,24]. Consistent with percentages found in previous works [23,25], nearly 70% of interval cancers appeared during the 12–23 months after the last screening participation. In addition, interval cancers represented 25% of all cancers in screened women in Copenhagen and 30% in Fyn, very similar to percentages reported from Norway [23], Sweden [9] and in a pooled data from six European countries [10], and quite lower than reported from The Netherlands and the UK [21,24]. The higher interval cancer rates reported in the current work reflect the high background incidence of breast cancer in Denmark, but not a poor quality of screening performance.

The trend in screen-detected cancers by women's screening round resembled data previously reported by Bordás et al. [9]. The interval cancer rate did, however, not increase in subsequent screenings as previously reported by these authors [9]. It could be expected that the sensitivity increased by women's screening round, as previous mammograms were here available for comparisons. This would result in a decreasing interval cancer rate, but we did not see such a pattern. A possible explanation could be that most interval cancers are fast growing tumors, not visible at previous screen and truly arousing in the screening interval. Some works that carried out a

radiological revision of both screening and diagnostic mammograms indicates that around 40% of interval cancers are true interval cancers [15,16].

As expected, pathological characteristics of screen-detected tumors were related to better prognosis. At diagnosis, screen-detected cancers were smaller, more frequently lymph node-negative and showed a higher percentage of tumors expressing ER than clinically-detected tumors. These findings are in line with several publications [10,15,21,23]. Although traditional prognostic factors as tumor size or lymph node involvement indicated worse prognosis among cancers in non-participants than for interval cancers, the closer percentages of tumors lacking ER may be a reflection of the more aggressive molecular pattern associated previously with these cancers, especially when true interval cancers were taken into account [15]. However, since we were not able to differentiate between false negative and true interval cancers, these differences could be attenuated given that all of them were considered together as interval cancers.

In conclusion, in spite of a high breast cancer incidence in Danish women, screening has allowed detection of breast cancers at early stages. Nonetheless, the proportion of breast cancers detected outside screening remained high, and these cases had less favorable prognostic characteristics than the

screen-detected cases. Our analysis demonstrated that high data quality is needed for evaluation of mammography screening. All data should be stored with personal identifiers, and all files, including the administrative files on targeted women and invitations, should be kept. While the organized programs in Copenhagen and Fyn operated in fairly similar ways there were nevertheless important differences. Our data illustrated that organizational aspects and screening setting of the programs, as invitation criteria and participation rates, strongly influenced the detection rate which is used as the standard indicator for the outcome of screening. It is therefore pertinent that evaluation of screening starts from the target population, and that the eventual selection in participating women is taken into account in the interpretation of the screening outcome.

### Acknowledgements

This study was partially funded by CIBER de Epidemiología y Salud Pública (CIBERESP).

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

### References

- [1] Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011;61:69–90.
- [2] Autier P, Boniol M, La Vecchia C, Vatten L, Gavin A, Hery C, et al. Disparities in breast cancer mortality trends between 30 European countries: Retrospective trend analysis of WHO mortality database. *BMJ* 2010;341:c3620.
- [3] Euler-Chelpin M. Breast cancer incidence and use of hormone therapy in Denmark 1978–2007. *Cancer Causes Control* 2011;22:181–7.
- [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;356:1670–4.
- [5] Bordas P, Jonsson H, Nystrom L, Lenner P. Survival from invasive breast cancer among interval cases in the mammography screening programmes of northern Sweden. *Breast* 2007;16:47–54.
- [6] Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L, editors. European guidelines for quality assurance in breast cancer screening and diagnosis. 4th ed. Luxembourg: Office for Official Publications of the European Communities; 2006.
- [7] Svendsen AL, Olsen AH, Euler-Chelpin M, Lynge E. Breast cancer incidence after the introduction of mammography screening: What should be expected? *Cancer* 2006;106:1883–90.
- [8] Bennett RL, Sellars SJ, Moss SM. Interval cancers in the NHS breast cancer screening programme in England, Wales and Northern Ireland. *Br J Cancer* 2011;104:571–7.
- [9] Bordas P, Jonsson H, Nystrom L, Lenner P. Interval cancer incidence and episode sensitivity in the Norrbotten Mammography Screening Programme, Sweden. *J Med Screen* 2009;16:39–45.
- [10] Tornberg S, Kemetli L, Ascunce N, Hofvind S, Anttila A, Seradour B, et al. A pooled analysis of interval cancer rates in six European countries. *Eur J Cancer Prev* 2010;19:87–93.
- [11] Bucchi L, Ravaioli A, Foca F, Colamartini A, Falcini F, Naldoni C. Incidence of interval breast cancers after 650,000 negative mammographies in 13 Italian health districts. *J Med Screen* 2008;15:30–5.
- [12] Utzon-Frank N, Vejborg I, Euler-Chelpin M, Lynge E. Balancing sensitivity and specificity: Sixteen year's of experience from the mammography screening programme in Copenhagen, Denmark. *Cancer Epidemiol* 2011;35:393–8.
- [13] Jensen A, Olsen AH, Euler-Chelpin M, Helle NS, Vejborg I, Lynge E. Do nonattenders in mammography screening programmes seek mammography elsewhere? *Int J Cancer* 2005;113:464–70.
- [14] Njor SH, Olsen AH, Bellstrom T, Dyreborg U, Bak M, Axelsson C, et al. Mammography screening in the county of Fyn. November 1993–December 1999. *APMIS Suppl* 2003; (110):1–33.
- [15] Domingo L, Sala M, Servitja S, Corominas JM, Ferrer F, Martinez J, et al. Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain. *Cancer Causes Control* 2010;21:1155–64.
- [16] Hofvind S, Geller B, Skaane P. Mammographic features and histopathological findings of interval breast cancers. *Acta Radiol* 2008;49:975–81.
- [17] Vitak B, Olsen KE, Manson JC, Arneson LG, Stal O. Tumour characteristics and survival in patients with invasive interval breast cancer classified according to mammographic findings at the latest screening: A comparison of true interval and missed interval cancers. *Eur Radiol* 1999;9:460–9.
- [18] Zackrisson S, Andersson I, Manjer J, Janzon L. Non-attendance in breast cancer screening is associated with unfavourable socio-economic circumstances and advanced carcinoma. *Int J Cancer* 2004;108:754–60.
- [19] Euler-Chelpin M, Olsen AH, Njor S, Vejborg I, Schwartz W, Lynge E. Women's patterns of participation in mammography screening in Denmark. *Eur J Epidemiol* 2006;21:203–9.
- [20] Njor SH, Olsen AH, Schwartz W, Vejborg I, Lynge E. Predicting the risk of a false-positive test for women following a mammography screening programme. *J Med Screen* 2007;14:94–7.
- [21] Fracheboud J, de Koning HJ, Beemsterboer PM, Boer R, Verbeek AL, Hendriks JH, et al. Interval cancers in the Dutch breast cancer screening programme. *Br J Cancer* 1999;81:912–7.
- [22] Sarkeala T, Hakama M, Saarenmaa I, Hakulinen T, Forsman H, Anttila A. Episode sensitivity in association with process indicators in the Finnish breast cancer screening program. *Int J Cancer* 2006;118:174–9.
- [23] Hofvind S, Bjurstam N, Sorum R, Bjorndal H, Thoresen S, Skaane P. Number and characteristics of breast cancer cases diagnosed in four periods in the screening interval of a biennial population-based screening programme. *J Med Screen* 2006;13:192–6.
- [24] Lawrence G, O'Sullivan E, Kearins O, Tappenden N, Martin K, Wallis M. Screening histories of invasive breast cancers diagnosed 1989–2006 in the West Midlands, UK: variation with time and impact on 10-year survival. *J Med Screen* 2009;16:186–92.
- [25] Nederend J, Duijm LE, Voogd AC, Groenewoud JH, Jansen FH, Louwman MW. Trends in incidence and detection of advanced breast cancer at biennial screening mammography in The Netherlands: A population based study. *Breast Cancer Res* 2012;14:R10.



## 2. Article 2

**Títol:** Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain.

**Autors:** Domingo L, Sala M, Servitja S, Corominas JM, Ferrer F, Martínez J, Macià F, Quintana MJ, Albanell J, Castells X

**Revista:** Cancer Causes Control 2010 Aug;21(8):1155-64.

**Factor d'impacte (2009):** 3.199 (Q1 Salut Pública, mediambiental i ocupacional; Q2 Oncologia)

**DOI:** 10.1007/s10552-010-9541-6

### **Resum:**

**OBJECTIVE:** To analyze phenotypic classification and other risk factors for interval breast cancer, focusing on true interval and false negative cancers.

**METHODS:** A nested case-control study was performed among 115 cancers detected between two screening mammograms (interval cancers) and 115 screen-detected cancers diagnosed between 1995 and 2008 in a population-based breast cancer screening program in Barcelona (Spain). Bivariate and multivariate analyses were performed to compare patient and tumor molecular characteristics among all interval cancers, true intervals and false negatives, and screen-detected cancers.

**RESULTS:** A total of 42.5% of interval cancers were true interval tumors and 16.2% were false negatives. High breast density and triple negative phenotype were more frequent in true interval cancers than in screen-detected cancers (57.6 and 34.1%, respectively for breast density,  $p = 0.023$ ; 28.1 and 7.5%, respectively for triple negative phenotype,  $p = 0.028$ ), while no statistically significant differences were observed between false negatives and screen-detected cancers. The main adjusted factors associated with true interval cancers compared with screen-detected cancers were high breast density and triple negative phenotype (OR = 3.1, 95% CI, 1.03-9.24 and OR = 8.9, 95% CI, 2.03-38.62, respectively).

**CONCLUSION:** A more aggressive molecular phenotype and high breast density were identified in breast tumors that truly arise in the interval between screenings





# Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain

Laia Domingo · Maria Sala · Sònia Servitja · Josep Maria Corominas ·  
Francisco Ferrer · Juan Martínez · Francesc Macià ·  
Maria Jesús Quintana · Joan Albanell · Xavier Castells

Received: 21 December 2009 / Accepted: 10 March 2010 / Published online: 28 March 2010  
© Springer Science+Business Media B.V. 2010

## Abstract

**Objective** To analyze phenotypic classification and other risk factors for interval breast cancer, focusing on true interval and false negative cancers.

**Methods** A nested case–control study was performed among 115 cancers detected between two screening mammograms (interval cancers) and 115 screen-detected cancers diagnosed between 1995 and 2008 in a population-based breast cancer screening program in Barcelona

(Spain). Bivariate and multivariate analyses were performed to compare patient and tumor molecular characteristics among all interval cancers, true intervals and false negatives, and screen-detected cancers.

**Results** A total of 42.5% of interval cancers were true interval tumors and 16.2% were false negatives. High breast density and triple negative phenotype were more frequent in true interval cancers than in screen-detected cancers (57.6 and 34.1%, respectively for breast density,  $p = 0.023$ ; 28.1 and 7.5%, respectively for triple negative phenotype,  $p = 0.028$ ), while no statistically significant differences were observed between false negatives and screen-detected cancers. The main adjusted factors associated with true interval cancers compared with screen-detected cancers were high breast density and triple negative phenotype (OR = 3.1, 95% CI, 1.03–9.24 and OR = 8.9, 95% CI, 2.03–38.62, respectively).

**Conclusion** A more aggressive molecular phenotype and high breast density were identified in breast tumors that truly arise in the interval between screenings.

**Keywords** Breast cancer screening · Interval cancer · False negatives · Phenotype · Radiology

---

L. Domingo · M. Sala (✉) · F. Macià · X. Castells  
Evaluation and Clinical Epidemiology Department, Hospital del  
Mar-IMIM, Passeig Marítim, 25-29, 08003 Barcelona, Spain  
e-mail: Msalaserra@imas.imim.es

L. Domingo · M. Sala · F. Macià · M. J. Quintana · X. Castells  
CIBER de Epidemiología y Salud Pública (CIBERESP),  
Barcelona, Spain

L. Domingo · M. Sala  
EHEA Doctoral Program in Public Health, Department  
of Pediatrics, Obstetrics and Gynecology, Preventive Medicine  
and Public Health, Universitat Autònoma de Barcelona (UAB),  
Bellaterra, Barcelona, Spain

S. Servitja · J. Albanell  
Medical Oncology, Hospital del Mar-IMIM, Passeig Marítim,  
25-29, 08003 Barcelona, Spain

J. M. Corominas  
Pathology Department, Hospital del Mar-IMIM, Passeig  
Marítim, 25-29, 08003 Barcelona, Spain

F. Ferrer · J. Martínez  
Radiology Department, CRC-Hospital del Mar, Passeig Marítim,  
25-29, 08003 Barcelona, Spain

M. J. Quintana  
Epidemiology Department, Hospital de la Santa Creu i Sant Pau,  
Sant Antoni Maria Claret, 171, 08041 Barcelona, Spain

## Introduction

Breast cancer is a biologically and clinically heterogeneous disease characterized by the accumulation of multiple molecular alterations. Tumors can show distinct proliferation patterns and outcomes. These characteristics may reflect wide variation in the biological characteristics and genetic profiles of breast cancer [1]. In the past two decades, widespread mammographic screening and effective systemic treatment modalities have led to a shift in stage at

presentation and have reduced mortality [2]. Despite the high sensitivity reported for most of population-based screening programs, ranging from 74.7 to 89.4% [3], a percentage of tumors remain undetected in screening rounds, which manifest clinically between a normal screening result and the following invitation for screening. These tumors, called interval breast cancers, are an expected part of any screening program. The detection rate of interval breast cancers is considered a quality measure of radiological performance and has been recognized as a valid indicator of screening quality and the sensitivity of the screening procedure. However, distinct categories of interval cancer may be distinguished by radiological findings in the latest screening mammogram, including false negative and true interval cancers [4]. While false negatives depend mainly on the radiologist's skills and occasional misinterpretations of mammograms, true interval cancers are mainly a consequence of the program's design, growth rate, and sojourn time in the detectable preclinical phase of the carcinoma [5].

There is consensus that interval breast cancers have a less-favorable prognosis than screen-detected cancers [6–10]. A delay in diagnosis could potentially worsen prognosis and consequently impair the efficiency of the screening program. Moreover, some biologic differences have been detected between interval cancers and screen-detected cancers. Interval breast cancers show increased tumor cell proliferation [8, 11, 12] and lower expression of both estrogen (ER) and progesterone receptors (PR) [8, 10, 13] than screen-detected cancers. In addition, higher levels of HER2 and p53 have been reported in interval cancers [10, 12, 14]. ER, PR, and HER2 status can be used to distinguish the triple negative phenotype, which lacks the benefit of specific adjuvant therapy and is associated with high grade and poor prognosis [15]. The triple negative phenotype was more frequently detected among interval cancers than among screen-detected cancers [10].

However, to our knowledge, phenotypic classification in true interval cancers and false negatives has not previously been examined. Analysis of these different interval cancer subgroups could contribute to greater understanding of breast cancer biology in relation to screening. The aim of this study was to characterize phenotypic classification and other risk factors for all interval breast cancers, true interval, and false negative cancers.

## Materials and methods

### Setting and study population

A nested case–control study was performed among women participating in a population-based breast cancer screening

program in an area of 300,000 inhabitants in the city of Barcelona, Spain (target population approximately 90,000 women). Women aged between 50 and 69 were invited to undergo mammography with a 2-year interval between screening rounds. Mediolateral oblique and craniocaudal views were available for each breast. All mammograms were read by two radiologists and when double reading led to different assessments, a third radiologist served as a tie breaker. From the beginning of the program in 1995 until September 2004, all the mammograms were obtained with a standard film-screen technique. In September 2004, a digital mammography technique was introduced, which has been the only system used since September 2004. No other major changes in the program took place during the study period. The program was based on the European Guidelines for Quality Assurance in Mammographic Screening [4] and its results met the Europe Against Cancer standards.

From the beginning of the program to September 2008, 206 interval cancers were detected by means of an active case search (consultation of hospital-based cancer registry and telephone contact with women who underwent mammography in the last scheduled screening but who did not attend the following screening mammogram), given that there is no population-based cancer registry in the city, covering 98% of women lost to follow-up from the program. A total of 115 women with histologically confirmed interval cancers who were diagnosed and treated in two hospitals in the city of Barcelona (98 in the Hospital del Mar, which runs the breast cancer screening program and 17 in a nearby public hospital with similar characteristics) were included in the study as cases. Three of these women had simultaneous bilateral tumors. In these women, both tumors were considered independently in the analysis. Therefore, 118 interval tumors from 115 women were assessed. All other interval cancers were detected in other public hospitals or private centers in the metropolitan area, hampering data collection. A total of 769 screen-detected cancers were reported for the same period. A control group matched by age and year of cancer diagnosis was selected at random among screen-detected cancers ( $n = 115$ ). All cases and controls had no other malignancy or severe underlying disorder.

### Data collection

Information from screening (first or successive screening round, date, additional examinations, etc.) and from women (age, menopausal status, hormone replacement therapy [HRT], previous benign breast disease, and first-degree familial breast cancer) was obtained from the program database. This information is routinely collected from a questionnaire administered at each screening round.

However, breast density was determined but this information was not routinely registered during the screening process. Tumor-related information (pathological tumor-node-metastasis [pTNM] status, histological grade, histological type, tumor size, lymph node involvement) was obtained from the hospital-based cancer registries. Information on biomarker expression (RE, RP, HER2, and p53) was obtained from the databases of the Department of Pathology. Further immunohistochemical analyses were performed in tumor samples in which biomarker expression had not been previously determined.

This study was approved by the ethics committee, and informed consent was available to supply tumor material of biopsy for pathologic evaluation and immunohistochemical analyses.

#### Interval breast cancer classification

For interval breast cancer classification, a retrospective review was performed by two radiologists with more than 10 years experience. This radiological review and breast density determination were carried out when screening or diagnostic mammograms ( $n = 80$ ) could be obtained from clinical records. The last screen mammograms before diagnosis were obtained from the clinical records of patients with interval cancer and were mixed with films with negative findings ( $n = 50$ ) randomly selected from women enrolled in the screening program. Screening mammograms were first reviewed independently and individually by the radiologists and were classified into three groups: positive (clearly malignant features detected), negative (no abnormality detected), and minimal signs (features difficult to detect, no clearly malignant features). In a second review, the radiologists assessed screening and diagnostic mammograms with the histological information to determine whether any abnormalities detected on the screening films corresponded to the site of the subsequent interval cancer.

Interval breast cancers were definitively classified into five groups, as recommended by the European Guidelines: (a) true interval cancers—those in which the previous screening mammogram showed normal or benign features; (b) false negatives—when an abnormality suspicious for malignancy is retrospectively seen on the previous mammogram; (c) minimal signs—when detectable but non-specific signs are identified at the latest screen; (d) occult cancers—that showed no mammographic abnormalities at diagnosis despite clinical signs; and (e) unclassifiable—mammograms not performed on presentation or not available [4]. Moreover, the radiologists determined the breast density of interval and screen-detected cancers included in the study into one of four groups, as defined by the four-point Breast Imaging Reporting and Data System (BI-RADS) system: (1) almost entirely fat, (2) scattered fibroglandular tissue, (3) heterogeneously dense, and (4) extremely dense. This four-point classification was collapsed into the categories of low breast density (categories 1 and 2) and high breast density (categories 3 and 4).

#### Laboratory methods

RE, RP, p53, and HER2 are routinely determined during the diagnostic process in the hospitals. However, not all biomarkers were analyzed during the entire study period. Samples lacking information on biomarker expression were reanalyzed following the same procedures as those used in clinical practice. Immunohistochemical staining was performed on sections of paraffin blocks with the tumors fixed in 10% neutral-buffered formalin for 24 h. The methods and antibodies used are described in Table 1. In accordance with standard guidelines, positivity for ER, PR, and p53 was based on more than 10% of the cells that tested positive. For HER2, scores of 0–1 in the HerceptTest were considered negative, whereas a HerceptTest score of three was considered positive. Equivocal scores of two were confirmed by fluorescent in situ hybridization (FISH) as

**Table 1** Immunohistochemistry: methods, antibodies, and dilutions for biomarker determination

Methods	ER	PR	HER2	P53
Retrieval method	Autoclave 120°C 1 Atm 3 min Citrate buffer pH 7.3	Autoclave 120°C 1 Atm 3 min Citrate buffer pH 7.3	Water bath 95–99°C 40 min Citrate buffer pH 6	Autoclave 120°C 1 Atm 3 min Citrate buffer pH 7.3
Primary antibody	Clone 1D5 (Dako) 1:50	Clone PgR636 (Dako) 1:200	Hercep-test (Dako)	Clone DO7 (Novocastra) 1:50
Detection system	Envision (Dako)	Envision (Dako)	Hercep-test (Dako)	Envision (Dako)
Automat	Autoestainer (Dako)	Autoestainer (Dako)	Autoestainer (Dako)	Autoestainer (Dako)

positive when *HER2/neu* oncogene amplification was detected.

Based on the expression of these biomarkers, the tumors were classified into four phenotypes. Tumors expressing hormone receptors were classified as Luminal (Luminal A for RE and RP expression, and Luminal B for RE expression only). The HER2 phenotype was determined by HER2 overexpression and tumors not expressing ER, PR, or HER2 were classified as triple negatives.

### Statistical analysis

Statistical analysis was performed using the SPSS (version 12.0) and R statistical software programs. A descriptive analysis to compare differences on patient's characteristics and biologic markers distributions between interval cancer subgroups and screen-detected cancers was done using Fisher's exact test. For interval cancers subgroups, only true interval cancers and false negatives are presented. Logistic regression analysis was performed to determine the effect of some women-related and tumor-related features on the odds of developing an interval cancer and a true interval cancer. Variables statistically significant in the bivariate analysis plus age were included in the regression model. All *p* values <0.05 were considered significant.

## Results

One-hundred and fifteen women with interval cancer (mean age 59.7) and 115 screen-detected cancers (mean age 59.2) were included in this study. The radiological classification of interval breast tumors is shown in Table 2. For 38 interval cancers, neither screening nor diagnostic mammograms could be obtained. In the remaining 80 cases, there were 34 (42.5%) true interval cancers, 13

(16.3%) false negatives, 14 (17.5%) occult tumors, 4 (5.0%) minimal signs, and 15 (18.8%) unclassifiable.

A total of 73 interval cancers (61.9%) developed during the second post-screening year, especially true interval cancers (26/34; 76.5%). Other subtypes were distributed homogeneously between the first and second post-screening years.

Patient-related data are shown in Table 3. The only statistically significant difference was observed for breast density, which was higher in women with interval cancer than in those with screen-detected cancers (34.1, 53.9, and 57.6% of high breast density in screen-detected, all interval, and true interval cancers, respectively). The proportion of interval cancers diagnosed in the first screening, premenopausal women, HRT users, current smokers, women with first-degree familial breast cancer, and those with a personal history of breast benign disease was higher than in the screen-detected group, although the differences were not significant. True interval cancers were the most frequently detected cancers in premenopausal women (17.6, 13.0, and 15.4% in true interval cancers, all interval cancers, and false negatives, respectively) and in women with high breast density (57.6, 53.9, and 54.5% in true interval cancers, all interval cancers, and false negatives, respectively).

Tumor-related data at diagnosis are shown in Table 4. The percentage of in situ cancers was 21.5% in screen-detected cancers and was 6.2% in interval cancers, while stage III was observed in 5.6% of screen-detected and in 24.8% of interval cancers. This percentage was higher in true interval cancers (29.4%) and false negatives (30.8%). All differences were statistically significant ( $p < 0.001$ ). Regarding tumor size, 51.0% of interval cancers measured >20 mm at diagnosis compared with 23.9% of screen-detected cancers ( $p < 0.001$ ). This percentage was higher in true interval tumors (66.7%) than in false negatives (60.0%). Interval cancers also had a higher histological grade than screen-detected tumors. Lymph node involvement was present in 41.6% of interval cancers and in 26.8% of screen-detected cancers ( $p = 0.024$ ), this proportion being higher in false negative tumors (66.7%). No metastasis was found in screen-detected cancer at diagnosis, while 2.6% of interval cancers showed tumoral spread.

Differences between groups in biomarker expression and phenotypic classification are shown in Table 5. The percentage of tumors expressing hormone receptors (especially estrogen receptor) was higher in screen-detected cancers (84.1%) than in all interval cancers (69.3%;  $p = 0.012$ ) and true interval cancers (60.6%;  $p = 0.007$ ). However, the percentage of tumors with ER expression was similar between false negatives and screen-detected cancers. The percentage of cancers with cell cycle dysregulation and potential genetic instability, as measured by p53 expression, was similar in the two study groups (21.8%

**Table 2** Radiological classification of interval cancers by months since last screening mammogram

	Overall	Time since screening	
		≤12 months	>12 months
No mammograms available <sup>a</sup>	38	15 (39.5)	23 (60.5)
Mammograms available, <i>n</i> (%)	80	30 (37.5)	50 (62.5)
True interval	34 (42.5)	8 (23.5)	26 (76.5)
False negative	13 (16.3)	6 (46.2)	7 (53.8)
Occult	14 (17.5)	7 (50.0)	7 (50.0)
Minimal signs	4 (5.0)	2 (50.0)	2 (50.0)
Unclassifiable	15 (18.8)	7 (46.7)	8 (53.3)

<sup>a</sup> Interval cancers with neither screening nor diagnostic mammograms available

**Table 3** Patient-related data according to cancer classification

Factor	Screen-detected cancers <i>n</i> = 115 (%)	Interval cancers		Interval cancers with available mammograms <i>n</i> = 80			
		<i>n</i> = 115 (%)	<i>p</i> value <sup>a</sup>	True interval cancers		False negatives	
				<i>n</i> = 34 (%)	<i>p</i> value <sup>b</sup>	<i>n</i> = 13 (%)	<i>p</i> value <sup>c</sup>
Age (mean; SD)	59.2 (5.7)	59.7 (5.7)		59.9 (5.4)		62.6 (5.4)	
Screening							
First screening	26 (22.6)	35 (30.4)		10 (29.4)		2 (15.4)	
Successive screening	89 (77.4)	80 (69.6)	0.232	24 (70.6)	0.494	11 (84.6)	0.732
Menopausal status							
Premenopausal	16 (15.1)	15 (13.0)		6 (17.6)		2 (15.4)	
Postmenopausal	90 (84.9)	100 (87.0)	0.701	28 (82.4)	0.788	11 (84.6)	1.000
Unknown <sup>d</sup>	9	0		0		0	
Hormone replacement therapy use							
No	109 (94.8)	105 (91.3)		31 (91.2)		12 (92.3)	
Former use	3 (2.6)	3 (2.6)		2 (5.9)		0 (0)	
Current use	3 (2.6)	7 (6.1)	0.475	1 (2.9)	0.567	1 (7.7)	0.537
Smoking							
Never	90 (78.3)	85 (73.9)		25 (73.5)		10 (76.9)	
Current or former	25 (21.7)	30 (26.1)	0.537	9 (26.5)	0.642	3 (23.1)	1.000
Family history of breast cancer							
No	100 (87.0)	92 (80.0)		27 (79.4)		12 (92.3)	
Yes	15 (13.0)	23 (20.0)	0.214	7 (20.6)	0.280	1 (7.7)	1.000
Personal history of benign breast disease							
No	93 (80.9)	77 (69.4)		23 (70.6)		9 (75.0)	
Yes	22 (19.1)	34 (30.6)	0.064	10 (29.4)	0.229	3 (25.0)	0.703
Unknown <sup>d</sup>	0	4		1		0	
Breast density							
Low	58 (65.9)	35 (46.1)		14 (42.4)		5 (45.5)	
High	30 (34.1)	41 (53.9)	0.012	19 (57.6)	0.023	6 (54.5)	0.201
Unknown <sup>d</sup>	27	39		1		2	

<sup>a</sup> Screen-detected cancers and all interval cancers are compared. Fisher's exact two-sided test

<sup>b</sup> Screen-detected cancers and true intervals are compared. Fisher's exact two-sided test

<sup>c</sup> Screen-detected cancers and false negatives are compared. Fisher's exact two-sided test

<sup>d</sup> Tumors with missing information were excluded for calculation of percentage

of screen-detected and 22.3% interval cancers;  $p = 1.000$ ), as was the percentage of HER2 overexpression (18.3% in screen-detected and 18.0% interval cancers;  $p = 1.000$ ). However, differences appeared when we compared screen-detected cancers with true interval cancers: true interval cancers showed higher p53 expression (21.8% of screen-detected and 34.4% of true interval cancers) and lower HER2 expression than screen-detected cancers (18.3% of screen-detected and 12.5% of true interval cancers). Indeed, p53 expression was associated with RE/RP-tumors ( $p < 0.001$ ) or HER2 + tumors ( $p < 0.001$ ) (data not shown). The distribution of phenotypes between interval and screen-detected cancers was heterogeneous, but the differences were only statistically significant when true interval tumors were compared with screen-detected

cancers. The most frequent phenotype in all four groups was Luminal A, but the highest percentage was observed in screen-detected cancers (55.7%), followed by all interval cancers (52.3%), false negatives (46.2%), and true interval cancers (40.6%). The HER2 phenotype occurred more frequently in screen-detected (18.9%) than in interval cancers (15.3%) and true interval cancers (12.5%). However, the percentage of triple negative phenotype was higher in interval tumors (17.1%) than in screen-detected cancers (7.5%), with the highest occurrence (28.1%) in true interval cancers. There were no significant differences in phenotypic classification between screen-detected and false negatives tumors.

The results of multiple logistic regression analysis are shown in Table 6. Tumors presenting triple negative

**Table 4** Tumor-related data according to cancer classification

Factor	Screen-detected cancers <i>n</i> = 115 (%)	Interval cancers		Interval cancers with available mammograms <i>n</i> = 80			
		<i>n</i> = 118 (%)	<i>p</i> value <sup>a</sup>	True interval cancers		False negatives	
				<i>n</i> = 34 (%)	<i>p</i> value <sup>b</sup>	<i>n</i> = 13 (%)	<i>p</i> value <sup>c</sup>
<b>TNM stage</b>							
0 (in situ)	23 (21.5)	7 (6.2)		0 (0)		0 (0)	
I	49 (45.8)	31 (27.4)		7 (20.6)		4 (30.8)	
II	29 (27.1)	44 (38.9)		16 (47.1)		4 (30.8)	
III	6 (5.6)	28 (24.8)		10 (29.4)		4 (30.8)	
IV	0 (0)	3 (2.7)	<0.001	1 (2.9)	<0.001	1 (7.7)	0.001
Unknown <sup>d</sup>	8	5		0		0	
<b>Tumor size (mm)</b>							
In situ and <10 mm	55 (48.7)	24 (23.1)		3 (9.09)		3 (30.0)	
10–20 mm	31 (27.4)	27 (26.0)		8 (24.2)		1 (10.0)	
>20 mm	27 (23.9)	53 (51.0)	<0.001	22 (66.7)	<0.001	6 (60.0)	0.080
Unknown <sup>d</sup>	2	14		1		3	
<b>Histologic grade</b>							
I	31 (33.3)	15 (15.6)		7 (22.6)		3 (25.0)	
II	37 (39.8)	45 (46.9)		14 (45.2)		5 (41.7)	
III	25 (26.9)	36 (37.5)	0.016	10 (32.3)	0.546	4 (33.3)	0.863
Unknown <sup>d</sup>	22	22		3		1	
<b>Morphology/histopathological type</b>							
Ductal cancer	92 (80.0)	84 (71.2)		29 (85.3)		11 (84.6)	
Lobular cancer	7 (6.1)	16 (13.6)		3 (8.8)		0 (0)	
Others	16 (13.9)	18 (15.3)	0.146	2 (5.9)	0.445	2 (1.8)	1.000
<b>Lymph node status</b>							
Negative	82 (73.2)	66 (58.4)		18 (52.9)		4 (33.3)	
Positive	30 (26.8)	47 (41.6)	0.024	16 (47.1)	0.035	8 (66.7)	0.008
Unknown <sup>d</sup>	3	5		4		1	
<b>Metastasis</b>							
Negative	110 (100)	112 (97.4)		33 (97.1)		12 (92.3)	
Positive	0 (0)	3 (2.6)	0.247	1 (2.9)	0.259	1 (7.7)	0.106
Unknown	5	3		0		0	

<sup>a</sup> Screen-detected cancers and all interval cancers are compared. Fisher's exact two-sided test

<sup>b</sup> Screen-detected cancers and true intervals are compared. Fisher's exact two-sided test

<sup>c</sup> Screen-detected cancers and false negatives are compared. Fisher's exact two-sided test

<sup>d</sup> Tumors with missing information were excluded for calculation of percentage

phenotype, dense breasts, and HRT use were independently associated with both all interval cancers and true interval cancers. The highest association effect was observed for HRT use (OR = 9.32) and presentation of triple negative phenotype (OR = 8.85) in true interval cancers.

## Discussion

An outstanding finding of this study is the high proportion of triple negative tumors among true interval cancers. This finding indicates that some interval cancers may represent a

specific subgroup of cancers that more frequently expresses a less-favorable predictive biomarker pattern and factors of worse prognosis. These factors would make these tumors less likely to be detected in regular screening programs.

To our knowledge, no previous study has evaluated phenotypic classification among true interval and false negative subgroups, and only a few studies have provided biomarker information on interval subgroups. This lack of studies may partly reflect difficulties in obtaining information on interval cancers and their classification. Despite these difficulties, radiological review of mammograms in the present study showed similar percentages of true

**Table 5** Biomarker expression and phenotypic classification according to cancer classification

Factor	Screen-detected cancers <i>n</i> = 115 (%)	Interval cancers		Interval cancers with available mammograms <i>n</i> = 80			
		<i>n</i> = 118 (%)	<i>p</i> value <sup>a</sup>	True interval cancers		False negatives	
				<i>n</i> = 34 (%)	<i>p</i> value <sup>b</sup>	<i>n</i> = 13 (%)	<i>p</i> value <sup>c</sup>
<b>ER status</b>							
Negative	18 (15.9)	35 (30.7)		13 (39.4)		3 (23.1)	
Positive	95 (84.1)	79 (69.3)	0.012	20 (60.6)	0.007	10 (76.9)	0.454
Unknown <sup>d</sup>	2	4		1		0	
<b>PR status</b>							
Negative	40 (35.4)	53 (46.5)		21 (63.6)		5 (38.5)	
Positive	73 (64.6)	61 (53.5)	0.106	12 (36.4)	0.005	8 (61.5)	1.000
Unknown <sup>d</sup>	2	4		1		0	
<b>p53 status</b>							
Negative	86 (78.2)	73 (77.7)		21 (65.6)		11 (84.6)	
Positive	24 (21.8)	21 (22.3)	1.000	11 (34.4)	0.166	2 (15.4)	0.734
Unknown <sup>d</sup>	5	24		2		0	
<b>HER2 status</b>							
Negative	89 (81.7)	91 (82.0)		28 (87.5)		11 (84.6)	
Positive	20 (18.3)	20 (18.0)	1.000	4 (12.5)	0.590	2 (15.4)	1.000
Unknown <sup>d</sup>	6	7		2		0	
<b>Phenotype</b>							
Luminal type A	59 (55.7)	58 (52.3)		13 (40.6)		6 (46.2)	
Luminal type B	19 (17.9)	17 (15.3)		6 (18.8)		3 (23.1)	
HER2	20 (18.9)	17 (15.3)		4 (12.5)		2 (15.4)	
Triple negative	8 (7.5)	19 (17.1)	0.193	9 (28.1)	0.028	2 (15.4)	0.645
Unknown <sup>d</sup>	9	7		2		0	

<sup>a</sup> Screen-detected cancers and all interval cancers are compared. Fisher's exact two-sided test

<sup>b</sup> Screen-detected cancers and true intervals are compared. Fisher's exact two-sided test

<sup>c</sup> Screen-detected cancers and false negatives are compared. Fisher's exact two-sided test

<sup>d</sup> Tumors with missing information were excluded for calculation of percentage

interval cancers and false negatives to other studies [5, 7, 16, 17] and to the European Guidelines, which recommend that false negative results represent less than 20% of reviewed mammograms. Most interval cancers were detected during the second year post-screening, which was particularly important for true interval cancers.

The mean age of women with interval cancer was 59.7, higher than the previously reported mean age, ranging from 54 to 59 [5, 7, 10, 13, 18]. One reason for this discrepancy could be differences in the age of the target population invited to participate in these programs. Some women [5, 13] entered the screening program at 40 or 45 years, which could reduce the mean age of women with interval cancers. We found no differences among interval subgroups, whereas Vitak et al. [5] reported that women with occult tumors were younger than those in other categories. Age is known as a determinant of cancer and is also associated with breast density and menopausal status, factors that have previously been reported to be strongly

related to interval breast cancer [18, 19]. Although controls were age matched to avoid the age confounding effect, interval cancers were more likely to occur in women with dense breasts, which is consistent with data reported in previous studies [10, 19–21]. Breast density has recently been shown not only to decrease mammogram sensitivity, explained by a masking effect, but also to be a variable of early biologic effect in epidemiological, clinical, and genetic studies of breast cancer [22, 23].

Distinct hypotheses have been suggested to explain the association between breast density and tumors that truly arise in the interval after screening; one possible mechanism that could link an increase in breast stroma to tumor aggressiveness could be the action of growth factors produced in the stroma [24], such as insulin-like growth factor (IGF)-I, an IGF that shows high mitotic activity and is closely related to density [25]. Other authors have suggested that there may be a relationship between breast density and the estrogen levels intrinsic to women's age

**Table 6** Factors associated with interval cancers and true interval cancers

Factor		All interval cancers		True interval cancers	
		Adjusted OR <sup>a</sup> (95% CI)	<i>p</i> value	Adjusted OR <sup>b</sup> (95% CI)	<i>p</i> value
Phenotype					
Others	Ref.				
Triple negative		4.70 (1.42–15.60)	0.011	8.85 (2.03–38.62)	0.004
Breast density					
Low	Ref.				
High		2.77 (1.26–6.12)	0.012	3.08 (1.03–9.24)	0.045
Hormone replacement therapy					
No	Ref.				
Yes <sup>c</sup>		4.76 (1.04–21.82)	0.045	9.32 (1.28–68.12)	0.028
Family history of breast cancer					
No	Ref.				
Yes		1.41 (0.47–4.27)	0.540	1.62 (0.36–7.17)	0.528
Personal history of breast disease					
No	Ref.				
Yes		1.97 (0.82–4.75)	0.129	1.92 (0.56–6.62)	0.303

CI confidence interval

<sup>a</sup> Odds ratio for occurrence of interval cancer, adjusted for age at screening mammogram, TNM stage and first or successive screening

<sup>b</sup> Odds ratio for occurrence of true interval cancer, adjusted for age at screening mammogram, TNM stage and first or successive screening

<sup>c</sup> Include current and former users

and associated with HRT use [26]. Finally, a genetic component has also been recognized [27]. However, this association between breast density and breast cancer risk has also been questioned [28] on the basis of the subjectivity of breast density interpretation and reader variability [29]. Therefore, the association between interval cancers and true interval cancers with dense breasts, the effect of high breast density on screening performance, as well as the relationship between breast density and biological processes are possible lines of future research. Harms and benefits that might suppose a routine breast density determination should be analyzed.

Hormone replacement therapy can affect the accuracy of breast cancer screening because of its association with breast density [30]. In our study, the high OR found was difficult to evaluate, because of very low number of cases. However, these results are in agreement with the hypothesis of a relationship between HRT and interval cancers, as reported by some authors [18, 31].

Metastases to regional lymph nodes were more frequent among interval breast cancers (especially false negative tumors), as observed by other authors [7, 8, 17, 32]. This finding is partly related to the longer time to diagnosis [32].

Some studies have compared biomarker expression between interval cancers and screen-detected cancers. All of these studies described lower ER and PR expression in interval cancers [8, 10, 12, 13]. The results on HER2

overexpression were more controversial, a higher proportion of HER2 overexpression being observed in the interval group in some studies [10, 13] and in screen-detected cancers in others [8]. We found no differences in HER2 expression between the two groups. The results on p53 expression were also contradictory. While Gilliland et al. [12] suggested that p53 may be an independent determinant for interval breast cancer (OR = 2.96), in our study, like others [7, 8], no association between p53 and interval cancers was observed. However, when distinct interval subgroups were analyzed, p53 was overexpressed in true interval cancers, especially in those with the triple negative phenotype. This result has not previously been reported, but Tan et al. [33] noted an association of p53 with triple negative phenotype tumors, which was also related to a high proportion of relapses. Studies of biomarker expression among interval subgroups also reported the lowest expression of ER and PR in true interval cancers [5, 16, 32], suggesting the existence of heterogeneity.

Phenotypic classification allowed us to detect a higher prevalence of the triple negative phenotype in interval cancers than in controls, especially in true interval cancers. These results are consistent with those reported by Collet et al. [10] in a study with a similar sample size (95 interval cancers) from a Norwegian Breast Cancer Screening Program that reported that basal-like phenotype was more frequent in interval cancers than in screen-detected



cancers. These results suggest that true interval cancers could be a particular subgroup with specific characteristics rendering these tumors less likely to be detected in regular screening and more likely to have the poorest prognosis, a possibility that is related to the length bias in the screening program. Efforts should be made to identify women at risk of developing triple negative breast cancers, including young women [34], BRCA1/2 mutation carriers [35], and women with high breast density [36]. These women could be offered shorter intervals between screening periods or be included in screening programs earlier.

The present study has some limitations, the main one being the small sample size. Cases included in a single screening program but diagnosed and treated at two hospitals were studied, reducing the sample size but ensuring the quality and homogeneity of the data. The sample size of the present study is too small to establish a high risk group except for women with dense breasts; however, some of the variables studied, such as HRT use and screening round, should be taken into consideration in further studies. Complete classification of all interval cancers was not possible because not all mammograms were available, in partly because of the long retrospective period analyzed. However, no differences were observed between interval cancers with and without mammograms, indicating the absence of selection bias. Few studies have reported a larger sample size, partly due to the low incidence of interval cancers and the difficulty of obtaining information on this type of tumor, which requires follow-up of the target population and can only be identified in population based, but not in opportunistic, breast cancer screening programs. In addition, radiological classification of interval cancers is not easy to perform, and well-established protocols are lacking. Consequently, comparison of the results on interval cancers is difficult, given that many studies follow different strategies of interval classification [37]. We performed two mammographic reviews with both blinded and unblinded components, as recommended by the European Guidelines [4].

In conclusion, a more aggressive molecular phenotype (triple negative phenotype) was identified in breast tumors that truly arise in the interval between screenings and most true interval cancers were detected in women with dense breasts. Although interval cancers were influenced by screening periodicity, this finding provides the opportunity to apply new knowledge on molecular and genetic breast cancer research to screening. The characterization of different subgroups of interval cancers, particularly true interval cancers, could help to improve the efficiency of screening programs by aiding the choice of optimal intervals between screenings or the use of screening technologies for specific subsets of women.

**Acknowledgments** The authors thank Cristina Hernández, Rubén Román, Paco Belvis, and Jose María Velarde for their assistance in analysis and data management and Teresa Baró for technical assistance in performing the immunohistochemical analysis. We also thank the Tumor Bank of the Department of Pathology of Hospital del Mar and Xarxa de Bancs de Tumors de Catalunya for providing tissue samples.

**Financial support** This study was partially supported by CIBER de Epidemiología y Salud Pública (CIBERESP) [AE08\_004], PS09/01285, RD06/0020/0109 (ISCiii/FEDER) and 2009SRG321 (Generalitat de Catalunya).

## References

1. Rakha EA, El Sayed ME, Reis-Filho JS, Ellis IO (2008) Expression profiling technology: its contribution to our understanding of breast cancer. *Histopathology* 52:67–81
2. Benson JR, Jatoi I, Keisch M, Esteva FJ, Makris A, Jordan VC (2009) Early breast cancer. *Lancet* 373:1463–1479
3. Tornberg S, Codd M, Rodrigues V, Segnan N, Ponti A (2005) Ascertainment and evaluation of interval cancers in population-based mammography screening programmes: a collaborative study in four European centers. *J Med Screen* 12:43–49
4. Perry N, Broeders M, de Wolf C, Törnberg C, Holland R, von Karsa L (2006) European guidelines for quality assurance in breast cancer screening and diagnosis, 4th edn. Office for Official Publications of the European Communities, Luxembourg
5. Vitak B, Olsen KE, Manson JC, Arnesson LG, Stal O (1999) Tumour characteristics and survival in patients with invasive interval breast cancer classified according to mammographic findings at the latest screening: a comparison of true interval and missed interval cancers. *Eur Radiol* 9:460–469
6. Brekelmans CT, van Gorp JM, Peeters PH, Collette HJ (1996) Histopathology and growth rate of interval breast carcinoma. Characterization of different subgroups. *Cancer* 78:1220–1228
7. Cowan WK, Angus B, Gray JC, Lunt LG, al Tamimi SR (2000) A study of interval breast cancer within the NHS breast screening programme. *J Clin Pathol* 53:140–146
8. Porter PL, El Bastawissi AY, Mandelson MT (1999) Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 91:2020–2028
9. Burrell HC, Sibbering DM, Wilson AR (1996) Screening interval breast cancers: mammographic features and prognosis factors. *Radiology* 199:811–817
10. Collett K, Stefansson IM, Eide J (2005) A basal epithelial phenotype is more frequent in interval breast cancers compared with screen detected tumors. *Cancer Epidemiol Biomarkers Prev* 14:1108–1112
11. Crosier M, Scott D, Wilson RG, Griffiths CD, May FE, Westley BR (1999) Differences in Ki67 and c-erbB2 expression between screen-detected and true interval breast cancers. *Clin Cancer Res* 5:2682–2688
12. Gilliland FD, Joste N, Stauber PM (2000) Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst* 92:743–749
13. Palka I, Kelemen G, Ormandi K (2008) Tumor characteristics in screen-detected and symptomatic breast cancers. *Pathol Oncol Res* 14:161–167
14. Anttinen J, Kuopio T, Nykanen M, Torkkeli H, Saari U, Juhola M (2003) Her-2/neu oncogene amplification and protein

- over-expression in interval and screen-detected breast cancers. *Anticancer Res* 23:4213–4218
15. Irvin WJ Jr, Carey LA (2008) What is triple-negative breast cancer? *Eur J Cancer* 44:2799–2805
  16. Bare M, Sentis M, Galceran J (2008) Interval breast cancers in a community screening programme: frequency, radiological classification and prognostic factors. *Eur J Cancer Prev* 17:414–421
  17. Hofvind S, Geller B, Skaane P (2008) Mammographic features and histopathological findings of interval breast cancers. *Acta Radiol* 49:975–981
  18. Hofvind S, Moller B, Thoresen S, Ursin G (2006) Use of hormone therapy and risk of breast cancer detected at screening and between mammographic screens. *Int J Cancer* 118:3112–3117
  19. Mandelson MT, Oestreicher N, Porter PL (2000) Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 92:1081–1087
  20. Ciatto S, Visioli C, Paci E, Zappa M (2004) Breast density as a determinant of interval cancer at mammographic screening. *Br J Cancer* 90:393–396
  21. Porter GJ, Evans AJ, Cornford EJ (2007) Influence of mammographic parenchymal pattern in screening-detected and interval invasive breast cancers on pathologic features, mammographic features, and patient survival. *AJR Am J Roentgenol* 188:676–683
  22. Boyd NF, Rommens JM, Vogt K (2005) Mammographic breast density as an intermediate phenotype for breast cancer. *Lancet Oncol* 6:798–808
  23. Harvey JA, Bovbjerg VE (2004) Quantitative assessment of mammographic breast density: relationship with breast cancer risk. *Radiology* 230:29–41
  24. Li T, Sun L, Miller N (2005) The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. *Cancer Epidemiol Biomarkers Prev* 14:343–349
  25. Guo YP, Martin LJ, Hanna W (2001) Growth factors and stromal matrix proteins associated with mammographic densities. *Cancer Epidemiol Biomarkers Prev* 10:243–248
  26. Pike MC, Pearce CL, Wu AH (2004) Prevention of cancers of the breast, endometrium and ovary. *Oncogene* 23:6379–6391
  27. Ziv E, Shepherd J, Smith-Bindman R, Kerlikowske K (2003) Mammographic breast density and family history of breast cancer. *J Natl Cancer Inst* 95:556–558
  28. Kopans DB (2008) Basic physics and doubts about relationship between mammographically determined tissue density and breast cancer risk. *Radiology* 246:348–353
  29. Nicholson BT, LoRusso AP, Smolkin M, Bovbjerg VE, Petroni GR, Harvey JA (2006) Accuracy of assigned BI-RADS breast density category definitions. *Acad Radiol* 13:1143–1149
  30. Kavanagh AM, Cawson J, Byrnes GB (2005) Hormone replacement therapy, percent mammographic density, and sensitivity of mammography. *Cancer Epidemiol Biomarkers Prev* 14:1060–1064
  31. Banks E (2001) Hormone replacement therapy and the sensitivity and specificity of breast cancer screening: a review. *J Med Screen* 8:29–34
  32. Porter GJ, Evans AJ, Burrell HC, Lee AH, Ellis IO, Chakrabarti J (2006) Interval breast cancers: prognostic features and survival by subtype and time since screening. *J Med Screen* 13:115–122
  33. Tan DS, Marchio C, Jones RL (2008) Triple negative breast cancer: molecular profiling and prognostic impact in adjuvant anthracycline-treated patients. *Breast Cancer Res Treat* 111:27–44
  34. Dolle JM, Daling JR, White E (2009) Risk factors for triple-negative breast cancer in women under the age of 45 years. *Cancer Epidemiol Biomarkers Prev* 18:1157–1166
  35. Haffty BG, Yang Q, Reiss M (2006) Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *J Clin Oncol* 24:5652–5657
  36. Ma H, Luo J, Press MF, Wang Y, Bernstein L, Ursin G (2009) Is there a difference in the association between percent mammographic density and subtypes of breast cancer? Luminal A and triple-negative breast cancer. *Cancer Epidemiol Biomarkers Prev* 18:479–485
  37. Ciatto S, Catarzi S, Lamberini MP (2007) Interval breast cancers in screening: the effect of mammography review method on classification. *Breast* 16:646–652

### 3. Article 3

**Títol:** Aggressiveness features and outcomes of true interval cancers: comparison between screen-detected and symptom-detected cancers.

**Autors:** Domingo L, Blanch J, Servitja S, Corominas JM, Murta-Nascimento C, Rueda A, Redondo M, Castells X, Sala M.

**Revista:** Eur J Cancer Prev 2013 Jan;22(1):21-8.

**Factor d'impacte** (2011): 2.130 (Q3 Oncologia)

**DOI:** 10.1097/CEJ.0b013e328354d324.

#### **Resum:**

**BACKGROUND:** The question of whether screen detection confers an additional survival benefit in breast cancer is unclear and subject to several biases. Our aim was to examine the role of the diagnostic method (screen-detected, symptom-detected, and true interval cancers) and the clinical-pathological features in relapse-free survival and overall survival in breast cancer patients.

**METHODS:** We included 228 invasive breast cancers diagnosed in Barcelona from 1996 to 2008 among women aged 50-69 years. Ninety-seven patients were screen detected within the screening, 34 truly arose between 2-year screening mammograms (true interval cancers), and 97 were symptom detected outside the screening. The clinical-pathological features at diagnosis were compared. The overall and disease-free survival probabilities were computed using the Kaplan-Meier method. Cox proportional hazard models were applied, with adjustment by clinical-pathological variables.

**RESULTS:** At diagnosis, symptom-detected and true interval cancers were in more advanced stages and were less differentiated. The highest proportion of triple-negative cancers was detected among true interval cancers ( $P=0.002$ ). At 5 years of follow-up, the disease-free survival rates for screen-detected, true interval, and symptom-detected cancers were 87.5% (95% confidence interval, 80.5-95.2%), 64.1% (46.4-88.5%), and 79.4% (71.0-88.8%), respectively, and the overall survival rates were 94.5% (89.3-99.9%), 65.5% (47.1-91.2%), and 85.6% (78.3-93.6%), respectively. True interval cancers had the highest hazard ratio for relapse prediction (1.89; 0.67-5.31) and a hazard ratio of death of 5.55 (1.61-19.15) after adjustment for tumor-node-metastasis stage and phenotype.

CONCLUSION: Clinically detected tumors, especially true interval cancers, more frequently showed biological features related to worse prognosis and were associated with poorer survival even after adjustment for clinical-pathological characteristics

# Aggressiveness features and outcomes of true interval cancers: comparison between screen-detected and symptom-detected cancers

Laia Domingo<sup>a,d,e</sup>, Jordi Blanch<sup>a,d</sup>, Sònia Servitja<sup>b</sup>, Josep Maria Corominas<sup>c</sup>, Cristiane Murta-Nascimento<sup>a</sup>, Antonio Rueda<sup>f</sup>, Maximino Redondo<sup>d,g</sup>, Xavier Castells<sup>a,d</sup> and Maria Sala<sup>a,d,e</sup>

The question of whether screen detection confers an additional survival benefit in breast cancer is unclear and subject to several biases. Our aim was to examine the role of the diagnostic method (screen-detected, symptom-detected, and true interval cancers) and the clinical-pathological features in relapse-free survival and overall survival in breast cancer patients. We included 228 invasive breast cancers diagnosed in Barcelona from 1996 to 2008 among women aged 50–69 years. Ninety-seven patients were screen detected within the screening, 34 truly arose between 2-year screening mammograms (true interval cancers), and 97 were symptom detected outside the screening. The clinical-pathological features at diagnosis were compared. The overall and disease-free survival probabilities were computed using the Kaplan–Meier method. Cox proportional hazard models were applied, with adjustment by clinical-pathological variables. At diagnosis, symptom-detected and true interval cancers were in more advanced stages and were less differentiated. The highest proportion of triple-negative cancers was detected among true interval cancers ( $P=0.002$ ). At 5 years of follow-up, the disease-free survival rates for screen-detected, true interval, and symptom-detected cancers were 87.5% (95% confidence interval, 80.5–95.2%), 64.1% (46.4–88.5%), and 79.4% (71.0–88.8%), respectively, and the overall survival rates were 94.5% (89.3–99.9%),

65.5% (47.1–91.2%), and 85.6% (78.3–93.6%), respectively. True interval cancers had the highest hazard ratio for relapse prediction (1.89; 0.67–5.31) and a hazard ratio of death of 5.55 (1.61–19.15) after adjustment for tumor-node-metastasis stage and phenotype. Clinically detected tumors, especially true interval cancers, more frequently showed biological features related to worse prognosis and were associated with poorer survival even after adjustment for clinical-pathological characteristics. *European Journal of Cancer Prevention* 22:21–28 © 2012 Wolters Kluwer Health | Lippincott Williams & Wilkins.

*European Journal of Cancer Prevention* 2013, 22:21–28

**Keywords:** breast cancer screening, interval cancer, phenotype, prognosis, survival

<sup>a</sup>Epidemiology and Evaluation Department, <sup>b</sup>Medical Oncology Department, <sup>c</sup>Pathology Department, Hospital del Mar-IMIM, <sup>d</sup>CIBER Epidemiology and Public Health (CIBERESP), <sup>e</sup>Department of Pediatrics, Obstetrics and Gynecology, EHEA Doctoral Program in Public Health, Preventive Medicine and Public Health, Autonomous University of Barcelona (UAB), Barcelona, <sup>f</sup>Medical Oncology Department and <sup>g</sup>Department of Biochemistry, Hospital Costa del Sol, University of Malaga, Málaga, Spain

Correspondence to Laia Domingo, MSc, Epidemiology and Evaluation Department, Hospital del Mar-IMIM, MAR Health Park, Pg. Maritim 25-29, 08003 Barcelona, Spain  
Tel: +34 933 160 794; fax: +34 932 483 496;  
e-mail: ldomingo@parcdesalutmar.cat

Received 22 February 2012 Accepted 6 April 2012

## Introduction

In the last two decades, breast cancer survival has improved. Reduced mortality from this disease has been attributed to a combination of widespread mammographic screening, the introduction of effective systemic treatment modalities, and progress in radiotherapy and surgery (Berry *et al.*, 2005; Benson *et al.*, 2009; Autier *et al.*, 2010). However, the contribution of screening to this reduction is unclear, as outcomes may be affected by several biases such as lead-time bias (screen-detected cancers are detected earlier in their natural course than those found outside of screening), length bias (screening tends to detect slow-growing tumors, which spend longer time in the asymptomatic phase), and selection bias (the screened population is not representative of the general population).

Previous works that have focused on breast cancer survival and detection mode have reported that screen detection confers an additional survival benefit beyond the stage shift (Shen *et al.*, 2005; Wishart *et al.*, 2008) and reduces the risk of systemic recurrence compared with symptomatic cancers at a similar stage (Joensuu *et al.*, 2004). However, part of this benefit remains unexplained as it is difficult to obtain molecular information from a large series of tumors for inclusion in statistical models (Vitak *et al.*, 1997; Immonen-Räihä *et al.*, 2005; Shen *et al.*, 2005; Bordas *et al.*, 2007; Zackrisson *et al.*, 2007; Dawson *et al.*, 2009; Lawrence *et al.*, 2009).

It is already known that tumors detected during screening are related to clinical-pathological features with a better prognosis, such as low grade or hormone-receptor

expression, compared with those detected by other means (i.e. outside of screening or as interval cancers; Bordas *et al.*, 2007). Wider differences have been described between screen-detected and interval cancers (tumors that manifest clinically between a normal screening result and the following invitation for screening), which are associated with a delay in diagnosis that could potentially worsen prognosis and consequently impair the program's efficiency. Nevertheless, the greatest differences have been found between screen-detected cancers and true interval cancers (tumors with a short preclinical phase and no suspicious finding after radiological review of the previous mammogram). Among this subgroup, increased tumor cell proliferation (Vitak *et al.*, 1997) and a higher occurrence of triple-negative tumors [tumors that lack expression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2)] have been reported (Domingo *et al.*, 2010; Rayson *et al.*, 2011). To date, this biomarker pattern lacks the benefit of specific adjuvant therapy and is associated with a poor prognosis (Chacón and Costanzo, 2010).

Although differences at diagnosis have been reported between screen-detected and true interval cancers, only two studies (Vitak *et al.*, 1997; Rayson *et al.*, 2011) have taken this latter group into consideration when evaluating clinical outcomes, and both have reported significantly poorer survival. However, neither of these studies included information on phenotype and a group of women who did not participate in the screening.

The aim of the present study was to evaluate the role of the diagnostic method and biological characteristics in relapse-free survival and overall survival in breast cancer patients.

## Methods

### Setting

The study was carried out among women diagnosed with breast cancer in a reference hospital in the city of Barcelona (Spain) between 1995 and 2008. This center is publicly funded, serves an area with around 300 000 inhabitants, and has run a population-based breast cancer screening program since 1995.

The screening program's target population (~90 000 women) is women aged 50–69 years, and involves women who are invited to undergo mammography every 2 years. Mediolateral, oblique, and craniocaudal views are available for each breast. All mammograms are read by two radiologists, and when double reading leads to different assessments, a third radiologist serves as a tiebreaker. The program is based on the European Guidelines for Quality Assurance in Mammographic Screening (Perry *et al.*, 2006) and its results meet the Europe Against Cancer standards, although the participation rate barely reached the standard level of 70%.

This study was approved by the Ethics Committee and informed consent was obtained to supply tumor biopsy material for pathologic evaluation.

### Study population

A total of 1432 patients with breast cancer aged 50–69 years were identified from the hospital-based cancer registry during the study period. Of these cancer cases studied, 740 were detected in women attending the breast cancer screening program (screen-detected cancers), 98 emerged as interval cancers between screening mammograms, and 594 cancers were diagnosed among women who did not participate in the screening program (symptom-detected cancers) and who were referred with breast abnormalities, typically palpable lesions, by their primary care physicians or were self-referred to the hospital.

Interval cancers were retrieved by merging the mammography register with the hospital-based cancer registry; in this process, we identified 98 interval cancers diagnosed and treated in our setting. For these, we aimed to perform a radiological review of both screening and diagnostic mammograms to classify them into the five subtypes, following European Guideline recommendations (Perry *et al.*, 2006). Both mammograms were available in 80 cases, and two experienced radiologists performed the retrospective review. Screening mammograms were first reviewed independently, and in a second review, the radiologists assessed screening and diagnostic mammograms with the histological information to determine whether any abnormalities detected on the screening films corresponded to the site of the subsequent interval cancer. Interval breast cancers were definitively classified into five subtypes: true interval cancer ( $n = 34$ ), false negative ( $n = 13$ ), occult tumors ( $n = 14$ ), minimal signs ( $n = 4$ ), and unclassifiable ( $n = 15$ ). A complete description of the interval cancer identification process has been published previously (Domingo *et al.*, 2010).

Because of difficulties in obtaining retrospective information on such a large number of cases, a random sample of screen-detected ( $n = 97$ ) and symptom-detected cancers ( $n = 97$ ) was selected. These two samples were compared with all true interval cancers ( $n = 34$ ). In-situ cancers and patients with stage IV breast cancer at diagnosis were excluded from the analyses.

### Data collection

Tumor-related data [pathological tumor–node–metastasis (TNM) status, histological type, histological grade], patient-related data (age), and vital status at the end of follow-up were obtained from the hospital-based cancer registry. Data on treatment, recurrences, and immunohistochemical information (ER, PR, HER, and p53 status) were obtained from the registry of pathology, if available, and from review of clinical records. Further immunohistochemical analyses were carried out in tumor

samples in which biomarker expression had not been determined previously. Information on adjuvant treatment (radiotherapy, chemotherapy, hormone therapy, and targeted therapies) and follow-up was obtained from the clinical records.

### Laboratory methods

ER, PR, p53, and HER2 are routinely determined during the diagnostic process in our hospital. Samples lacking information on biomarker expression were analyzed following the same procedures as those used in clinical practice. Immunohistochemical staining was performed on paraffin block sections with the tissue specimens fixed in 10% neutral-buffered formalin for 24 h [ER clone ID5, 1:50 (Dako, Glostrup, Denmark); PR clone PgR636, 1:200 (Dako); p53 clone DO7, 1:50 (Novocastra Laboratories, Newcastle, UK); for Her2/neu protein overexpression (HerceptTest; Dako)]. In accordance with standard guidelines, positivity for ER, PR, and p53 was based on more than 10% of the cells testing positive. For HER2, scores of 0–1 in the HerceptTest kit were considered negative, whereas a HerceptTest score of 3 was considered positive. Equivocal scores of 2 were confirmed by fluorescent in-situ hybridization as positive when *HER2/neu* oncogene amplification was detected.

On the basis of the expression of ER, PR, and HER2, the tumors were classified into four phenotypes: (a) luminal A (ER+, PR+, HER-); (b) luminal B (ER or PR+, HER- or ER+, PR+/-, HER+); (c) HER2 (ER-, PR-, HER2+); and (d) triple negative (ER-, PR-, HER-) (Perou *et al.*, 2000; Cheang *et al.*, 2009).

### Follow-up of cancer cases

Locoregional recurrence was defined as disease recurrence within the ipsilateral breast or chest wall, in the ipsilateral axillary nodes, internal mammary nodes, or supraclavicular nodes. Distant recurrence was defined as disease recurrence in sites other than the breast or regional lymph nodes (bone, skin, or visceral metastasis).

Disease-free survival was defined as the time from diagnosis to the first occurrence of one or more of the following: a local or regional recurrence, cancer in the contralateral breast, distant metastasis, and second primary carcinoma, whichever occurred first. Overall survival was defined from the date of diagnosis to death from any cause.

When disease-free survival was computed, women lost to follow-up or those who died were censored either at last visit or at death. For overall survival, patients were censored at the date of their last hospital visit. The median follow-up period was 5.13 years.

### Statistical analysis

Contingency tables were calculated to compare possible differences in patient and histopathological characteristics

among the study groups. Statistical significance was assessed using  $\chi^2$ -tests. Survival curves were generated using the Kaplan–Meier method and were compared by the log-rank test. Kaplan–Meier estimates of 5-year disease-free and overall survival rates after diagnosis were computed with 95% confidence intervals (CI).

Cox proportional hazard regression analyses were carried out to evaluate survival differences between screen-detected, true interval, and symptom-detected cancers, controlling for known prognostic and predictive factors such as age, TNM stage, and phenotype in an attempt to control lead time and length biases related to screening. Unadjusted and adjusted hazard ratios and 95% CI were computed and are shown for our main variable of interest (detection method). We computed a baseline regression model that included only the detection method, and gradually, all other study variables were added to control for their potential effect on survival times. The proportional hazards assumption was ascertained by assessment of log–log survival plots. All calculations were carried out using the statistical software SPSS, version 12.0 (SPSS Inc., Chicago, Illinois, USA) and R, version 2.12.2 (R Development Core Team, 2011). All *P*-values were two-sided and values less than 0.05 were considered statistically significant.

### Results

The analyses included 228 patients with breast cancer, of which 97 were screen-detected, 34 were true interval cancers, and 97 were cancers detected symptomatically outside of the screening program. The clinical and pathological characteristics for the three detection groups are shown in Table 1. At diagnosis, clinically detected tumors (symptom-detected and true interval cancers) were at more advanced stages, larger, more frequently classified as lymph node positive, and poorly differentiated compared with screen-detected cancers. The most frequent phenotype in all three groups was luminal A, but the highest percentage was observed among screen-detected cancers (66.3%). However, the most frequent occurrence of the triple-negative phenotype was found among true interval cancers (28.1%), in contrast to the proportions observed among screen-detected and symptom-detected cancers (3.5 and 10.7%, respectively; *P* = 0.002).

The median follow-up was 6.7 years (range, 0.1–14.0) for women with screen-detected cancer, 3.9 years (range, 0.5–11.7) for women with true interval cancer, and 6.2 years (range, 0.5–15.2) for symptomatic women. Figure 1 shows the disease-free survival curves by detection mode and molecular phenotypes. Figure 2 shows the overall survival curves by these same factors. Disease-free survival and overall survival were worse in true interval cancers than in screen-detected and symptomatic cancers. Disease-free survival was longer in luminal cancers (both A and B) than in HER2 and triple-negative tumors

**Table 1 Clinical-pathological characteristics of screen-detected cancers, true interval cancers, and symptom-detected cancers**

	Screen-detected cancers n=97 (%)	True interval cancers n=34 (%)	Symptom-detected cancers n=97 (%)	P-value <sup>a</sup>
Age group (years)				
50–59	45 (46.4)	18 (52.9)	44 (45.4)	0.741
60–69	52 (53.6)	16 (47.1)	53 (54.6)	
pTNM stage				
I	56 (58.9)	7 (20.6)	34 (37.0)	<0.001
II	34 (35.8)	18 (52.9)	36 (39.1)	
III	5 (5.3)	9 (26.5)	22 (23.9)	
Unknown <sup>b</sup>	2	0	5	
Tumor size (mm)				
<20	69 (72.6)	13 (40.6)	41 (50.6)	0.003
20–50	22 (23.2)	13 (40.6)	31 (38.3)	
>50	4 (4.2)	6 (18.8)	9 (11.1)	
Unknown <sup>b</sup>	2	2	16	
Lymph node involvement				
Negative	71 (74.0)	16 (47.1)	47 (50.5)	0.001
Positive	25 (26.0)	18 (52.9)	46 (49.5)	
Unknown <sup>b</sup>	1	0	4	
Histological grade				
I	45 (47.4)	9 (26.5)	27 (29.7)	0.023
II	36 (37.9)	14 (41.2)	36 (39.6)	
III	14 (14.7)	11 (32.3)	28 (30.8)	
Unknown <sup>b</sup>	2	0	6	
ER status				
Negative	12 (13.0)	13 (39.4)	20 (24.4)	0.005
Positive	80 (87.0)	20 (60.6)	62 (75.6)	
Unknown <sup>b</sup>	5	1	15	
PR status				
Negative	29 (31.5)	21 (63.6)	34 (42.0)	0.005
Positive	63 (68.5)	12 (36.4)	47 (58.0)	
Unknown <sup>b</sup>	5	1	16	
HER2 status				
Negative	76 (87.4)	28 (87.5)	56 (71.8)	0.023
Positive	11 (12.6)	4 (12.5)	22 (28.2)	
Unknown <sup>b</sup>	10	2	19	
p53 status				
Negative	75 (84.3)	21 (65.6)	59 (76.6)	0.081
Positive	14 (15.7)	11 (34.4)	18 (23.4)	
Unknown <sup>b</sup>	8	2	20	
Phenotype				
Luminal type A	57 (66.3)	12 (37.5)	34 (45.3)	0.002
Luminal type B	19 (22.1)	7 (21.9)	23 (30.7)	
HER2	7 (8.1)	4 (12.5)	10 (13.3)	
Triple negative	3 (3.5)	9 (28.1)	8 (10.7)	
Unknown <sup>b</sup>	11	2	22	
Systemic treatment				
Chemotherapy alone	13 (13.4)	14 (41.2)	20 (20.6)	0.006
Hormonal therapy alone	37 (38.1)	4 (11.8)	29 (29.9)	
Chemotherapy and hormonal therapy and/or trastuzumab <sup>c</sup>	36 (37.1)	15 (44.1)	41 (42.3)	
Local treatment only <sup>d</sup>	11 (11.3)	1 (2.9)	7 (7.2)	
Relapses				
No relapses	85 (87.6)	26 (76.5)	75 (77.3)	0.067
Local relapses	2 (2.1)	0 (0)	1 (1.0)	
Distant relapses	6 (6.2)	8 (23.5)	19 (19.6)	
Second malignances	4 (4.1)	0 (0)	2 (2.1)	
Death				
No	89 (91.8)	26 (76.5)	74 (75.5)	0.009
Yes	8 (8.2)	8 (23.5)	23 (23.7)	

ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor; TNM, tumor–node–metastasis.

<sup>a</sup>Screen-detected, true interval, and symptom-detected cancers are compared.  $\chi^2$  two-sided test.

<sup>b</sup>Tumors with missing information were excluded for the calculation of percentage.

<sup>c</sup>Trastuzumab for HER2-positive breast cancer since 2006.

<sup>d</sup>Includes only local therapies (such as surgery and/or radiotherapy).

(log-rank test = 12.1;  $P = 0.007$ ). Kaplan–Meier estimates of the 5-year disease-free survival rates after diagnosis for screen-detected, true interval, and symptom-detected cancers were 87.5% (95% CI, 80.5–95.2%), 64.1% (95% CI, 46.4–88.5%), and 79.4% (95% CI, 71.0–88.8%), respectively. Kaplan–Meier estimates of the 5-year overall survival rates after diagnosis for screen-detected, true interval, and symptom-detected cancers were 94.5% (95% CI, 89.3–99.9%), 65.5% (95% CI, 47.1–91.2%), and 85.6% (95% CI, 78.3–93.6%), respectively.

Estimates of the Cox regression models to evaluate the determinants associated with relapse-free survival and overall survival are presented in Tables 2 and 3, respectively. The unadjusted model showed that true interval cancers and symptom-detected cancers had a higher risk of relapse than screen-detected cancers, the estimated hazard ratios being statistically significant for true interval cancers. However, in the adjusted model, the prognostic effect of the detection method was attenuated, especially when biological factors were included in the multivariate model. The hazard risk of death was 5.02 (95% CI, 1.87–13.48) times higher among true interval cancers, and was 2.78 (95% CI, 1.24–6.24) times higher among symptom-detected cancers, compared with screen-detected cancers. Adjustment for TNM stage attenuated both values, which nevertheless remained statistically significant. Finally, in the model including phenotype, only true interval cancers were associated with a worse prognosis.

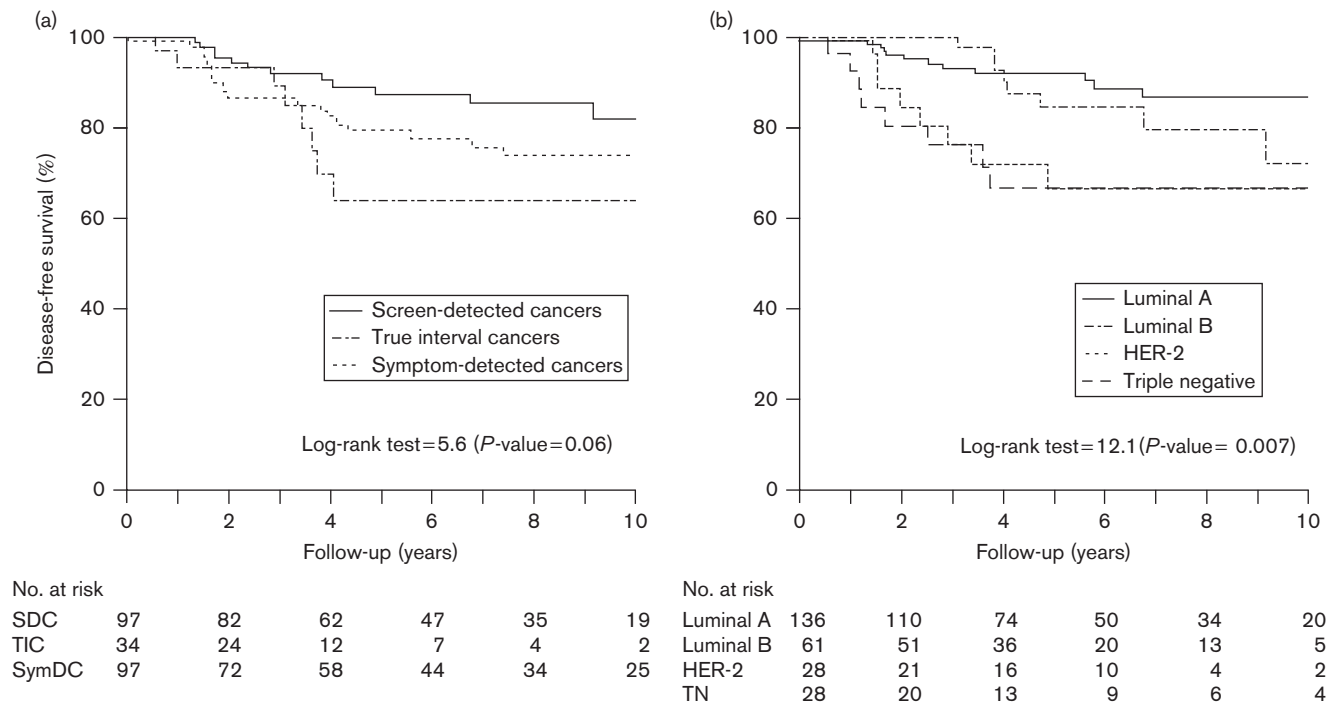
## Discussion

Our results suggest that tumors detected clinically, especially true interval cancers, encompass a subgroup of tumors with different biological characteristics leading to poorer prognosis and worse outcomes than in screen-detected cancers. When adjustment for age, TNM stage, and phenotype was carried out, the detection mode remained as an independent factor for overall survival.

As expected, the pathological characteristics of screen-detected tumors were related to a better prognosis. At diagnosis, screen-detected cancers were smaller, more frequently lymph node-negative, and of lower grade than clinically detected tumors. These findings are in agreement with several publications (Vitak *et al.*, 1997; Joensuu *et al.*, 2004; Collett *et al.*, 2005; Shen *et al.*, 2005; Bordas *et al.*, 2007; Pálka *et al.*, 2008; Chiarelli *et al.*, 2012). Among clinically detected tumors, 18.8% of true interval cancers were larger than 50 mm, whereas this percentage was 11.1% for symptom-detected cancers. This finding is especially important, given that true interval cancers have a short preclinical phase (sojourn time) and truly arise in less than 2 years. ER and PR positivity was more frequent in screen-detected than in clinically detected cancers, as observed in other studies (Sihto *et al.*, 2008; Dawson *et al.*, 2009; Mook *et al.*, 2011; Nagtegaal *et al.*, 2011), but not in



Fig. 1



Kaplan–Meier curves comparing disease-free rates by detection mode and phenotypes. Kaplan–Meier curves for disease-free survival by detection method (a) and phenotypic classification (b) of breast tumors. The number of women at risk every 2 years is presented on the time axis. HER2, human epidermal growth factor receptor 2; SDC, screen-detected cancers; SymDC, symptom-detected cancers; TIC, true interval cancers; TN, triple negative.

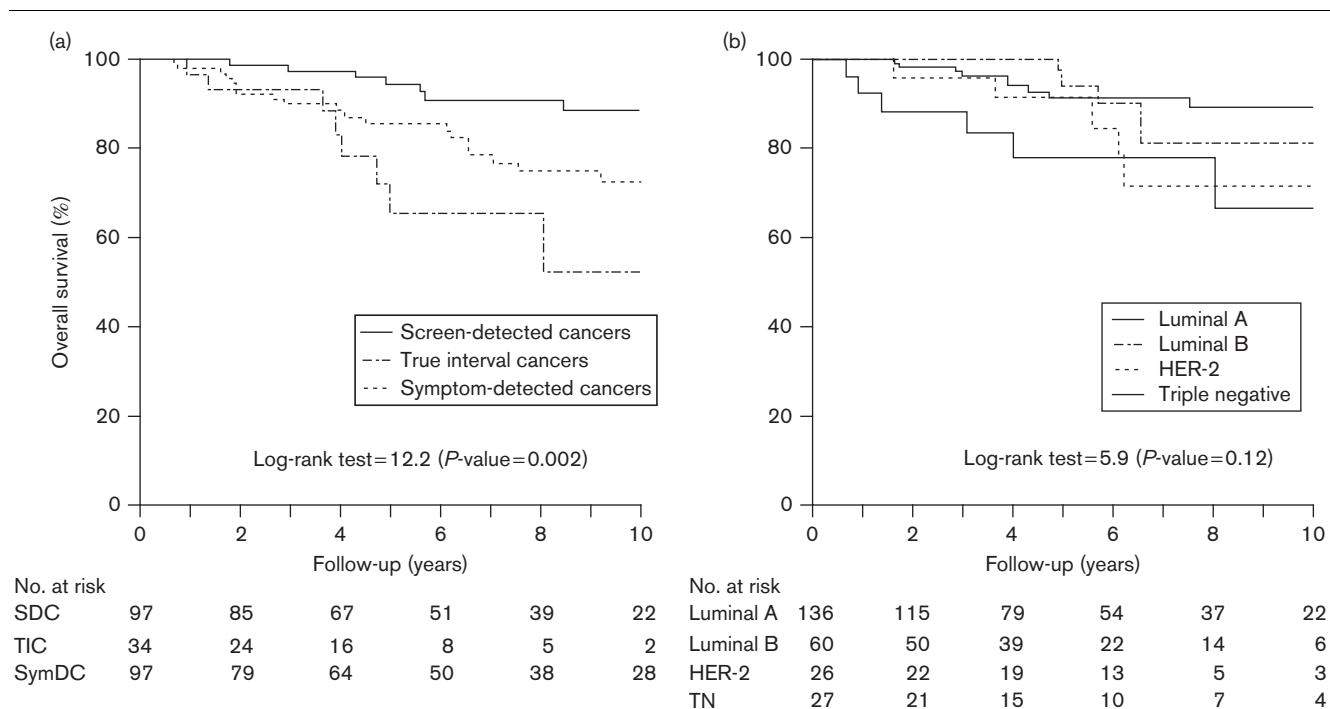
all (Joensuu *et al.*, 2004). In agreement with previous series (Domingo *et al.*, 2010; Van der Vegt *et al.*, 2010; Rayson *et al.*, 2011), we found the highest percentage of triple-negative tumors among true interval cancers, with substantial differences when compared with both screen-detected and symptom-detected cancers.

These results, together with the findings of other researchers that relate true interval cancers with tumors with high mitotic capacity and a high cell-proliferation phase (Ki-67 and S-phase fraction; Vitak *et al.*, 1997; Crosier *et al.*, 1999; Kirsh *et al.*, 2011), support the hypothesis that true interval cancers constitute a subgroup of breast cancers with rapid growth and high aggressiveness that are less likely to be detected on screening mammograms. These features may determine the probability of detection by screening programs; tumors with a shorter sojourn time (which is the time taken by tumors to grow from a mammographically detectable size to a clinically detectable size) are more likely to be detected between screening intervals (Weedon-Fekjaer *et al.*, 2005). This is in part a reflection of the length bias associated with screening practices, in agreement with the idea that tumors detected on routine screening are not simply tumors diagnosed earlier, before becoming symptomatic, but also show biological differences. However, a percentage

of screen-detected cancers also show features of worse prognosis (41.1% in stages II–III, 26.3% lymph node positive, and 3.5% presenting a triple-negative phenotype) that could benefit from screening.

Several studies have shown that screen detection remains an independent prognostic factor after adjustment for disease stage (Joensuu *et al.*, 2004; Shen *et al.*, 2005; Wishart *et al.*, 2008) and some biological characteristics (Gill *et al.*, 2004; Sihto *et al.*, 2008; Dawson *et al.*, 2009). Most of these studies, which do not differentiate interval cancers, recommend taking the detection method into account when estimating individual prognosis (Joensuu *et al.*, 2004; Mook *et al.*, 2011). In agreement with previous series, our results showed better survival outcomes among tumors detected by mammography than among those detected by other means (Vitak *et al.*, 1997; Joensuu *et al.*, 2004; Bordas *et al.*, 2007; Zackrisson *et al.*, 2007). Nonetheless, most of the studies considered all interval cancers together, which may have attenuated the worse outcome in the true interval cancer subgroup. Studies revealed that true interval cancers (Vitak *et al.*, 1997; Van der Vegt *et al.*, 2010; Rayson *et al.*, 2011) showed a trend toward decreased relapse-free and overall survival when compared with cancers detected by other means. However, only Vitak *et al.* (1997) compared true interval with both screen-detected

Fig. 2



Kaplan–Meier curves comparing overall survival rates by detection mode and phenotypes. Kaplan–Meier curves for overall survival by detection method (a) and phenotypic classification (b) of breast tumors. The number of women at risk every 2 years is presented on the time axis. HER2, human epidermal growth factor receptor 2; SDC, screen-detected cancers; SymDC, symptom-detected cancers; TIC, true interval cancers.

Table 2 Comparison of disease-free survival regression models, including the detection method, adjusted by different factors

Model variables	Categories	Number		HR adj (95% CI)
		Patients	Events	
Detection method	SDC	97	12	Ref.
	TIC	34	8	2.75 (1.12–6.77)
	SymDC	97	22	1.85 (0.91–3.76)
Detection method + age + TNM stage	SDC	95	12	Ref.
	TIC	34	8	2.27 (0.90–5.72)
	SymDC	92	22	1.84 (0.89–3.82)
Detection method + age + TNM stage + phenotype	SDC	84	11	Ref.
	TIC	32	8	1.89 (0.67–5.31)
	SymDC	72	12	0.91 (0.37–2.27)
Detection method + age + TNM stage + phenotype + treatment	SDC	84	11	Ref.
	TIC	32	8	1.95 (0.66–5.79)
	SymDC	72	12	0.93 (0.37–2.35)

CI, confidence interval; HR adj, adjusted hazard ratio; ref., reference; SDC, screen-detected cancers; SymDC, symptom-detected cancers; TIC, true interval cancers; TNM, tumor–node–metastasis.

Table 3 Comparison of overall survival regression models, including the detection method, adjusted by different factors

Model variables	Categories	Number		HR adj (95% CI)
		Patients	Events	
Detection method	SDC	97	8	Ref.
	TIC	34	8	5.02 (1.87–13.48)
	SymDC	97	23	2.78 (1.24–6.24)
Detection method + age + TNM stage	SDC	95	8	Ref.
	TIC	34	8	3.87 (1.40–10.75)
	SymDC	92	23	2.73 (1.20–6.21)
Detection method + age + TNM stage + phenotype	SDC	84	6	Ref.
	TIC	32	8	5.55 (1.61–19.15)
	SymDC	72	12	2.14 (0.72–6.30)
Detection method + age + TNM stage + phenotype + treatment	SDC	84	6	Ref.
	TIC	32	8	7.67 (2.07–28.50)
	SymDC	72	12	2.44 (0.79–7.54)

CI, confidence interval; HR adj, adjusted hazard ratio; ref., reference; SDC, screen-detected cancers; SymDC, symptom-detected cancers; TIC, true interval cancers; TNM, tumor–node–metastasis.

and symptom-detected cancers, whereas the remaining studies did not include this last group.

Unadjusted Cox models for disease-free survival showed the method of detection as an independent factor for relapse prediction, consistent with previous series including true interval cancers (Vitak *et al.*, 1997; Rayson *et al.*, 2011). However, when tumor size and biological characteristics were included in the multivariate analyses

to adjust for lead and length bias, as expected, the prognostic value of the detection mode lost its effect for recurrences. Nevertheless, for overall survival, our trends suggest that true interval cancers trigger a subgroup of breast cancers with an independent unfavorable prognostic significance beyond that explained by the conventional factors included in the present study. If confirmed, this finding may merit further investigation into its underlying biological mechanisms. Recently, some

genetic and epigenetic mechanisms have been related to interval malignancies, such as the methylation process of specific genes (Suijkerbuijk *et al.*, 2011) or the action of growth factors produced in the breast stroma in response to tumor aggressiveness (Li *et al.*, 2005). Further efforts should be made in the future to improve understanding of the genetic mechanisms related to cancer aggressiveness, tumor cell proliferation, and the process of carcinogenesis to improve the early detection of fast-growing tumors.

This study has some limitations, the main one being the small sample size. Given the difficulties in identifying and classifying interval cancers, to date, few studies have focused on true interval cancers. However, the current study adds to previous analyses limited by incomplete pathological data or a failure to restrict the analyses to true interval cancers. Second, misclassification of the detection mode cannot be excluded. Some interval cancers could be classified as screen-detected if symptomatic women waited for the screening visit instead of making an appointment with a physician. This misclassification would attenuate the effect studied. Thus, the survival difference in favor of screened women might be greater than that observed because of the inclusion of some women with symptomatic cancers in the screened cohort. In addition, the sociodemographic characteristics of women participating in the screening program, whose interval cancer was diagnosed and treated in other hospitals, did not differ from those of women treated in our setting. However, as participation in screening practices may be affected by several selection factors, we considered a group of women not affected by screening in the analyses.

This study has some strengths. To our knowledge, this is the first work focused on true interval cancer characterization that analyzes prognosis by considering cancers detected both inside and outside screening, in addition to the molecular profile. This design allows us to control the main biases that affect the outcomes of the screening practices. Moreover, all analyses were carried out with information on phenotype and other biological markers from recently diagnosed patients (1996–2008) who received homogeneous adjuvant treatment on the basis of the oncology protocol of our institution: chemotherapy based on anthracyclines or anthracyclines and taxanes, hormone therapies (tamoxifen or aromatase inhibitors), and targeted therapies (trastuzumab has been prescribed to patients with HER2 overexpression since 2006) in the modern era.

## Conclusion

These results suggest that screening programs detect a major proportion of tumors with more favorable biological characteristics, which could partly explain the better survival of patients with screening-detected tumors.

Moreover, our results show that true interval cancers share a subset of features of worse prognosis, related both to their growth rate and short sojourn time, and to their clinical course. Further understanding of their biologic features and individual determinants could increase the benefits of screening: in the short term, by increasing the sensitivity of the programs and aiding the choice of optimal screening interval for specific subsets of women at high risk, and in the long term, by increasing the early detection of tumors with a less favorable natural history.

## Acknowledgements

The authors acknowledge the contribution of the Hospital del Mar Tumor Registry (Barcelona) in providing tumor-related data. They also thank Cristina Hernández and Marta Román for their assistance in data management and Teresa Baró for her technical assistance in performing the immunohistochemical analysis.

This study was partially supported by CIBER de Epidemiología y Salud Pública (AE08\_004).

## Conflicts of interest

There are no conflicts of interest.

## References

- Autier P, Boniol M, La Vecchia C, Vatten L, Gavin A, Héry C, *et al.* (2010). Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database. *BMJ* **341**:1–7.
- Benson JR, Jatoi I, Keisch M, Esteva FJ, Makris A, Jordan VC (2009). Early breast cancer. *Lancet* **373**:1463–1479.
- Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, *et al.* (2005). Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* **353**:1784–1792.
- Bordas P, Jonsson H, Nystrom L, Lenner P (2007). Survival from invasive breast cancer among interval cases in the mammography screening programmes of northern Sweden. *Breast* **16**:47–54.
- Chacón RD, Costanzo MV (2010). Triple-negative breast cancer. *Breast Cancer Res* **12** (Suppl 2):1–9.
- Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, *et al.* (2009). Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst* **101**:736–750.
- Chiarelli AM, Edwards SA, Sheppard AJ, Mirea L, Chong N, Paszat L, *et al.* (2012). Favourable prognostic factors of subsequent screen-detected breast cancers among women aged 50–69. *Eur J Cancer Prev* **23**:1–8.
- Collett K, Stefansson IM, Eide J, Braaten A, Wang H, Eide GE, *et al.* (2005). A basal epithelial phenotype is more frequent in interval breast cancers compared with screen detected tumors. *Cancer Epidemiol Biomarkers Prev* **14**:1108–1112.
- Crosier M, Scott D, Wilson RG, Griffiths CD, May FE, Westley BR (1999). Differences in Ki67 and c-erbB2 expression between screen-detected and true interval breast cancers. *Clin Cancer Res* **5**:2682–2688.
- Dawson SJ, Duffy SW, Blows FM, Driver KE, Provenzano E, LeQuesne J, *et al.* (2009). Molecular characteristics of screen-detected vs. symptomatic breast cancers and their impact on survival. *Br J Cancer* **101**:1338–1344.
- Domingo L, Sala M, Servitja S, Corominas JM, Ferrer F, Martínez J, *et al.* (2010). Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain. *Cancer Causes Control* **21**:1155–1164.
- Gill PG, Farshid G, Luke CG, Roder DM (2004). Detection by screening mammography is a powerful independent predictor of survival in women diagnosed with breast cancer. *Breast* **13**:15–22.
- Immonen-Räihä P, Kauhava L, Parvinen I, Holli K, Kronqvist P, Pylkkänen L, *et al.* (2005). Mammographic screening reduces risk of breast carcinoma recurrence. *Cancer* **103**:474–482.

- Joensuu H, Lehtimäki T, Holli K, Elomaa L, Turpeenniemi-Hujanen T, Kataja V, *et al.* (2004). Risk for distant recurrence of breast cancer detected by mammography screening or other methods. *JAMA* **292**:1064–1073.
- Kirsh VA, Chiarelli AM, Edwards SA, O'Malley FP, Shumak RS, Yaffe MJ, *et al.* (2011). Tumor characteristics associated with mammographic detection of breast cancer in the Ontario Breast Screening Program. *J Natl Cancer Inst* **103**:942–950.
- Lawrence G, Wallis M, Allgood P, Nagtegaal ID, Warwick J, Cafferty FH, *et al.* (2009). Population estimates of survival in women with screen-detected and symptomatic breast cancer taking account of lead time and length bias. *Breast Cancer Res Treat* **116**:179–185.
- Li T, Sun L, Miller N, Nicklee T, Woo J, Hulse-Smith L, *et al.* (2005). The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. *Cancer Epidemiol Biomarkers Prev* **14**:343–349.
- Mook S, van't Veer LJ, Rutgers EJ, Ravdin PM, van de Velde AO, van Leeuwen FE, *et al.* (2011). Independent prognostic value of screen detection in invasive breast cancer. *J Natl Cancer Inst* **103**:585–597.
- Nagtegaal ID, Allgood PC, Duffy SW, Kearins O, Sullivan EO, Tappenden N, *et al.* (2011). Prognosis and pathology of screen-detected carcinomas: how different are they? *Cancer* **117**:1360–1368.
- Pálka I, Kelemen G, Ormandi K, Lazar G, Nyari T, Thurzo L, *et al.* (2008). Tumor characteristics in screen-detected and symptomatic breast cancers. *Pathol Oncol Res* **14**:161–167.
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, *et al.* (2000). Molecular portraits of human breast tumours. *Nature* **406**:747–752.
- Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (2006). *European guidelines for quality assurance in breast cancer screening and diagnosis*. 4th ed. Luxembourg: Office for Official Publications of the European Communities.
- Rayson D, Payne JI, Abdolell M, Barnes PJ, MacIntosh RF, Foley T, *et al.* (2011). Comparison of clinical-pathologic characteristics and outcomes of true interval and screen-detected invasive breast cancer among participants of a Canadian breast screening program: a nested case-control study. *Clin Breast Cancer* **11**:27–32.
- R Development Core Team (2011). *R: A language and environment for statistical computing*. ISBN 3-900051-07-0. Vienna, Austria: R Foundation for Statistical Computing. Available at: <http://www.R-project.org/>.
- Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA (2005). Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. *J Natl Cancer Inst* **97**:1195–1203.
- Sihto H, Lundin J, Lehtimäki T, Sarlomo-Rikala M, Bützow R, Holli K, *et al.* (2008). Molecular subtypes of breast cancers detected in mammography screening and outside of screening. *Clin Cancer Res* **14**:4103–4110.
- Suijkerbuijk KP, van Diest PJ, van der WE (2011). Improving early breast cancer detection: focus on methylation. *Ann Oncol* **22**:24–29.
- Van der Vegt B, Wesseling J, Pijnappel RM, Dorrius MD, den Heeten GJ, de Roos MA, *et al.* (2010). Aggressiveness of 'true' interval invasive ductal carcinomas of the breast in postmenopausal women. *Mod Pathol* **23**: 629–636.
- Vitak B, Stal O, Manson JC, Thomas BA, Arnesson LG, Ekelund L, *et al.* (1997). Interval cancers and cancers in non-attenders in the Ostergotland Mammographic Screening Programme. Duration between screening and diagnosis, S-phase fraction and distant recurrence. *Eur J Cancer* **33**: 1453–1460.
- Weedon-Fekjaer H, Vatten LJ, Aalen OO, Lindqvist B, Tretli S (2005). Estimating mean sojourn time and screening test sensitivity in breast cancer mammography screening: new results. *J Med Screen* **12**:172–178.
- Wishart GC, Greenberg DC, Britton PD, Chou P, Brown CH, Purushotham AD, *et al.* (2008). Screen-detected vs. symptomatic breast cancer: Is improved survival due to stage migration alone? *Br J Cancer* **98**:1741–1744.
- Zackrisson S, Janzon L, Manjer J, Andersson I (2007). Improved survival rate for women with interval breast cancer – results from the breast cancer screening programme in Malmö, Sweden 1976–1999. *J Med Screen* **14**:138–143.

#### 4. Article 4

**Títol:** Differences in radiological patterns, tumour characteristics and diagnostic precision between digital mammography and screen-film mammography in four breast cancer screening programmes in Spain.

**Autors:** Domingo L, Romero A, Belvis F, Sánchez M, Ferrer J, Salas D, Ibáñez J, Vega A, Ferrer F, Laso MS, Macià F, Castells X, Sala M.

**Revista:** Eur Radiol. 2011 Sep;21(9):2020-8.

**Factor d'impacte** (2010): 3.594 (Q1 Radiologia, medicina nuclear i imatge mèdica)

**DOI:** 10.1007/s00330-011-2143-1

#### **Resum:**

**OBJECTIVES:** To compare tumour characteristics between cancers detected with screen-film mammography (SFM) and digital mammography (DM) and to evaluate changes in positive predictive values (PPVs) for further assessments, for invasive procedures and for distinct radiological patterns in recalled women.

**METHODS:** 242,838 screening mammograms (171,191 SFM and 71,647 DM) from 103,613 women aged 45-69 years, performed in four population-based breast cancer screening programmes in Spain, were included. The tumour characteristics and PPVs of each group were compared. Radiological patterns (masses, calcifications, distortions and asymmetries) among recalled women were described and PPVs were evaluated.

**RESULTS:** The percentages of ductal carcinoma *in situ* (DCIS) were higher in DM than in SFM both in the first [18.5% vs. 15.8%(p=0.580)] and in successive screenings [23.2% vs. 15.7%(p=0.115)]. PPVs for masses, asymmetries and calcifications were higher in DM, being statistically significant in masses (5.3% vs. 3.9%; proportion ratio: 1.37 95%CI: 1.08-1.72). Among cancers detected by calcifications, the percentage of DCIS was higher in DM (60.3% vs. 46.4%, p=0.060).

**CONCLUSIONS:** PPVs were higher when DM was used, both for further assessments and for invasive procedures, with similar cancer detection rates and no statistically significant differences in tumour characteristics. The greatest improvements in PPVs were found for masses



# Differences in radiological patterns, tumour characteristics and diagnostic precision between digital mammography and screen-film mammography in four breast cancer screening programmes in Spain

Laia Domingo · Anabel Romero · Francesc Belvis · Mar Sánchez · Joana Ferrer · Dolores Salas · Josefa Ibáñez · Alfonso Vega · Francesc Ferrer · M. Soledad Laso · Francesc Macià · Xavier Castells · Maria Sala

Received: 28 January 2011 / Revised: 16 March 2011 / Accepted: 27 March 2011 / Published online: 11 May 2011  
© European Society of Radiology 2011

## Abstract

**Objectives** To compare tumour characteristics between cancers detected with screen-film mammography (SFM) and digital mammography (DM) and to evaluate changes in positive predictive values (PPVs) for further assessments, for invasive procedures and for distinct radiological patterns in recalled women.

**Methods** 242,838 screening mammograms (171,191 SFM and 71,647 DM) from 103,613 women aged 45–69 years, performed in four population-based breast cancer screening programmes in Spain, were included. The tumour characteristics and PPVs of each group were compared. Radiological patterns (masses, calcifications, distortions and asymmetries) among recalled women were described and PPVs were evaluated.

**Results** The percentages of ductal carcinoma in situ (DCIS) were higher in DM than in SFM both in the first [18.5% vs.

15.8% ( $p=0.580$ )] and in successive screenings [23.2% vs. 15.7% ( $p=0.115$ )]. PPVs for masses, asymmetries and calcifications were higher in DM, being statistically significant in masses (5.3% vs. 3.9%; proportion ratio: 1.37 95%CI: 1.08–1.72). Among cancers detected by calcifications, the percentage of DCIS was higher in DM (60.3% vs. 46.4%,  $p=0.060$ ).

**Conclusions** PPVs were higher when DM was used, both for further assessments and for invasive procedures, with similar cancer detection rates and no statistically significant differences in tumour characteristics. The greatest improvements in PPVs were found for masses.

**Keywords** Digital mammography · Screen-film mammography · Screening · Breast cancer · Breast imaging · Ductal carcinoma in situ

L. Domingo · A. Romero · F. Belvis · F. Macià · X. Castells · M. Sala (✉)  
Department of Epidemiology and Evaluation,  
IMIM-Hospital del Mar, Passeig Marítim 25-29,  
08003 Barcelona, Spain  
e-mail: Msalaserra@parcdesalutmar.cat

L. Domingo  
e-mail: Ldomingo@parcdesalutmar.cat

A. Romero  
e-mail: Aromeromoraes@parcdesalutmar.cat

F. Belvis  
e-mail: fbelvis@terrassa.uned.es

F. Macià  
e-mail: Fmacia@parcdesalutmar.cat

X. Castells  
e-mail: Xcastells@parcdesalutmar.cat

L. Domingo · A. Romero · F. Belvis · F. Macià · X. Castells · M. Sala  
CIBER de Epidemiología y Salud Pública (CIBERESP),  
Barcelona, Spain

L. Domingo · M. Sala  
EHEA Doctoral Program in Public Health. Department of  
Pediatrics, Obstetrics and Gynecology, Preventive Medicine and  
Public Health, Universitat Autònoma de Barcelona (UAB),  
Bellaterra,  
Barcelona, Spain

M. Sánchez  
General Directorate of Public Health. Department of Health,  
Government of Cantabria,  
C/ Federico Vial, 13,  
39009 Santander, Spain  
e-mail: sanchez\_mm@gobcantabria.es

## Introduction

Full-field digital mammography (DM) systems have been commercially available since 2000. This new technology is widely used in breast screening programmes given its many technical advantages, such as image quality and dose optimisation, archiving and image transfer capability, processing after image acquisition, immediate digital magnification, reduced radiation exposure, and computer-assisted detection [1].

Trials comparing the two techniques have reported that detection rates in DM are at least as high as those for screen-film mammography (SFM) [2]. Some of these studies have shown higher accuracy in pre- or perimenopausal women, young women (<50 years old) and women with dense breast tissue [3]. However, some studies have reported an overall higher cancer detection rate with a higher recall rate [4] with DM than with SFM while others have described lower recall rates and false-positive results with DM [5–8] leading to differences in the positive predictive values (PPV) of breast cancer among recalled women. Some differences in the proportion of ductal carcinoma in situ (DCIS) and invasive cancer have also been reported [9, 10]. One study reported a high detection rate for cancers depicted as calcifications with DM, most of which were ductal carcinomas in situ (DCIS) [11], while another suggested that DM has the potential to increase the rate of invasive cancers detected on the basis of isolated calcifications [12]. Hambly et al. reported that the detection rate of DCIS due to calcifications was significantly higher for DM, and the invasive cancers rate detected due to architectural distortion was also higher in DM [13].

However, few studies are available with somehow contradictory results concerning differences in the radiological findings for women recalled for diagnostic work-up in SFM versus DM in population-based screening pro-

grammes. Although the available information suggests that differences, if any, are few, complete evaluation of DM in breast cancer screening programmes is required to allow the balance of risks and benefits of these programmes to be accurately evaluated at a time when the value of screening is being questioned [14].

Our objective was to compare tumour characteristics (proportion of invasive cancers or DCIS, size, and lymph node status) between cancers detected with SFM and DM and to evaluate changes in diagnostic precision for all further assessments (whether invasive or non-invasive), for invasive procedures and for distinct radiological patterns in recalled women by comparing positive predictive values (PPVs) in SFM and DM.

## Methods

### Population and methods

The project enrolled five radiology units from four different population-based breast cancer screening programmes in Spain (Cantabria, Barcelona, Girona and Valencia) covering a population of 1,300,000 inhabitants. All the programmes are based on the European guidelines for quality assurance in mammography screening [15] and their results meet the required standards.

The study was approved by the ethics committee and informed consent was not required. The selection criterion for inclusion in the screening programmes was having completed at least one screening round performed with DM by December 2007. Women in the target population received information on screening and were invited to undergo mammography with a 2-year interval between screening rounds. Screening age was 50 to 69 years in three programmes and 45 to 69 years in one programme. All radiology units began screening activities between 1996

J. Ferrer  
Radiology Unit, Hospital Santa Caterina,  
C/ Doctor Castany s/n,  
17190 Girona, Spain  
e-mail: joana\_ferrer@ias.scs.es

D. Salas · J. Ibáñez  
General Directorate Public Health & Centre for Public Health  
Research (CSISP),  
Avd. Catalunya, 21,  
46002 Valencia, Spain

D. Salas  
e-mail: salas\_dol@gva.es

J. Ibáñez  
e-mail: ibanyez\_jos@gva.es

A. Vega  
Radiology Unit, Hospital Universitario Marqués de Valdecilla,  
Avda. Valdecilla, 25,  
39008 Santander, Spain  
e-mail: avegab@telefonica.net

F. Ferrer  
Radiology and Nuclear Medicine Service, Hospital del Mar,  
Passeig Marítim,  
08003 Barcelona, Spain  
e-mail: fferrer@parcdesalutmar.cat

M. S. Laso  
Breast Cancer Screening Unit Burjassot, Valencia (Spain) & Centre  
for Public Health Research (CSISP),  
C/ Sixto Cámara, 12. 46100 Burjassot,  
Valencia, Spain  
e-mail: laso\_sol@gva.es



and 1998 using screen-film radiographic technology and switched to full-field digital mammography between September 2004 (one programme) and January 2005 (three programmes).

More detailed information on screening performance in the programmes involved in the study and on database construction is described in a previous work [16].

The final database covered information on the women, on the screening technique (screen-film or digital mammography), information on further assessments after a positive screening mammogram (non-invasive or invasive procedures, and early recalls), and information on the characteristics of the tumour (DCIS or invasive cancer, tumour size and lymph node status, according to the TNM classification). A definitive diagnosis of breast cancer was always histopathologically confirmed. Moreover, the mammographic findings of recalled women were collected and classified as 1) tumour-like mass, 2) parenchymal asymmetry, 3) calcification, 4) parenchymal distortion, 5) other patterns, and 6) multiple patterns. To assess the PPV of each mammographic pattern, multiple patterns were excluded from the analysis.

#### Study population

This study included women participating in at least one screening in any of the four screening programmes from the beginning of each programme to December 2007. Mammograms performed with DM during the first 3 months after the switch in technology, considered a learning period, were excluded and only soft copy reading mammograms were included in the analysis. A total of 242,838 screening mammograms from 103,613 women were included in the analysis, of which 171,191 were SFM and 71,647 were DM.

#### Screening results

Three possible outcomes of a screening test were considered: normal findings (for which a follow-up at 2 years is recommended), a positive result (abnormal findings requiring recall for further assessment to rule out malignancy), and early recall (i.e., requiring an intermediate mammogram before the normal 2-year interval). A positive result was considered a true-positive result if, after further assessments, breast cancer was found (DCIS or invasive cancer). Otherwise, the result was considered a false-positive. Early recalls were not considered false-positives if they did not involve further procedures.

Further assessments could include both noninvasive (magnetic resonance imaging, ultrasonography, additional mammography) and invasive procedures (fine-needle aspiration cytology, core-needle biopsy and open biopsy).

#### Statistical analysis

Cancer detection was defined as the percentage of screened women with a true-positive result, i.e. a final diagnosis of DCIS or invasive cancer. The PPV was the proportion of patients with positive results who were correctly diagnosed. PPV was computed in DM and SFM for women whom at least one further assessment was performed (whether invasive or non-invasive) ( $PPV_1$ ) and for women recalled for invasive procedures ( $PPV_2$ ). PPVs were compared using a formal test for comparison of proportions.

Cancer detection rates and tumour characteristics were described for the SFM and DM groups by using simple proportions and were compared using Pearson's chi-square test.

The evolution of cancer detection rates (for overall cancers, DCIS and invasive cancers) was represented using a combined time/technique variable, which allowed us to exclude potential confounding due to time trends. We divided the screening history in each radiology unit in six consecutive time intervals. The SFM period was divided in four equal time intervals (quartiles) and, similarly, the DM period was divided into two intervals (median). In each radiology unit, the first digital period was preceded by four screen-film periods.

PPVs were also calculated according to radiological findings, i.e. the number of cancers detected in each radiological pattern in both SFM and DM. The PPVs in each radiological pattern were compared in SFM and DM by the proportional ratio (PR).

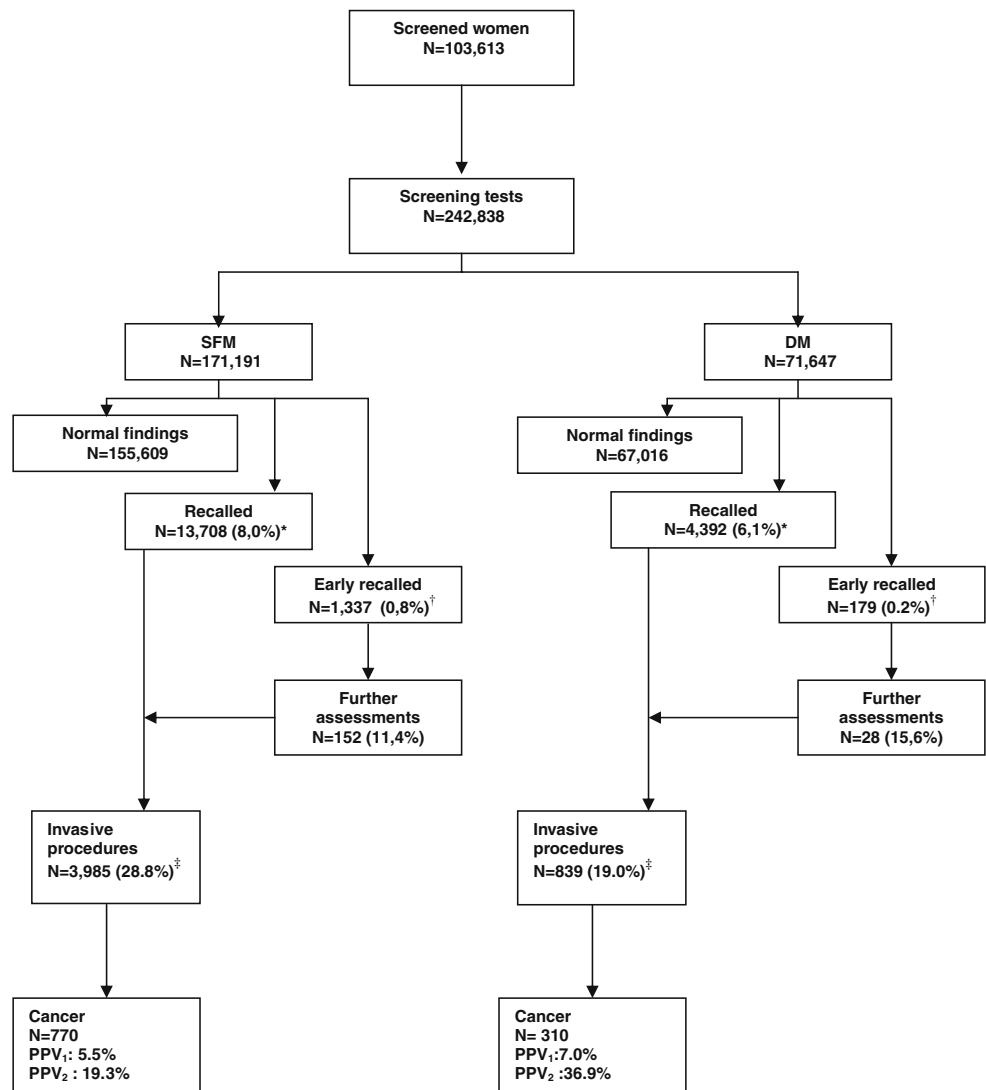
The percentages of invasive cancers and DCIS for each radiological pattern were compared between the two techniques. For these latter two analyses, we only included information of screening mammograms requiring further assessments. Multiple radiological patterns were excluded in an attempt to establish direct links between suspicious patterns and cancer diagnoses. PPVs were compared by estimating the PR.

Statistical analyses were performed using the SPSS (version 12.0) and R statistical software programmes. All p values <0.05 were considered significant.

#### Results

During the study period, 242,838 screening mammograms were performed in 103,613 women: 171,191 with SFM and 71,647 with DM. Figure 1 summarizes the screening mammograms performed, the number of women who were recalled for further assessments, and the invasive procedures performed in the SFM and DM groups. The number of women with an early recall who required an intermediate mammogram, and the number of cancers detected in each group are also shown. The percentage of recalled and early-recalled women was lower in the DM group. For recalled women, the percentages of invasive procedures were lower in the DM group. As a result,

**Fig. 1** Flowchart of the screening tests performed in Screen-Film Mammography (SFM) and Digital Mammography (DM), number of recalled or early recalled women, invasive procedures performed and cancers detected at each group



\*Percentage of recalled women for further assessments among the total of screening tests

†Percentage of early recalled women among the total of screening tests

‡Percentage of invasive procedures performed among the total of recalled women and early recalled women with further assessments

PPV: Predictive Positive Value

PPV<sub>1</sub>: Percentage of cancers among the total of women whom at least one further assessment was performed

PPV<sub>2</sub>: Percentage of cancers among the total of women recalled for invasive procedures performed

The total number of screening tests in SFM and DM was not the sum of tests with normal findings, recalled for further assessments and early recalled since some recalled and early-recalled women did not attend the invitation for further assessments (N=537 in SFM and N=60 in DM).

PPV1 and PPV2 were higher in DM than in SFM: [PPV1: 7.0% (310 of 4,420) and 5.5% (770 of 13,860) ( $p < 0.001$ ); PPV2: 36.9% (310 of 839) and 19.3% (770 of 3,985) ( $p < 0.001$ ) in DM and SFM respectively].

Cancer detection rates and tumour characteristics in the SFM and DM groups are shown in Table 1. We found no statistically significant differences in cancer detection rates or tumour characteristics between the two groups. In the first and successive screenings, the percentages of DCIS were higher in the DM than in the SFM group: 18.5% (15 of 81) vs 15.8% (51 of 322) ( $p = 0.580$ ), respectively, in the

first screening and 23.2% (43 of 185) vs 15.7% (52 of 332) ( $p = 0.115$ ), respectively, in successive screenings. Tumour sizes were similar in both groups, most of them having a maximum diameter of 2 cm or less, both for the first and successive screenings. Most of the tumours detected showed no lymph node involvement at diagnosis. The percentage of negative lymph node status was slightly higher in the DM group: 79.3% (65 of 82) vs. 74.9% (239 of 319) ( $p = 0.412$ ) in first screening and 83.6% (153 of 183) vs. 78.7% (263 of 334) ( $p = 0.182$ ) in successive screenings.

**Table 1** Cancer detection and tumour characteristics according to screen-film mammography or digital mammography

	Screen-film mammography		Digital mammography		p	Total	
	n	%	n	%		n	%
Screened women <sup>a</sup>	84,871	—	61,795	—	—	103,613	—
Screening mammograms	171,191	70.5	71,647	29.5	—	242,838	100
Cancer detection <sup>b</sup>	770	0.45	310	0.43	0.592	1,080	0.44
Age <50	48	0.31	22	0.32	0.904	70	0.31
Age ≥50	722	0.46	288	0.44	0.578	1,010	0.46
Tumour size (mm) <sup>c</sup>							
First screening							
DCIS	51	15.8	15	18.5		66	14.1
T1 (≤2 cm)	190	59.0	46	56.8	0.580	236	50.4
T2 (2–5 cm)	67	20.8	18	22.2		85	18.2
T3 (>5 cm)	5	1.6	1	1.2		6	1.3
T4	9	2.8	1	1.2		10	2.1
Unknown size	61	—	4	—		65	13.9
Successive screening							
DCIS	52	15.7	43	23.2		95	15.8
T1 (≤2 cm)	211	63.6	117	63.2	0.115	328	54.4
T2 (2–5 cm)	56	16.9	22	11.9		78	12.9
T3 (>5 cm)	8	2.4	3	1.6		11	1.8
T4	5	1.5	0	0.0		5	0.8
Unknown size	52	—	34	—		86	14.3
Lymph node status							
First screening							
Negative	239	74.9	65	79.3	0.412	304	75.8
Positive	80	25.1	17	20.7		97	24.2
Unknown	67	—	4	—		71	—
Successive screening							
Negative	263	78.7	153	83.6	0.182	416	80.5
Positive	71	21.3	30	16.4		101	19.5
Unknown	50	—	41	—		91	—

<sup>a</sup> The total number of screened women was not the sum of screen-film and digital mammography screened women, since some women were screened by both techniques

<sup>b</sup> The percentage of cancers detected was computed as follows: (number of cancers / screening mammograms performed) × 100

<sup>c</sup> N=1,071 tumours. In 6 cancers in SFM and in 3 in DM the type of tumour (invasive or in situ) was unknown

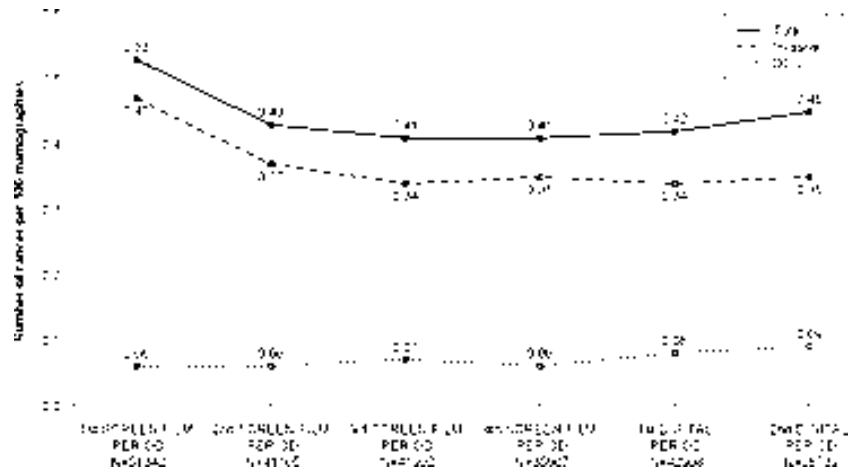
DCIS: Ductal carcinoma in situ

Cancer detection rates by technique-period are shown in Fig. 2. After an initial decrease from the first SFM period (which included prevalent screening for all screening programmes) to the second SFM period, there was a slightly rising temporal trend ( $p=0.049$ ), mainly for overall and DCIS rates.

Out of the 242,838 screening mammograms performed (whether screening mammograms or intermediate mammograms), 18,280 (7.5%) required women to be recalled for further assessments. Of these 18,280 mammograms, radiological information was available in 15,327 (83.8%; 11,413 in SFM and 3,914 in DM).

Among these 15,327 mammograms, multiple patterns were observed in 3,848 (25.1%), which were excluded for PPV comparisons. Consequently, Table 2 provides information on 11,479 (74.9%, 11,479 of 15,327) mammograms (8,297 in SFM and 3,182 in DM) showing a single radiological pattern. For masses, asymmetries and calcifications, PPVs were higher in the DM group and were statistically significant in masses [5.3% (97 of 1,826) vs. 3.9% (223 of 5,736); proportion ratio: 1.37; confidence interval 95%: 1.08–1.72]. The PPV for distortions was higher among SFM, although this difference was not statistically significant.

**Fig. 2** Cancer detection rates (total, invasive and ductal carcinoma in situ) by technique-period



The types of cancer detected (invasive or DCIS) by SFM and DM, based on the radiological pattern, are shown in Table 3. No statistically differences were observed among the types of tumour detected in each radiological pattern in the two study groups. Most tumours detected on the basis of a suspicious tumour-like mass were invasive in both groups: 94.6% (210 of 222) in SFM and 98.9% (92 of 93) in DM ( $p=0.118$ ), as were tumours detected by distortions or asymmetries. All tumours detected by asymmetries were invasive. In cancers detected by observation of calcifications, the percentage of DCIS was higher in the DM group: 60.3% (41 of 68) vs. 46.4% (64 of 138) ( $p=0.060$ ).

**Discussion**

This study retrospectively analysed a cohort of 103,613 women screened between 1996 and 2007 using SFM and/or DM. Our results show that PPV increases when DM is used, both for all further assessments and for invasive

procedures, with similar cancer detection rates and without statistically significant differences in tumour characteristics. The greatest improvements in PPV were detected in the radiological pattern of masses.

Increases in PPV were expected, given the reduction in recall rates with DM, with similar detection rates for both techniques observed in our setting [16]. Our results with DM are closer to the recommendations in the European guidelines, which set an acceptable level of recall for further assessments of <7% of women in the initial screening and of <5% in subsequent screening examinations, and percentages of early-recalled women at <1% [15].

Although consensus is lacking [13, 17], several publications report higher PPV with DM than with SFM [9, 11, 18]. Del Turco et al. reported PPVs of 14.7% for SFM and 15.9% for DM ( $p=0.65$ ). Similarly, Vigeland et al. reported PPVs of 13.5% for SFM and 16.6% for DM ( $p=0.014$ ). Lipasti et al. also reported a higher PPV for DM, with a relative rate of 1.42 (confidence interval 95%: 1.11–1.82). However, there were some methodological

**Table 2** Positive predictive values (PPV) according to radiological pattern in screen-film and digital mammography

Radiological patterns (N)	Screen-film mammography			Digital mammography				
	Radiological findings N (%)	Cancer N	PPV %	Radiological findings N (%)	Cancer N	PPV %	PR	95% CI
Mass	5,736 (69.1)	223	3.9	1,826 (57.4)	97	5.3	1.37	1.08–1.72
Microcalcifications	994 (12.0)	140	14.1	403 (12.7)	69	17.1	1.22	0.93–1.58
Distortion	512 (6.2)	103	20.1	230 (7.2)	34	14.8	0.73	0.52–1.05
Asymmetry	1,013 (12.2)	26	2.6	707 (22.2)	22	3.1	1.21	0.69–2.12
Other	42 (0.5)	3	7.1	16 (0.5)	2	12.5	1.75	0.32–9.52
Total	8,297 (100.0)	495	6.0	3,182 (100.0)	224	7.0	1.18	1.01–1.37

$N=11479$ , Single radiological patterns observed in women recalled for further assessments ( $N=8,297$  in screen-film mammography and  $N=3,182$  in digital mammography)

$N=719$  detected cancers (495 in screen-film mammography and 224 in digital mammography)

$$PPV\% = \left( \frac{\text{Cancers}}{\text{Radiological findings}} \right) \times 100$$

$$PR(\text{proportion ratio}) = \left( \frac{PPV_{dm}}{PPV_{sfm}} \right)$$

**Table 3** Invasive or ductal carcinoma in situ detected by screen-film mammography and digital mammography, based on the radiological pattern observed

	Screen film mammography N (%)	Digital mammography N (%)	p
<b>Masses</b>			
All cancers	223	97	0.118
Invasive	210 (94.6)	92 (98.9)	
DCIS	12 (5.4)	1 (1.1)	
<b>Calcifications</b>			
All cancers	140	69	0.060
Invasive	74 (53.6)	27 (39.7)	
DCIS	64 (46.4)	41 (60.3)	
<b>Distortions</b>			
All cancers	103	34	0.092
Invasive	101 (98.1)	30 (90.9)	
DCIS	2 (1.9)	3 (9.1)	
<b>Asymmetries</b>			
All cancers	26	22	na
Invasive	26 (100.0)	22 (100.0)	
DCIS	0 (0.0)	0 (0.0)	
<b>Other</b>			
All cancers	3	2	0.400
Invasive	2 (66.7)	0 (0.0)	
DCIS	1 (33.3)	2 (100.0)	

DCIS: Ductal carcinoma in situ  
Only cancers based on a single radiological pattern ( $N=719$ ; 495 in screen film mammography and 224 in digital mammography) are included  
The number of all cancers may not be the sum of invasive cancers and DCIS, since the type of tumour was sometimes unknown: 3 cases in screen film mammography and 6 cases in digital mammography

differences between these studies and our own, mainly in the design of the study groups, thus hampering comparisons. The PPV values reported in above mentioned studies were probably higher due to the lower recall rate in their settings.

In the final results of the Oslo II randomised trial, the PPVs computed for invasive procedures were higher when DM was used [17]. To date there are no data from other population-based programmes. The need from data from population-based programmes is especially pressing because a false-positive result leading to an invasive procedure has a greater negative physical impact on women and involves a higher cost than imaging procedures [19]. Implementation of digital technologies seems to decrease the adverse effects related to recall rate and false positive results whereas detection performance is at least as good as with SFM [16]. Because most European screening programmes are population-based, changes in screening results would prevent a large number of women from experiencing the adverse effects of this preventive modality and would reduce related costs. In the USA, where recall rates could be at least two or three times higher than in Europe, the impact of these findings could represent a greater improvement [20].

Although not statistically significant, tumour characteristics at diagnosis seem to be in an earlier stage when digital technology is used. As reported in other studies, lymph node involvement tends to be lower [9] and invasive tumours tend to be smaller [11] when DM is used, both for

first (prevalent cancers) and successive screenings (incident cancers). In agreement with previous works [9–11], in our study the proportion of DCIS was higher in DM than SFM, although this difference was nonsignificant. Vigeland et al. found that the proportion of DCIS in their DM study population was 27.1% compared with 16.3% in women screened with SFM ( $p<0.001$ ) [9]. The European guidelines state that the proportion of screen-detected invasive cancers should be 80–90%, meaning that 20% is the maximum limit for DCIS [15]; however, the optimal proportion of DCIS in a screening programme is still controversial.

Detection of DCIS is highly important because this type of tumour might be the immediate precursor of invasive breast cancers, which are potentially lethal [21], although some DCIS—around 25%–30%—may never progress to invasive disease [22]. At this time, we cannot determine which tumours will and which tumours will not progress [23]. Thus, screening mammography may currently be benefiting some women whose DCIS will progress to invasive cancer but potentially harming other women whose DCIS would never be associated with subsequent invasive cancer and who, for lack of good prognostic indicators, are almost always treated with surgery and adjuvant therapies [24]. Therefore, detection of DCIS reduces the subsequent incidence of invasive ductal carcinoma, and early diagnosis may save lives by preventing the development of invasive carcinoma with metastatic potential [25].

Time trends for the incidence of in situ and invasive cancers in the cohort studied should be carefully interpreted because of the relatively short DM period. However, in the last digital period, the incidence of both invasive cancers and DCIS increased, reaching a similar DCIS incidence rate to that reported in 2009 by Weigel et al. in their population, screened with DM only [26]. The slightly rising temporal trend observed in our study should be confirmed in future works.

The PPVs for radiological patterns were similar with both techniques, but were higher with DM, reaching statistical significance only for masses. Only a few studies have compared the PPVs of DM and SFM according to radiological patterns [11, 18]. Our relative rate for masses was very similar to that in a Finnish study, which also found that the PPV specific for calcifications was higher with DM, although this difference was not statistically significant (53.4% vs. 41.1%). Another study performed in a digital screening unit of a German nationwide programme reported a PPV specific for calcifications of 18.5%, a percentage very similar to our data (17.1%) [12]. Therefore, whereas there are no conclusive data indicating better diagnostic precision in DM for some mammographic findings (calcifications, distortions and asymmetries), DM is clearly more precise in detecting masses. This result is also consistent with the findings of a phantom study by Yang et al., who found that DM was significantly superior to SFM in detecting and characterizing small masses in mixed and dense breast backgrounds [27]. In a recent study, Pinker et al. reported that DM was significantly superior to SFM in the conspicuity of screen-detected breast cancers for all histological types and breast densities regardless of manifestation as mass; in addition, architectural distortion, asymmetric density or microcalcifications were also better visualized by DM [28].

When cancer detection based on radiological patterns was examined, we found that, among malignancies detected on the basis of calcifications, the percentage of DCIS was higher in the DM than in the SFM group, although this difference was not statistically significant. This finding is consistent with those of a previous study [13] reporting that the percentage of DCIS among the cancers detected on the basis of suspicion of calcifications was higher when DM was used.

Our study has several limitations, one of the most important being the relatively short period in which DM was used, leading to a lack of opportunity to draw conclusions about trends in the incidence of DCIS and invasive cancers. Another limitation is that some information was missing for the variable of radiological patterns, which lowered the statistical power. However this lack of information was similar in both the SFM and DM groups. Information about BIRADS was not included since some of

the programmes participating in the study report this data in a different way. Finally, we have no information on breast density, which is an important factor, given that some studies have shown that DM is more accurate in women with radiographically dense breasts [3].

Data provided in this work provide insight into the impact of the introduction of DM. Information on PPVs based on the radiological patterns and differences in tumour-related data could be useful and necessary to evaluate the risks and benefits of breast cancer screening, given the large number of programmes that are using this technology.

In conclusion, the present study reinforces the message of previous works that DM has a similar diagnostic precision to SFM and fewer adverse effects. The differences observed in tumour characteristics and the slightly higher rates of DCIS observed with DM, although not statistically significant, suggest an advance in early detection. However, a longer period of using DM is needed to confirm the increase in DCIS detection.

**Acknowledgments** We are grateful to Sonia Sánchez, Paula Merino, and Cristina Hernández for their contribution in the data collection of the screening programs, as well as to Jordi Blanch and Rubén Román for their assistance in analysis and data management. This study was supported by grants from Instituto de Salud Carlos III FEDER (PI07/90293).

## References

1. Bick U, Diekmann F (2007) Digital mammography: what do we and what don't we know? *Eur Radiol* 17:1931–1942. doi:10.1007/s00330-007-0586-1
2. Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N, Dos Santos Silva IM (2005) Full-field digital versus screen-film mammography: comparison within the UK breast screening programme and systematic review of published data. *Radiology* 251:347–358. doi:10.1148/radiol.2512081235
3. Pisano ED, Gatsonis C, Hendrick E et al (2005) Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 353:1773–1783
4. Skaane P (2009) Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. *Acta Radiol* 50:3–14. doi:10.1080/02841850802563269
5. Sala M, Comas M, Macia F, Martinez J, Casamitjana M, Castells X (2009) Implementation of digital mammography in a population-based breast cancer screening programme: effect of screening round on recall rate and cancer detection. *Radiology* 252:31–39. doi:10.1148/radiol.2521080696
6. Lewin JM, Hendrick RE, D'Orsi CJ et al (2001) Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations. *Radiology* 218:873–880
7. Vanovcanova L, Lehotska V, Rauova K (2010) Digital mammography—a new trend in breast carcinoma diagnostics. *Bratisl Lek Listy* 111:510–513
8. Hedden B, Ronnow K, Olsson M, Miller D (2007) Digital versus screen-film mammography: a retrospective comparison in a

- population-based screening programme. *Eur J Radiol* 64:419–425. doi:10.1016/j.ejrad.2007.02.030
9. Vigeland E, Klaasen H, Klingen TA, Hofvind S, Skaane P (2008) Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: The Vestfold county study. *Eur Radiol* 18:183–191. doi:10.1007/s00330-007-0730-y
  10. Karssemeijer N, Bluekens AM, Beijerinck D et al (2009) Breast cancer screening results 5 years after introduction of digital mammography in a population-based screening programme. *Radiology* 253:353–358. doi:10.1148/radiol.2532090225
  11. Del Turco MR, Mantellini P, Ciatto S et al (2007) Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. *AJR Am J Roentgenol* 189:860–866. doi:10.2214/AJR.07.2303
  12. Weigel S, Decker T, Korsching E, Hungermann D, Bocker W, Heindel W (2010) Calcifications in digital mammographic screening: improvement of early detection of invasive breast cancers? *Radiology* 255:738–745. doi:10.1148/radiol.10091173
  13. Hambly NM, McNicholas MM, Phelan N, Hargaden GC, O'Doherty A, Flanagan FL (2009) Comparison of digital mammography and screen-film mammography in breast cancer screening: a review in the Irish breast screening programme. *AJR Am J Roentgenol* 193:1010–1018. doi:10.2214/AJR.08.2157
  14. Kalager M, Zelen M, Langmark F, Adami HO (2010) Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med* 363:1203–1210
  15. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L (2008) European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition—summary document. *Ann Oncol* 19:614–622. doi:10.1093/annonc/mdm481
  16. Sala M, Salas D, Belvis F et al (2011) Reduction in false-positive results after the introduction of digital mammography: analysis from four population-based breast cancer screening programs in Spain. *Radiology* 258:388–395
  17. Skaane P, Hofvind S, Skjennald A (2007) Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening programme: follow-up and final results of Oslo II study. *Radiology* 244:708–717. doi:10.1148/radiol.2443061478
  18. Lipasti S, Anttila A, Pamilo M (2010) Mammographic findings of women recalled for diagnostic work-up in digital versus screen-film mammography in a population-based screening programme. *Acta Radiol* 51:491–497. doi:10.3109/02841851003691961
  19. Román R, Sala M, Salas D, Ascunce N, Zubizarreta R, Castells X, et al. (2011) Effect of protocol-related variables and women's characteristics on the cumulative false-positive risk in breast cancer screening. *Ann Oncol* doi:10.1093/annonc/mdr032
  20. Venkatesan A, Chu P, Kerlikowske K, Sickles EA, Smith-Bindman R (2009) Positive predictive value of specific mammographic findings according to reader and patient variables. *Radiology* 250:648–657. doi:10.1148/radiol.2503080541
  21. Ernster VL, Barclay J, Kerlikowske K, Wilkie H, Ballard-Barbash R (2000) Mortality among women with ductal carcinoma in situ of the breast in the population-based surveillance, epidemiology and end results programme. *Arch Intern Med* 160:953–958
  22. Allred DC (2010) Ductal carcinoma in situ: terminology, classification, and natural history. *J Natl Cancer Inst Monogr* 2010:134–138. doi:10.1093/jncimonographs/lgq035
  23. D'Orsi CJ (2010) Imaging for the diagnosis and management of ductal carcinoma in situ. *J Natl Cancer Inst Monogr* 2010:214–217. doi:10.1093/jncimonographs/lgq037
  24. Kerlikowske K (2010) Epidemiology of ductal carcinoma in situ. *J Natl Cancer Inst Monogr* 2010:139–141. doi:10.1093/jncimonographs/lgq027
  25. Feig SA (2000) Ductal carcinoma in situ. Implications for screening mammography. *Radiol Clin North Am* 38:653–668, vii
  26. Weigel S, Batzler WU, Decker T, Hense HW, Heindel W (2009) First epidemiological analysis of breast cancer incidence and tumour characteristics after implementation of population-based digital mammography screening. *Rofo* 181:1144–1150. doi:10.1055/s-0028-1109831,10.1055/s-0028-1109831
  27. Yang WT, Lai CJ, Whitman GJ, Murphy WA, Dryden MJ, Kushwaha AC, Sahin AA, Johnston D, Dempsey PJ, Shaw CC (2006) Comparison of full-field digital mammography and screen-film mammography for detection and characterization of simulated small masses. *AJR Am J Roentgenol* 187:W576–581. doi:10.2214/AJR.05.0126
  28. Pinker K, Perry N, Vinnicombe S, Shiel S, Weber M (2011) Conspicuity of breast cancer according to histopathological type and breast density when imaged by full-field digital mammography compared with screen-film mammography. *Eur Radiol* 21(1):18–25. doi:10.1007/s00330-010-1906-4





## **IV. DISCUSSIÓ**

---



## 1. Resultats principals

Els resultats dels quatre treballs presentats donen suport a la hipòtesi principal de la tesi, posant de manifest que tant el disseny, com l'organització i les característiques de la prova de cribratge poden influenciar tant en els beneficis i com en els efectes adversos dels programes de detecció precoç del càncer de mama.

De manera resumida, les principals troballes d'aquesta recerca són:

- 1) La introducció del cribratge poblacional ha suposat un impacte sobre la detecció de càncer de mama en la població no participant als programes.
- 2) Les diferències organitzatives i en els criteris d'elegibilitat de la població diana poden modificar les taxes de detecció de càncer de mama dels programes de cribratge.
- 3) Els tumors detectats entre intervals de cribratge presenten característiques biològiques associades amb un pitjor pronòstic i són més comuns en dones amb elevada densitat mamària.
- 4) Entre els veritables càncers d'interval, tumors que no presentaven signes de sospita en l'última mamografia de cribratge, hi ha un percentatge elevat de tumors amb fenotip triple negatiu.
- 5) La via diagnòstica (detecció en el cribratge, entre intervals de cribratge o fora del cribratge) és un factor rellevant en el pronòstic de la malaltia. Les dones amb veritables càncers d'interval presenten un major risc de morir que les dones amb càncers detectats en el cribratge i les dones no participants.
- 6) La mamografia digital té un valor predictiu positiu superior que la mamografia analògica. Les sospites radiològiques per les quals ha millorat més el valor predictiu positiu són les masses i les microcalcificacions.
- 7) La introducció de la mamografia digital apunta a un augment de la detecció de tumors *in situ*.

## 2. Discussió conjunta dels articles

La discussió conjunta dels articles s'articula en base a les preguntes de recerca en les quals es centra la tesi. La discussió específica de cada treball es troba detallada en cada un dels articles, on hi figuren també les limitacions i fortaleces específiques.

L'observació de les taxes de detecció del càncer de mama segons la via diagnòstica en el context danès, han demostrat l'existència d'un pic de prevalença en la primera ronda de cribratge, tant entre la cohort de dones participants en el cribratge com entre les no participants. Aquesta troballa suggereix que la introducció del cribratge poblacional també té un impacte sobre la població no participant. Una de les principals problemàtiques en l'estimació dels beneficis del cribratge resideix en la comparabilitat dels resultats de la població cribrada amb població no cribrada<sup>139</sup>. Generalment, quan es duen a terme aquestes comparacions, l'efecte de la detecció precoç queda restringit al grup de dones participants al cribratge, i els canvis que s'observen entre les dones no participants s'atribueixen a aspectes globals referents a la detecció i el tractament de la malaltia, que per altra banda també afecten al grup de dones cribrades<sup>22</sup>. El fet de considerar que l'efecte del cribratge queda restringit a la població participant, pot suposar una certa subestimació dels efectes de la detecció precoç sobre la reducció de la morbiditat i la mortalitat, quan es comparen ambdues poblacions. Aquesta reflexió es veu reforçada per un treball recentment publicat, realitzat a Noruega, on s'apunta que al voltant del 50% de casos detectats en dones no participants al cribratge es detecten en fase asimptomàtica mitjançant cribratge oportunístic i descriuen com aquests tumors presenten unes característiques clíniques similars a les dels tumors detectats en els programes poblacionals<sup>140</sup>.

Per altra banda hem observat que diferències organitzatives i en els criteris d'elegibilitat tenen un efecte important en les taxes de detecció del càncer de mama entre programes. Normalment en la comparació d'indicadors de qualitat dels programes no es tenen en compte les diferències organitzatives. En aquest aspecte, s'ha observat com les diferències en els criteris d'elegibilitat de la població diana poden explicar diferències en els valors d'alguns indicadors i esbiaixar-ne els resultats. Concretament, si els criteris d'elegibilitat exclouen dones amb factors de risc del càncer de mama, serà esperable que entre la cohort de dones no participants al cribratge existeixi una major incidència de la malaltia. En aquest cas, la comparació dels indicadors de programa no serà reflex únicament d'aspectes lligats en la seva qualitat, sinó també relacionats amb les característiques de la població participant.

El principal interès d'estudiar els càncers d'interval recau en poder aprofundir en les característiques dels falsos negatius, tumors potencialment evitables des del punt de vista dels programes, i les dels veritables càncers d'interval, tumors inevitables amb les actuals estratègies<sup>69</sup>. Al realitzar la classificació radiològica s'han pogut comparar les característiques biològiques dels falsos negatius i els veritables càncers d'interval amb les dels càncers detectats en el cribratge. Els tumors falsos negatius, que representaren prop del 20% dels càncers d'interval radiològicament classificats, foren biològicament equivalents als de cribratge, tot i que es diagnosticaren en estadis més avançats. Els veritables càncers d'interval, en canvi, representaren un grup amb característiques biològiques relacionades amb un creixament més ràpid i una major agressivitat. Entre aquestes característiques destaca una elevada proporció de tumors amb fenotip triple negatiu, tumors que s'han relacionat amb una proliferació més ràpida i una pitjor resposta al

tractament<sup>141</sup>. Així, aquesta troballa és rellevant ja que ajuda a comprendre millor la història natural de la malaltia i en part explica perquè certs càncers presenten un *sojourn time* més curt i, per tant, són menys susceptibles a beneficiar-se de les actuals estratègies de cribratge<sup>142</sup>. Alguns treballs ja havien reportat una major proporció de tumors triples negatius entre els càncers d'interval<sup>125</sup>, però aquest treball posa de manifest que la majoria de tumors triples negatius es troben entre els veritables càncers d'interval, en línia amb els resultats de publicacions més recents<sup>143,144</sup>. És evident que el fet que entre els veritables càncers d'interval hi hagi una major proporció de tumors de característiques agressives explica, en part, el seu pitjor pronòstic en comparació amb càncers detectats per altres vies. Tanmateix, els models multivariants apunten que la supervivència de les dones amb veritables càncers d'interval és menor que la de dones amb tumors diagnosticats dins i fora del cribratge, tot i tenint en compte el patró molecular i l'estadiatge al moment del diagnòstic.

De cara a definir estratègies de cribratge més eficients, és important el fet que entre els càncers d'interval el percentatge de dones amb mames denses sigui major. Tot i que no existeix consens sobre el mecanisme que relaciona elevada densitat mamària i risc de càncer d'interval<sup>119,145</sup>, existeixen diferents hipòtesis que apunten a un efecte enmascarador sobre la mamografia així com a un efecte de tipus biològic que per diferents mecanismes estimularia la proliferació dels tumors<sup>146-148</sup>. Sigui com sigui, aquesta associació és rellevant ja que pot ajudar a definir grups de dones susceptibles de presentar tumors de ràpid creixement.

Finalment, la introducció de la mamografia digital estava relativament poc avaluada des del punt de vista poblacional<sup>129</sup>, tot i haver-se provat en assajos clínics la igualtat en la capacitat diagnòstica<sup>128</sup>. En el context espanyol, s'ha observat que la mamografia digital està suposant una disminució dels falsos positius i un augment del valor predictiu positiu de la prova, tant de manera global com específicament per les principals sospites radiològiques. Concretament, l'augment del valor predictiu positiu per masses s'atribueix a una disminució del nombre de proves addicionals davant d'aquestes sospites, mentre que per les calcificacions seria degut a un augment dels diagnòstics de càncer. A més, la introducció d'aquesta tècnica apunta a un augment de la detecció de tumors menys avançats, encara que cal un període més llarg d'observació per confirmar aquests resultats en el nostre context. Treballs publicats potseriorment també reporten un augment de la detecció de tumors *in situ*, però amb interpretacions contràries sobre el possible sobrediagnòstic que pot comportar<sup>132,149</sup>.

Dels quatre treballs es pot concloure que l'avaluació dels programes poblacionals de detecció precoç de càncer de mama és complexa i susceptible a biaixos i limitacions difícils de controlar. Els programes de cribratge s'adrecen a milions de dones arreu del món i, per tant, l'impacte que pugui tenir sobre certs indicadors, encara que sigui petit en termes relatius, pot ser potencialment gran quan es tradueix al conjunt de la població.

### 3. Limitacions

Les principals limitacions d'aquesta tesi, són les pròpies dels dissenys epidemiològics de cadascun dels treballs que la componen. L'estudi de cohorts retrospectives, tot i permetre disposar d'informació d'un gran número de subjectes i amplis períodes temporals, queda limitat a la quantitat d'informació disponible en les fonts de dades. En l'article 1, la informació procedia de registres poblacionals de tumors, les bases de dades dels propis programes, juntament amb un registre específic de càncer de mama. Aquest fet suposà que només es disposés d'informació de les principals variables clíniques dels tumors, impedit aprofundir en la seva caracterització molecular. Tanmateix, per respondre els objectius de l'estudi no es requeria específicament l'anàlisi d'aquestes variables.

La situació fou semblant en l'article 4. En aquest cas sí que es contemplava l'estudi de les sospites radiològiques en les mamografies. Aquesta informació estava disponible en la majoria dels episodis on es realitzaren proves addicionals, on el percentatge de desconeguts representava el 16.2% dels casos. De totes maneres, els resultats es centraren en l'estudi sospites per patrons únics i es descartaren els patrons múltiples, per evitar la dispersió dels resultats. Fou també limitada la informació referent als tumors, quedant restringida a la invasivitat i la grandària dels tumors, i a l'afectació limfàtica. Tanmateix, algunes variables com la densitat mamària o la categoria BI-RADS no estaven disponibles o presentaven importants mancances, fet que no permeté incorporar-les a les anàlisis ni estudiar-ne el seu efecte.

La principal limitació de l'article 2, un disseny cas-control, fou el baix nombre de subjectes inclosos a l'estudi. Malgrat es partia d'un nombre de dones cribrades important (57.467 dones cribrades a l'Hospital del Mar, des de l'inici del programa fins l'any 2008), cal tenir present que els càncers d'interval són relativament poc freqüents (representen el 20-25% del total de càncers diagnosticats en dones participants al cribratge)<sup>81,150</sup> i la seva classificació en subtipus encara en redueix més el nombre. A més, com s'apunta en el treball, l'obtenció de les mamografies per poder fer la classificació del càncer d'interval fou un dels aspectes limitants en el nombre de casos finals inclosos a les anàlisis. En el nostre cas, tan sols es pogueren recuperar les dues mamografies en el 56% dels casos de càncer d'interval identificats. Tot i les limitacions, però, els resultats del treball són clars i consistents amb els d'estudis publicats amb posterioritat<sup>120,143</sup>.

L'article 3, que neix com a continuació del segon, comparteix amb aquest la limitació de la grandària de la mostra. En aquest cas, a més de la informació dels tumors, es requeria informació de seguiment, fet que afegí un grau de complexitat a l'obtenció de la informació ja que no totes les dones es segueixen sempre en el mateix centre on s'han diagnosticat.

Una limitació inherent en els treballs que estudien els càncers d'interval és la possible mala classificació. Per la pròpia definició de càncer d'interval, que ve determinada per la periodicitat de la prova de cribratge, és probable que dones amb simptomatologia entre dues mamografies de cribratge s'esperin a obtenir un

diagnòstic a la següent invitació, contribuint en que alguns tumors detectats per mamografia siguin simptomàtics. I, de manera contrària, tampoc es pot descartar que en l'interval entre cribratges es diagnostiquin càncers de mama en dones asimptomàtiques per la realització, per exemple, de mamografies fora del circuit del cribratge. Ambdues situacions podrien suposar una atenuació de les diferències biològiques i clíniques entre els càncers d'interval i els càncers diagnosticats al cribratge.

Una altra limitació que dificulta l'estudi dels càncers d'interval en el nostre context és l'absència de registres de càncer poblacionals que cobreixin tota la població. En el cas de Catalunya, n'existeixen tan sols a Tarragona i Girona. Els registres poblacionals són l'eina principal per l'avaluació del cribratge i la font més fiable per la identificació dels càncers d'interval. Tanmateix, des de la posada en marxa del programa de detecció precoç del càncer de mama de l'Hospital del Mar s'ha fet un esforç per seguir a les dones que no retornen al cribratge i per identificar si el motiu de no retorn es deu, entre d'altres, a un càncer d'interval. Aquest seguiment permet posar-se en contacte amb més del 90% de dones que no retornen al programa, i per tant assegura una elevada exhaustivitat en la identificació de dones amb càncer d'interval.

#### **4. Fortaleses**

L'anàlisi de cohorts retrospectives en els article 1 i 4 ha permès disposar d'informació d'un gran nombre de dones en diferents contextos, fet que dóna consistència als resultats. En els dos treballs, a més, les fonts d'informació garanteixen la robustesa de les dades: per l'article 1, s'utilitzà el registre civil, el registre poblacional de tumors i les bases de dades dels programes de cribratge de dues regions daneses. En l'article 4, la informació procedia de les bases de dades de 4 programes de cribratge espanyols, fet que el convertí en un dels treballs que avaluaren l'impacte de la mamografia digital sobre una població més gran.

En els treballs amb disseny cas-control, es disposà d'informació exhaustiva de variables clíniques i moleculars referents als tumors i segons la classificació radiològica dels subtipus de càncer d'interval, aspecte no analitzat fins aquell moment. Així mateix, per l'article 3 s'obtingué informació sobre el tractament i el seguiment clínic, en termes de recaigudes i mortalitat. La recollida d'informació es realitzà a partir de la revisió d'històries clíniques i mamografies, amb el suport d'un equip multidisciplinari compost per radiòlegs, oncòlegs, patòlegs i epidemiòlegs.

En els quatre treballs es disposà d'informació de llargs períodes temporals, des de la posada en marxa dels programes fins a finals de la dècada del 2000. En el primer article, el període temporal estudiat fou de 17 anys, fet que el convertí en l'estudi que avaluà la detecció del càncer de mama dins i fora del cribratge en un període més ampli. Pels articles 2 i 3 s'incorporaren subjectes de l'inici del cribratge fins a

l'any 2008, i el període de seguiment s'amplià fins a finals del 2010. El període d'estudi del treball 4 també anava des de l'inici dels programes, entre el 1996 i el 1998, fins el desembre del 2007. En aquest cas, a més, tots els programes comptaren amb al menys dues rondes completes realitzades amb la tècnica digital.

Els quatre treballs exploraren aspectes controvertits en l'avaluació del cribratge, i aportaren respostes inèdites que es sumaren al coneixement sobre l'impacte de les pràctiques de detecció precoç del càncer de mama. Tot i no concloure amb solucions definitives i deixar interrogants oberts, els diferents articles aprofundiren en tres qüestions rellevants de l'avaluació del cribratge de càncer de mama, com l'impacte del cribratge sobre la població no participant, l'estudi de l'expressió de biomarcadors en els principals subtipus de càncers d'interval i l'efecte sobre la precisió diagnòstica de la tècnica digital. Altres treballs realitzats en posterioritat han corroborat els resultats inicials, donant crèdit a les troballes que es presenten<sup>120,132</sup>.

## **5. Continuïtat i futures línies de recerca**

Com s'ha comentat, aquesta tesi s'emmarca en la línia de recerca en avaluació del cribratge del Servei d'Epidemiologia i Avaluació de l'Hospital del Mar-IMIM. En el context del CIBER i actualment de la REDISSEC, s'han demanat i concedit diferents projectes per aprofundir en diferents aspectes del cribratge, que al seu temps donen continuïtat a les qüestions presentades.

L'experiència adquirida en l'estudi dels càncers d'interval en el marc del programa de l'Hospital del Mar fou clau per iniciar un projecte coordinat de caràcter estatal que ha permès agrupar informació de més de 600.000 dones cribrades en 7 programes espanyols (Projecte INCA). S'identificaren més de 1.600 càncers d'interval, els quals s'han classificat radiològicament, i s'ha obtingut informació clínica i patològica de més de 2.500 càncers de mama. Durant els anys 2009-2012 es realitzà la recollida d'informació, la construcció de la cohort i actualment ja s'estan analitzant les dades i estan en marxa diferents articles. Les principals qüestions a les quals es vol donar resposta, són: a) la descripció de les taxes de càncer d'interval a nivell espanyol; b) l'estudi dels determinants del càncer d'interval, tant relacionats amb la dona com amb el protocol dels programes; c) la caracterització dels diferents subtipus de càncer d'interval des del punt de vista biològic; i, d) l'impacte de la introducció de la mamografia digital sobre els càncers d'interval.

Amb l'objectiu d'aprofundir en el coneixement de les lesions benignes i les sospites radiològiques identificades en el cribratge i la seva posterior evolució a càncer, l'any 2012 s'inicià el BELLE Project. Es tracta d'un projecte de caràcter estatal, on hi participen 8 programes de detecció precoç, i s'espera que els resultats aportin noves dades útils per l'adequació del cribratge en funció dels perfils de risc de desenvolupar càncer.



La continuïtat del treball centrat en l'evolució clínica dels tumors en funció de la via diagnòstica, ha derivat en un projecte coordinat recentment concedit pel FIS (PI12/00387), per avaluar de forma integrada el conjunt dels serveis sanitaris per atendre el càncer de mama, des del diagnòstic precoç fins a les diferents estratègies de tractament. En el subprojecte centrat en la detecció precoç, es preveu recollir informació de l'evolució clínica de les dones que formen part de la cohort INCA. Aquest projecte té com objectiu explorar el període lliure de malaltia i la supervivència global dels càncers de cribratge i dels càncers d'interval, diferenciant entre subtipus i tenint en compte tant les característiques clíniques com biològiques dels tumors. Per altre banda, també es vol estudiar el pronòstic dels tumors en funció de si han estat diagnosticats amb la tècnica analògica o la digital.

Aquesta experiència, juntament amb les col·laboracions iniciades amb grups europeus i la necessitat de treballar en un marc més ampli, fa que actualment s'estigui treballant per un futur projecte col·laboratiu amb grups de recerca europeus. La necessitat de disposar de grans cohorts per l'estudi d'esdeveniments de baixa freqüència i d'informació procedent de diferents contextos, pot contribuir en millorar l'actual informació sobre el cribratge i estudiar l'efecte sobre els beneficis i els efectes adversos de diferents estratègies.

Com en d'altres branques de la medicina, el futur de la recerca sobre el cribratge del càncer de mama, sembla estar encaminat cap al replantejament de les actuals estratègies universals '*one size fits all*' per estratègies més personalitzades que permetin maximitzar els beneficis entre les dones més susceptibles de desenvolupar un càncer i minimitzar els efectes adversos i els costos en global. La manera òptima de valorar l'efectivitat de diferents estratègies de cribratge seria realitzar assajos clínics aleatoritzats. Tanmateix, l'esforç que es requereix per dur-los a terme és considerable i caldria disposar de certa evidència dels seus avantatges abans de proposar nous escenaris. Com alternativa per valorar el potencial efecte de noves estratègies de cribratge, fa anys que els models matemàtics basats en la simulació s'estan consolidant com una eina amb moltes potencialitats. Tot i la seva complexitat, en el nostre context i dins del nostre propi grup de recerca, s'ha acumulat una experiència sòlida demostrant la seva utilitat tant en l'àmbit hospitalari<sup>151,152</sup> com poblacional<sup>153-156</sup>. Les dades procedents dels estudis RAFP i INCA actualment s'estan utilitzant en combinació amb dades sobre costos, per alimentar models matemàtics que pretenen valorar l'impacte de canvis en la periodicitat del cribratge o en l'edat d'inici en els resultats globals, tant pel que fa a beneficis (detecció precoç, reducció de la mortalitat), efectes adversos (falsos positius, càncers d'interval, sobrediagnòstic) i costos de les diferents estratègies. Davant l'enorme ventall d'escenaris alternatius de cribratge, les eines de simulació representen una alternativa molt útil per tenir resultats a curt termini i amb un cost reduït, de l'impacte a nivell poblacional de canvis organitzatius en les pràctiques de cribratge del càncer de mama.



## **V. CONCLUSIONS I IMPLICACIONS**

---



## 1. Conclusions

1. La implementació del cribratge poblacional ha suposat un augment de la detecció del càncer de mama entre la població no participant als programes de detecció precoç. Aquest efecte serà important de quantificar i considerar quan es comparen els resultats i els beneficis del cribratge entre la població participant i la població no participant.
2. Aspectes organitzatius com els criteris d'elegibilitat de la població diana tenen impacte en el comportament de les taxes de detecció del càncer de mama dins i fora dels programes de cribratge. Uns criteris basats en el risc de la malaltia poden influenciar en les taxes d'incidència de la malaltia en la població participant i no participant. Així, excloure dones amb factors de risc potencials de la població diana a cribar, pot fer augmentar la taxa de detecció entre les dones no participants.
3. Els tumors diagnosticats en l'interval entre mamografies de cribratge són majoritàriament tumors invisibles en l'anterior mamografia, o veritables càncers d'interval, i prop d'un 20% són tumors falsos negatius.
4. Els veritables càncers d'interval presenten un perfil molecular més agressiu, amb un percentatge elevat de tumors triples negatius.
5. El percentatge de dones amb mames denses amb veritables càncers d'interval és superior al descrit entre les dones amb tumors detectats en el cribratge.
6. Els veritables càncers d'interval presenten característiques associades amb un pitjor pronòstic, mostrant una menor supervivència global respecte els tumors detectats en el cribratge, després de controlar per variables clinocopatològiques,.
7. La introducció de la mamografia digital en el context espanyol ha suposat una millora de la precisió diagnòstica, derivant en un augment significatiu del valor predictiu positiu de la prova.
8. Les sospites radiològiques per masses i microcalcificacions són les que experimenten majors millores del valor predictiu positiu amb la introducció de la mamografia digital. En el primer cas, s'atribueix a una disminució de les proves addicionals davant d'aquestes sospites, i en el segon cas, a un augment dels diagnòstics de càncer.

9. Les diferències observades en les característiques dels tumors apunten a un augment de la detecció de tumors *in situ* i de tumors menys avançats amb la tècnica digital.

## 2. Recomanacions i implicacions en Salut Pública

Els resultats dels treballs que conformen la tesi aporten més informació per avaluar el balanç entre riscos i beneficis del cribratge. Quan es parla de beneficis del cribratge caldrà tenir en compte que la població no participant també és permeable a la influència del cribratge, i que aquesta influència, en part, pot diluir les diferències entre ambdues poblacions.

Donat que els criteris organitzatius poden esbiaixar les taxes de detecció dels programes de cribratge, resultarà pertinent que en l'avaluació d'aquestes pràctiques es descrigui acuradament la població diana, i que en la interpretació i comparació dels resultats es tinguin en compte els criteris de selecció de la població elegible i les taxes de participació.

Els resultats dels diferents treballs poden contribuir en adequar la informació sobre el cribratge que s'ofereix a les dones que són convidades a participar als programes poblacionals de detecció precoç del càncer de mama. Com s'apuntava a la introducció, un dels aspectes positius dels programes de detecció és la sensació de tranquil·litat i el refermament en el programa, davant d'un resultat negatiu en el test de cribratge. Aquest aspecte pot representar una falsa sensació de seguretat en el cas dels càncers d'interval i contribuir al retard diagnòstic. Cal informar a les dones amb resultat negatiu, que existeix la possibilitat de que aparegui un tumor en l'interval entre mamografies i per tant, que cal estar alerta davant de possibles signes de la malaltia entre els intervals de cribratge.

La informació sobre la caracterització biològica dels càncers d'interval pot ser útil per valorar futures estratègies d'optimització del cribratge, revisant, per exemple, la periodicitat de les proves o la utilització de diferents tecnologies per grups específics de dones amb elevada densitat mamària.

La classificació radiològica dels càncers d'interval, tot i suposar un gran esforç en la recuperació de les mamografies i en el propi procés de revisió, aporta informació molt rellevant per comprendre les particularitats dels tumors diagnosticats entre mamografies de cribratge. Així mateix, pot ser útil per millorar el coneixement de la història natural del càncer de mama, tant pel que fa al desenvolupament de tumors de fase preclínica curta com a factors pronòstic de la malaltia.

Cal continuar promovent la creació de registres de càncer de base poblacional que cobreixin la totalitat de la població catalana i espanyola. Aquests registres són l'eina principal per l'avaluació del cribratge, imprescindibles per poder conèixer l'efectivitat d'aquestes pràctiques. L'estudi d'efectes rellevants com el sobrediagnòstic i els càncers d'interval en el context espanyol es veu limitat per l'absència de registres

poblacionals, dificultant la comparació entre països i impedit conèixer l'impacte global de la detecció precoç en el nostre àmbit.

Cal afavorir el desplegament de la tecnologia digital en el conjunt dels programes de l'Estat, ja que més enllà de les avantatges en l'emmagatzematge i l'adquisició de les imatges, en el nostre context està suposant una disminució de la taxa de reconvocatòria a igualtat de capacitat diagnòstica, és a dir, un augment de la precisió del cribratge. La reducció dels resultats falsos positius es tracta d'un fet rellevant, ja que implica que un menor nombre de dones experimentin l'angoixa davant d'un possible diagnòstic de càncer, i representa una disminució de la càrrega assistencial en el nombre de proves complementàries. Davant d'aquests resultats, es considera important seguir amb la implementació d'aquesta tecnologia arreu del territori.

L'augment de tumors menys avançats amb la tècnica digital significa que les dones es puguin beneficiar de tractaments menys agressius, fet que s'associa amb la millora de la qualitat de vida. Tanmateix, es necessita un període d'estudi més llarg per concloure si l'augment de la detecció de tumors menys avançats es pot atribuir a un avanç diagnòstic o bé a un augment de la detecció de tumors, que aletraria pel possible sobrediagnòstic. Serà convenient que en el futur es realitzin treballs que explorin els resultats globals de cohorts de dones cribrades només amb la mamografia digital.





## **VI. ANNEX**

---





**Universitat Autònoma de Barcelona (UAB)**

Faculty of Medicine

PhD in Paediatrics, Obstetrics and Gynaecology, Preventive Medicine and  
Public Health

**THREE ISSUES IN ASSESSING POPULATION-BASED SCREENING  
FOR BREAST CANCER:**

**Effect on unscreened women,  
association between mode of diagnosis and tumour characteristics,  
and the impact of the imaging modality on results**

**PhD Dissertation - Extended Summary in English-**

**Laia Domingo Torrell**

**PhD Supervisor**

**Dr. Maria Sala Serra**

This dissertation was carried out in the Department of Epidemiology and Evaluation of the Hospital del Mar, in Barcelona, under the direction of PhD Maria Sala. It is presented as a compendium of publications.

Barcelona, May 2013

## SUMMARY

**Introduction.** Early detection of breast cancer has been recognized as the best strategy to improve the prognosis of breast cancer and decrease mortality in the long term. Following the recommendations of the Council of the European Union, several European countries have implemented population-based screening programmes. Most of these countries offer two-yearly mammography to women aged 50-69 years. However, there is some disagreement on the effectiveness of screening, based on the possibility that the reduction in breast cancer mortality may be smaller than generally assumed and that the harms may be greater. In view of the current controversy, this thesis examines three issues that form part of the debate. Because estimation of screening benefits is based on comparisons between screened and unscreened women, we focused on cancer detection rates among unscreened women in a region where screening is population-based. We next examined interval cancers, which are a short-term indicator of screening quality. Few studies have provided data on the association between molecular patterns and prognosis. Finally, we studied the impact of the switch to digital mammography on recall rates, focusing on the role of radiological patterns and tumour characteristics.

**Objectives.** This thesis aims to provide greater knowledge of the evaluation of population-based breast cancer screening programmes, specifically by i) assessing the impact of the introduction of screening programmes on detection rates in the entire target population, ii) studying the association of the mode of diagnosis (detected at screening or in the interval between screenings) with tumour features and cancer prognosis, and iii) evaluating the impact of the imaging modality on performance indicators.

**Methods.** Four studies have been carried out in 3 cohorts of women targeted for screening in Barcelona, Spain, and Denmark. For the first objective, we assessed the breast cancer incidence rates of the target populations of two Danish regions, from the introduction of population-based screening in 1991 to 2008. We identified breast cancers diagnosed at screening, during the screening interval, or among unscreened women ( $n=3,290$ ,  $n=1,257$  and  $n=1,548$ , respectively). We computed rates by mode of diagnosis and compared tumour-related characteristics and performance indicators between the two regions.

For the second objective, we focused on the study of interval cancers in the cohort of women screened between 1995 and 2008 at Hospital del Mar, Barcelona. We designed a case-control study to compare the biological characteristics of interval cancers ( $n=115$ ) and screen-detected cancers ( $n=115$ ). Interval cancers were radiologically classified following the European Guidelines into true interval cancers, false negatives, occult tumours, and minimal-sign cancers. Bivariate and multivariate analyses were performed to compare patient and molecular characteristics among all interval cancers, true interval cancers, and false negatives vs. screen-detected cancers. We carried out a follow-up study in the same cohort of women to evaluate overall survival and disease-free survival in women with cancer detected by screening,

outside screening, and those that truly appeared between two screening mammograms. Hence, we restricted this analysis to screen-detected cancers (n=98) and true interval cancers (n=34), and we incorporated cases from unscreened women (n=98). The probabilities of overall survival (based on mortality) and disease-free survival (based on relapses) were computed using the Kaplan–Meier method. Cox proportional hazard models were applied, with adjustment by clinical–pathological variables.

For the third objective, we evaluated the impact of the transition from screen-film mammography to digital mammography among a cohort of 103,613 women aged 45-69 screened in Spain between 1996 and 2007. Tumour characteristics (invasiveness, tumour size, lymph node involvement) and radiological patterns were compared among study groups. We also compared the positive predictive values (PPV= cancers diagnosed among the total number of women recalled for further assessments) obtained during the screen-film and digital periods.

Both database management and the statistical analyses were performed with SPSS (version 12.0 and 18.0) and SAS (version 9.2). Graphs were generated with R statistical software (version 2.12.2). An alpha error of 5% is assumed in the whole dissertation, and therefore p-values of <0.05 were considered statistically significant.

**Results.** i) The introduction of population-based screening in Denmark resulted in a prevalence peak of breast cancer diagnoses not only among screening participants (5.79 per 1000 persons-year in Copenhagen and 4.47 in Fyn), but also among unscreened women (4.23 and 5.93 in Copenhagen and Fyn, respectively). In Fyn, the breast cancer incidence among unscreened women was even higher than among screened women. The two programmes differed in eligibility criteria and participation patterns. ii) The study of interval cancers revealed differences in stage at diagnosis and molecular status. In addition, women with interval cancers were more likely to have dense breasts. Specifically, true interval cancers showed a strong association with the triple-negative phenotype (lack of hormone receptor and HER2 expression) [OR= 8.85 (95%CI: 2.03–38.62)] and high breast density [3.08 (95%CI: 1.03–9.24)]. True interval cancers had the highest hazard ratio for relapse prediction (1.89; 0.67–5.31) and a hazard ratio of death of 5.55 (1.61–19.15) after adjustment for tumour–node–metastasis stage and phenotype. iii) PPV was higher for digital mammography than for screen-film mammography (7.0% vs. 5.5%, respectively;  $p < 0.001$ ). The greatest improvements were found for masses and calcifications. A slightly rising temporal trend for the detection rate of *in situ* carcinomas was observed in the digital period.

**Conclusions.** i) Data from two long-standing population-based screening programmes conducted in Denmark showed that the implementation of screening programmes led to a prevalence peak of breast cancer detection among screened and unscreened women, and that performance indicators could be strongly influenced by eligibility criteria. ii) True interval cancers were associated with a more aggressive molecular phenotype and high breast density. These cancers were also associated with poorer survival

even after adjustment for clinical–pathological characteristics. iii) The introduction of digital mammography in Spain improved positive predictive value and suggested increased detection of *in situ* carcinomas.

**Further research.** The experience gained in the study of interval cancers in Hospital del Mar was key to the initiation of a nationwide project for the study of rates and determinants of interval cancers, with information from more than 600,000 screened women in Spain and over 1,000 radiologically classified interval cancers (INCA Study). We also aimed to determine the impact of digital mammography on the number of interval cancers and false negatives and have recently obtained funding to complete the information with clinical follow-up. Furthermore, in the current debate about the balance of benefits and risks of breast cancer screening, and the technological improvements occurring over time, there is a need to consider screening optimization to offset higher costs with better results. To evaluate more personalized screening strategies, data from large cohorts of screened women will be needed to determine the impact of strategies aiming to maximize benefits and decrease harms.

## OBJECTIVES

The general objective of this dissertation is to examine the following issues related to population-based breast cancer screening:

- The impact of the introduction of screening on the detection of breast cancer
- The association of the mode of diagnosis with tumour features and clinical outcomes
- The impact of the characteristics of the imaging modality on quality indicators

The specific objectives are as follows:

The impact of the introduction of screening on breast cancer detection:

1. To describe the detection rates of breast cancer in the entire target population, according to mode of diagnosis (detected at screening, in the interval between screenings, or outside screening).
2. To analyse the impact of organizational differences in screening programmes on breast cancer detection rates inside and outside screening.

The association of mode of diagnosis with tumour features and clinical outcomes

3. To compare the biological characteristics of interval cancers and their subtypes with those of screen-detected cancers.
4. To estimate the impact of the mode of diagnosis and other variables, such as biological characteristics and biomarkers, on survival in women with breast cancer.

The impact of the characteristics of the imaging modality on quality indicators:

5. To compare the PPV of screening mammography before and after the implementation of digital mammography.
6. To evaluate changes in the PPVs for distinct radiological patterns in recalled women with screen-film mammography and digital mammography.
7. To compare tumour characteristics between cancers detected with screen-film mammography and digital mammography.

## METHODS

The four studies were carried out within the framework of population-based screening, but each analysed different populations. To facilitate understanding of the complete thesis, the table below lists the specific objectives of the thesis, the article that answers each research question, its study design and study period, and the main features of the study population.

Objectives	Articles	Study design	Study period	Study population
<b>The impact of the introduction of screening on breast cancer detection</b>				
1. Detection rates inside and outside screening	Domingo et al. <i>Acta Oncol.</i> 2013 Jan;52(1):48-56.	Retrospective cohort	April 1991 –April 2008	Target population of population-based screening programmes in Copenhagen and Fyn, Denmark Women aged 50-69 years 3,290 screen-detected cancers 1,257 interval cancers 1,548 cancers in unscreened women
2. Impact of organizational differences on detection rates			Follow-up until January, 2011 for interval cancer assessment	
<b>The association of mode of diagnosis with tumour features and clinical outcomes</b>				
3. Biological features of interval cancers vs. screen-detected cancers	Domingo et al. <i>Cancer Causes Control</i> 2010 Aug;21(8):1155-64.	Nested case-control	November 1995-September 2008	Women screened in the population-based screening programme of Hospital del Mar, Barcelona, Spain Women aged 50-69 years 115 screen-detected cancers 115 interval cancers
4. Prognosis according to setting and time of diagnosis	Domingo et al. <i>Eur J Cancer Prev</i> 2013 Jan;22(1):21-8.	Follow-up study	Cases diagnosed between November 1995 and September 2008 Follow-up until December 2010	Breast cancers diagnosed and treated at Hospital del Mar, Barcelona, Spain Women aged 50-69 years 97 screen-detected cancers 34 true interval cancers 97 cancers detected outside screening
<b>The impact of the characteristics of the imaging modality on quality indicators</b>				
5. PPV of screen-film mammography and digital mammography	Domingo et al. <i>Eur Radiol.</i> 2011 Sep;21(9):2020-8.	Retrospective cohort	January 1996-December 2007	Women screened in 4 population-based breast cancer screening programmes in Spain Women aged 45/50-69 years 171,191 screen-film mammograms 91,647 digital mammograms
6. PPV according to radiological patterns with screen-film mammography and digital mammography				
7. Comparison of tumour characteristics between cancers detected with screen-film mammography and digital mammography				



## ARTICLES

The following four articles were published as Original Articles in international journals indexed in the subject category of 'Public, Environmental and Occupational Health', 'Oncology' and 'Radiology, Nuclear Medicine and Medical Imaging'.

**Article 1.** Seventeen-years overview of breast cancer inside and outside screening in Denmark.

Laia Domingo, Katja K. Jacobsen, My von Euler-Chelpin, Ilse Vejborg, Walter Schwartz, Maria Sala, Elsebeth Lynge.

Acta Oncol. 2013 Jan;52(1):48-56.

**Article 2.** Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain.

Laia Domingo, Maria Sala, Sònia Servitja, Josep Maria Corominas, Francesc Ferrer, Juan Martínez, Francesc Macià, Maria Jesús Quintana, Joan Albanell, Xavier Castells.

Cancer Causes Control 2010 Aug;21(8):1155-64.

**Article 3.** Aggressiveness features and outcomes of true interval cancers: comparison between screen-detected and symptom-detected cancers.

Laia Domingo, Jordi Blanch, Sònia Servitja, Josep Maria Corominas, Cristiane Murta-Nascimento, Antonio Rueda, Maximino Redondo, Xavier Castells, Maria Sala.

Eur J Cancer Prev 2013 Jan;22(1):21-8.

**Article 4.** Differences in radiological patterns, tumour characteristics and diagnostic precision between digital mammography and screen-film mammography in four breast cancer screening programmes in Spain.

Laia Domingo, Anabel Romero, Francesc Belvis, Mar Sánchez, Joana Ferrer, Dolores Salas, Josefa Ibáñez, Alfonso Vega, Francesc Ferrer, Maria Soledad Laso, Francesc Macià, Xavier Castells, Maria Sala.

Eur Radiol. 2011 Sep;21(9):2020-8.

## 1. Article 1

**Title:** Seventeen-years overview of breast cancer inside and outside screening in Denmark. Seventeen-years overview of breast cancer inside and outside screening in Denmark

**Authors:** Domingo L, Jacobsen KK, Euler-Chelpin M, Vejborg I, Schwartz W, Sala M, Lynge E.

**Journal:** Acta Oncol. 2013 Jan;52(1):48-56.

**Impact factor** (2011): 3.330 (Q2 Oncology)

**DOI:** 10.3109/0284186X.2012.698750

### **Abstract:**

**BACKGROUND:** Long-term data on breast cancer detection in mammography screening programs are warranted to better understand the mechanisms by which screening changes the breast cancer pattern in the population. We aimed to analyze 17 years of breast cancer detection rates inside and outside screening in two Danish regions, emphasizing the influence of organizational differences of screening programs on the outcomes.

**MATERIAL AND METHODS:** We used data from two long-standing population-based mammography screening programs, Copenhagen and Fyn, in Denmark. Both programs offered biennial screening to women aged 50-69 years. We identified targeted, eligible, invited and participating women. We calculated screening detection and interval cancer rates for participants, and breast cancer incidence in non-screened women (= targeted women excluding participants) by biennial invitation rounds. Tumor characteristics were tabulated for each of the three groups of cancers.

**RESULTS:** Start of screening resulted in a prevalence peak in participants, followed by a decrease to a fairly stable detection rate in subsequent invitation rounds. A similar pattern was found for breast cancer incidence in non-screened women. In Fyn, non-screened women even had a higher rate than screening participants during the first three invitation rounds. The interval cancer rate was lower in Copenhagen than in Fyn, with an increase over time in Copenhagen, but not in Fyn. Screen-detected cancers showed tumor features related with a better prognosis than tumors detected otherwise, as more than 80% were smaller than 20 mm and estrogen receptor positive.

**CONCLUSION:** Data from two long-standing population-based screening programs in Denmark illustrated that even if background breast cancer incidence and organization were rather similar, performance indicators of screening could be strongly influenced by inclusion criteria and participation rates. Detection rates should be interpreted with caution as they may be biased by selection into the screening population.

## 2. Article 2

**Title:** Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain.

**Authors:** Domingo L, Sala M, Servitja S, Corominas JM, Ferrer F, Martínez J, Macià F, Quintana MJ, Albanell J, Castells X

**Journal:** Cancer Causes Control 2010 Aug;21(8):1155-64.

**Impact factor** (2009): 3.199 (Q1 Public, Environmental & Occupational Health; Q2 Oncology)

**DOI:** 10.1007/s10552-010-9541-6

### **Abstract:**

**OBJECTIVE:** To analyze phenotypic classification and other risk factors for interval breast cancer, focusing on true interval and false negative cancers.

**METHODS:** A nested case-control study was performed among 115 cancers detected between two screening mammograms (interval cancers) and 115 screen-detected cancers diagnosed between 1995 and 2008 in a population-based breast cancer screening program in Barcelona (Spain). Bivariate and multivariate analyses were performed to compare patient and tumor molecular characteristics among all interval cancers, true intervals and false negatives, and screen-detected cancers.

**RESULTS:** A total of 42.5% of interval cancers were true interval tumors and 16.2% were false negatives. High breast density and triple negative phenotype were more frequent in true interval cancers than in screen-detected cancers (57.6 and 34.1%, respectively for breast density,  $p = 0.023$ ; 28.1 and 7.5%, respectively for triple negative phenotype,  $p = 0.028$ ), while no statistically significant differences were observed between false negatives and screen-detected cancers. The main adjusted factors associated with true interval cancers compared with screen-detected cancers were high breast density and triple negative phenotype (OR = 3.1, 95% CI, 1.03-9.24 and OR = 8.9, 95% CI, 2.03-38.62, respectively).

**CONCLUSION:** A more aggressive molecular phenotype and high breast density were identified in breast tumors that truly arise in the interval between screenings.

### 3. Article 3

**Title:** Aggressiveness features and outcomes of true interval cancers: comparison between screen-detected and symptom-detected cancers.

**Authors:** Domingo L, Blanch J, Servitja S, Corominas JM, Murta-Nascimento C, Rueda A, Redondo M, Castells X, Sala M.

**Journal:** Eur J Cancer Prev 2013 Jan;22(1):21-8.

**Impact factor** (2011): 2.130 (Q3 Oncology)

**DOI:** 10.1097/CEJ.0b013e328354d324.

#### **Abstract:**

**BACKGROUND:** The question of whether screen detection confers an additional survival benefit in breast cancer is unclear and subject to several biases. Our aim was to examine the role of the diagnostic method (screen-detected, symptom-detected, and true interval cancers) and the clinical-pathological features in relapse-free survival and overall survival in breast cancer patients.

**METHODS:** We included 228 invasive breast cancers diagnosed in Barcelona from 1996 to 2008 among women aged 50-69 years. Ninety-seven patients were screen detected within the screening, 34 truly arose between 2-year screening mammograms (true interval cancers), and 97 were symptom detected outside the screening. The clinical-pathological features at diagnosis were compared. The overall and disease-free survival probabilities were computed using the Kaplan-Meier method. Cox proportional hazard models were applied, with adjustment by clinical-pathological variables.

**RESULTS:** At diagnosis, symptom-detected and true interval cancers were in more advanced stages and were less differentiated. The highest proportion of triple-negative cancers was detected among true interval cancers ( $P=0.002$ ). At 5 years of follow-up, the disease-free survival rates for screen-detected, true interval, and symptom-detected cancers were 87.5% (95% confidence interval, 80.5-95.2%), 64.1% (46.4-88.5%), and 79.4% (71.0-88.8%), respectively, and the overall survival rates were 94.5% (89.3-99.9%), 65.5% (47.1-91.2%), and 85.6% (78.3-93.6%), respectively. True interval cancers had the highest hazard ratio for relapse prediction (1.89; 0.67-5.31) and a hazard ratio of death of 5.55 (1.61-19.15) after adjustment for tumor-node-metastasis stage and phenotype.

**CONCLUSION:** Clinically detected tumors, especially true interval cancers, more frequently showed biological features related to worse prognosis and were associated with poorer survival even after adjustment for clinical-pathological characteristics.

#### 4. Article 4

**Title:** Differences in radiological patterns, tumour characteristics and diagnostic precision between digital mammography and screen-film mammography in four breast cancer screening programmes in Spain.

**Authors:** Domingo L, Romero A, Belvis F, Sánchez M, Ferrer J, Salas D, Ibáñez J, Vega A, Ferrer F, Laso MS, Macià F, Castells X, Sala M.

**Journal:** Eur Radiol. 2011 Sep;21(9):2020-8.

**Impact factor** (2010): 3.594 (Q1 Radiology, Nuclear Medicine and Medical Imaging)

**DOI:** 10.1007/s00330-011-2143-1

#### **Abstract:**

**OBJECTIVES:** To compare tumour characteristics between cancers detected with screen-film mammography (SFM) and digital mammography (DM) and to evaluate changes in positive predictive values (PPVs) for further assessments, for invasive procedures and for distinct radiological patterns in recalled women.

**METHODS:** 242,838 screening mammograms (171,191 SFM and 71,647 DM) from 103,613 women aged 45-69 years, performed in four population-based breast cancer screening programmes in Spain, were included. The tumour characteristics and PPVs of each group were compared. Radiological patterns (masses, calcifications, distortions and asymmetries) among recalled women were described and PPVs were evaluated.

**RESULTS:** The percentages of ductal carcinoma *in situ* (DCIS) were higher in DM than in SFM both in the first [18.5% vs. 15.8%(p=0.580)] and in successive screenings [23.2% vs. 15.7%(p=0.115)]. PPVs for masses, asymmetries and calcifications were higher in DM, being statistically significant in masses (5.3% vs. 3.9%; proportion ratio: 1.37 95%CI: 1.08-1.72). Among cancers detected by calcifications, the percentage of DCIS was higher in DM (60.3% vs. 46.4%, p=0.060).

**CONCLUSIONS:** PPVs were higher when DM was used, both for further assessments and for invasive procedures, with similar cancer detection rates and no statistically significant differences in tumour characteristics. The greatest improvements in PPVs were found for masses

## CONCLUSIONS AND RECOMMENDATIONS

### 1. Conclusions

1. The implementation of population-based screening programmes has increased breast cancer detection among women not participating in the screening programmes. This effect should be considered when measuring and comparing the benefits and outcomes of screening among the participating and non-participating populations.
2. The organizational features and eligibility criteria of the target population affect breast cancer detection rates inside and outside screening. Criteria based on disease risk may influence the incidence of the disease in the participating and non-participating population. Thus, excluding women in the target population with potential risk factors for breast cancer can increase the detection rate among non-participants.
3. Cancers diagnosed during the interval between 2 screening mammograms were mostly undetectable at the last mammography (ie, were true interval cancers), and almost a 20% in our setting were false-negative tumours.
4. True interval cancers showed a molecular profile associated with aggressiveness, with a high proportion of triple-negative tumours.
5. The proportion of women with dense breasts and true interval cancers was higher than that described among women with tumours detected in screening.
6. True interval cancers were more likely to show features related to a worse prognosis and were associated with poorer survival than screen-detected tumours, even after adjustment for clinicopathological characteristics.
7. The introduction of digital mammography in Spanish screening programmes has improved diagnostic accuracy, resulting in a statistically significant improvement in the PPV of the screening test.
8. The greatest improvement in PPVs in the digital era was found for radiologically suspicious masses and calcifications. The improved PPV for masses was attributed to decreased recall due to this radiological pattern, and that for calcifications to an increase in cancer diagnoses.

9. The differences in tumour characteristics indicated increased detection of *in situ* carcinomas and of less advanced cancers with digital mammography.

## 2. Recommendations

The results of this dissertation provide further information that could be useful in the assessment of the balance of benefits and harms of screening. When evaluating the benefits, an important consideration is that non-participants are also influenced by screening, which may partially dilute the differences between the two populations.

Because organizational features may bias the detection rates of screening programmes, evaluation of screening should start with an accurate description of the target population, and eligibility criteria and participation rates should be included in the interpretation and comparison of results between screening programmes.

The results presented herein could contribute to providing adequate information to women invited to participate in population-based breast cancer screening programs. As noted, one of the positive effects of screening programmes is that they reassure women with a negative result. This effect, however, may lead to a false sense of security and contribute to diagnostic delay. Women with a negative screening result should be instructed to remain vigilant for the possible signs of disease between two screening mammograms.

Information on the biological characterization of interval cancers may be useful to assess future strategies aiming to optimize screening, by reviewing, for instance, the frequency of screening mammograms, or the use of different imaging modalities for specific subsets of women with high breast density.

The time and effort spent on reviewing interval cancers is offset by the information provided, which is useful to understand the characteristics of tumours diagnosed between two screening mammograms. Such classification also may also provide new insights into the development of tumours with a short sojourn time and into prognostic factors of the disease.

The promotion of population-based cancer registries to cover the entire Catalan and Spanish population should be encouraged. Population-based registries are the main tool for screening evaluation and are essential to estimate the effectiveness of screening. In many Spanish regions, the lack of such registries limits the study of important issues such as overdiagnosis or the occurrence of interval cancer and also hampers comparisons between countries and prevents assessment of the global impact of early detection in our area.

The implementation of digital mammography in all screening programmes in Spain should be promoted. This modality has technical advantages for image acquisition and storage and, although it has not improved detection rates in Spain, it has reduced recall rates. A reduction in the number of false-positive results would prevent a large number of women from experiencing anxiety and would reduce related costs.

The increase in the detection of less advanced cancers with digital mammography could lead to these women receiving less aggressive treatments, which would improve their quality of life. Nonetheless, a long period is needed to determine whether the increase in the detection of *in situ* carcinomas leads to an early diagnosis or to a worrisome increase in overdiagnosis. We encourage further works exploring the outcomes of cohorts of women screened only with digital mammography.



## **VII. BIBLIOGRAFIA**

---



1. Frame PS, Carlson SJ. A critical review of periodic health screening using specific screening criteria. Part 1: Selected diseases of respiratory, cardiovascular, and central nervous systems. *J Fam Pract* 1975; 2(1):29-36.
2. Frame PS, Carlson SJ. A critical review of periodic health screening using specific screening criteria. Part 2: Selected endocrine, metabolic and gastrointestinal diseases. *J Fam Pract* 1975; 2(2):123-129.
3. Frame PS, Carlson SJ. A critical review of periodic health screening using specific screening criteria. Part 3. Selected diseases of the genitourinary system. *J Fam Pract* 1975; 2(3):189-194.
4. Frame PS, Carlson SJ. A critical review of periodic health screening using specific screening criteria. Part 4: selected miscellaneous diseases. *J Fam Pract* 1975; 2(4):283-289.
5. Commission of the European Communities. Council Recommendation on Cancer Screening 2003/0093. Brussels: Commission of the European Communities, 2003.
6. Ascunce N, Salas D, Zubizarreta R, Almazan R, Ibanez J, Ederria M. Cancer screening in Spain. *Ann Oncol* 2010; 21 Suppl 3:iii43-iii51.
7. Castells X, Sala M, Ascunce N, Salas D, Zubizarreta R, Casamiñjana M. Descripción del cribado del Cáncer en España. Proyecto DESCRIC. Madrid: Plan Nacional para el SNS del MSC. Agència d'Avaluació de Tecnologia i Recerca Mèdiques. Informes de Evaluación de Tecnologías Sanitarias, 2007.
8. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011; 61(2):69-90.
9. Adami H-O, Hunter D, Trichopoulos D. Textbook of cancer epidemiology, 2nd Edition. Oxford University Press, 2008.
10. Ferlay J, Shin HR, Forman D, Mathers C, Parkin DM. GLOBOCAN 2008 v2.0, Cancer Incidence and Mortality Worldwide: IARC. Lyon, France: International Agency for Research on Cancer; 2010. Available from: <http://globocan.iarc.fr>, accessed on 15/11/2012.
11. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010; 127(12):2893-2917.
12. Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiol Biomarkers Prev* 2010; 19(8):1893-1907.

13. Luengo-Matos S, Polo-Santos M, Saz-Parkinson Z. Mammography use and factors associated with its use after the introduction of breast cancer screening programmes in Spain. *Eur J Cancer Prev* 2006; 15(3):242-248.
14. 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):641-644.
15. Hulka BS, Moorman PG. Breast cancer: hormones and other risk factors. *Maturitas* 2008; 61(1-2):203-213.
16. Parkin DM. Is the recent fall in incidence of post-menopausal breast cancer in UK related to changes in use of hormone replacement therapy? *Eur J Cancer* 2009; 45(9):1649-1653.
17. Pollán 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):1584-1591.
18. 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):1670-1674.
19. Seradour B, Allemand H, Weill A, Ricordeau P. Changes by age in breast cancer incidence, mammography screening and hormone therapy use in France from 2000 to 2006. *Bull Cancer* 2009; 96(4):E1-E6.
20. 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* 2012.
21. Autier P, Boniol M, La Vecchia C, Vatten L, Gavin A, Hery C et al. Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database. *BMJ* 2010; 341:c3620.
22. Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med* 2010; 363(13):1203-1210.
23. Bennett IC, Gattas M, Teh BT. The genetic basis of breast cancer and its clinical implications. *Aust N Z J Surg* 1999; 69(2):95-105.
24. Warner E. Clinical practice. Breast-cancer screening. *N Engl J Med* 2011; 365(11):1025-1032.
25. Oza AM, Boyd NF. Mammographic parenchymal patterns: a marker of breast cancer risk. *Epidemiol Rev* 1993; 15(1):196-208.
26. 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):227-236.

27. Martin LJ, Boyd NF. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. *Breast Cancer Res* 2008; 10(1):201.
28. 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.
29. Wellings SR, Jensen HM. On the origin and progression of ductal carcinoma in the human breast. *J Natl Cancer Inst* 1973; 50(5):1111-1118.
30. 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):231-273.
31. Allred DC, Mohsin SK, Fuqua SA. Histological and biological evolution of human premalignant breast disease. *Endocr Relat Cancer* 2001; 8(1):47-61.
32. Zagouri F, Sergentanis TN, Zografos GC. Precursors and preinvasive lesions of the breast: the role of molecular prognostic markers in the diagnostic and therapeutic dilemma. *World J Surg Oncol* 2007; 5:57.
33. Reis-Filho JS, Lakhani SR. The diagnosis and management of pre-invasive breast disease: genetic alterations in pre-invasive lesions. *Breast Cancer Res* 2003; 5(6):313-319.
34. Ma XJ, Salunga R, Tuggle JT, Gaudet J, Enright E, McQuary P et al. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci U S A* 2003; 100(10):5974-5979.
35. Allred DC, Wu Y, Mao S, Nagtegaal ID, Lee S, Perou CM et al. Ductal carcinoma in situ and the emergence of diversity during breast cancer evolution. *Clin Cancer Res* 2008; 14(2):370-378.
36. 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):229-237.
37. Allred DC. Ductal carcinoma in situ: terminology, classification, and natural history. *J Natl Cancer Inst Monogr* 2010; 2010(41):134-138.
38. 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):280-286.
39. Evans A, Pinder S, Wilson R, Sibbering M, Poller D, Elston C et al. Ductal carcinoma in situ of the breast: correlation between mammographic and pathologic findings. *AJR Am J Roentgenol* 1994; 162(6):1307-1311.
40. Evans A. The diagnosis and management of pre-invasive breast disease: radiological diagnosis. *Breast Cancer Res* 2003; 5(5):250-253.
41. Ernster VL, Barclay J. Increases in ductal carcinoma in situ (DCIS) of the breast in relation to mammography: a dilemma. *J Natl Cancer Inst Monogr* 1997;(22):151-156.

42. Perou CM, Sorlie T, Eisen MB, van de RM, Jeffrey SS, Rees CA et al. Molecular portraits of human breast tumours. *Nature* 2000; 406(6797):747-752.
43. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* 2011; 22(8):1736-1747.
44. Allred DC. Molecular biomarkers of risk in premalignancy and breast cancer prevention. *Cancer Prev Res (Phila)* 2011; 4(12):1947-1952.
45. Bick U, Diekmann F. Digital mammography: what do we and what don't we know? *Eur Radiol* 2007; 17(8):1931-1942.
46. Hauge IH, Pedersen K, Sanderud A, Hofvind S, Olerud HM. Patient doses from screen-film and full-field digital mammography in a population-based screening programme. *Radiat Prot Dosimetry* 2012; 148(1):65-73.
47. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Bohm-Velez M et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA* 2008; 299(18):2151-2163.
48. Corsetti V, Houssami N, Ferrari A, Ghirardi M, Bellarosa S, Angelini O et al. Breast screening with ultrasound in women with mammography-negative dense breasts: evidence on incremental cancer detection and false positives, and associated cost. *Eur J Cancer* 2008; 44(4):539-544.
49. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA* 2012; 307(13):1394-1404.
50. Kilburn-Toppin F, Barter SJ. New Horizons in Breast Imaging. *Clin Oncol (R Coll Radiol)* 2013; 25(2):93-100.
51. Abramson RG, Arlinghaus LR, Weis JA, Li X, Dula AN, Chekmenev EY et al. Current and emerging quantitative magnetic resonance imaging methods for assessing and predicting the response of breast cancer to neoadjuvant therapy. *Breast Cancer (London)* 2012; 2012(4):139-154.
52. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991; 19(5):403-410.
53. Giuliano AE, McCall L, Beitsch P, Whitworth PW, Blumencranz P, Leitch AM et al. Locoregional recurrence after sentinel lymph node dissection with or without axillary dissection in patients with

- sentinel lymph node metastases: the American College of Surgeons Oncology Group Z0011 randomized trial. *Ann Surg* 2010; 252(3):426-432.
54. Verry H, Lord SJ, Martin A, Gill G, Lee CK, Howard K et al. Effectiveness and cost-effectiveness of sentinel lymph node biopsy compared with axillary node dissection in patients with early-stage breast cancer: a decision model analysis. *Br J Cancer* 2012; 106(6):1045-1052.
  55. Ragaz J, Olivotto IA, Spinelli JJ, Phillips N, Jackson SM, Wilson KS et al. Locoregional radiation therapy in patients with high-risk breast cancer receiving adjuvant chemotherapy: 20-year results of the British Columbia randomized trial. *J Natl Cancer Inst* 2005; 97(2):116-126.
  56. Jahanzeb M. Adjuvant trastuzumab therapy for HER2-positive breast cancer. *Clin Breast Cancer* 2008; 8(4):324-333.
  57. Paci E. 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.
  58. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012; 380(9855):1778-1786.
  59. 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):349-355.
  60. 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):1490-1499.
  61. 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):943-948.
  62. 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):909-919.
  63. Bjurstam N, Bjorneld L, Duffy SW, Smith TC, Cahlin E, Erikson O et al. The Gothenburg Breast Cancer Screening Trial: preliminary results on breast cancer mortality for women aged 39-49. *J Natl Cancer Inst Monogr* 1997;(22):53-55.
  64. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. 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):2053-2060.
  65. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2006;(4):CD001877.
  66. Hakama M, Coleman MP, Alexe DM, Auvinen A. Cancer screening: evidence and practice in Europe 2008. *Eur J Cancer* 2008; 44(10):1404-1413.

67. Holland W, Steward S. Policy brief: Screening in Europe. European Observatory on Health Systems and Policies, World Health Organization 2006.
68. Europe Against Cancer Programme of the European Community. European Code Against Cancer. Third version 2003.
69. Perry N, Broeders M, de Wolf C, Törnberg S, Hollande R, von Karsa L, Puthaar E (eds). European guidelines for quality assurance in breast cancer screening and diagnosis - Fourth edition. Luxembourg: Office for Official Publications of the European Communities; 2006.
70. von Karsa L, Anttila A, Ronco G, Ponti A, Malila N, Arbyn M et al. Cancer Screening in the European Union. Report on the implementation of the Council Recommendation on cancer screening. International Agency for Research on Cancer (IARC) 2008.
71. Red de Programas de Cribado. <http://www.programascancerdemama.org/>. Accessed on 13-2-2013.
72. 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):516-526.
73. Stomper PC, Kopans DB, Sadowsky NL, Sonnenfeld MR, Swann CA, Gelman RS et al. Is mammography painful? A multicenter patient survey. *Arch Intern Med* 1988; 148(3):521-524.
74. 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.
75. 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):104-111.
76. Brett J, Bankhead C, Henderson B, Watson E, Austoker J. The psychological impact of mammographic screening. A systematic review. *Psychooncology* 2005; 14(11):917-938.
77. Alamo-Junquera D, Murta-Nascimento C, Macia F, Bare M, Galceran J, Ascunce N et al. Effect of false-positive results on reattendance at breast cancer screening programmes in Spain. *Eur J Public Health* 2012; 22(3):404-408.
78. 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.
79. Andersen SB, Vejborg I, Euler-Chelpin M. Participation behaviour following a false positive test in the Copenhagen mammography screening programme. *Acta Oncol* 2008; 47(4):550-555.



80. Burman ML, Taplin SH, Herta DF, Elmore JG. Effect of false-positive mammograms on interval breast cancer screening in a health maintenance organization. *Ann Intern Med* 1999; 131(1):1-6.
81. Törnberg S, Kemetli L, Ascunce N, Hofvind S, Anttila A, Seradour B et al. A pooled analysis of interval cancer rates in six European countries. *Eur J Cancer Prev* 2010; 19(2):87-93.
82. Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J Med Screen* 2012; 19 Suppl 1:42-56.
83. Raffle A, Gray M. *Screening. Evidence and practice*. New York, USA: Oxford University Press, 2007.
84. Paci E, Warwick J, Falini P, Duffy SW. Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? *J Med Screen* 2004; 11(1):23-27.
85. Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. *BMJ* 2004; 328(7445):921-924.
86. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009; 339:b2587.
87. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan B, Nygren P et al. Screening for Breast Cancer: An Update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009; 151(10):727-737.
88. Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ* 2011; 343:d4411.
89. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010; 17(1):25-30.
90. Jorgensen KJ, Zahl PH, Gotzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. *BMC Womens Health* 2009; 9:36.
91. Tabar L, Yen MF, Vitak B, Chen HH, Smith RA, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *Lancet* 2003; 361(9367):1405-1410.
92. Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000; 355(9198):129-134.
93. Gøtzsche PC. *Mammography Screening: Truth, Lies and Controversy*. Radcliff e Publishing, 2012.
94. Hackshaw A. The benefits and harms of mammographic screening for breast cancer: building the evidence base using service screening programmes. *J Med Screen* 2012; 19 Suppl 1:1-2.

95. Gøtzsche P. Why the results from the EUROSCREEN Working Group are false. Nordic Cochrane Centre. 2012.
96. Dubin N, Friedman DR, Toniolo PG, Pasternack BS. Breast cancer detection centers and case-control studies of the efficacy of screening. *J Chronic Dis* 1987; 40(11):1041-1050.
97. Duffy SW, Cuzick J, Tabar L, Vitak B, Chen THH, Yen MF. Correcting for non-compliance bias in case-control studies to evaluate cancer screening programmes. *Appl Statist* 2002;(51):235-243.
98. Autier P, Boniol M. Pitfalls in using case-control studies for the evaluation of the effectiveness of breast screening programmes. *Eur J Cancer Prev* 2012.
99. Duffy SW, Nagtegaal ID, Wallis M, Cafferty FH, Houssami N, Warwick J et al. Correcting for lead time and length bias in estimating the effect of screen detection on cancer survival. *Am J Epidemiol* 2008; 168(1):98-104.
100. Fletcher RH, Fletcher SW. *Clinical Epidemiology. The essentials.* 4th Edition ed. 2005.
101. Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. *J Natl Cancer Inst* 2005; 97(16):1195-1203.
102. Dawson SJ, Duffy SW, Blows FM, Driver KE, Provenzano E, LeQuesne J et al. Molecular characteristics of screen-detected vs symptomatic breast cancers and their impact on survival. *Br J Cancer* 2009; 101(8):1338-1344.
103. Kafadar K, Prorok PC. Effect of length biased sampling of unobserved sojourn times on the survival distribution when disease is screen detected. *Stat Med* 2009; 28(16):2116-2146.
104. Mahnken JD, Chan W, Freeman DH, Jr., Freeman JL. Reducing the effects of lead-time bias, length bias and over-detection in evaluating screening mammography: a censored bivariate data approach. *Stat Methods Med Res* 2008; 17(6):643-663.
105. Joensuu H, Lehtimäki T, Holli K, Elomaa L, Turpeenniemi-Hujanen T, Kataja V et al. Risk for distant recurrence of breast cancer detected by mammography screening or other methods. *JAMA* 2004; 292(9):1064-1073.
106. Wishart GC, Greenberg DC, Britton PD, Chou P, Brown CH, Purushotham AD et al. Screen-detected vs symptomatic breast cancer: is improved survival due to stage migration alone? *Br J Cancer* 2008; 98(11):1741-1744.
107. Hofvind S, Yankaskas BC, Bulliard JL, Klabunde CN, Fracheboud J. Comparing interval breast cancer rates in Norway and North Carolina: results and challenges. *J Med Screen* 2009; 16(3):131-139.

108. Jensen A, Olsen AH, Euler-Chelpin M, Helle NS, Vejborg I, Lynge E. Do nonattenders in mammography screening programmes seek mammography elsewhere? *Int J Cancer* 2005; 113(3):464-470.
109. Puig-Vives M, Pollan M, Rue M, Osca-Gelis G, Saez M, Izquierdo A et al. Rapid increase in incidence of breast ductal carcinoma in situ in Girona, Spain 1983-2007. *Breast* 2012; 21(5):646-651.
110. Euler-Chelpin M, Lynge E, Rebolj M. Register-based studies of cancer screening effects. *Scand J Public Health* 2011; 39(7 Suppl):158-164.
111. Olsen J. Register-based research: some methodological considerations. *Scand J Public Health* 2011; 39(3):225-229.
112. Hofvind S, Wang H, Thoresen S. Do the results of the process indicators in the Norwegian Breast Cancer Screening Program predict future mortality reduction from breast cancer? *Acta Oncol* 2004; 43(5):467-473.
113. Olsen AH, Agbaje OF, Myles JP, Lynge E, Duffy SW. Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program. *Breast J* 2006; 12(4):338-342.
114. Elena PM, Nehmat H, Ermes M, Piera C, Maria Q, Guia M et al. Quality of mammography screening in the Milan programme: Evidence of improved sensitivity based on interval cancer proportional incidence and radiological review. *Breast* 2009; 18(3):208-10.
115. 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):447-455.
116. Törnberg S, Codd M, Rodrigues V, Segnan N, Ponti A. Ascertainment and evaluation of interval cancers in population-based mammography screening programmes: a collaborative study in four European centres. *J Med Screen* 2005; 12(1):43-49.
117. Ciatto S, Catarzi S, Lamberini MP, Risso G, Saguatti G, Abbattista T et al. Interval breast cancers in screening: the effect of mammography review method on classification. *Breast* 2007; 16(6):646-652.
118. Hofvind S, Geller B, Skaane P. Mammographic features and histopathological findings of interval breast cancers. *Acta Radiol* 2008; 49(9):975-981.
119. Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 2000; 92(13):1081-1087.

120. Kirsh VA, Chiarelli AM, Edwards SA, O'Malley FP, Shumak RS, Yaffe MJ et al. Tumor characteristics associated with mammographic detection of breast cancer in the Ontario breast screening program. *J Natl Cancer Inst* 2011; 103(12):942-950.
121. Jalalian A, Mashohor SB, Mahmud HR, Saripan MI, Ramli AR, Karasfi B. Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. *Clin Imaging* 2012 Nov 13. pii: S0899-7071(12)00293-8.
122. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA* 2012; 307(13):1394-1404.
123. Lord SJ, Lei W, Craft P, Cawson JN, Morris I, Walleser S et al. A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. *Eur J Cancer* 2007; 43(13):1905-1917.
124. Bordas P, Jonsson H, Nystrom L, Lenner P. Interval cancer incidence and episode sensitivity in the Norrbotten Mammography Screening Programme, Sweden. *J Med Screen* 2009; 16(1):39-45.
125. Collett K, Stefansson IM, Eide J, Braaten A, Wang H, Eide GE et al. A basal epithelial phenotype is more frequent in interval breast cancers compared with screen detected tumors. *Cancer Epidemiol Biomarkers Prev* 2005; 14(5):1108-1112.
126. Pisano ED, Gatsonis CA, Yaffe MJ, Hendrick RE, Tosteson AN, Fryback DG et al. American College of Radiology Imaging Network digital mammographic imaging screening trial: objectives and methodology. *Radiology* 2005; 236(2):404-412.
127. Skaane P. Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. *Acta Radiol* 2009; 50(1):3-14.
128. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005; 353(17):1773-1783.
129. Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N, Dos SS, I. Full-field digital versus screen-film mammography: comparison within the UK breast screening program and systematic review of published data. *Radiology* 2009; 251(2):347-358.
130. Sala M, Comas M, Macia F, Martinez J, Casamitjana M, Castells X. Implementation of digital mammography in a population-based breast cancer screening program: effect of screening round on recall rate and cancer detection. *Radiology* 2009; 252(1):31-39.

131. Sala M, Salas D, Belvis F, Sanchez M, Ferrer J, Ibanez J et al. Reduction in false-positive results after introduction of digital mammography: analysis from four population-based breast cancer screening programs in Spain. *Radiology* 2011; 258(2):388-395.
132. Nederend J, Duijm LE, Louwman MW, Groenewoud JH, Donkers-van Rossum AB, Voogd AC. Impact of transition from analog screening mammography to digital screening mammography on screening outcome in The Netherlands: a population-based study. *Ann Oncol* 2012; 23(12):3098-3103.
133. Hoff SR, Abrahamsen AL, Samset JH, Vigeland E, Klepp O, Hofvind S. Breast cancer: missed interval and screening-detected cancer at full-field digital mammography and screen-film mammography-- results from a retrospective review. *Radiology* 2012; 264(2):378-386.
134. Ascunce N, Ederra M, Delfrade J, Baroja A, Erdozain N, Zubizarreta R et al. Impact of intermediate mammography assessment on the likelihood of false-positive results in breast cancer screening programmes. *Eur Radiol* 2012; 22(2):331-340.
135. Castells X, Roman M, Romero A, Blanch J, Zubizarreta R, Ascunce N et al. Breast cancer detection risk in screening mammography after a false-positive result. *Cancer Epidemiol* 2012.
136. Sala M, Salas D, Zubizarreta R, Ascunce N, Rue M, Castells X. [Current status of research in breast cancer screening in Spain: implications for prevention]. *Gac Sanit* 2012; 26(6):574-581.
137. Salas D, Ibanez J, Roman R, Cuevas D, Sala M, Ascunce N et al. Effect of start age of breast cancer screening mammography on the risk of false-positive results. *Prev Med* 2011; 53(1-2):76-81.
138. Zubizarreta R, Fernández-Llanes AB, Almazán A, Román R, Velarde J, Queiro T et al. Effect of radiologist experience on the risk of false-positive results in breast cancer screening programs. *Eur Radiol* 2011; 21(10):2083-2090.
139. Welch HG. Screening mammography--a long run for a short slide? *N Engl J Med* 2010; 363(13):1276-1278.
140. Hoff SR, Klepp O, Hofvind S. Asymptomatic breast cancer in non-participants of the national screening-programme in Norway: a confounding factor in evaluation? *J Med Screen* 2013.
141. Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL et al. Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. *J Natl Cancer Inst* 2011; 103(3):250-263.
142. Malorni L, Shetty PB, De Angelis C, Hilsenbeck S, Rimawi MF, Elledge R et al. Clinical and biologic features of triple-negative breast cancers in a large cohort of patients with long-term follow-up. *Breast Cancer Res Treat* 2012; 136(3):795-804.

143. Rayson D, Payne JI, Abdolell M, Barnes PJ, MacIntosh RF, Foley T et al. Comparison of clinical-pathologic characteristics and outcomes of true interval and screen-detected invasive breast cancer among participants of a Canadian breast screening program: a nested case-control study. *Clin Breast Cancer* 2011; 11(1):27-32.
144. van d, V, Wesseling J, Pijnappel RM, Dorrius MD, den Heeten GJ, de Roos MA et al. Aggressiveness of 'true' interval invasive ductal carcinomas of the breast in postmenopausal women. *Mod Pathol* 2010; 23(4):629-636.
145. Pollan M, Ascunce N, Ederra M, Murillo A, Erdozain N, Ales-Martinez JE et al. Mammographic density and risk of breast cancer according to tumor characteristics and mode of detection: a Spanish population-based case-control study. *Breast Cancer Res* 2013; 15(1):R9.
146. Boyd NF, Rommens JM, Vogt K, Lee V, Hopper JL, Yaffe MJ et al. Mammographic breast density as an intermediate phenotype for breast cancer. *Lancet Oncol* 2005; 6(10):798-808.
147. Guo YP, Martin LJ, Hanna W, Banerjee D, Miller N, Fishell E et al. Growth factors and stromal matrix proteins associated with mammographic densities. *Cancer Epidemiol Biomarkers Prev* 2001; 10(3):243-248.
148. Li T, Sun L, Miller N, Nicklee T, Woo J, Hulse-Smith L et al. The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. *Cancer Epidemiol Biomarkers Prev* 2005; 14(2):343-349.
149. Bluekens AM, Holland R, Karssemeijer N, Broeders MJ, den Heeten GJ. Comparison of digital screening mammography and screen-film mammography in the early detection of clinically relevant cancers: a multicenter study. *Radiology* 2012; 265(3):707-714.
150. Domingo L, Sala M, Servitja S, Corominas JM, Ferrer F, Martinez J et al. Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain. *Cancer Causes Control* 2010; 21(8):1155-1164.
151. Comas M, Castells X, Hoffmeister L, Roman R, Cots F, Mar J et al. Discrete-event simulation applied to analysis of waiting lists. Evaluation of a prioritization system for cataract surgery. *Value Health* 2008; 11(7):1203-1213.
152. Mar J, Arrospe A, Comas M. Budget impact analysis of thrombolysis for stroke in Spain: a discrete event simulation model. *Value Health* 2010; 13(1):69-76.
153. Comas M, Roman R, Cots F, Quintana JM, Mar J, Reidy A et al. Unmet needs for cataract surgery in Spain according to indication criteria. Evaluation through a simulation model. *Br J Ophthalmol* 2008; 92(7):888-892.
154. Comas M, Roman R, Quintana JM, Castells X. Unmet needs and waiting list prioritization for knee arthroplasty. *Clin Orthop Relat Res* 2010; 468(3):789-797.

155. Rue M, Vilaprinyo E, Lee S, Martinez-Alonso M, Carles MD, Marcos-Gragera R et al. Effectiveness of early detection on breast cancer mortality reduction in Catalonia (Spain). *BMC Cancer* 2009; 9:326.
156. Vilaprinyo E, Puig T, Rue M. Contribution of early detection and adjuvant treatments to breast cancer mortality reduction in Catalonia, Spain. *PLoS One* 2012; 7(1):e30157.