

Mortalitat per càncer de mama a Europa l'any 2005 i projeccions per a l'any 2015

Joan Valls i Marsa

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

UNIVERSITAT DE BARCELONA
FACULTAT DE MEDICINA

TESI DOCTORAL

Mortalitat per càncer de mama a Europa l'any 2005
i projeccions per a l'any 2015

Memòria de la tesi doctoral presentada per **Joan Valls i Marsal** a la Universitat de Barcelona. El treball s'ha realitzat en el marc del Programa de Doctorat de Medicina, sota la direcció del **Dr. Ramon Clèries i Soler** i de la **Dra. Josepa Ribes i Puig**.

Barcelona, 2010

Índex

Agraïments	xi
Resum	xv
Llista d'acrònims	xvii
1 Introducció	1
1.1 Càncer	1
1.2 Càncer de mama	2
1.3 Objectius de la tesi i hipòtesis de treball	5
1.4 Contribucions de la tesi	7
1.5 Estructura de la tesi	9
2 Epidemiologia del càncer de mama	13
2.1 Biologia del càncer de mama	14
2.2 Incidència i mortalitat del càncer de mama	16
2.3 Supervivència del càncer de mama	18
2.4 Factors de risc associats al càncer de mama	19
2.4.1 Factors de risc ambientals	19
2.4.2 Factors de risc endògens	26
2.5 Prevenció del càncer de mama	27
3 Mètodes estadístics	33
3.1 Taxes d'incidència o mortalitat	34
3.2 Comparació de taxes	40

3.3	Anàlisi estadística de tendències de taxes	44
3.4	Projeccions de taxes	51
3.5	Avaluació de factors de riscos i demogràfics	57
4	Mortalitat per càncer de mama a Europa l'any 2005 i projeccions per a l'any 2015	61
4.1	Font de les dades	62
4.2	Anàlisis Estadístiques	63
4.3	Models per a les projeccions	63
4.4	Resultats	65
4.4.1	La mortalitat per càncer de mama a Europa els anys 2000 i 2005	66
4.4.2	Anàlisi de la tendència de la mortalitat per càncer de mama a Europa durant els 10 darrers anys	67
4.4.3	La mortalitat per càncer de mama projectada a l'any 2015	76
4.4.4	Anàlisi dels factors demogràfics i de risc en les diferències entre la mortalitat de l'any 2015 respecte l'any 2005	83
5	Eines computacionals desenvolupades	89
5.1	Eines per a l'anàlisi de tendències i projeccions de taxes d'incidència o mortalitat	90
5.1.1	Funcions implementades	90
5.1.2	Mortalitat a Espanya els anys 1996-2005 i projecció l'any 2015	91
5.2	Riskdiff: anàlisi de factors de risc i demogràfics en l'evolució de la incidència o mortalitat d'una malaltia	93
5.2.1	Funcions implementades	93
5.2.2	Comparació de la mortalitat a Espanya l'any 2015 respecte el 2005 i avaluació de factors demogràfics	94
5.2.3	Interfície web <i>RiskDiff</i>	96

5.3	Codi de les anàlisis presentades en el capítol 4	99
6	Discussió	101
6.1	Metodologia estadística	101
6.2	Anàlisis epidemiològiques	106
7	Conclusions	115
7.1	Metodologia estadística	115
7.2	Anàlisis epidemiològiques	116
A	Codi de les funcions	119
A.1	Funcions per a l'anàlisi de tendències i projeccions	119
A.2	Funcions per a l'anàlisi les diferències degudes a risc i a factors demogràfics en dades d'incidència o mortalitat	121
B	RiskDiff: a web tool for the analysis of the difference due to risk and demographic factors for incidence or mortality data	127
C	Differences due to risk and demographic factors for the pro- jections and burden of breast cancer mortality in Europe from 2005 to 2015	137
	Bibliografia	153

Índex de figures

1.1	Incidència del càncer de mama al món	3
1.2	Mortalitat per càncer de mama al món	4
2.1	Taxes d'incidència anuals mitjanes de càncer de mama per 100.000 dones per grups d'edat	21
4.1	Rànking de les taxes ajustades de mortalitat per càncer de mama a Europa	68
4.2	Tendència de la mortalitat per càncer de mama a l'Europa del Nord	77
4.3	Tendència de la mortalitat per càncer de mama a l'Europa de l'Oest	77
4.4	Tendència de la mortalitat per càncer de mama a l'Europa del Sud	78
4.5	Tendència de la mortalitat per càncer de mama a l'Europa del Centre i de l'Est	79
4.6	Percentatges de canvi en la mortalitat de càncer de mama a Europa l'any 2015 respecte l'any 2005	85
4.7	Rànking del canvi en la mortalitat per càncer de mama de l'any 2015 respecte l'any 2005	88
5.1	Comparació de la mortalitat per càncer de mama a Espanya l'any 2015 respecte l'any 2005	97
5.2	Aparença de l'eina web RiskDiff	98

Índex de taules

2.1	Supervivència del càncer de mama a Europa	20
2.2	Factors de risc associats al càncer de mama	31
3.1	Estructura dels grups d'edat per poblacions de referència	37
4.1	Taxes de mortalitat per càncer de mama a Europa del Nord . . .	69
4.2	Taxes de mortalitat per càncer de mama a Europa de l'Oest . . .	70
4.3	Taxes de mortalitat per càncer de mama a Europa del Sud . . .	71
4.4	Taxes de mortalitat per càncer de mama a Europa Central i de l'Est	72
4.5	Taxes de mortalitat per càncer de mama a Europa Central i de l'Est (cont.)	73
4.6	Resultats dels models, període de 10 anys	81
4.7	Resultats dels models, període inferior a 10 anys	82
4.8	Taxes de mortalitat per càncer de mama a les regions d'Europa .	84

Agraïments

Aquesta tesi és la suma de esforços, iniciatives, inquietuds, idees, converses, vivències, i molts altres fets que m’han estat envoltant en els darrers tres anys. Arriba doncs ara el moment de mostrar el meu agraïment a totes aquelles persones que han fet possible aquest treball i que m’han acompanyat i donat suport durant aquest temps. Entre ells, doncs estic especialment agraït als meus directors, la Josepa Ribes i el Ramon Clèries, per haver-me donat l’oportunitat d’iniciar aquest treball i estar al meu costat en tot moment, compartint idees, dubtes, problemes, solucions, èxits i també alguns fracassos. Moltes gràcies per haver cregut en mi i per tot el que he après al vostre costat. Més enllà de la tesi, també em sento molt afortunat per haver compartit amb la Pitus i el Ramon el meu dia a dia, especialment en els darrers 5 anys. Heu estat al meu costat alhora de superar els petits tràngols de la vida i també en els molts bons moments que hem compartit.

Especialment al Ramon vull agrair-li la relació d’amistat que tenim d’ençà ja fa molt de temps (uf! potser més de 15 anys?) quan ens vam conèixer a la Facultat d’Estadística. No han estat poques les excursions que hem fet, les estades a Finlàndia i tots els moments viscuts a Lleida, Igualada i Barcelona. Moltes gràcies, company!

A la Pitus vull agrair-li la seva atenció, positivisme i, sobretot, la seva generositat humana. Pitus, ets una persona molt especial amb un do excepcional per a ficar-te a la pell de l’altre. Moltes gràcies pels consells que m’has donat i espero, tot i que ara estigui a Lleida i ja no siguis la meva *jefa*, poder continuar tenint-te a prop. Penso que el món seria millor si hi haguessin més Pitus... com ho podríem fer per a que et multipliquessis?

De la meva estada a l’Institut Català d’Oncologia (ICO), vull agrair en primer lloc al Josep Maria Borràs pel suport institucional que sempre ens ha donat en el projecte que ha desembocat en aquesta tesi i, especialment, per apostar per la recerca dins del Pla Director d’Oncologia, així com la

seva disponibilitat personal. També voldria agrair al Victor Moreno per la meva primera etapa a l'ICO, quan em va introduir a l'estadística aplicada a la genòmica funcional, així com al Miquel Àngel Pujana pel seu interès i involucració en tot allò que fèiem. Moltes gràcies a tots dos per haver-me iniciat en el món de la recerca i per tot el treball que vaig realitzar sota la vostra direcció.

Vull també agrair especialment al Xavier Matias-Guiu, de l'Institut de Recerca Biomèdica de Lleida (IRB), les facilitats que m'ha donat per a poder continuar i acabar els treballs d'aquesta tesi, compaginant-ho amb la meva activitat actual a l'IRB. Aprofito per reconèixer els esforços que ha fet per a tirar endavant aquest nou institut de recerca així com, més particularment, el suport en la creació de la Unitat de Bioestadística i per haver confiat en la meva persona per a contribuir, modestament, a aquest projecte.

Especialment voldria agrair al Timo Hakulinen i al Tadek Dyba per la fantàstica acollida que em van donar durant les estades que vaig realitzar al Registre del Càncer de Finlàndia. Vull agrair al Tadek la seva implicació en la part metodològica d'aquesta tesi. Moltes gràcies per ajudar-me tant generosament en la part de modelització estadística i compartir amb el Ramon i jo mateix els teus coneixements. També vull agrair al Tadek totes les converses, sopars, nits al Poseidon, ..., que hem passat plegats a Helsinki així com les excursions pels parcs nacionals. *Kiitos!*

Pel que fa al mètode RiskDiff voldria agrair, en primer lloc, a Jacques Estève per la lectura crítica del primer text de l'article i pels comentaris que em va adreçar que van millorar-lo substancialment. També voldria agrair a la Rosa Gispert la seva contribució en aquest mètode.

De la UAB, voldria agrair al Pere Puig, per resoldre'm alguns dubtes estadístics d'aquesta tesi, així com al Llorenç Badiella, amb els qui vaig poder treballar conjuntament en el Servei d'Estadística des de la seva creació. A més, al Llorenç vull agrair-li una bona amistat (ets el meu mestre!).

Dels 7 anys que he estat a l'ICO vull agrair a totes les persones amb qui he compartit qüestions de feina, però també cafès, sopars, converses i altres experiències que m'han enriquit enormement. La meva estada a l'ICO m'ha permès conèixer un grup ampli de persones i marxar, finalment, amb la motxilla ben carregada. Especialment, voldria agrair a la gent del Plà Director d'Oncologia (PDO) el dia a dia viscut en els anys 2006-2009. Especialment, a la Laura Esteban, Laura Pareja, Jordi Gàlvez, Xavi Sanz, Josep Maria Escribà, Àngels Melià, Laura Roca, Meri Nomen, Lluïsa Aliste

i Rebeca Font per tots els moments viscuts i també a la Tarsi Ferro i Josep Alfons Espinàs pel suport mostrat, així com la resta de personal del PDO. També a l'ICO he tingut el plaer de conèixer moltes persones que m'han aportat aspectes positius. Gràcies a la Mònica Grau, Antònia Obrador i Xavier Solé pels primers moments en l'ICO en que em vàreu ajudar tant a situar-me en el món de la genètica. Moltes gràcies també als membres del *clúster d'estadístics* de l'ICO (crec que és la institució catalana amb un major nombre d'estadístics per $m^2!$), entre ells la Yolanda Benavente, Eli Guinó, Mireia Díaz, JR González, Guillem Pera, Toni Berenguer, Gina Alberó, Sara Tous amb qui he pogut compartir moltes converses estadístiques (i alguna que altra més normal, eh!). Més recentment, agrair a la Noémie Travier (ja saps que el francès i el català no són incompatibles!) per les converses i sortides que hem fet plegats, així com al grup de muntanya de l'ICO (especialment a l'Eva González i al Xavi Vallés), per les sortides que heu organitzat i que m'han permès desconnectar. Vull també agrair a totes les persones amb qui hem compartit estones a l'ICO, dinars a la cafeteria, ... entre elles, al Carlos Alberto González, per totes les converses tant interessants que hem compartit, a la Silvia de Sanjosé, per haver-me orientat més d'un cop en diverses qüestions en recerca, al Jesús Muñoz, per donar-me suport informàtic (i digital... ets una font de recursos!), la Laia Bruni, la Laia Alemany, la Leila Luján, la Núria Monfuleda, en Carles Miralles, el Toni Agudo, la Genevieve Buckland, la Michelle Méndez, el Raúl Mauricio, la Cristina Rajo, el Xavier Bosch, el Xavier Castellsagué, la Liz Allison i molts altres amb qui he anat compartint petits moments de la meua vida diària a l'ICO.

Més recentment, vull agrair a tota la gent de l'IRB Lleida per fer tant fàcil el meu retorn a Lleida. I és que, ara ja, poc a poc, torno a recuperar el meu parlar lleidatà (i menjo més caragols!). Especialment vull agrair a la Montse Martínez, la meua nova companya de despatx, per haver-me ajudat a situar-me en un primer moment, però també pel caràcter tant positiu i ser una persona tant íntegra (així com els cafès amb llet del matí!). També vull agrair a la Maria Ruiz (gràcies per fer-me de llebre amb la tesi!), Oriol Garrofé (aviam si acabem l'anàlisi de l'Impact Factor!) i a la Sònia Fernández, Josep Maria Bosch, Irene Rossell, Katia Segovia i Anna Giménez pel suport mostrat. Voldria també agrair a l'Aleix Vilalta, amb qui ens hem animat mútuament per a acabar la tesi (recorda, Aleix que tenim una aposta: el darrer que la llegeixi paga un sopar!). El meu agraïment també per l'acollida que he tingut per part de la gent de la Facultat de Medicina de la UdL, entre ells

la Montse Rué, l'Albert Sorribes, el Rui Alves i l'Esther VilaprinYO. Moltes gràcies també a totes les persones que he conegut en aquests darrers mesos a Lleida, a l'Arnau de Vilanova... Estic molt content d'estar amb vosaltres i d'estar novament a la meva ciutat!

També vull dedicar un espai d'agraïment als meus amics personals que, en certa manera, m'han ajudat a tirar endavant aquesta tesi. Al David Vives, per ser amic meu des de fa ja molts anys, a la Yolanda Garcia, per escoltar-me i estar sempre disponible, a la Mireia Petanàs pels *beures* que fem plegats, als *nens* del grup d'esplai: Miqui, Roger, Sergi, ... (que ara ja teniu 25 anys... i és que ja fa més de 15 anys que m'aguanteu com a monitor!), a la Mònica Bécue (per les trobades gastronòmiques, entre moltes altres coses), a l'Anna Huguet (per compartir amb mi molts bons moments de lleure, de manera especial en els darrers mesos...) i moltes altres persones amb qui nodreixo el meu dia a dia.

Especialment voldria agrair al grup de famílies que, des de ja fa una colla d'anys ens trobem per organitzar sortides de caps de setmana i colònies d'estiu amb els nostres fills, ... Cada vegada som més i cada cop ens ho passem millor... Si és que, què seria un estiu sense les colònies?

Moltes gràcies a la meva família per tot el que m'han donat. Crec que mai us podré tornar tot el que he rebut. Especialment a la meva mare Neus, per estar tant incondicionalment al meu costat i també a les meves tres germanes Magda, Dolors i Montse. També als meus cunyats Jordi i Octavi, amb qui compartim tantes trobades familiars i al Castany, per ser tant autènticament granadellenc!. I, com no, als meus 5 nebots: Neus, Marta, Ainhoa, Xavier i Martí,... sou tots molt *pispires*! També vull expressar el meu record emotiu cap al meu cunyat Ismael, que ens va abandonar l'estiu del 2009 i que va ser com un germà per a mi (*Isi* pervius!).

I, finalment, a tu Pau, que ets un gripau, que em fas riure tant, quan fas de pallasso *Kalilo*, i que em sorprens amb els teus trucs de *mag Baldiri*. Fa cinc anys que vas néixer, t'he vist créixer i jo he crescut amb tu. Gràcies a tu, he recuperat l'infant que porto a dins. Espero continuar jugant amb tu i passant-nos-ho tant bé ... creixent i aprenent plegats. Pau, endavant i ja saps que qualsevol nit pot sortir el sol!

I res més, solament afegir que em sento molt afortunat de haver-vos conegut a tots. Moltes gràcies!

Resum

El treball presentat en aquesta tesi es centra en l'anàlisi de la mortalitat per càncer de mama Europa. Concretament, hem realitzat una anàlisi estadística i epidemiològica de la tendència de la mortalitat, per a un conjunt de 32 països europeus, en els darrers 10 anys disponibles (fins a l'any 2005) i també hem proporcionat les projeccions per a les taxes de mortalitat per a l'any 2015. Els resultats obtinguts mostren que, en general, la tendència de la mortalitat a Europa és decreixent, situant-se al voltant d'un decrement del 2% anual. Tanmateix, en la majoria de països de l'Europa del Centre i de l'Est, només s'observa aquesta disminució per als grups de dones més joves, podent-se constatar tendències fins i tot creixents per a les dones d'edat més avançada. Hem estimat que a Europa, l'any 2015 hi haurà 127.983 defuncions per càncer de mama en dones majors de 35 anys, el que representa un augment del 4,68% respecte les dades absolutes de l'any 2005. D'aquestes defuncions, 10.288 (8,04%) correspondran a dones entre 35 i 49 anys, 54.226 (42,37%) a dones entre 50 i 69 anys i, finalment, 63.469 (49,59%) a dones majors de 70 anys. El major nombre de defuncions s'observarà en països de l'Europa de l'Est i del Centre (38,69%), seguit per dones de l'Europa de l'Oest (30,19%) i, en una magnitud menor, dones del Sud i del Nord d'Europa (17,98% i 13,14% respectivament). Pel que fa a l'augment del nombre de defuncions de l'any 2015 respecte l'any 2005 en un 4,68%, aquest pot descomposar-se en tres parts: un augment del 5,28% degut a un creixement de la població europea, un augment del 6,20% degut a un envelliment d'aquesta població i, finalment, un decrement del 6,8% atribuïble

a una disminució del risc de morir per aquesta malaltia. D'aquesta forma hem pogut constatar que el risc de morir per càncer de mama disminuirà de forma important, essent aquest decrement del 14% al Nord d'Europa, del 10,61% i 7,37% a l'Europa de l'Oest i del Sud respectivament, i per a l'Europa del Centre i Est es constata un augment del risc del 0,49%.

Pel que fa a la metodologia estadística, en aquesta tesi aportem un marc de modelització que permet analitzar eficientment dades de mortalitat o d'incidència durant un període de temps i, mitjançant models edat-període, obtindre projeccions en el futur. L'estratègia de modelització ha donat lloc a uns resultats raonablement acceptables per a analitzar la mortalitat per càncer de mama a Europa. Hem desenvolupat un conjunt de funcions en el programa R que són lliurement accessibles i que permeten reproduir els resultats o bé aplicar la metodologia en altres àmbits i dades. També s'ha implementat una eina web que permet analitzar quina part de les diferències observades en la mortalitat o incidència d'una malaltia entre dos moments són degudes a factors demogràfics respecte el risc de desenvolupar o morir per la malaltia.

Llista d'acrònims

AAAP Additive Age-specific Age Period model

APC Percentatge de canvi annual

DNA Àcid desoxiribonucleic

CM Càncer de mama

IGF-I Insulin-like growth factor

IMC Índex de massa corporal

MAP Multiplicative Age-specific Period model

MAAP Multiplicative Age-specific Age-Period model

RR Risc Relatiu

TA Taxa ajustada a la població de mundial de mortalitat per 100.000 dones

TC Taxa crua de mortalitat per 100.000 dones

Capítol 1

Introducció

El treball desenvolupat en aquesta memòria es situa en el marc de l'anàlisi descriptiva de dades epidemiològiques en càncer de mama (CM). Més particularment, en l'anàlisi de les taxes de mortalitat per CM actuals a Europa i les seves projeccions en el futur. En aquest capítol contextualitzarem, en primer lloc, el marc d'anàlisi, aportant una breu introducció a l'epidemiologia del càncer, en general, i del CM, en particular. En segon lloc, descriurem els objectius que persegueix la tesi. A continuació, presentarem un resum de les contribucions aportades i, finalment, descriurem els diferents capítols en els que està estructurada aquesta memòria.

1.1 Càncer

El càncer és un grup de malalties caracteritzades per un creixement descontrolat de cèl·lules anòmales i la seva propagació. Si aquesta no es pot aturar, el càncer pot arribar a provocar la mort. De forma general, el càncer està causat per factors externs o ambientals (tabac, agents químics, radiació, organismes infecciosos, ...) i per factors interns (mutacions genètiques heretades, hormones, condicions d'immunitat i mutacions provocades pel metabolisme), que poden actuar conjuntament o de forma seqüencial per a iniciar i progressar en la carcinogènesi. El desenvolupament de la majoria

de càncers requereix múltiples passes que poden tenir lloc durant molts anys. Per aquesta raó el càncer es pot controlar en diferents moments de la carcinogènesi. La prevenció primària es pot realitzar en alguns tipus de càncer, eliminant o reduint l'exposició a alguns factors ambientals que n'acceleren el procés, com el tabac. D'altres, es poden detectar abans que les cèl·lules esdevinguin canceroses o bé en una fase primerenca en la que la malaltia és sovint tractable, el que constitueix la prevenció secundària del càncer. El càncer pot ser tractat a partir de cirurgia, radiació, quimioteràpia, tractament hormonal i immunoteràpia, per intentar disminuir el seu impacte, en el context ja de la prevenció terciària.

Tenint en compte totes les localitzacions tumorals, s'ha estimat que l'any 2007 hi van haver 12 milions de casos nous de càncer al món i 7,6 milions de morts van ser degudes a aquesta malaltia [Garcia et al., 2007]. A Europa, es varen estimar 3 milions de casos incidents de càncer i 1,8 milions de defuncions a causa d'aquesta malaltia. El càncer és la segona causa de mort al món, després de les malalties cardiovasculars [Garcia et al., 2007].

1.2 Càncer de mama

Epidemiologia descriptiva: incidència, mortalitat i supervivència

S'estima que l'any 2007 hi van haver 1,3 milions de casos nous de CM en dones al món [Garcia et al., 2007], essent el tipus de càncer més diagnosticat en dones. Pel que fa a la mortalitat, s'estima que 465.000 dones van morir per aquesta causa l'any 2007, representant la primera causa de mort per càncer. A Europa s'estima que l'any 2006 es van diagnosticar 429.900 dones i que van morir per aquesta causa 131.900 [Ferlay et al., 2007].

En els darrers 25 anys, la incidència del CM ha augmentat al món en un 30%, degut a canvis en els models reproductius i, més recentment, degut a la implantació de sistemes de cribatge poblacionals [Colditz et al., 2006b]. Per l'altra banda, la mortalitat per CM s'ha mantingut estable en els darrers 25 anys o fins i tot ha disminuït, fet que s'ha atribuït a la detecció precoç

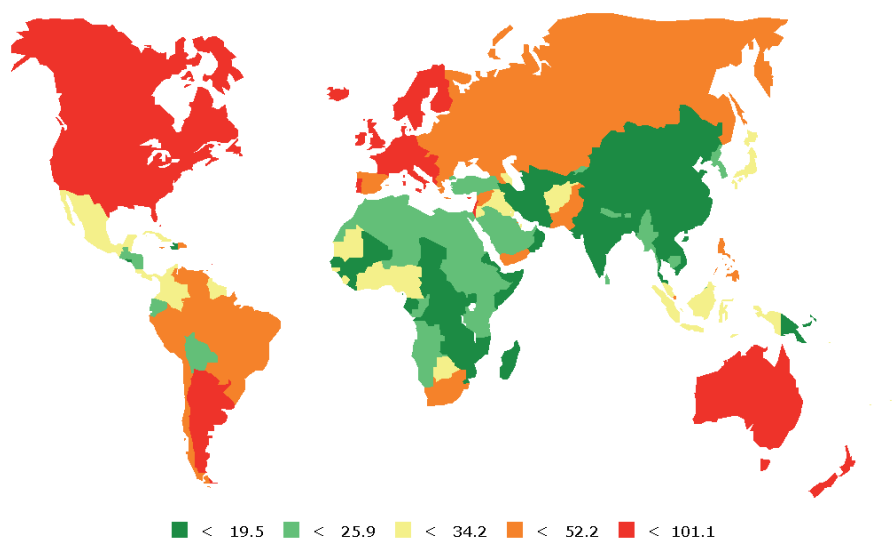


Figura 1.1: Incidència del càncer de mama al món (casos per 100.000 habitants). Taxes ajustades per edat. Dades extretes de GLOBOCAN 2002 [Ferlay et al., 2004].

amb mamografies i a millores del tractament [Althuis et al., 2005]. Tant les taxes d'incidència com de mortalitat per CM presenten un alt nivell de variabilitat geogràfica entre els diferents països del món (figures 1.1 i 1.2, informació extreta de l'eina web GLOBOCAN 2002 [Ferlay et al., 2004]). Pel que fa a la incidència de CM, Amèrica del nord, Austràlia i el Nord i Oest d'Europa presenten les taxes més elevades, Europa de l'Est i Amèrica del Sud presenten taxes d'incidència intermèdies mentre que molts països d'Àfrica i Àsia presenten les taxes més baixes. Aquest fet posa d'evidència les grans diferències socio-demogràfiques que es troben actualment entre els països desenvolupats i els països en vies de desenvolupament, on el diagnòstic, tractament, però també els factors de riscos del CM i característiques de la població són molt diferents. Pel que fa a la mortalitat, les taxes més elevades es troben a Europa del Nord i de l'Oest, mentre que països de l'Europa del Sud i de l'Est així com Estats Units i Austràlia presenten taxes mitjanes, i finalment, les taxes més baixes s'observen en molts països d'Àsia, Àfrica i Amèrica Central i del Sud.

La supervivència relativa del CM, mesurada com el percentatge de dones

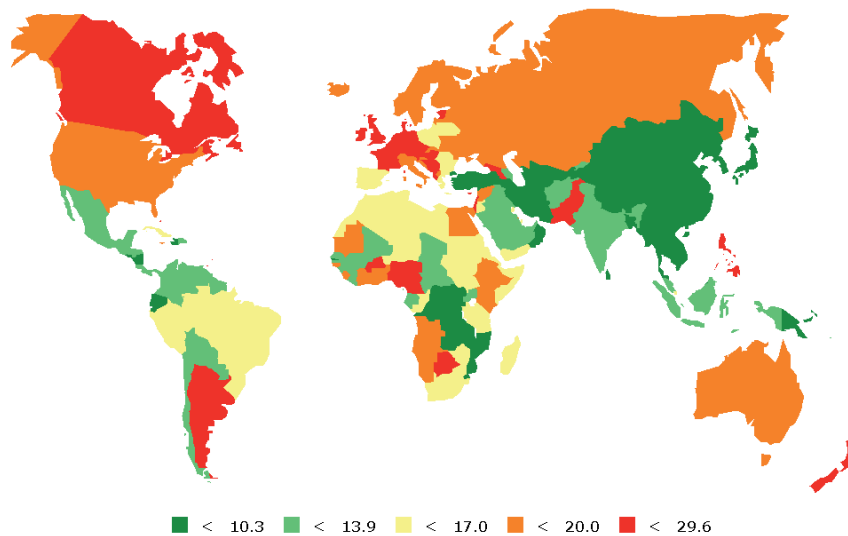


Figura 1.2: Mortalitat per càncer de mama al món (casos per 100.000 habitants). Taxes ajustades per edat. Dades extretes de GLOBOCAN 2002 [Ferlay et al., 2004].

que continuen estant vives 5 anys després del diagnòstic, és del 79% a Europa [Sant et al., 2009] i del 89% als Estats Units [Jemal et al., 2006]. La detecció precoç del CM és clau per a millorar-ne el pronòstic i fer més efectius els tractaments. Per exemple, als Estats Units la supervivència relativa del càncer de mama entre les dones americanes és del 98% quan aquest es troba en un estadi primerenc, comparat amb el 84% i el 28% quan la malaltia s'ha estès a ganglis limfàtics regionals o ha metastatitzat en òrgans a distància, respectivament [Jemal et al., 2006].

Factors de risc

L'edat és el factor de risc més associat al CM, augmentant la probabilitat de desenvolupar-lo. El risc associat a l'edat es pot incrementar amb la presència de dues mutacions genètiques (BRCA1 i BRCA2), una història familiar amb casos de càncer de mama, una densitat elevada en el teixit mamari, una hiperplàsia en el teixit mamari confirmada per biòpsia o una exposició alta a radiació en el pit deguda a procediments mèdics. D'altra banda, existeixen

també un conjunt factors reproductius que incrementen el risc de desenvolupar CM, tals com una menarquia prematura i una menopausa tardana, no haver tingut fills, ús de contraceptius orals i tenir el primer fill després dels 30 anys. Altres factors que poden incrementar el risc de patir CM en dones menopàusiques són el sobrepès, l'obesitat, l'ús de teràpies hormonals substituïtives (especialment combinat amb estrògens i teràpia amb progestina), inactivitat física i consum d'una o més begudes alcohòliques per dia. Finalment, factors com donar el pit, activitat física moderada o vigorosa i el manteniment del cos en un bon estat de salut estan associats amb un risc menor de CM.

1.3 Objectius de la tesi i hipòtesis de treball

La mortalitat per CM a Europa presenta un nivell alt d'heterogeneïtat, tant pel que fa a la variabilitat observada entre els països com per l'evolució històrica de cada país. Els motius d'aquestes diferències es poden explicar, en part, per la variabilitat en les exposicions a factors ambientals, en el període d'introducció de programes de cribatge, en l'accés sanitari dels malalts, o en els factors socio-econòmics, entre altres. L'objectiu principal d'aquesta tesi és analitzar la mortalitat per CM en les dones europees en un període recent i estimar les projeccions per l'any 2015, aportant una interpretació que tingui en compte els factors ambientals esmentats. Per complir aquest objectiu és necessari realitzar una anàlisi estadística de les taxes de mortalitat per CM conjunta que permeti adequar els models estadístics a cada país i que proporcionï inferència sobre projeccions futures. En aquest context, els objectius específics d'aquesta tesi són:

- Desenvolupar un marc de modelització estadística que permeti analitzar eficientment les dades històriques de mortalitat per CM dels diferents països d'Europa.
- Identificar diferents patrons d'evolució de la tendència de la mortalitat

per CM en els països europeus i proporcionar una estimació comparable per a l'any 2005.

- Proporcionar una predicció del nombre de morts per CM que hi hauran l'any 2015 als diferents països europeus, en base als models estadístics ajustats.
- Analitzar l'efecte de factors demogràfics sobre els increments o decrements del nombre de morts per CM de l'any 2015 respecte el 2005.

Aquests objectius tenen, principalment, un interès a nivell epidemiològic i de salut pública, doncs la càrrega futura del CM és rellevant en termes de planificació sanitària i consum de recursos sanitaris.

D'altra banda, aquesta tesi té també com objectiu més metodològic el de desenvolupar un conjunt de facilitats computacionals, en programari obert i lliure, que permeti realitzar les anàlisis estadístiques descrites prèviament. Els objectius metodològics específics són:

- Desenvolupar un conjunt de funcions en \mathbb{R} per a l'anàlisi de l'evolució de taxes de incidència o mortalitat d'una malaltia determinada i la seva projecció futura.
- Desenvolupar un conjunt de funcions en \mathbb{R} per a l'anàlisi de les diferències degudes a factors de risc i a factors demogràfics en l'evolució de la incidència o mortalitat d'una malaltia determinada.
- Desenvolupar una eina web per la implementació de les funcions anteriors en una interfície amigable que sigui aplicable en qualsevol camp d'investigació.

Finalment, els objectius enunciats anteriorment permetran avaluar un conjunt d'hipòtesis de treball, epidemiològiques, que enunciem a continuació:

- Es preveu un augment del nombre de casos morts per CM a Europa l'any 2015, amb el que augmentarà considerablement la càrrega del

càncer de mama, degut principalment a factors demogràfics (augment i envelliment de la població europea).

- Es preveu, en general a Europa, una disminució de les taxes de mortalitat per CM, influïda majoritàriament per causes epidemiològiques.
- En algunes zones de Europa, especialment països de l'Est, es preveu que la disminució de la mortalitat per CM serà menys marcada o nul·la.

1.4 Contribucions de la tesi

La contribució principal d'aquesta tesi [Clèries et al., 2010, submitted] presenta els resultats obtinguts a partir de la modelització estadística de les taxes de mortalitat per CM a 32 països d'Europa, fins a l'any 2005 i la seva projecció futura l'any 2015. A més, també es realitza un estudi que permet quantificar quina part dels increments i decrements previstos per l'any 2015 són deguts a canvis en el risc de morir respecte aquells que ho són degut a canvis en factors demogràfics, tals com l'envelliment de la població o l'augment del nombre d'habitants.

La segona contribució d'aquesta tesi [Valls et al., 2009] presenta les eines computacionals creades per a l'anàlisi de les diferències degudes a factors de risc i demogràfics per a dades d'incidència o mortalitat d'una malaltia, que s'ha traduït en la implementació de l'eina web que hem anomenat *RiskDiff*. Els mètodes implementats estan basats en un mètode proposat l'any 2000 [Bashir and Esteve, 2000]. En l'article que presentem, es realitza una descripció de les funcions en R i eina web desenvolupades, lliurement accessibles, i s'utilitzen dades de mortalitat per càncer a Catalunya per tal d'il·lustrar el seu ús.

Part dels resultats d'aquesta tesi han estat comunicats a diferents reunions científiques, que es detallen a continuació:

- Valls J., Clèries R., Ribes J., Hakulinen T., Borràs J.M., Dyba T. Mortalidad por càncer de mama en Europa en 2005 y proyecciones para

el 2015.. *XXVII Congreso de la Sociedad Española de Epidemiología*. Octubre 2009. Saragossa. Comunicació oral.

- Valls J., Clèries R., Gálvez J., Sanz X., Esteban L., Pareja L., Ribes J., Borràs J.M. Análisis de los cambios en la incidencia debidos a riesgo y a factores demográficos: aplicación a las diferencias en la incidencia del cáncer proyectada en Cataluña para los años 2005 y 2015. *XXVI Congreso de la Sociedad Española de Epidemiología*. Octubre 2008. Girona. Comunicació oral.
- Valls J., Clèries R., Dyba T., Hakulinen T., Borràs J.M., Ribes J. Modelling breast cancer mortality in Europe during 1990-2004 among women older than 35. *XXXIII Meeting of the Group for the Epidemiology and Cancer Registration in Latin Countries*. Maig 2008. Parma. Itàlia. Comunicació oral.
- Valls J., Clèries R., Ribes J., Gisper R., Borràs J.M. Riskdiff: a web tool for the analysis of the difference due to risk and demographic factors for incidence or mortality data. *XXXIV Meeting of the Group for the Epidemiology and Cancer Registration in Latin Countries*. Maig 2009. Lugano. Suïssa. Pòster.
- Valls J., Clèries R., Ribes J., Borràs J.M., Hakulinen T., Dyba T. The burden of breast cancer mortality in Europe in 2005 and projection for 2015. *XXXIV Meeting of the Group for the Epidemiology and Cancer Registration in Latin Countries*. Maig 2009. Lugano. Suïssa. Pòster.

També, en aquest punt, destaquem que una part important del treball realitzat en aquesta tesi es va realitzar durant les estades de recerca que el doctorand va realitzar al Registre del Càncer de Finlàndia, a Helsinki, sota la direcció dels doctors Timo Hakulinen i Tadek Dyba durant els anys 2008 i 2009. En aquest marc, el doctorand va obtindre una beca d'ajuts per estades a l'estranger, concedida per l'Agència de Gestió d'Ajuts Universitaris i de

Recerca de la Generalitat de Catalunya (Referència 2007BE-1 00046) que ho va fer possible.

Per últim, com a contribucions complementàries destaquem un conjunt d'articles que varen ser publicats l'any 2008 dins d'un suplement monogràfic a la revista *Medicina Clínica* amb el títol "El impacto del cáncer en Cataluña". Aquesta monografia presenta un conjunt d'anàlisis que permeten donar una visió actualitzada i global sobre la situació de la incidència, mortalitat, supervivència i projeccions del càncer a Catalunya. Un parell dels articles d'aquest suplement [Borras et al., 2008, Ribes et al., 2008] presenten una anàlisi de la tendència del càncer i la seva projecció l'any 2015, que estan relacionats metodològicament amb les anàlisis presentades en aquesta tesi.

1.5 Estructura de la tesi

Aquesta memòria està estructurada en set capítols, el contingut dels quals es descriu a continuació.

Capítol 1. Introducció

Es contextualitzen els treballs realitzats en aquesta tesi. A continuació, s'estableixen els objectius i es descriuen les contribucions de la mateixa, així com l'estructura de la memòria.

Capítol 2. Epidemiologia del càncer de mama

Es presenta de manera àmplia l'epidemiologia actual del CM, recollint aspectes com la incidència, mortalitat i supervivència d'aquesta malaltia, especialment a nivell europeu. També es realitza una petita introducció a la biologia del CM així com els factors de risc associats i qüestions de salut pública relacionades amb la prevenció d'aquest càncer.

Capítol 3. Mètodes estadístics

S'enuncia formalment el problema de l'anàlisi de taxes d'incidència o mortalitat, des del punt de vista estadístic. Es presenten diferents estratègies de modelització estadística per al càlcul de tendències i projeccions de les taxes d'incidència o mortalitat. Per últim, es descriu la metodologia que permet avaluar quina part dels increments o decrements observats en dades d'incidència o mortalitat són deguts a factors de risc respecte els que ho són deguts a factors demogràfics.

Capítol 4. Mortalitat per càncer de mama a Europa l'any 2005 i projeccions per a l'any 2015

Es presenta en primer lloc els antecedents existents en anàlisis similars així com les hipòtesis que planteja aquesta tesi. En segon lloc es descriuen les fonts de dades emprades i, tot seguit, es descriuen formalment els models estadístics emprats per analitzar les dades així com l'estratègia de modelització seguida. Finalment, es presenten els resultats obtinguts en les anàlisis estadístiques de forma estructurada. D'aquesta forma, es comença per descriure la mortalitat per CM a Europa els anys 2000 i 2005, per continuar amb l'anàlisi dels patrons de tendència observats i, finalment, descriure la mortalitat projectada per l'any 2015 així com el paper dels factors demogràfics i de risc en les diferències observades respecte l'any 2005.

Capítol 5. Eines computacionals desenvolupades

Es descriu inicialment quins han estat els programes emprats per a implementar les eines que ha aportat aquesta tesi. En segon lloc es descriuen els dos conjunts de funcions desenvolupats, el primer per a l'anàlisi de tendències i projeccions i el segon per a l'anàlisi de factors de risc i demogràfics en l'evolució de l'incidència o mortalitat. En tercer lloc, es descriu la interfície web que hem implementat per a l'anàlisi dels factors demogràfics i de risc en l'avaluació de les diferències observades en dades d'incidència o mortalitat d'una malaltia determinada.

Finalment, es descriuen el fitxer que conté el codi amb les anàlisis i el fitxer que conté les dades analitzades en aquesta tesi, així com s'indica la pàgina web des d'on es poden descarregar.

Capítol 6. Discussió

Es realitza una valoració crítica sobre els mètodes estadístics emprats, apuntant algunes de les seves limitacions, així com les fonts de les dades, i altres aspectes, també epidemiològics, relacionats amb els resultats obtinguts.

Capítol 7. Conclusions

S'enuncien les conclusions més rellevants que es desprenen del treball realitzat en aquesta tesi, realitzant una distinció entre aspectes primordialment epidemiològics respecte d'altres essencialment estadístics.

Apèndix A. Codi de les funcions

Es presenta el codi de les funcions desenvolupades.

Apèndix B. Article publicat

Còpia de l'article ja publicat [Valls et al., 2009] que presenta l'eina *RiskDiff* que hem implementat per a l'anàlisi dels factors de risc i demogràfics en l'evolució de la incidència o mortalitat d'una malaltia determinada.

Apèndix C. Article sotmès en revisió

Còpia del text de l'article sotmès [Cleries et al., 2010, submitted] que presenta els resultats principals d'aquesta tesi pel que fa a l'anàlisi de la mortalitat per CM a Europa i la seva projecció.

Apèndix D. Llista d'acrònims

Es presenta la llista d'acrònims emprats en el text d'aquesta tesi.

Capítol 2

Epidemiologia del càncer de mama

L'any 2007 hi van haver 1,3 milions de dones al món que van ser diagnosticades de CM i 465.000 van morir per aquesta malaltia [Garcia et al., 2007], constituint la principal causa de mort i el tipus de càncer més diagnosticat en dones. A Europa, l'any 2006 es van diagnosticar 429.900 casos de CM i es van registrar 131.900 defuncions [Ferlay et al., 2007]. La detecció precoç del CM continua sent clau per tal de garantir l'èxit del tractament. Es tracta doncs d'un càncer que, si es diagnostica en un estadi primerenc, presenta una supervivència relativa del 98% a 5 anys [Jemal et al., 2006].

En aquest capítol presentem en primer lloc una introducció a la biologia bàsica del CM. A continuació es descriu, des del punt de vista epidemiològic, la incidència i mortalitat actuals per CM a Europa, posant èmfasi en els treballs més recents i en l'estat de la qüestió. Finalment presentem característiques referents a la supervivència del CM a Europa, així com els factors de risc més destacats i qüestions relatives a la prevenció i cribatge.

2.1 Biologia del càncer de mama

El càncer es desenvolupa a través d'un procediment en el que cèl·lules normals i sanes del cos es converteixen en anòmales i es multipliquen sense control. Les cèl·lules normals del cos es comuniquen entre elles i regulen la seva proliferació, és a dir, divisió. Quan el càncer té lloc, les cèl·lules tenen un creixement diferent al de les cèl·lules normals i proliferen de manera descontrolada. En la majoria de casos, aquest procediment té lloc durant molt anys i implica diferents passes, entre les quals, mutacions dels gens que regulen la proliferació de la cèl·lules. De fet, un conjunt permanent de mutacions en els *gens supressors de tumor* i en els *proto-oncogens* són necessàries per a que el càncer es desenvolupi [Colditz et al., 2006a].

En cadascuna dels trilions de cèl·lules del cos humà hi han molècules de DNA, que regulen el comportament cel·lular. El DNA és un codi dividit en gens que ha de ser transcrit i traduït per tal que la cèl·lula sintetitzi les proteïnes indicades per cadascun dels gens. Aquestes proteïnes són les que finalment regeixen el comportament i la funció de la cèl·lula. Una mutació és un canvi en el codi genètic. Aquestes mutacions poden ocórrer per atzar o bé poder ser induïdes. Els mutagens són les substàncies que indueixen aquestes mutacions, que poden interferir amb la funcionalitat de la cèl·lula i codificar una proteïna que enviï una senyal incorrecta. La majoria de mutacions són reparades per la mateixa cèl·lula, però en alguns casos, rars, això no és així. Quan la cèl·lula es divideix, aquesta mutació és heretada per les cèl·lules filles i esdevé permanent. La majoria de mutacions en el codi del DNA no tenen cap relació amb el fet que una cèl·lula esdevingui cancerosa. En canvi, quan els gens afectats són els que regulen la divisió i creixement cel·lular (proto-oncogens i gens supressors de tumors), poden aleshores ocasionar el càncer [Colditz et al., 2006a].

El teixit mamari és particularment sensible, per diversos motius, al desenvolupament del càncer. L'estrogen hormonal femení estimula la divisió de les cèl·lules mamàries. Aquesta divisió pot augmentar el risc de malmetre el DNA. A més, les cèl·lules mamàries no estan prou madures en noies o dones

joves que no hagin tingut el primer embaràs complet. D'aquesta forma, aquestes cèl·lules estan lligades a *carcinògens* (agents que causen càncer) i no són suficientment eficients per a reparar el possible DNA malmès en les cèl·lules mamàries [Colditz et al., 2006a].

Les dones joves i noies tenen un teixit mamari que és especialment sensible a carcinògens. A diferència d'altres tipus de teixits del cos, com el fetge o el cor que ja estan formats al naixement, el teixit mamari en el moment del naixement és un minúscul conducte que, a partir de la pubertat i com a resposta a l'activitat hormonal que es desenvolupa en la menarquia (principalment estrogen), creix ràpidament en una estructura semblant a la d'un arbre. El desenvolupament del pit ocorre a la pubertat i també durant el primer embaràs de les dones. Les cèl·lules mamàries immadures, també anomenades cèl·lules mare, es divideixen ràpidament durant aquest període. La qüestió és que aquestes cèl·lules, immadures, no són molt eficients en la reparació de mutacions i, per tant, són més sensibles a carcinògens. Així doncs, és important reduir l'exposició en dones joves i noies a aquests carcinògens. Com a exemple, s'ha vist que noies japoneses que van estar exposades a radiació per la bomba atòmica durant la segona guerra mundial han tingut una major taxa d'incidència de CM en l'edat adulta [Preston et al., 1987]. També esdevé important reduir l'exposició a estrògens ambientals, que són agents químics que poden actuar com l'estrogen humà i estimular la divisió de les cèl·lules en el teixit mamari. D'altra banda, després d'un embaràs complet, una proporció elevada de les cèl·lules mamàries maduren, estimulades hormonalment, per a produir llet. Aquest tipus de cèl·lules són menys sensibles a carcinògens i la susceptibilitat a mutacions es redueix [Colditz et al., 2006a].

Així doncs, la història natural del CM conté un conjunt d'etapes fisiològiques i patològiques que, en definitiva, comença amb una proliferació epitelial anòmala i progressa cap a carcinomes in situ i invasius i, finalment, culmina amb metàstasi [Burstein et al., 2004]. La metàstasi és la causa del 90% de les morts per tumors sòlids [Gupta and Massague, 2006] i, actualment, el CM

detectat en un estadi primerenc es considera una malaltia curable. Com a conseqüència, molts esforços en la recerca actual van destinats a aprofundir en els coneixements biològics i, més concretament genètics, que es tenen sobre el procediment pel qual aquest tipus de tumors avancen en la progressió i arriben fins a la metastasi [Hu et al., 2008].

2.2 Incidència i mortalitat del càncer de mama

S'estima que l'any 2007 hi van haver 1,3 milions de casos nous de CM en dones al món [Garcia et al., 2007], essent el tipus de càncer més diagnosticat en dones. Pel que fa a la mortalitat, s'estima que 465.000 dones van morir per aquesta causa l'any 2007, el que representa la primera causa de mort per càncer en les dones. Pel que fa a Europa s'estima que l'any 2006 es van diagnosticar 429.900 dones i que van morir per aquesta causa 131.900 dones [Ferlay et al., 2007].

Tant les taxes d'incidència com les taxes de mortalitat per CM presenten un alt nivell de variabilitat quan aquestes s'analitzen a nivell mundial (figures 1.1 i 1.2) i, com hem destacat en el capítol anterior, es constaten diferències importants entre els països desenvolupats respecte els països en desenvolupament, on aquestes taxes són inferiors. A nivell global, la incidència del CM ha augmentat al món en un 30% en els darrers 25 anys, degut a canvis en els models reproductius i, més recentment, degut a la implantació de sistemes de criatge [Colditz et al., 2006b].

A Europa, hi han hagut molts canvis en la implementació de programes de detecció precoç i en el tractament aplicat en el CM des de la meitat dels anys 80 [Veronesi et al., 2005]. Les taxes de mortalitat per CM van créixer des de 1951, experimentant un canvi de tendència a partir de l'any 1990, quan van començar a estabilitzar-se i seguidament a disminuir fins a l'actualitat. Tanmateix les taxes de mortalitat per CM continuen augmentant en alguns països de l'Europa de l'Est [Veronesi et al., 2005], amb l'excepció de la República Txeca [Hery et al., 2008a,b]. Els motius que han influït

en el descens de la mortalitat per CM són, principalment, la implantació de mamografies en programes de cribatge i les millores substancials en el diagnòstic i tractament, sobretot l'ús de tamoxifè [Peto et al., 2000].

L'anàlisi de les tendències temporals de la incidència i mortalitat del CM és d'interès a nivell de salut pública ja que ofereix la possibilitat d'avaluar els efectes dels canvis sanitaris implementats en els països en un període. Les tendències de la incidència són el resultat de la variació en el temps de l'exposició a factors de risc en les poblacions (com la fertilitat decreixent, l'augment de l'alçada i el pes), millores en la diagnosi deguda a la implementació de programes de cribatge o simplement una major consciència pel que fa als temes de salut [Botha et al., 2003]. Les tendències en la mortalitat per CM són el resultat de les variacions en la incidència del CM, implementació de la detecció precoç i el tractament de la malaltia. En les darreres dècades hi ha hagut un interès creixent en l'anàlisi estadística d'aquest tipus de dades per tal d'avaluar, des del punt de vista epidemiològic, l'evolució de la incidència i mortalitat del CM, en un intent de *monitoritzar* l'evolució de la malaltia respecte els recursos sanitaris implementats.

En els treballs més recents [Hery et al., 2008a,b] s'ha pogut constatar que, des dels anys 1960 l'augment en la incidència de CM es va situar entre el 2% i 3% anual, sobretot en dones amb una edat compresa entre els 50 i els 69 anys. Aquest estudi es va realitzar a països desenvolupats on predominava la població caucàsica. També es va constatar, en aquest estudi, que la disminució de la mortalitat per CM, a partir de l'any 2000, va ser del 2% anual per a dones entre 50 i 60 anys. Tot i això aquest estudi també apunta la heterogeneïtat present a Europa que fa que alguns països tinguin disminucions en la incidència (o augments en la mortalitat) molt majors o menors que la resta, com és el cas de la majoria de països de l'Europa de l'Est, en els que es detecten uns patrons marcadament diferenciats.

En qualsevol cas, la interpretació de les tendències de les dades d'incidència de CM i mortalitat per CM requereix un coneixement profund de les dades mateixes, amb les seves limitacions. D'una banda, les dades d'in-

cidència s'obtenen dels registres de càncer poblacionals que emmagatzemen i analitzen informació sobre tots els nous casos de càncer en poblacions ben definides. Tanmateix, a data d'avui els registres només cobreixen el 20% de població mundial [Bray and Moller, 2006], percentatge que augmenta fins al 70% en àrees desenvolupades. Per l'altra banda, les dades de mortalitat habitualment provenen de les oficines nacionals o regionals que registren tots els casos de mort a nivell nacional o regional. D'aquesta forma les dades d'incidència, procedents dels registres de càncer poden no ser totalment exhaustius però les dades tenen una elevada qualitat. Per l'altre costat, les dades de mortalitat, malgrat ser exhaustives i, referir-se a tota la població, són de menor qualitat que les de incidència, degut a la qualitat associada als certificats de defunció, que es considera menor que la qualitat en el diagnosi de la malaltia.

2.3 Supervivència del càncer de mama

La supervivència relativa del CM, mesurada com el percentatge de dones que continuen estant vives 5 anys després del diagnòstic, és del 79% a Europa [Sant et al., 2009] i del 89% als Estats Units [Jemal et al., 2006]. La detecció precoç del CM és clau per a millorar-ne el pronòstic i fer més efectius els tractaments. Per exemple, als Estats Units la supervivència del CM entre les dones americanes és del 98% quan aquest es troba en un estadi primerenc, comparat amb el 84% i el 28% quan la malaltia s'ha estès a ganglis limfàtics regionals o ha metastatitzat en òrgans a distància, respectivament [Jemal et al., 2006].

A Europa, l'estudi *EUROCORE* que, des de l'any 1990, constitueix el major estudi cooperatiu d'Europa sobre supervivència del càncer, i inclou un gran nombre de registres europeus [Berrino, 2003], conté informació sobre tres milions de casos de càncer, provinents de 82 registres de càncer en 23 països europeus. La darrera actualització, *EUROCORE-4*, inclou els individus diagnosticats per càncer en el període 1995-99 que es van seguir fins al desembre

de l'any 2003 [Sant et al., 2009]. La supervivència relativa del CM, reportada per aquest estudi, s'estima que és del 79% a Europa, el que representa un augment respecte el 76% que es va reportar pels individus diagnosticats en 1990-94 (estudi *EUROCORE-3*). Pel que fa a la variabilitat dins d'Europa, els països del Nord així com França, Itàlia i Suïssa van presentar la major supervivència relativa, per sobre del 82%. El Regne Unit va presentar una supervivència relativa del 77.3%, lleugerament inferior a la mitjana, tot i que l'evolució històrica està clarament en augment. Països de l'Est, com Polònia, la República Txeca i Eslovènia tenen supervivències més baixes, inferiors al 73%, degut segurament a la menor implantació de programes de cribatge de càncer i de millores del tractament. La taula 2.1 mostra la supervivència relativa estimada per a cadascun dels països europeus inclosos en l'estudi [Sant et al., 2009].

2.4 Factors de risc associats al càncer de mama

2.4.1 Factors de risc ambientals

Edat

L'edat és el factor de risc més associat al CM, a l'igual que al càncer en general. L'increment en l'edat està clarament associat amb els canvis moleculars, cel·lulars i fisiològics que influeixen en la carcinogènesi i el creixement tumoral subseqüent [Balducci and Ershler, 2005]. La incidència de CM augmenta de manera exponencial amb l'edat, particularment durant la premenopausa, independentment de la població estudiada [Moolgavkar et al., 1980]. La figura 2.1 mostra l'augment de la incidència de CM amb l'edat en quatre poblacions de dones (USA, Àustria, Finlàndia i Girona), segons les estimacions proporcionades per l'Agència Internacional sobre el Càncer [IARC, 2002a]. En aquesta figura es pot constatar l'augment exponencial en la incidència del CM i l'estabilització a partir de la menopausa, independentment de la població de dones analitzada. Les taxes de mortalitat per CM també aug-

País	Supervivència relativa (Interval de confiança al 95%)	
	A 1 any	A 5 anys
Dinamarca	93,4 (92,9 - 94,0)	77,5 (76,5 - 78,6)
Finlàndia	95,7 (95,1 - 96,2)	83,5 (82,4 - 84,6)
Islàndia	97,4 (95,1 - 99,7)	88,0 (82,8 - 93,5)
Noruega	95,6 (95,0 - 96,1)	82,4 (81,3 - 83,6)
Suècia	96,6 (96,2 - 96,9)	84,6 (83,9 - 85,4)
Irlanda	91,4 (90,5 - 92,2)	73,8 (72,2 - 75,3)
Anglaterra	91,8 (91,6 - 92,0)	77,3 (77,0 - 77,7)
Irlanda del Nord	92,5 (91,5 - 93,6)	77,6 (75,6 - 79,6)
Escòcia	91,7 (91,1 - 92,3)	74,9 (73,8 - 76,0)
Gal·les	92,1 (91,4 - 92,9)	76,9 (75,6 - 78,3)
Austria	92,4 (91,9 - 92,8)	80,0 (79,2 - 80,9)
Bèlgica	93,1 (92,5 - 93,7)	77,3 (76,1 - 78,4)
França	96,0 (95,5 - 96,6)	83,1 (82,0 - 84,2)
Alemanya	92,9 (91,8 - 94,0)	78,3 (76,1 - 80,6)
Holanda	95,7 (95,3 - 96,2)	81,3 (80,4 - 82,3)
Suïssa	95,5 (94,8 - 96,3)	82,0 (80,5 - 83,5)
Itàlia	95,8 (95,6 - 96,1)	82,7 (82,2 - 83,2)
Malta	93,1 (90,8 - 95,4)	76,1 (71,8 - 80,7)
Portugal	92,6 (91,5 - 93,8)	77,2 (75,1 - 79,3)
Eslovènia	92,3 (91,1 - 93,5)	71,9 (69,7 - 74,1)
Espanya	95,4 (94,8 - 95,9)	80,3 (79,2 - 81,5)
República Txeca	90,8 (88,9 - 92,8)	69,2 (65,6 - 73,0)
Polònia	90,7 (89,5 - 91,9)	73,7 (71,7 - 75,8)
Europa	93,8 (93,7 - 94,0)	79,4 (79,1 - 79,7)

Taula 2.1: Supervivència del càncer de mama a Europa a 1 any i 5 anys (intervalls de confiança al 95%) segons dades de EURO CARE-4 [Sant et al., 2009]

menten amb l'edat, des de 9 casos per 100.000 dones entre 35-39 anys fins a 98 casos per cada 100.000 dones entre 70 i 74 anys [Ries et al., 2005].

Edat de la menarquia i característiques del cicle menstrual

Una edat precoç en la menarquia s'ha associat consistentment amb un risc augmentat de desenvolupar CM. La majoria d'estudis epidemiològics suggereixen que l'edat precoç de la menarquia està associada tant amb el CM

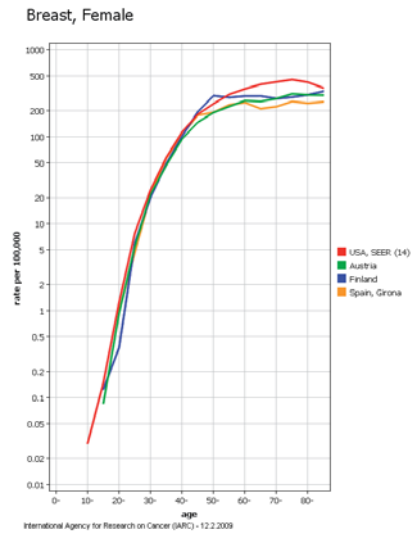


Figura 2.1: Incidència anual mitjana de CM per 100.000 dones per grups d'edat. Dades extretes de la pàgina web de la IARC [Ferlay et al., 2007].

pre com post-menopàusic [Kelsey et al., 1993]. El risc de desenvolupar CM disminueix entre un 10% i un 24% per cada any de retard en l'aparició de la menarquia [Bernstein, 2002].

Paritat, edat en el primer embaràs complet i lactància

Les dones que no han tingut fills tenen un risc de desenvolupar CM superior respecte a aquelles que han completat, almenys, un embaràs. Aquest risc és particularment evident en el CM diagnosticat en dones més grans de 40-45 anys, però no en dones més joves. La reducció del risc de CM després de l'embaràs no és immediat sinó que es manifesta al cap de 10 a 15 anys [Bruzzi et al., 1988]. Un nombre elevat d'embarassos complets també està associat amb un risc inferior de CM, especialment si els embarassos es produeixen en espais de temps propers [Rosner et al., 1994], fet que seria degut a les menors possibilitat d'acumular mutacions en el DNA.

La lactància és un factor protector conegut del CM. S'ha descrit que aquelles dones amb paritat alta que hagin alletat als seus fills durant 2 anys o més evidencien una disminució del risc de patir CM en un 50% [Romieu

et al., 1996]. És a dir, el risc relatiu de CM disminueix en un 4,3% per cada 12 mesos de lactància, i en un 7% per cada naixement [Collaborative Group on Hormonal Factors in Breast Cancer, 2002]

Edat en la menopausa

La taxa d'incidència del CM s'estabilitza a partir de la menopausa, que marca la fi del cicle mensual hormonal que indueix la proliferació de les cèl·lules mamàries. Diversos estudis epidemiològics han indicat que una edat precoç de la menopausa està associat amb un menor risc de desenvolupar CM [Trichopoulos et al., 1972]. De mitjana, s'ha establert que el risc de CM augmenta en un 3% per cada any de retard en l'inici de la menopausa [Colditz and Rosner, 2000].

Contraceptius orals

Inicialment, es va creure que el contraceptius orals augmentaven el risc de desenvolupar CM, ja que contenen concentracions d'estrogen i progestina que podrien ser majors que els nivells d'aquestes hormones produïdes per una dona durant el cicle ovulatori fisiològic [Bernstein, 2002]. Tot i això, estudis epidemiològics recents han mostrat que l'augment de risc de desenvolupar CM és petit o fins i tot nul. En una meta-anàlisi es va veure que el risc de desenvolupar CM per a usuàries actuals de contraceptius orals respecte les dones que mai els havien utilitzat era de 1,24, mentre que el risc relatiu per a dones que no els havien emprat des de feia més de 4 anys i 9 anys era de 1,16 i 1,07 respectivament [Collaborative Group on Hormonal Factors in Breast Cancer, 1996].

Tractament de substitució hormonal posterior a la menopausa

La possible relació entre l'ús d'estrògens després de la menopausa i el risc de desenvolupar CM s'ha investigat en un nombre important d'estudis epidemiològics en els darrers 20 anys. Una meta-anàlisi basada en 51 estudis

epidemiològics va estimar una associació estadísticament significativa entre l'ús actual o recent d'hormones substitutives en la postmenopausa i el risc de desenvolupar CM, identificant-se una relació dosi-resposta, ja que l'associació va ser més important en les dones que van emprar més temps aquest tipus de tractament [Collaborative Group on Hormonal Factors in Breast Cancer, 1997]. No es va observar, tanmateix, cap augment significatiu en el risc de CM per aquelles dones que havien deixat el tractament de substitució hormonal en un període superior als 5 anys, independentment del temps de duració del tractament.

D'altra banda, darrerament l'ús de progestina afegit als estrògens ha esdevingut cada cop més freqüent, ja que minimitza el risc d'una hiperplàsia en l'endometri. Uns primers estudis epidemiològics varen apuntar que l'impacte de la progestina afegida als estrògens podria disminuir el risc de CM [Gambrell et al., 1983], tot i que aquests estudis no tenien en compte altres factors associats a aquest càncer com l'edat o la paritat. Estudis epidemiològics posteriors varen indicar un efecte protector d'aquest tractament combinat [Collaborative Group on Hormonal Factors in Breast Cancer, 1997]. Per una altra banda, algunes dades suggereixen que l'augment en el risc de CM associat a l'estrogen més progestina podria ser superior respecte l'ús d'estrogen aïlladament. Un estudi va determinar que el risc de CM augmentava al voltant d'un 1% per cada any d'ús d'estrogen i al voltant d'un 8% quan aquest s'acompanyava de progestina [Schairer et al., 2000]. Estudis més recents semblen indicar que la teràpia de substitució hormonal augmenta el risc de recurrència en CM [Holmberg and Anderson, 2004, Holmberg et al., 2008]. L'ús creixent d'estrògens i progestina és molt recent, hi ha poques dades disponibles i, a més, donat les diferents formulacions, dosis i esquemes d'ús, fa que encara que no es disposin de dades concloents que permetin avaluar-ne l'efecte sobre el risc de desenvolupar aquesta malaltia.

Greixos en la dieta

La relació entre la ingesta de greix i el risc de CM ha centrat l'atenció de diversos estudis epidemiològics. Alguns estudis de casos i controls han indicat una feble associació entre la ingesta de greix i el risc d'aquest càncer [Howe et al., 1990]. En canvi, una anàlisi combinada que va incloure estudis prospectius no va observar cap associació entre la ingesta de greix i el CM, diferenciant segons si es tractava de greixos monosaturats, poliinsaturats o saturats [Hunter et al., 1996]. En els estudis amb animals, la majoria conclouen una major incidència de tumors mamaris deguda a una dieta elevada en greixos poliinsaturats.

Fibra

S'ha hipotetitzat que un consum elevat de fibra està associat a un menor risc de CM, degut a que la fibra inhibeix la re-absorció d'estrògens [Goldin et al., 1982]. Estudis epidemiològics prospectius, tanmateix, han mostrat una associació lleu o nul·la entre la ingesta de fibra i el risc de desenvolupar aquesta malaltia.

Micronutrients, fruites i verdures

S'ha examinat el paper de les vitamines A, C i E i els carotenoides sobre el CM, degut a que aquests nutrients tenen una funció antioxidant que podria implicar un dany menor a l'ADN. En aquest sentit, hi ha de moment una feble evidència en quant al paper protector del retinol en el CM. Pel que fa a la ingesta de β -carotens la majoria d'estudis epidemiològics han detectat que el risc de desenvolupar CM es redueix amb majors ingestes de β -carotens. En canvi, majors ingestes de vitamines C i E no semblen ser factors protectors del CM [Willett, 2001].

Les fruites i les verdures, junt amb els cereals i els suplementes de vitamines, són les fonts principals d'ingesta de la majoria d'aquests nutrients. Hi ha actualment alguna evidència del paper protector de la ingesta de fruites

i verdures sobre el risc de càncer de mama, tot i que en un estudi combinat prospectiu no es va observar cap associació significativa [Smith-Warner et al., 2001]. Així doncs, avui en dia, roman encara en qüestió el paper d'aquests nutrients en l'etiologia de la neoplàsia de mama.

Alcohol

L'associació entre el consum d'alcohol i el risc de CM s'ha examinat en més de 100 estudis, conclouent una clara associació causal. En una meta-anàlisi de 38 estudis de casos i controls es varen estimar uns riscos relatius de 1,1 per una beguda alcohòlica al dia, de 1,2 per a dues begudes i de 1,4 per a tres, essent tots ells estadísticament significatius [Longnecker, 1994].

Factors antropomètrics

Estudis epidemiològics realitzats en diferents poblacions han determinat que l'alçada està positivament relacionada amb el risc de CM. En una meta-anàlisi, el risc relatiu per cada increment en l'alçada de 5 centímetres, després d'ajustar per altres factors de risc, va ser de 1,07, essent aquest estadísticament significatiu [van den Brandt et al., 2000]. De fet, l'alçada es pot entendre com el resultat d'un conjunt de factors genètics i ambientals, sobretot amb el que té a veure amb la ingesta d'energia durant la infància. Així doncs, l'associació entre l'alçada i el risc de CM sembla menor en les poblacions en les que el creixement infantil està sotmès a limitacions en el consum d'energia, el que suggereix que la ingesta d'energia en la infància tindria un paper en la carcinogènesi del CM [Friedenreich, 2001].

D'altra banda, s'ha constatat una relació inversa entre el pes corporal (generalment a partir de l'índex de massa corporal, IMC) i la incidència del CM en la premenopausa [Ursin et al., 1995]. L'associació entre l'IMC i el risc de desenvolupar CM entre dones en la postmenopausa és dèbilment positiu [Howe et al., 1990].

Finalment, l'activitat física pot prevenir o retardar el CM, ja que redueix la freqüència dels cicles ovulatoris i disminueix el nivell d'estrògens, dismi-

nuint l'obesitat en dones en la postmenopausa [Brinton et al., 1998].

Tabac

Tot i que hi han hagut nombrosos estudis epidemiològics per avaluar l'efecte del tabac sobre el risc de desenvolupar CM, les dades no proporcionen una evidència clara a favor d'una possible associació, tal i com es desprèn dels resultats obtinguts en l'estudi més nombrós de casos i controls realitzat fins a data d'avui [Baron et al., 1996].

2.4.2 Factors de risc endògens

Hormones sexuals endògenes

Les diferències en les taxes d'incidència de CM en les dones pre i postmenopàusiques suggereixen el paper central que desenvolupen les hormones sexuals en la etiologia d'aquest càncer. Després de la menopausa, el teixit adipós és la major font d'estrògen, i les dones postmenopàusiques obeses tenen elevats nivells d'estrògens endògens i un risc elevat de CM [Harris et al., 1992]. També, manipulacions hormonals basades en anti-estrògens (tals com tamoxifè i raloxifè) redueixen la incidència d'aquest càncer [Fisher et al., 1998]. Recentment, anàlisis combinades d'estudis epidemiològics han mostrat que l'estrògen i la testosterona contribueixen significativament a l'increment del risc de desenvolupar CM [Collaborative Group on Hormonal Factors in Breast Cancer, 2002].

IGF-I

El factor de creixement anàleg a la insulina (*Insulin-like growth factor, IGF-I*) és una hormona amb una homologia similar a la insulina i es regula principalment a partir de l'hormona de creixement. Existeix a data d'avui una evidència creixent que aquesta hormona estimula la proliferació del CM [Yang et al., 1996]. De fet, el pes i l'alçada al naixement estan positivament correlacionats amb els nivells de IGF-I [Lassarre et al., 1991].

Susceptibilitat genètica i familiar

Una història familiar que inclogui casos de CM constitueix un factor de risc ben establert per aquest tumor. Més concretament, les dones amb un cas d'aquest càncer en un familiar de primer grau presenten un increment del risc del CM d'un 50%. De mitjana, entre el 5% i el 10% dels casos de CM tenen el seu origen en mutacions genètiques heretades [Bennett et al., 1999]. Els CM d'origen hereditari es caracteritzen per ser diagnosticats en edats primerenques.

Entre el 2% i 5% dels CM tenen el seu origen en una mutació en BRCA1 o bé BRCA2 [Easton et al., 1993], ja que s'ha vist que aquests gens tenen un paper important en la estabilitat genòmica, reparació del DNA i en el control del cicle cel·lular [Unger and Weber, 2000]. D'altra banda, el gen p53 és un gen supressor associat també amb el risc de desenvolupar aquest càncer. Alguns altres gens que han estat associats al CM són PTEN, ATM, CYP17, CYP19, 17 β -HSD, CYP1A1 i CYP1B1.

2.5 Prevenció del càncer de mama

Prevenció primària

Malauradament, el coneixement actual sobre l'etiologia del CM no es tradueix fàcilment en mesures per a la prevenció primària d'aquest tumor [Colditz et al., 2006a]. La majoria dels factors de risc que s'han establert o que es sospita que puguin estar associats al CM estan resumits a la taula 2.2. Alguns d'aquests factors de risc, com per exemple la menarquia i la menopausa, tot i està prou establerts són difícils de modificar. Altres, en canvi estan molt lligats al paper de la dona en la societat desenvolupada (retard en el primer embaràs, no lactància o menor nombre d'embarassos). En canvi, tot i no està molt establerts, alguns factors presenten beneficis importants que en justifiquen la seva modificació (per exemple, substituir una dieta en greixos saturats per greixos monosaturats amb suplementes vitamínics). A més, la

majoria de factors de risc establerts presenten un risc relatiu de magnitud moderada, normalment dins el rang 1,3-1,8. Tanmateix, els factors de risc lligats a comportaments tant penetrats en una societat són difícils de modificar.

Una edat primerenca en el primer embaràs complet, redueix substancialment el risc de CM, tot i que les tendències actuals en la societat dels països desenvolupats van en la direcció contrària. Mesures socials com la compatibilitat de l'horari familiar amb el laboral podrien afavorir embarassos en edats més joves i lactàncies naturals, durant més temps que podrien conduir a una disminució de la incidència del CM.

El consum d'alcohol comporta un conjunt d'efectes que són adversos i d'altres de beneficiosos. El consum d'alcohol està associat amb un major risc de CM. D'aquesta forma, dones de mitjana edat amb un consum diari d'alcohol podrien reduir el risc de CM reduint el consum d'alcohol i, també estaria indicat un suplement de vitamines i àcid fòlic [Zhang et al., 2003].

L'ús d'hormones en la postmenopausa sembla està associat també a un major risc d'aquest càncer, pel que la estratègia òptima seria deixar aquesta pràctica. Tot i això, donat que es tracta d'un tractament nou, amb moltes variants, tot indica que caldrà esperar a recollir noves evidències.

Evitar un guany de pes excessiu també sembla influenciar el risc de desenvolupar CM. Substituir els greixos saturats per monosaturats conjuntament amb un augment de l'activitat física podrien tenir una influència en la disminució no només del risc de desenvolupar CM de mama sinó també malalties cardiovasculars i diabetis.

L'aparició de la menarquia en edats precoces és degut, en els països més desenvolupats, al ràpid creixement i guany de pes de les nenes degut a una alimentació abundant, bon estat de salut i baix nivell d'activitat física. Òbviament, aquesta situació es majoritàriament desitjable i no hi ha motius per intentar retardar l'edat en la menarquia, com per exemple s'ha pogut observar en algunes àrees rurals de la Xina. D'altra banda, el temps que els nens miren diàriament la televisió és un determinant del pes en excés dels nens

[Berkey et al., 2003] i per tant un factor a reduir per tal de disminuir el risc de desenvolupar aquest càncer així com les malalties cardiovasculars i diabetis. En aquest sentit, la societat i l'escola haurien de promoure l'activitat física en els menors especialment.

Prevenió secundària: cribatge

Avui en dia existeix un ampli consens en que un cribatge de CM poblacional i planificat és efectiu per a reduir la mortalitat per CM. Tant les mamografies com els exàmens clínics tenen un bon nivell de sensibilitat, que es veu augmentada en el cas que s'apliquin en paral·lel. Així doncs, un programa de mamografies per a dones majors de 50 anys és d'una utilitat indiscutible. De fet, les dones que han seguit programes de cribatge de CM tenen una reducció entre el 25 i el 30 % en el risc de morir per aquesta malaltia respecte les dones que no han seguit programes de cribatge [Smith et al., 2003].

Hi han també alguns estudis que demostren que la mamografia realitzada en dones de més de 40 anys pot disminuir el risc de mort per CM. Això podria ser explicat en part per que el teixit mamari en les dones joves tendeix a tenir major densitat, factor que influència que les mamografies en aquestes edats tinguin una limitació major per a distingir teixit normal respecte de l'anòmal i, per tant una menor sensibilitat i especificitat.

Prevenió terciària

El tercer nivell de prevenció per al CM, un cop aquest ha estat diagnosticat, inclou el tractament, la rehabilitació i el control del dolor. Sempre que sigui possible, la cirurgia és la primera línia de tractament per a qualsevol dona diagnosticada amb un CM. L'operació pot consistir en eliminar el tumor (tumorectomia), o el pit sencer (mastectomia) i també en alguns casos eliminar els ganglis limfàtics de l'aixella. Aleshores, depenent de l'estadi en que es troba el càncer, es pot indicar una segona línia de tractament, tal com quimioteràpia, radiació o teràpia hormonal, amb l'objectiu d'evitar que el càncer es torni a desenvolupar. Aquests tractaments es coneixen com a adju-

vants, ja que s'apliquen conjuntament amb el tractament principal quirúrgic. Darrerament, també s'estan emprant tractament neoadjuvants, consistents en optar primer per la quimioteràpia i després per la cirurgia. Habitualment, la teràpia basada en radiació s'empra només en els tractaments quirúrgics en els que es conservi el pit (o una part), mentre que la quimioteràpia s'utilitza en els altres casos, tot i que també és força habitual emprar els dos tractaments de forma seqüencial. La teràpia hormonal, basada habitualment en l'ús del tamoxifè, s'utilitza habitualment en càncers precoços i avançats, específicament per a dones amb els receptors hormonals d'estrògens positius (ER+). En aquests casos, les cèl.lules canceroses requereixen estrògens per a proliferar, essent el tamoxifè un tractament adequat ja que inhibeix la producció d'estrògens en el teixit mamari. Per altra banda, el tractament basat en quimioteràpia consisteix en l'administració de fàrmacs citostàtics amb l'objectiu d'eliminar les cèl.lules tumorals del cos. Per això, aquests fàrmacs actuen de forma que pertorben el desenvolupament de les cèl.lules (tant les tumorals com les normals). Les cèl.lules normals poden normalment reparar els efectes de la quimioteràpia. Habitualment, s'utilitzen combinacions de fàrmacs seguint diferents esquemes d'administració.

Factors reproductius	RR	Factors hormonals	RR	Factors nutricionals i antropomètrics	RR	Altres factors	RR
Edat a la menarquia (≥ 15 vs. 11)	-	Substitució d'èstrogens (<5 anys vs. cap)	+	Greixos monosaturats	-	Història familiar (Familiar en primer grau)	+++
Edat al primer embaràs complet (≥ 35 vs. ≤ 20)	++	Ús de contraceptius orals (actual vs. mai)	+	Greixos saturats	+	Hiperplàsia mama	++
Nombre de naixements (0 o 1 fills)	+	Èstrogens elevats sang (postmenopausa)	+++	Alcohol (>1 beguda/dia vs. cap)	+	Radiacions ionitzants (si/no)	+
Edat a la menopausa (increment de 5 anys)	+	Nivell elevat IGF-1 sang (premenopausa)	+++	Alçada ($>5,7''$)	+		
Lactància (>1 any vs. cap)	-	Nivell prolactina elevat en sang	++	Obesitat premenopausa (>27 vs. <21 IMC)	-		
				Obesitat postmenopausa (>27 vs. <21 IMC)	+		
				Activitat física ≥ 3 hores setmana	-		

Taula 2.2: Factors de risc associats al càncer de mama i força de l'associació

RR: risc relatiu $+1,1 \leq RR \leq 1,4$; $++1,5 \leq RR \leq 2,9$; $+++3,0 \leq RR \leq 6,9$; $-0,7 \leq RR \leq 0,8$

Capítol 3

Mètodes estadístics

Els mètodes estadístics en l'epidemiologia descriptiva es van caracteritzar inicialment per tenir objectius primordialment exploratoris. Sovint, l'epidemiologia analítica pretén principalment confirmar hipòtesis, generades en l'epidemiologia descriptiva [Esteve et al., 1994]. En aquest sentit, els mètodes emprats per a l'anàlisi de dades d'incidència o mortalitat d'una malaltia, són similars als que s'utilitzen habitualment en demografia, i van compartir una mateixa base en el seu origen. Tanmateix, l'acumulació de dades de mortalitat i incidència ha motivat l'aparició de models estadístics sofisticats, tant amb usos explicatius com predictius. D'aquesta forma, avui en dia, l'epidemiologia descriptiva moderna disposa doncs d'unes eines estadístiques formals per a l'anàlisi de les dades, formulació d'hipòtesis i el seu contrast inferencial, sense deixar de banda els mètodes més exploratoris, que continuen sent útils i necessaris a la pràctica.

En aquest capítol descrivim els mètodes estadístics emprats en aquesta tesi, que tenen a veure amb la modelització de les taxes d'incidència o mortalitat, ja sigui per a inferir i descriure característiques de la població analitzada com per a predir l'evolució d'aquestes poblacions en el futur, pel que fa al comportament de la incidència o mortalitat d'una malaltia determinada, tenint en compte els aspectes demogràfics inherents a les poblacions estudiades.

3.1 Taxes d'incidència o mortalitat

Les taxes d'incidència o mortalitat d'una malaltia, tal com el càncer, han esdevingut els indicadors poblacionals en epidemiologia per a analitzar l'evolució de les malalties en una població, i avaluar l'impacte de nous tractaments i programes de salut implementats.

Habitualment, la incidència (o mortalitat) d'una malaltia es recull de forma agregada per grups d'edat. Denotarem per i els diferents grups d'edat, $i = 1, \dots, I$. Considerarem, i aquesta és la situació per defecte, que hi han 18 grups d'edat ($I = 18$) que corresponen als grups d'edat quinquennals següents: $\{0-4 \text{ anys}\}$, $\{5-9 \text{ anys}\}$, \dots , $\{80-84 \text{ anys}\}$, $\{85 \text{ o més anys}\}$. Igualment, denotarem per t el període en el temps en el que s'han observat les dades d'incidència o mortalitat. Habitualment, i també assumirem aquesta situació per defecte, t és un enter que fa referència a un període annual sencer.

D'aquesta forma, denotarem per c_{it} el nombre de casos (incidents o morts) d'una determinada malaltia per al grup d'edat i en el període t . Assumirem que c_{it} és un enter positiu. D'altra banda, denotarem per n_{it} el nombre de persones a risc durant el període observat pel mateix grup d'edat, és a dir la mida de la població total estudiada, en el grup d'edat corresponent. Formalment, la incidència (o mortalitat) d'una malaltia es defineix com la raó entre el nombres de casos i el nombre de persones a risc, que denotarem per $M_{it} = \frac{c_{it}}{n_{it}}$. Assumirem que $0 \leq M_{it} \leq 1$, el que requereix suposar que una mateixa persona no pot ser cas més d'un cop en un mateix període, el que és obvi pels casos morts, però pot ser controvertit en algunes situacions particulars en casos incidents, tot i que la situació és poc freqüent a la pràctica i, en el cas del càncer, rara.

Suposem ara que λ_{it} és la probabilitat de desenvolupar una malaltia o de morir a causa seva per un grup d'edat i durant el transcurs del període t . Aquesta probabilitat, que en el cas del càncer sovint s'anomena *risc de càncer*, és la probabilitat per a una persona pertanyent a una determinada població de ser diagnosticada o de morir per aquesta causa durant el període

concret (un any habitualment). Des del punt de vista estadístic, c_{it} és una variable aleatòria amb esperança $E[c_{it}] = \lambda_{it}n_{it} = \theta_{it}$. Suposarem doncs, que la població a risc en un període concret és un valor fixat no aleatori. Aquesta és una suposició discutible, ja les poblacions són dinàmiques, però que, des del punt de vista estadístic, ens permetrà caracteritzar la distribució de probabilitat del nombre de casos segons lleis de Poisson independents [Breslow and Day, 1987], tal com segueix:

$$c_{it} \sim \mathcal{P}(\theta_{it} = \lambda_{it}n_{it}) \quad (3.1)$$

La incidència observada per un grup d'edat durant un període qualsevol, M_{it} , és doncs l'estimador del risc de càncer λ_{it} , el que es pot denotar $M_{it} = \hat{\lambda}_{it}$. Tot i això, aquesta probabilitat és sovint poc interpretable des del punt de vista epidemiològic, doncs per malalties com el càncer la proporció de persones que es diagnostiquen o moren per un tipus de tumor és molt petita. D'aquesta forma, es defineixen les taxes d'incidència o mortalitat, com el nombre de casos incidents o morts per cada 100.000 habitants, que denotem per $r_{it} = M_{it} \cdot 100000 = \frac{c_{it}}{n_{it}} \cdot 100000$. A la pràctica, resulta convenient donar un estimador de la taxa que faci referència a tota la població, independentment del grup d'edat, que es denota simplement per r_t i que s'obté agregant el nombre de casos i de persones a risc dels grups. Aquesta taxa es coneix habitualment com a *taxa crua* o *taxa bruta* i es calcula tal com segueix:

$$r_t = \frac{\sum_{i=1}^I c_{it}}{\sum_{i=1}^I n_{it}} \cdot 100000 \quad (3.2)$$

Les taxes d'incidència o mortalitat crues són doncs els indicadors poblacionals habitualment emprats en epidemiologia per a avaluar l'evolució i càrrega d'una malaltia en una població. Tanmateix, sovint es desitja comparar les taxes en diferents regions, grups, anys, entre altres. En aquests casos, l'efecte de l'edat podria dificultar la comparació [Esteve et al., 1994], ja que diferents poblacions o grups poden tenir unes estructures de grups d'edat molt diferenciades i, per tant, provocar diferències entre les taxes. Per exem-

ple, en el cas del càncer, donat que l'edat és el factor de risc més associat a la malaltia, pot ser que les diferències observades entre dues poblacions, tals com dos països, siguin degudes majoritàriament a que, per exemple, en una d'elles hi hagi una població més envellida i, per tant, presenti unes taxes d'incidència i mortalitat per càncer superiors. Aquest és el motiu pel que aquestes taxes, habitualment, es solen estandaritzar (o ajustar) per una població de referència, de tal forma que les taxes facin referència a una mateixa població i les diferències observades entre dues poblacions puguin ser atribuïbles a diferències en l'exposició a factors de risc (pel que fa a la incidència) o en diferències en la implementació de programes de cribatge o tractament (pel que fa a la mortalitat).

Habitualment s'empra el *mètode directe* per a ajustar les taxes d'incidència o mortalitat crues. Suposem que disposem d'una població de referència, més específicament, els pesos proporcionals de cadascun del grups d'edat per a una població teòrica, que denotarem per w_i , de tal forma que $\sum_{i=1}^I w_i = 1$. Aleshores les *taxes ajustades per edat* segons aquesta població, que denotarem per r_t^{adj} , es poden calcular tal com segueix:

$$r_t^{adj} = \sum_{i=1}^I w_i r_{it} = \sum_{i=1}^I w_i \frac{c_{it}}{n_{it}} \cdot 100000 = \sum_{i=1}^I w_i M_{it} \cdot 100000 \quad (3.3)$$

Anàlogament, la incidència (o mortalitat) ajustada per edat es defineix com $M_t^{adj} = \sum_{i=1}^I w_i M_{it}$, que s'interpreta doncs com el risc associat a la malaltia en el supòsit que la població de referència tingui una piràmide de població segons els pesos per als grups d'edat w_i .

Habitualment es sol emprar, com a població de referència, una estimació de la població mundial realitzada en la dècada del 1950 [Segi, 1960], tot i que també es habitual l'ús de la població europea o bé, l'africana quan es volen comparar dades de països en desenvolupament. La taula 3.1 recull els pesos w_i per diferents poblacions de referència [Esteve et al., 1994]. Sovint es calculen el que s'anomenen les *taxes ajustades truncades*, amb l'objectiu d'analitzar la incidència o mortalitat d'una malaltia per només alguns grups

Grup d'edat	Població de referència			Truncada
	Mundial	Africana	Europea	35 a 64 anys
0-4	0,12	0,1	0,08	0
5-9	0,1	0,1	0,07	0
10-14	0,09	0,1	0,07	0
15-19	0,09	0,1	0,07	0
20-24	0,08	0,1	0,07	0
25-29	0,08	0,1	0,07	0
30-34	0,06	0,1	0,07	0
35-39	0,06	0,1	0,07	$0,06/0,31 = 0,194$
40-44	0,06	0,05	0,07	$0,06/0,31 = 0,194$
45-49	0,06	0,05	0,07	$0,06/0,31 = 0,194$
50-54	0,05	0,03	0,07	$0,05/0,31 = 0,161$
55-59	0,04	0,02	0,06	$0,04/0,31 = 0,129$
60-64	0,04	0,02	0,05	$0,04/0,31 = 0,129$
65-69	0,03	0,01	0,04	0
70-74	0,02	0,01	0,03	0
75-79	0,01	0,005	0,02	0
80-84	0,005	0,003	0,01	0
85 o més	0,005	0,002	0,01	0
Total	1	1	1	1

Taula 3.1: Estructura dels grups d'edat (a partir dels seus pesos w_i) per diferents poblacions de referència habitualment emprades [Esteve et al., 1994]

d'edat, com per exemple aquells en els que hi ha major incidència o mortalitat de la malaltia estudiada. En el cas del càncer s'acostuma a utilitzar la taxa truncada per edat entre els 35 i 64 anys. En qualsevol cas, les taxes truncades s'obtenen doncs emprant uns pesos w_i que sumin la unitat i que, en alguns grups d'edat tindran un valor nul. La taula 3.1 mostra aquests pesos per al càlcul de taxes ajustades per la població mundial, truncades entre els 35 i 64 anys, com és habitual en el cas del càncer i altres malalties associades fortament amb l'edat. En aquests casos, el motiu d'obviar els grups d'edat majors (65 anys i més) és el d'eliminar la contribució d'aquests ja que sovint es tracta de casos infra-registrats, és a dir, que s'assumeix que podrien observar-se menys casos dels que realment hi han. Òbviament, a

la pràctica, pot ser raonable emprar altres poblacions de referència, com per exemple emprar la piràmide de població d'un país sencer per ajustar les taxes d'incidència de les diverses regions del mateix país. Convé remarcar que en el cas d'emprar la població mundial estimada en la dècada dels 1950, un estudi recent va concloure que, per qüestions pràctiques, no era necessari actualitzar-la amb una població mundial més recent [Bray et al., 2002].

Les taxes ajustades es calculen doncs a partir del nombre de casos observats en cadascun dels grups d'edat, c_{it} , i, per tant, es tracta de l'observació d'una variable aleatòria sotmesa a variació. Per tal d'avaluar aquest error d'estimació convé doncs presentar l'estimació de les taxes juntament amb el seu error estàndard i/o l'interval de confiança corresponent. Com hem exposat prèviament, el nombre de casos segueix una distribució de Poisson, $c_{it} \sim \mathcal{P}(\lambda_{it}n_{it})$, amb el que la seva variància equival a l'esperança,

$$E(c_{it}) = \text{Var}(c_{it}) = \lambda_{it}n_{it} \quad (3.4)$$

A partir d'aquí es pot obtindre la variància de la incidència (o mortalitat) també anomenat risc com:

$$\text{Var}(M_{it}) = \text{Var}\left(\frac{c_{it}}{n_{it}}\right) = \frac{\text{Var}(c_{it})}{n_{it}^2} = \frac{\lambda_{it}}{n_{it}} \quad (3.5)$$

Aleshores, la variància de la incidència (o mortalitat) ajustada per edat s'obté de la forma següent,

$$\text{Var}(M_t^{adj}) = \sum_{i=1}^I \text{Var}(w_i M_{it}) = \sum_{i=1}^I w_{it}^2 \text{Var}(M_{it}) = \sum_{i=1}^I w_{it}^2 \frac{\lambda_{it}}{n_{it}} \quad (3.6)$$

Donat que λ_{it} és desconegut, aquest es pot estimar a partir de les dades, mitjançant el seu estimador descrit prèviament, $\hat{\lambda}_{it} = \frac{c_{it}}{n_{it}}$, i, aleshores es pot obtenir l'expressió per la variància estimada següent,

$$\widehat{\text{Var}}\left(M_t^{adj}\right) = \sum_{i=1}^I w_{it}^2 \frac{c_{it}}{n_{it}} = \sum_{i=1}^I \left(\frac{w_{it}}{n_{it}}\right)^2 c_{it} \quad (3.7)$$

En aquest punt, s'assumeix que l'estimador del risc té una distribució asimptòticament Normal, ja que es tracta d'una suma ponderada de variables aleatòries independents [Boyle and Parkin, 1991]. Així doncs, si denotem el risc ajustat poblacional per $\lambda_t^{adj} = E(M_t^{adj})$, es pot obtenir la distribució per a M_t^{adj} tal com segueix,

$$M_t^{adj} \sim N\left(\lambda_t^{adj}, \widehat{\text{Var}}\left(M_t^{adj}\right)\right) \quad (3.8)$$

La taxa centrada i reduïda té aleshores una distribució Normal estandaritzada, que s'expressa,

$$\frac{M_t^{adj} - \lambda_t^{adj}}{\sqrt{\widehat{\text{Var}}\left(M_t^{adj}\right)}} \sim N(0, 1) \quad (3.9)$$

Aleshores, es pot calcular un interval de confiança per al risc ajustat poblacional, amb un nivell de confiança $1 - \alpha$ tal com segueix,

$$\lambda_t^{adj} \in \left[M_t^{adj} - z_{\alpha/2} \sqrt{\widehat{\text{Var}}\left(M_t^{adj}\right)}, M_t^{adj} + z_{\alpha/2} \sqrt{\widehat{\text{Var}}\left(M_t^{adj}\right)} \right] \quad (3.10)$$

A la pràctica, però, interessa expressar els intervals de confiança en termes de les taxes ajustades, en comptes del risc que sol ser una quantitat propera a zero. Simplement, cal multiplicar la variància per 10^{10} tal com segueix,

$$\widehat{\text{Var}}\left(r_t^{adj}\right) = \widehat{\text{Var}}\left(M_t^{adj} \cdot 100000\right) = \widehat{\text{Var}}\left(M_t^{adj}\right) \cdot 10^{10} \quad (3.11)$$

D'aquesta forma, seguint un raonament anàleg a l'emprat per λ_t^{adj} podem obtenir l'interval de confiança per a la taxa ajustada poblacional, que denotem per $\rho_t^{adj} = \lambda_t^{adj} \cdot 100000$, tal com segueix,

$$\rho_t^{adj} \in \left[r_t^{adj} - z_{\alpha/2} 10^5 \sqrt{\hat{\text{Var}}(M_t^{adj})} \quad , \quad r_t^{adj} + z_{\alpha/2} 10^5 \sqrt{\hat{\text{Var}}(M_t^{adj})} \right] \quad (3.12)$$

Convé remarcar de nou que el procediment que permet construir els intervals de confiança, a partir de l'error estàndard estimat, implica que la distribució de l'estimador de les taxes és aproximadament Normal. Això, a la pràctica, és assumible quan el nombre de casos és suficientment gran. Com és habitual, en estadística, resulta difícilment objectivable definir quan es pot considerar que una mostra és *suficientment gran*. Una altra assumpció que cal fer és que el numerador de les taxes, el nombre de casos, segueixi una distribució de Poisson. En aquest punt, tot i que menys conflictiu des del punt de vista empíric, poden sorgir problemes pel que fa a la sobre-dispersió. Aquest és un fet prou estudiat en dades discretes, tals com lleis de Poisson, en les que la variabilitat empíricament observada és superior a la esperada, que coincideix amb l'esperança. Aquestes situacions es poden avaluar estadísticament, per exemple a partir del coeficient de variació, i s'han desenvolupat mètodes estadístics per a esmenar aquestes situacions. A la pràctica, però, la majoria de cops s'assumeixen aquestes dues hipòtesis pel que fa a les dades, tot i que cal ésser conscient que els intervals de confiança podrien ser més amples, degut a un major error estàndard. En qualsevol cas, emprar aquest escenari més simple representa una solució conservadora i, per tant, que és assumible des del punt de vista empíric, malgrat que la potència estadística es vegi limitada.

3.2 Comparació de taxes

Quan es desitgen comparar les taxes d'incidència o mortalitat d'una malaltia en diferents poblacions, és preferible emprar les taxes ajustades per edat a una població de referència prèviament escollida. Aleshores, per avaluar si les diferències observades són reals o bé degudes a l'atzar, en una primera

aproximació es podrien comparar els intervals de confiança, tot i que no és sempre suficient per avaluar la significació estadística. A continuació descriuim mètodes estadístics senzills per a resoldre aquest problema específic.

Comparació de taxes crues

En el cas que es vulguin comparar dues taxes d'incidència o mortalitat crues, siguin aquestes r_1 i r_2 , podem construir un test estadístic senzill a partir de les seves variàncies, que es poden obtenir emprant l'assumpció sobre la distribució del nombre de casos descrita en l'expressió 3.1. Si denotem per n_1 i n_2 el nombre de persones a risc en les dues poblacions, per un any concret i per tots els grups d'edat, aleshores la distribució de probabilitat de la diferència de taxes té com a variància,

$$\text{Var}(r_1 - r_2) = \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \Phi \quad (3.13)$$

on Φ representa la taxa teòrica comuna a ambdós poblacions, és a dir, $\Phi = \frac{\lambda_1 n_1 + \lambda_2 n_2}{n_1 + n_2} \cdot 100000$, i λ_1 i λ_2 són els riscos teòrics en les poblacions respectives. Sota la hipòtesi nul·la, que ambdues poblacions comparteixen el mateix risc ($H_o : \lambda_1 = \lambda_2$), aquesta es pot estimar tal com segueix.

$$\hat{\Phi} = \frac{n_1 r_1 + n_2 r_2}{n_1 + n_2} \quad (3.14)$$

el que ens permet estimar la variància anterior com a,

$$\hat{\text{Var}}(r_1 - r_2) = \frac{n_1 r_1 + n_2 r_2}{n_1 n_2} \quad (3.15)$$

Finalment, podem calcular un estadístic que ens permeti avaluar la hipòtesi nul·la d'igualtat,

$$Z = \frac{r_1 - r_2}{\sqrt{\frac{n_1 r_1 + n_2 r_2}{n_1 n_2}}} \quad (3.16)$$

D'aquesta forma, l'estadístic Z es distribueix segons una normal estandaritzada sota la hipòtesi nul·la, el que ens permet resoldre la prova d'hipòtesi

$H_o : \lambda_1 = \lambda_2$ versus $H_a : \lambda_1 \neq \lambda_2$ mitjançant el càlcul del p-valor que segueix,

$$p = 2\Pr(N(0, 1) \geq |Z|) \quad (3.17)$$

Òbviament, també es pot emprar la distribució de l'estadístic, emprant la variància expressada en 3.15, per a construir intervals de confiança o per a dur a terme altres procediments d'inferència estadística.

Comparació de taxes ajustades

En el cas de comparar dues taxes ajustades per a un any concret, que denotarem per r_1^{adj} i r_2^{adj} , també es pot obtenir una estimació per a la variància de la seva diferència, que és la següent,

$$\hat{\text{Var}}(r_1^{adj} - r_2^{adj}) = \sum_{i=1}^I w_i \frac{n_1 r_{1i} + n_2 r_{2i}}{n_1 n_2} \quad (3.18)$$

on, r_{1i} i r_{2i} representen les taxes específiques per grups d'edat per cadascun dels dos grups comparats en un any qualsevol. Alternativament, es pot donar una segona estimació per a aquesta variància, aplicant l'expressió 3.13 en cadascun dels grups d'edat. D'aquesta forma s'obté,

$$\hat{\text{Var}}(r_1^{adj} - r_2^{adj}) = \sum_{i=1}^I w_i^2 \frac{c_{1i} + c_{2i}}{n_1 n_2} \quad (3.19)$$

on c_{1i} i c_{2i} representen el nombre de casos observats pels dos grups en cadascun dels grups d'edat, en un any qualsevol. En qualsevol cas, es pot emprar l'estadístic següent,

$$Z = \frac{r_1^{adj} - r_2^{adj}}{\sqrt{\hat{\text{Var}}(r_1^{adj} - r_2^{adj})}} \quad (3.20)$$

per a resoldre la prova d'hipòtesi $H_o : \lambda_1 = \lambda_2$ versus $H_a : \lambda_1 \neq \lambda_2$ de forma anàloga a la descrita anteriorment per a les taxes crues. Les dues expressions descrites en 3.18 i 3.19 donen resultats, a la pràctica, molt similars. Tot i això és preferible la segona expressió, basada en les dades específiques per

grups d'edat, en el cas que aquestes siguin disponibles [Esteve et al., 1994].

Prova de Mantel-Haenzel

La prova de Mantel-Haenzel, es va desenvolupar en el marc de les proves d'independència en taules de contingència (2×2) específicament per a estudis epidemiològics de casos i controls [Mantel and Haenzel, 1959], com una millora a la clàssica prova d'independència de la χ^2 de Pearson. La prova de Mantel-Haenzel també pot ser aplicada, amb una petita modificació, per a la comparació de taxes d'incidència o mortalitat per càncer (o una altra malaltia).

La prova de Mantel-Haenzel per a la comparació de taxes avalua la hipòtesi nul·la que les taxes són iguals per a tots els grups d'edat respecte la hipòtesi alternativa de que aquestes són proporcionals. En primer lloc es defineix el nombre de casos esperats en cadascun dels dos grups. Per això, considerarem c_{1i} i c_{2i} els casos observats en els dos grups per a cadascun dels grups d'edat i n_{1i} i n_{2i} el nombre de persones a risc. Aleshores, es defineix el nombre de casos esperats per al grup 1, sota la hipòtesi nul·la, tal com segueix

$$c_{1i}^{exp} = \frac{c_{1i} + c_{2i}}{n_{1i} + n_{2i}} n_{1i} \quad (3.21)$$

i, simètricament, es defineix c_{2i}^{exp} per al segon grup. Es tracta doncs de l'estimació del nombre de casos que esperaríem en el cas que ambdues poblacions compartissin el mateix risc, estimat com $\frac{c_{1i} + c_{2i}}{n_{1i} + n_{2i}}$.

A partir d'aquí, la prova es basa en la diferència global entre els casos observats i esperats, pel que es calcula l'estadístic següent,

$$T = \sum_{i=1}^I (c_{i1} - c_{1i}^{exp}) \quad (3.22)$$

Aquest estadístic es pot formular de forma totalment simètrica per al segon grup i, és fàcil veure, que el resultat és el mateix valor per a l'estadístic però amb signe contrari. D'altra banda, l'estadístic T està desenvolupat de tal forma que detecti diferències sistemàtiques del mateix signe entre els

casos observats i esperats. Per tal d'avaluar la variabilitat d'aquest estadístic, la prova de Mantel-Haenzel assumeix que el nombre de casos observats segueixen una distribució binomial, tal com segueix,

$$c_{1i} \sim \text{binomial} \left(c_{1i} + c_{2i}, \frac{n_{1i}}{n_{1i} + n_{2i}} \right) \quad (3.23)$$

definició que pot enunciar-se simètricament per a c_{2i} . D'aquesta forma es pot donar una expressió per a la variància de c_{1i} tal com segueix,

$$\text{Var}(c_{1i}) = (c_{1i} + c_{2i}) \frac{n_{1i}n_{2i}}{(n_{1i} + n_{2i})^2} \quad (3.24)$$

i, consegüentment, es pot determinar la variància per a l'estadístic T ,

$$\text{Var}(T) = \sum_{i=1}^i \text{Var}(c_{1i} - c_{1i}^{exp}) = \sum_{i=1}^i (c_{1i} + c_{2i}) \frac{n_{1i}n_{2i}}{(n_{1i} + n_{2i})^2} \quad (3.25)$$

La prova de Mantel-Haenzel es basa finalment en el càlcul de l'estadístic $Z = T/\sqrt{\text{Var}(T)}$ que es distribueix aproximadament segons una normal estandarditzada i, per tant podem resoldre la prova d'hipòtesis sobre la igualtat de taxes en ambdues poblacions. Aquest estadístic té una potència estadística limitada en el cas que la hipòtesi alternativa no sigui la especificada prèviament. Per exemple, podria donar-se el cas que les taxes d'incidència fossin majors per les edats joves i menors per les edats més avançades, el que provocaria que el valor de T fos proper a zero, encara que la hipòtesi nul·la no fos certa [Esteve et al., 1994].

3.3 Anàlisi estadística de tendències de taxes

En el context de l'epidemiologia descriptiva es donen diverses circumstàncies que motiven l'estudi de tendències de taxes d'incidència o mortalitat a través del temps. En primer lloc, l'evolució històrica del risc (incidència o mortalitat) pot generar hipòtesis etiològiques. El fet d'estudiar l'evolució d'una

malaltia en una mateixa població al llarg del temps permet definir possibles associacions entre els canvis de la introducció, desaparició o modificació de l'exposició a factors de risc ambientals i una malaltia. En el cas del càncer aquestes anàlisis són fonamentals i han permès avaluar com, per exemple, la mortalitat per càncer de pulmó ha augmentat a mesura que s'introduïa el consum de tabac en una població i com aquesta mortalitat disminuïa quan la proporció de fumadors es reduïa. Malauradament, la relació entre el càncer de pulmó i el tabac és una excepció, ja que el consum de tabac és el factor causal per la majoria de casos. En canvi, per alguns càncers, l'etiologia és multicausal i és més difícil avaluar la influència dels canvis en les exposicions als factors de risc i a la incidència d'aquests tumors [Bray and Moller, 2006]. La monitorització d'una malaltia al llarg del temps, mitjançant l'anàlisi de les tendències de les taxes d'incidència o mortalitat temporals, és també un instrument per als epidemiòlegs que pot tenir com a resultat la generació d'hipòtesis de noves exposicions a factors de risc. En termes de prevenció secundària, aquestes anàlisis han servit per a avaluar els programes de cribatge de càncer i han servit per a demostrar la disminució de la mortalitat, com en el cas del CM. Finalment, l'anàlisi estadística de tendències observades fins a un cert moment, pot tenir com a objectiu el de predir l'evolució de la malaltia en el futur. Les anàlisis de projeccions, són doncs una passa posterior a l'anàlisi de tendències i descrivim la seva metodologia en un apartat posterior.

Un dels models estadístics més àmpliament emprat per analitzar les tendències de taxes d'incidència o mortalitat és el model anomenat *edat-període* [Hakulinen and Dyba, 1994], que té en compte l'evolució d'aquestes taxes al llarg del temps, segons l'edat de la població estudiada i el període on s'avalua la tendència. Tanmateix, sovint es vol analitzar l'efecte generacional, és a dir la data de naixement de diferents cohorts (poblacions) i com les seves taxes difereixen segons l'època en què una persona va nàixer. En aquests casos es poden emprar el que s'han anomenat models *edat-cohort*. Finalment, també s'han desenvolupat els anomenats models *edat-període-cohort* [Holford, 1983]

que tenen en compte tots tres factors, que han de ser correctament introduïts en els models per tal de tenir en compte possibles confusions d'efectes, degut a la relació lineal existent entre ells. En aquesta memòria hem centrat l'interès en els models edat-període, ja que amb un nombre de paràmetres petit, solen oferir una solució parsimoniosa per a l'anàlisi de tendències de les taxes observades, doncs es tracta de models log-lineals simples. A continuació descrivim aquest model i la seva utilització en aquest context.

Model edat-període

Reprenem c_{it} i n_{it} com el nombre de casos i població a risc, respectivament, per al grup d'edat i durant el període t . El model edat-període habitualment emprat [Dyba and Hakulinen, 2000] és el que s'especifica a continuació,

$$\log \left(E \left(\frac{c_{it}}{n_{it}} \right) \right) = \alpha_i + \beta_i t \quad (3.26)$$

Aquest model assumeix que el risc de desenvolupar o morir d'una malaltia, $\lambda_{it} = E \left(\frac{c_{it}}{n_{it}} \right)$, es relaciona log-linealment amb el temps. Pel que fa al predictor lineal s'assumeix un efecte específic per a cadascun dels grups d'edat, α_i , amb una pendent de creixement (o decreixement) específica també per cadascun dels grups d'edat, β_i . Aquest model pot ser ajustat en el marc dels models lineals generalitzats [McCullagh and Nelder, 1989]. A l'igual que hem descrit prèviament, s'assumeix que els casos c_{it} es distribueixen segons una llei de Poisson i que la població a risc està fixada. D'aquesta forma, es pot definir la funció de versemblança i emprar la teoria d'estimació màxim versemblant per a ajustar els models i obtenir les estimacions dels paràmetres i dels seus errors estàndards. Emprant la notació habitual dels models lineals generalitzats, la funció *link* és la logarítmica i la família exponencial corresponent és la de Poisson.

Per a avaluar la tendència de les taxes d'incidència o de mortalitat d'una malaltia concreta es sol emprar una mesura del canvi percentual anual, habitualment denotada per l'acrònim APC (*Annual Percent Change*). Aquest

indicador mesura la intensitat anual en que augmenten o bé disminueixen les taxes i es defineix tal com segueix,

$$APC_i = \frac{\lambda_{it} - \lambda_{i(t-1)}}{\lambda_{i(t-1)}} \quad (3.27)$$

assumint doncs que l'APC és constant al llarg del temps i diferent per a cada grup d'edat. Del model descrit en 3.26 es pot deduir que $\lambda_{it} = e^{\alpha_i + \beta_i t}$, el que ens permet reescriure l'APC, a partir dels paràmetres del model, tal com segueix,

$$APC_i = \frac{e^{\alpha_i + \beta_i t} - e^{\alpha_i + \beta_i (t-1)}}{e^{\alpha_i + \beta_i (t-1)}} = \frac{e^{\alpha_i + \beta_i t}}{e^{\alpha_i + \beta_i (t-1)}} - 1 = e^{\beta_i} - 1 \quad (3.28)$$

D'aquesta forma, emprant les propietats dels estimadors màxim versemblants, es pot obtenir un estimador, també màxim versemblant, per al canvi percentual anual en la tendència, $\widehat{APC}_i = e^{\hat{\beta}_i} - 1$, on $\hat{\beta}_i$ és l'estimació màxim versemblant per al paràmetre del model 3.26. A més, si denotem la seva variància per $\sigma_{\hat{\beta}_i}^2 = \text{Var}(\hat{\beta}_i)$, aquesta es pot estimar emprant també els procediments habitual d'estimació màxim versemblant, de tal forma que es pot derivar la distribució asimptòtica per a $\hat{\beta}_i$ tal com segueix,

$$\hat{\beta}_i \sim N(\beta_i, \hat{\sigma}_{\beta_i}^2) \quad (3.29)$$

Així doncs es pot construir un interval de confiança per a l'APC a partir dels paràmetres ajustats pel model tal com segueix,

$$APC_i \in \left[e^{\hat{\beta}_i - z_{\alpha/2} \hat{\sigma}_{\beta_i}} - 1, e^{\hat{\beta}_i + z_{\alpha/2} \hat{\sigma}_{\beta_i}} - 1 \right] \quad (3.30)$$

Sovint interessa resoldre la prova d'hipòtesi $H_o : APC_i = 0$ versus $H_a : APC_i \neq 0$, el que és equivalent a resoldre $H_o : \beta_i = 0$ versus $H_a : \beta_i \neq 0$, per tal d'avaluar si la tendència, per a cadascun dels grups d'edat, és significativament creixent o decreixent. Per això, es pot emprar la distribució asimptòtica expressada en 3.29, el que permet calcular un p-valor associat a

la prova d'hipòtesi tal com segueix,

$$p = 2\Pr\left(N(0, 1) \geq \left|\frac{\hat{\beta}_i}{\hat{\sigma}_{\beta_i}}\right|\right) \quad (3.31)$$

A més, es pot construir un test per avaluar la significació global de l'efecte conjunt del temps en els grups d'edat, amb hipòtesi nul.la $H_o : \beta_1 = \beta_2 = \dots \beta_I = 0$, emprant el test de Wald o bé un test basat en la raó de versemblances. En aquest segon cas, cal ajustar primer el model nul, que s'especifica tal com segueix,

$$\log\left(\mathbb{E}\left(\frac{c_{it}}{n_{it}}\right)\right) = \alpha_i \quad (3.32)$$

i aleshores, procedir a comparar les deviances del model mitjançant un test clàssic basat en la raó de versemblances. Concretament, si denotem per D_M i D_{M_o} les deviances associades als models 3.26 i 3.32 respectivament, aleshores la seva diferència es distribueix aproximadament segons una distribució de khi-quadrat amb I graus de llibertat, que es pot expressar tal com segueix,

$$D_{M_o} - D_M \sim \chi_I^2 \quad (3.33)$$

Pel que fa a la deviança, aquesta es defineix per a un model donat com -2 vegades el logaritme de la raó de versemblances entre el model estimat i el model complet, és a dir, aquell per a que els valors predits equivalen als observats. En el cas anterior, on es pretén comparar un model respecte un sub-model (amb menys paràmetres) es construeix la prova de raó de versemblances entre els models de tal forma que els graus de llibertat de la khi-quadrat equivalen a la diferència en termes del nombre de paràmetres dels models. D'aquesta forma, l'estadístic resultant es distribueix, sota la hipòtesi que els paràmetres específics del model gran són nuls, segons una khi-quadrat amb I graus de llibertat, ja que el model 3.26 té $2 \cdot I$ paràmetres (I paràmetres per als efectes individuals dels grups d'edat més I paràmetres per a les seves pendents) mentre que el sub-model 3.32 només té I paràmetres (per

als efectes individuals dels grups).

Mitjançant el model proposat (3.26) es poden doncs estimar els APC per cadascun dels grups d'edat i procedir a la inferència, emprant els mecanismes habituals per als models lineals generalitzats. A la pràctica, pot interessar aportar una estimació de l'APC conjunta per a tots els grups d'edat. Per fer-ho, cal definir quins pesos tenen cadascun dels grups d'edat, que denotarem per w_i . Es poden emprar els pesos corresponents a la població mundial, europea, o uns altres qualssevol, tal com hem exposat prèviament en la taula 3.1 amb el que es calcularia una APC *ajustada per edat* o bé es poden emprar els pesos observats de cadascun dels grups, és a dir, $w_i = \frac{\sum_t n_{it}}{\sum_{i,t} n_{it}}$. En qualsevol cas, l'estimador conjunt de l'APC es pot expressar tal com segueix,

$$\text{APC}^{adj} = \sum_{i=1}^I w_i \text{APC}_i \quad (3.34)$$

Anàlogament, es pot definir $\beta^{adj} = \sum_{i=1}^I w_i \beta_i$ i, emprant les propietats del estimadors màxim versemblants, es pot deduir una distribució asimptòtica pel seu estimador,

$$\hat{\beta}^{adj} \sim N(\beta^{adj}, \hat{\sigma}_{\beta^{adj}}^2) \quad (3.35)$$

on

$$\hat{\sigma}_{\beta^{adj}}^2 = \sum_{i=1}^I w_i^2 \hat{\sigma}_{\beta_i}^2 + 2 \sum_{i < j} w_i w_j \hat{\text{Cov}}(\beta_i, \beta_j) \quad (3.36)$$

A la pràctica pot també resultar convenient ajustar el model que assumeix que tots els grups d'edat comparteixen un únic paràmetre, el que es pot interpretar que la tendència és la mateixa per a tots els grups d'edat, β . Aquest model, anomenat sovint *age-drift*, s'especifica de la forma següent,

$$\log \left(\text{E} \left(\frac{c_{it}}{n_{it}} \right) \right) = \alpha_i + \beta t \quad (3.37)$$

D'altra banda, també s'ha proposat un model lineal en escala aritmètica [Hakulinen and Dyba, 1994], tal com segueix

$$E\left(\frac{c_{it}}{n_{it}}\right) = \alpha_i + \beta_i t \quad (3.38)$$

Des del punt de vista empíric i de parsimònia, és sensat emprar el model lineal (3.38) per aquells casos en que la tendència d'incidència o mortalitat sigui creixent, doncs altrament podria donar lloc a prediccions fins i tot negatives, mentre que els models log-lineals (3.26 i 3.32) poden ser emprats en ambdós casos.

Finalment, també s'ha proposat un model no lineal [Dyba et al., 1997] que, conceptualment, considera que tots els grups d'edat comparteixen una mateixa pendent de creixement, β , que es veu amplificada o reduïda segons la magnitud de les taxes de cadascun dels grups d'edat en el moment inicial, α_i , i s'especifica tal com segueix:

$$E\left(\frac{c_{it}}{n_{it}}\right) = \alpha_i(1 + \beta t) \quad (3.39)$$

Un treball desenvolupat pels creadors del model [Dyba and Hakulinen, 2000], on es van comparar els models anteriors a partir de dades reals i dades simulades va concloure que el model log-lineal sencer (3.26) és el que presentava millors resultats tant des del punt de vista de bondat d'ajust com des del punt de vista de capacitat explicativa del model. El model no lineal (3.39), en el cas que ajusti, sol donar uns errors estàndards més elevats, mentre que l'estimació puntual és pràcticament la mateixa que pel model 3.26.

En qualsevol cas, davant del repte de modelar la tendència temporal d'unes taxes d'incidència o mortalitat donades, es pot plantejar l'ajust dels quatre models exposats anteriorment. Per tal de comparar la seva bondat d'ajust, es pot emprar l'estadístic de la khi-quadrat de Pearson, que es pot escriure tal com segueix,

$$X^2 = \sum_{i=1}^I \sum_{t=1}^T \frac{(c_{it} - \hat{c}_{it})^2}{\text{Var}(\hat{c}_{it})} \quad (3.40)$$

on \hat{c}_{it} és el nombre de casos estimats a partir del model. L'estadístic X^2 es distribueix, sota la hipòtesi nul·la que el model s'ajusta a les dades, segons una khi-quadrat amb els graus de llibertat corresponents als graus de llibertat residuals del model, és a dir, el nombre d'observacions menys el nombre de paràmetres.

Pel que fa a programari, les anàlisis estadístiques anteriors poden ser resoltes en qualsevol paquet estadístic que disposi d'ajust de models lineals generalitzats i, en particular, per a models de Poisson. En que cas del paquet estadístic **R**, es poden emprar còmodament les funcions disponibles a la llibreria **glm**.

3.4 Projeccions de taxes

Les prediccions de les taxes d'incidència o mortalitat d'una malaltia en el futur, a partir d'un conjunt de dades històriques observades, ajustades a un determinat model estadístic, i a una població futura determinada prèviament, es coneixen normalment com a projeccions. Des del punt de vista estadístic es tracta no únicament d'una predicció efectuada a partir del model, sinó que aquesta té lloc fora del rang de les dades observades, és a dir, en un temps superior a l'observat. Per aquest motiu, des del punt de vista estadístic, els mètodes per a projectar taxes presenten unes particularitats tant teòriques com d'aplicació. Per una altra banda, la predicció (o projecció) de les taxes d'incidència o de mortalitat d'una malaltia, és important en tant que constitueix la base per a la planificació sanitària [Einhorn, 1989, Schaubel et al., 1998, Dyba and Hakulinen, 2000]. En el cas del càncer, la càrrega de la malaltia és molt important en termes econòmics i de salut pública, fet que motiva la necessitat de disposar d'unes estimacions futures acurades en els processos de decisió. Les projeccions s'empren també quan es pretén avaluar l'impacte d'una intervenció o bé d'un programa de cribatge d'una malaltia en una població. Per això, es pot comparar la projecció del que s'hagués observat en l'absència del programa, emprant les dades històriques, respecte

les taxes d'incidència o de mortalitat finalment observades amb el programa d'intervenció. Per exemple, en el criatge del CM, aquests tipus d'estudis s'han realitzat, observant-se com la incidència d'aquest càncer era molt superior en les dones cribades d'una població respecte la incidència projectada en el mateix grup de dones sense criatge [Hakulinen and Hakama, 1991]. Des del punt de vista científic i metodològic, resulta interessant avaluar quan aquestes projeccions esdevenen veritables o no, de tal forma que es pugui discutir l'ús dels models i la seva validesa [Prior et al., 1996]. Des del punt de vista estadístic, cal doncs oferir un interval de predicció que proveeixi un marge versemblant per a una observació futura, sota l'assumpció que el model escollit és el correcte [Hakulinen and Dyba, 1994].

La fiabilitat d'una projecció d'una malaltia determinada depèn directament de la tria del model i de les variables que hi estan incloses. En la predicció del càncer, de vegades, no existeix una clara evidència sobre la prevalença poblacional dels factors etiològics per a cada localització tumoral. A més, en els casos que els factors etiològics estan ben establerts, sovint no es poden introduir en els models degut a que les dades de la prevalença d'aquests factors en la població i, més encara, la seva evolució en el temps no estan suficientment recollits. Un exemple d'això és la dieta que es coneix que està molt associada al càncer colorectal. D'aquesta forma, els models emprats per a la projecció són els que s'han enunciat en l'apartat previ, basats en l'edat i el període d'observació. També convé remarcar que, a més, existeixen models complexos, com per exemple models edat-període-cohort polinomials [Coleman et al., 1993], que tenen una càrrega en el nombre de paràmetres excessiva i que habitualment produeixen intervals de confiança per a les prediccions massa amples, fet que en limita el seu ús a la pràctica [Dyba and Hakulinen, 2000]. Així doncs, aplicant el principi de parsimònia, els models simples presentats en l'apartat anterior són els que habitualment s'empren en la projecció de taxes, per tal d'obtindre una predicció sobre la càrrega futura de malalties, tal com el càncer. És cert que un bon ajust no garanteix una bona projecció, però també és cert que un molt bon ajust no

sempre durà a una projecció dolenta [Dyba and Hakulinen, 2000], i la pràctica d'aquests models indica que els models edat-període són prou satisfactoris.

Aportar prediccions fiables per a l'ocurrència d'una malaltia és una qüestió doncs complexa. Per a diverses localitzacions tumorals, es desconeix llargament l'etiologia i a més, aquests factors etiològics poden variar enormement al llarg del temps. En qualsevol cas, aquestes projeccions es realitzen a partir de models estadístics que estan sotmesos a dues fonts d'error. En primer lloc, la incertesa sobre els paràmetres del model, ja que les dades històriques a partir de les que s'han estimat estan sotmeses a variació aleatòria. En aquest punt, com s'ha exposat prèviament, s'assumeix que el nombre de casos segueixen distribucions de Poisson independents. Els errors estàndards estimats per als paràmetres reflecteixen aquest error, sota l'assumpció que el model és correcte. Aquesta primera font d'error està associada a la incertesa de les futures taxes, però fins i tot en el cas que es coneguessin les veritables futures taxes, els nombres observats no serien exactament els mateixos, degut a que les taxes futures també són variables aleatòries, el que constitueix la segona font d'error. D'aquesta forma, cal doncs tenir en compte que el nombre futur de casos incidents o morts es distribueix també segons una llei de Poisson. El càlcul d'interval de confiança per a projeccions a partir de models basats en distribucions normals està descrit en la majoria de llibres d'estadística [Weisberg, 1985]. En el cas de projeccions de taxes d'incidència o mortalitat basades en la llei de Poisson, l'estimació de la variabilitat de la projecció està resolta [Hakulinen and Dyba, 1994] tot i que és una mica més complex i, actualment no es troba implementat en els programes estadístics més habituals. En qualsevol cas, l'estimació puntual no es veu afectada, doncs no hi ha una estimació de la precisió en la estimació. A continuació descriurem l'estimació d'aquesta variabilitat en els models edat-període [Hakulinen and Dyba, 1994].

Suposem el nombre de casos observats d'una malaltia en un determinat moment per a un grup d'edat, c_{it} , com una variable aleatòria regida per una llei de Poisson amb esperança θ_{it} . Donat que θ_{it} és desconegut, aquest pot ser

estimat per $\hat{\theta}_{it}$. Emprant un argument condicional podem expressar [Chiang, 1968],

$$\text{Var}(c_{it}) = \text{Var}_i \left[\text{E} \left(c_{it} | \hat{\theta}_{it} \right) \right] + \text{E}_i \left[\text{Var} \left(c_{it} | \hat{\theta}_{it} \right) \right] \quad (3.41)$$

on Var_i i E_i denoten la variància i esperança respecte a la distribució de $\hat{\theta}_{it}$, seguint la notació proposada per [Hakulinen and Dyba, 1994]. L'assumpció de Poisson també permet obtenir

$$\text{Var}(c_{it}) = \text{Var} \left(\hat{\theta}_{it} \right) + \text{E} \left(\hat{\theta}_{it} \right) \quad (3.42)$$

Per al nombre total de casos, $c_t = \sum_{i=1}^I c_{it}$ tenim

$$\text{Var}(c_t) = \sum_{i=1}^I \text{Var}(c_{it}) + \sum_{i=1}^I \sum_{j \neq i} \text{Cov}(c_{it}, c_{jt}) \quad (3.43)$$

on podem expressar la covariància de la forma següent,

$$\text{Cov}(c_{it}, c_{jt}) = \text{E}_{ij} \left[\text{Cov}(c_{it}, c_{jt} | \hat{\theta}_{it}, \hat{\theta}_{jt}) \right] + \text{Cov}_{ij} \left[\text{E}(c_{it} | \hat{\theta}_{it}, \hat{\theta}_{jt}), \text{E}(c_{jt} | \hat{\theta}_{it}, \hat{\theta}_{jt}) \right] \quad (3.44)$$

on E_{ij} i Cov_{ij} denoten l'esperança i covariància, respectivament, pel que fa a la distribució conjunta de $\hat{\theta}_{it}$ i $\hat{\theta}_{jt}$. El primer terme és zero degut a la assumpció d'independència de les observacions distribuïdes segons una Poisson, donant per exemple,

$$\text{Cov}(c_{it}, c_{jt} | \hat{\theta}_{it} = \theta_{it}, \hat{\theta}_{jt} = \theta_{jt}) = 0 \quad (3.45)$$

D'aquesta forma,

$$\text{Cov}(c_{it}, c_{jt}) = \text{Cov}(\hat{\theta}_{it}, \hat{\theta}_{jt}) \quad (3.46)$$

Finalment,

$$\text{Var}(c_t) = \sum_{i=1}^I \text{Var}(c_{it}) + \sum_{i=1}^I \sum_{j \neq i} \text{Cov}(\hat{\theta}_{it}, \hat{\theta}_{jt}) \quad (3.47)$$

Emprant aleshores l'equació 3.42, obtenim

$$\text{Var}(c_t) = \sum_{i=1}^I \text{Var}(\hat{\theta}_{it}) + \sum \sum_{i \neq j} \text{Cov}(\hat{\theta}_{it}, \hat{\theta}_{jt}) + \sum_{i=1}^I \text{E}(\hat{\theta}_{it}) \quad (3.48)$$

Si denotem $\hat{\theta}_t = \sum_{i=1}^I \hat{\theta}_{it}$,

$$\text{Var}(c_t) = \text{Var}(\hat{\theta}_t) + \text{E}(\hat{\theta}_t) \quad (3.49)$$

Així doncs, el resultat és l'equivalent a considerar c_t com una variable Poisson amb esperança θ_t . Resulta convenient remarcar que per als models edat-període que tenen en compte l'efecte del grups d'edat (3.26 i 3.38), $\text{Cov}(\hat{\theta}_{it}, \hat{\theta}_{jt}) = 0$ per a tots $i \neq j$ donat que els diferents grups d'edat no comparteixen paràmetres comuns.

Les dues components de $\text{Var}(c_t)$ en l'equació 3.49 corresponen a les dues fonts de variació. $\text{Var}(\hat{\theta}_t)$ reflecteix la incertesa present sobre el paràmetre futur Poisson i pot ser utilitzat per a intervals de confiança per al valor esperat sobre el futur nombre de casos, $\hat{\theta}_t$. $\text{E}(\hat{\theta}_t)$ reflecteix la mida del error aleatori per a una observació distribuïda segons una variable Poisson amb una esperança equivalent a $\text{E}(\hat{\theta}_t)$.

Per al model lineal específic pera grups d'edat (3.38) es poden aproximar les quantitats $\text{Var}(\hat{\theta}_{it})$ i $\text{Cov}(\hat{\theta}_{it}, \hat{\theta}_{jt})$, doncs tenint en compte que $\hat{\theta}_{it} = n_{it}(\hat{\alpha}_i + \hat{\beta}_i t)$, obtenim

$$\text{Var}(\hat{\theta}_{it}) = n_{it}^2 \left[\text{Var}(\hat{\alpha}_i) + t^2 \text{Var}(\hat{\beta}_i) + 2t \text{Cov}(\hat{\alpha}_i, \hat{\beta}_i) \right] \quad (3.50)$$

Per als models log-lineals específics per a grups d'edat o amb una tendència comuna (3.26 i 3.37 específicament) es pot emprar el mètode delta o la fórmula de Gauss basada en l'expansió de funcions no lineals basada en sèries de Taylor. Per al model log-lineal amb tendència comuna per als grups d'edat (3.37) s'obté,

$$\hat{\theta}_{it} = n_{it} e^{\hat{\alpha}_i + \hat{\beta}_i t} \quad (3.51)$$

$$\begin{aligned}\widehat{\text{Var}}(\hat{\theta}_{it}) &\approx (\partial\hat{\theta}_{it}/\partial\hat{\alpha}_i)^2\widehat{\text{Var}}(\hat{\alpha}_i)+(\partial\hat{\theta}_{it}/\partial\hat{\beta})^2\widehat{\text{Var}}(\hat{\beta})+2(\partial\hat{\theta}_{it}/\partial\hat{\alpha}_i)(\partial\hat{\theta}_{it}/\partial\hat{\beta})\widehat{\text{Cov}}(\hat{\alpha}_i, \hat{\beta}) \\ &= \hat{\theta}_{it}^2 \left[\widehat{\text{Var}}(\hat{\alpha}_i) + t^2\widehat{\text{Var}}(\hat{\beta}) + 2t\widehat{\text{Cov}}(\hat{\alpha}_i, \hat{\beta}) \right]\end{aligned}\quad (3.52)$$

i

$$\begin{aligned}\widehat{\text{Cov}}(\hat{\theta}_{it}, \hat{\theta}_{jt}) &\approx (\partial\hat{\theta}_{it}/\partial\hat{\alpha}_i)(\partial\hat{\theta}_{jt}/\partial\hat{\alpha}_j)\widehat{\text{Cov}}(\hat{\alpha}_i, \hat{\alpha}_j)+(\partial\hat{\theta}_{it}/\partial\hat{\alpha}_i)(\partial\hat{\theta}_{jt}/\partial\hat{\beta})\widehat{\text{Cov}}(\hat{\alpha}_i, \hat{\beta}) \\ &\quad +(\partial\hat{\theta}_{jt}/\partial\hat{\alpha}_j)(\partial\hat{\theta}_{jt}/\partial\hat{\beta})\widehat{\text{Cov}}(\hat{\alpha}_j, \hat{\beta}) + (\partial\hat{\theta}_{it}/\partial\hat{\beta})(\partial\hat{\theta}_{jt}/\partial\hat{\beta})\widehat{\text{Var}}(\hat{\beta}) \\ &= \hat{\theta}_{it}\hat{\theta}_{jt} \left[\widehat{\text{Cov}}(\hat{\alpha}_i, \hat{\alpha}_j) + t\widehat{\text{Cov}}(\hat{\alpha}_i, \hat{\beta}) + t\widehat{\text{Cov}}(\hat{\alpha}_j, \hat{\beta}) + t^2\widehat{\text{Var}}(\hat{\beta}) \right]\end{aligned}\quad (3.53)$$

D'aquesta forma, els intervals de confiança asimptòtics poden doncs calcular-se a partir dels errors estàndards emprant l'assumpció de normalitat (estimació $\pm 1.96 \times$ error estàndard de l'estimació). Aquestes fórmules varen ser implementades pels autors [Hakulinen and Dyba, 1994] en el paquet estadístic GLIM. Per al model log-lineal específic per a grups d'edat (3.26) es pot emprar el mètode Delta per a obtenir,

$$\widehat{\text{Var}}(\hat{\theta}_{it}) \approx (\partial\hat{\theta}_{it}/\partial\hat{\eta}_{it})^2\widehat{\text{Var}}(\log\hat{\theta}_{it}) = \hat{\theta}_{it}^2\widehat{\text{Var}}(\log\hat{\theta}_{it}) \quad (3.54)$$

on

$$\hat{\eta}_{it} = \hat{\alpha}_i + \hat{\beta}_i t \quad (3.55)$$

Així doncs, podem concloure que l'estimació del nombre futur de casos (incidents o morts) es pot separar en dues passes:

- Obtenir un model que ajusti raonablement bé a les dades observades dins del seu rang de temps. Per això es poden emprar els models lineals i log-lineals, que assumeixen que el nombre de casos es distribueixen segons una Poisson, tant amb tendència específica per a grups d'edats com conjunta i seleccionar aquell amb un bon ajust, seguin els principis enunciats en aquest capítol.

- Obtenir l'estimació del nombre de casos futurs, emprant el model ajustat. Per això és necessari establir el moment en el futur en el que es vol obtenir l'estimació i disposar, a més, d'una estimació de la població futura segons els grups d'edat. Cal doncs remarcar que es tracta d'emprar un model ajustat per a obtenir una predicció més enllà del rang d'observació de les dades. Es poden construir intervals de confiança asimptòtics i/o resoldre proves d'hipòtesi a partir dels errors estàndards estimats que s'han presentat.

3.5 Avaluació de factors de riscos i demogràfics

L'anàlisi estadística de les diferències observades en la incidència o mortalitat d'una malaltia donada pot ser de gran interès, tant des del punt de vista científic com de planificació sanitària [Hakulinen and Hakama, 1991]. Sovint, es centra l'interès en la comparació del nombre de casos incidents o morts d'una malaltia en dues situacions donades, amb l'objectiu de quantificar les diferències observades, per a ser interpretades epidemiològicament, però també per a donar suport a la presa de decisions en salut pública. En aquest sentit, les tendències temporals s'utilitzen habitualment per a analitzar l'evolució històrica del risc de desenvolupar o morir per una malaltia i per a avaluar l'ocurrència d'aquesta malaltia en un cert període de temps [Esteve et al., 1994]. En aquests casos, s'han desenvolupat mètodes que avaluen quina part d'aquestes diferències són atribuïbles a canvis en el risc de desenvolupar la malaltia (incidència) o de morir per la malaltia (mortalitat) respecte aquelles diferències que podrien ser degudes a canvis demogràfics en les poblacions comparades. En aquest apartat descrivim un mètode [Bashir and Esteve, 2000] que té en compte aquests factors, i que nosaltres fem en aquesta tesi per tal d'atribuir el seu paper en els canvis entre la malaltia observada i la projectada.

Suposarem doncs que es desitja comparar la incidència o mortalitat d'una malaltia en dos moments del temps diferents. Sovint, aquestes comparacions

es fan a partir de la diferència absoluta en el nombre de casos observats, la taxa crua (usualment, per 100.000 persones), i alguns cops es calcula el percentage de canvi [Breslow and Day, 1987]. Tanmateix, tot i que les taxes brutes d'incidència o mortalitat podrien ser emprades per a comparar malalties en una mateixa població, podrien no ser útils per a comparar taxes de la mateixa malaltia en poblacions diferents o al llarg del temps [Sasieni and Adams, 1999]. Per a esmenar aquesta qüestió es calculen les taxes ajustades, que empren una mateixa població de referència, tal i com hem esmentat en aquest mateix capítol. Es calcula el percentatge de canvi per a mesurar el canvi percentual degut a canvis en el risc de desenvolupar o morir per una malaltia. Aquests canvis poden ser atribuïbles a l'efecte de factors demogràfics de forma important, sobretot si les piràmides de població respectives mostren diferències substancials. Així doncs, una variació en la mida de la població al llarg del temps podria explicar canvis en el nombre de casos, degut al consegüent increment (o decrement) del nombre de persones a risc. Canvis en l'estructura de la piràmide, com l'envelliment, que està clarament associat a canvis moleculars, cel·lulars i fisiològics que influeixen en la carcinogènesi i el subseqüent creixement tumoral [Balducci and Ershler, 2005], podria augmentar el nombre de casos entre els grups d'edat més avançada [Bray and Moller, 2006] i, per tant, influir en la diferència observada entre les taxes brutes i ajustades.

Bashir i Estève van desenvolupar un mètode per quantificar la variació en la incidència o mortalitat d'una malaltia entre dos poblacions, atribuint el percentatge de canvi explicat per factors demogràfics (mida de la població i estructura) i per factors epidemiològics associats al risc de desenvolupar o morir per una determinada malaltia [Bashir and Esteve, 2000]. Aquest mètode està basat en la idea de calcular, en primer lloc, la incidència o mortalitat que s'esperaria observar si la mida i estructura de la població fossin iguals per als dos grups i, en segon lloc, atribuir la diferència respecte el canvi net a factors estrictament demogràfics. A més, el canvi atribuït a factors demogràfics es pot aleshores separar entre el causat per canvis en la

mida de la població i el causat per canvis en l'estructura de la població, com l'envelliment.

El mètode proposat per Bashir i Estève, assumeix que s'han observat el nombre de casos (incidents o morts) per als dos grups que anomenarem grup basal (o de referència) i grup de comparació, i que denotarem respectivament per $c_{i|0}$ i $c_{i|1}$ on i fa referència al grup d'edat (habitualment 18 grups d'edats quinquennals). A més, denotarem per $n_{i|0}$ i $n_{i|1}$ el nombre de persones a risc durant el període observat en cadascun dels grups d'edat. D'aquesta forma, les taxes crues (per 100.000 habitants) $r_{|0}$ i $r_{|1}$ poden ser calculades tal com segueix (seguint la fórmula 3.2):

$$r_{|0} = \frac{\sum_{i=1}^I c_{i|0}}{\sum_{i=1}^I n_{i|0}} \cdot 100000 \quad (3.56)$$

i, anàlogament per al grup de comparació, $r_{|1}$.

El mètode de Bashir i Estève es basa en el càlcul d'una taxa intermèdia, que denotarem per $r_{|2}$ i que consisteix en la taxa ajustada per al grup basal, emprant com a població de referència la població del grup de comparació, és a dir,

$$r_{|2} = \sum_{i=1}^I w_i r_{i|0} = \sum_{i=1}^I \frac{n_{i|1}}{\sum_i n_{i|1}} r_{i|0} = \sum_{i=1}^I \frac{n_{i|1}}{\sum_i n_{i|1}} \frac{c_{i|0}}{n_{i|0}} \cdot 100000 \quad (3.57)$$

D'aquesta forma, resulta immediat veure que es pot fer una partició de la diferència relativa entre la taxa crua del grup comparació, tal i com s'expressa en la fórmula següent:

$$\frac{r_{|1} - r_{|0}}{r_{|0}} = \frac{r_{|2} - r_{|0}}{r_{|0}} + \frac{r_{|1} - r_{|2}}{r_{|0}} \quad (3.58)$$

Així doncs, $\frac{r_{|1} - r_{|0}}{r_{|0}}$ representa el canvi proporcional entre les taxes observades entre els dos grups, que es pot particionar en

1. el canvi proporcional degut a diferències en l'estructura de la població,

representat per $\frac{r_{|2}-r_{|0}}{r_{|0}}$

2. el canvi proporcional degut a diferències en aspectes epidemiològics (risc de desenvolupar o morir per càncer), representat per $\frac{r_{|1}-r_{|2}}{r_{|0}}$.

En conclusió, la quantificació de les diferències segons si són causades per canvis en factors demogràfics o a canvis epidemiològics permeten obtenir una millor informació sobre la càrrega d'una malaltia, el que esdevé relevant pel que fa a la planificació sanitària.

Capítol 4

Mortalitat per càncer de mama a Europa l'any 2005 i projeccions per a l'any 2015

Durant l'any 2006, el CM va representar la causa de mort per càncer més freqüent entre les dones europees. En aquest any, es va estimar que 131.900 dones europees varen morir per aquesta causa [Ferlay et al., 2007]. Durant l'any 2002, les taxes de mortalitat per CM varen assolir uns nivells màxims a l'Oest i Nord d'Europa, mentre que als països de l'Est i Centre d'Europa, en general, la mortalitat per aquesta causa es va mostrar sensiblement menor [Hery et al., 2008a,b, Tyczynski et al., 2004].

A Europa, des de la meitat dels anys 80 s'han experimentat canvis pel que fa a la detecció precoç i el tractament envers el CM que van conduir primer a una estabilització de la mortalitat per aquesta malaltia i a un descens a partir de la meitat dels anys 90. No obstant, aquesta tendència observada al Nord, Oest i Sud d'Europa no és va constatar a la majoria de països de l'Est i Centre d'Europa, amb l'excepció de la República Txeca, on el decreixement de la mortalitat per CM no va ser significatiu fins a l'any 2002 [Hery et al., 2008a,b]. D'aquesta forma, donat que la implementació de les intervencions per a la reducció de la mortalitat per CM no ha estat homogènia entre els

països europeus [Hendrick et al., 2002], resulta d'interès analitzar-la a Europa en els anys futurs.

En aquest capítol, presentem els resultats obtinguts que ens han permès estimar la mortalitat per CM a Europa l'any 2005 i projectar-la l'any 2015, avaluant el paper dels factors demogràfics i epidemiològics per a cadascun dels països considerats.

4.1 Font de les dades

Les dades de mortalitat per CM han estat extretes de la base de dades de mortalitat que proporciona l'Organització Mundial de la Salut (OMS, pàgina web: www.who.int/whosis, últim accés novembre 2008), i que també està disponible a la pàgina web de l'Agència Internacional per a la Recerca en el Càncer (IARC, pàgina web: www.iarc.fr, últim accés novembre de 2008). Les dades de població europea per als anys analitzats i per a les projeccions s'han obtingut de la pàgina web de la Secció d'Estadística de les Nacions Unides, específicament a partir de la revisió de l'any 2006 sobre la població mundial (extretes de <http://esa.un.org/unpp/>, últim accés novembre de 2008). D'aquesta forma, hem inclòs un total de 32 països europeus, per als que hem generat una base de dades que conté el nombre de morts per CM (agrupats per grups quinquennals) en un període variable, des de l'any en el que es va iniciar el registre de la mortalitat fins al darrer any disponible. Així doncs, disposem de dades fins a l'any 2007 per a Romania; fins a l'any 2006 per a Finlàndia, Islàndia, Irlanda, Regne Unit (que inclou Anglaterra, Gal·les i Escòcia), Àustria, Alemanya, Suïssa, Grècia, Eslovènia, Letònia, Lituània, Polònia i la Federació Russa; fins a l'any 2005 per a Noruega, Suècia, França, Luxemburg, Holanda, Malta, Espanya, República Txeca, Estònia, Hongria, Eslovàquia i Ucraïna; fins a l'any 2004 per a Bulgària; fins a l'any 2003 per Itàlia i Portugal; fins a l'any 2001 per a Dinamarca i fins a l'any 1997 per a Bèlgica. Les dades de població s'han registrat a la mateixa base de dades, incloent la població en els anys pels que es disposa de dades de mortalitat,

però també completant-ho fins a la població de l'any 2015. Hem dividit els països europeus en quatre regions, que detallem a continuació.

1. Nord d'Europa: Dinamarca, Finlàndia, Irlanda, Islàndia, Noruega, Regne Unit, Suècia.
2. Oest d'Europa: Alemanya, Àustria, Bèlgica, França, Luxemburg, Països Baixos, Suïssa.
3. Sud d'Europa: Eslovènia, Espanya, Grècia, Itàlia, Malta, Portugal.
4. Est-Centre d'Europa: Bielorrússia, Bulgària, Eslovàquia, Estònia, Federació Russa, Hongria, Letònia, Lituània, Polònia, República Txeca, Romania, Ucraïna.

4.2 Anàlisis Estadístiques

Hem calculat les taxes de mortalitat per CM crues i ajustades segons la població estàndard mundial [Segi, 1960] per a 100.000 dones per a tots els grups d'edat quinquennals, i pels següents grups d'edat: de 35 a 49 anys; de 50 a 69 anys, més de 70 anys i de forma agrupada per a les dones majors de 35 anys. Hem mesurat el canvi percentual anual (APC) de les taxes de mortalitat per CM, emprant per això els 10 darrers anys disponibles per a cadascun dels països. Per a estimar les APCs hem optat per un model edat-període log-linear amb distribució Gaussiana, que correspon al model especificat a l'expressió 3.26, però assumint que les taxes $\frac{c_{it}}{n_{it}}$ es distribueixen normalment. Els paràmetres ajustats pel propi model i la seva variabilitat estimada ens han permès avaluar la significació estadística de les APCs i calcular el seu interval de confiança al 95%.

4.3 Models per a les projeccions

Seguint la metodologia descrita per Dyba i Hakulinen [Dyba and Hakulinen, 2000, Hakulinen and Dyba, 1994, Dyba, 2000], hem ajustat per a cadascun

dels 32 països els tres models que s'exposen a continuació i que han estat descrits en el capítol anterior (3.38, 3.37 i 3.26 respectivament):

$$\text{Model 'AAAP'} \quad E\left(\frac{c_{it}}{n_{it}}\right) = \alpha_i + \beta_i t \quad (4.1)$$

$$\text{Model 'MAP'} \quad \log\left(E\left(\frac{c_{it}}{n_{it}}\right)\right) = \alpha_i + \beta t \quad (4.2)$$

$$\text{Model 'MAAP'} \quad \log\left(E\left(\frac{c_{it}}{n_{it}}\right)\right) = \alpha_i + \beta_i t \quad (4.3)$$

L'acrònim AAAP del primer model indica que es tracta d'un model edat-període additiu, que modelitza un efecte individual per a cadascun dels grup d'edat (*Additive Age-specific Age Period*). La segona expressió, defineix un model multiplicatiu amb efectes individuals pels grups d'edat i que assumeix el mateix ritme de creixement o decreixement per a tots els grups d'edat (*Multiplicative Age-specific Period*). Aquest model també sovint es coneix com *age-drift*. Finalment, la tercera expressió representa al clàssic model edat-període, que especifica efectes individuals per als grups d'edat així com tendències diferents (*Multiplicative Age-specific Age-Period*).

Per a cadascun dels 32 països analitzats hem emprat les dades de mortalitat per CM (c_{it}) i població (n_{it}) per als 11 grups d'edat quinquennals corresponents a dones majors de 35 anys, doncs es va considerar que la mortalitat per CM per sota d'aquesta edat era massa petita i podia influir els model finals. Per escollir quin és el model ajustat per a cadascun dels països, hem emprat els criteris exposats en el capítol anterior. D'aquesta forma, per als països amb tendències de mortalitat per CM creixents, es va optar preferencialment pel model AAAP, degut a que els altres dos models podien aportar projeccions excessivament elevades. Els països amb una tendència de la mortalitat per CM estable o decreixent, es va donar preferència als dos models MAP i MAAP, emprant l'estadístic de la khi-quadrat de Pearson per avaluar-me la bondat d'ajust i escollir aquell model amb un valor menor per aquest estadístic. En el cas de models amb bondats d'ajust similars es va

emprar el criteri de la parsimònia, és a dir el que requeria un nombre menor de paràmetres ajustats.

Cadascun dels models ha estat doncs ajustat emprant les dades de mortalitat per CM dels últims 10 anys disponibles. Quan el model no presentava un ajust acceptable, es va observar que en la majoria de països la tendència de la mortalitat era canviant en els darrers 10 anys. Per aquest motiu, es van repetir les modelitzacions restringint el període d'ajust als 9, 8, 7, 6 o 5 darrers anys, fins que el model oferís un ajust acceptable.

Un cop establert un model vàlid per a cadascun dels 32 països, es va obtenir la projecció dels casos morts per CM per a cada grup d'edat a l'any 2015 i es va emprar el mètode proposat per Bashir i Estève [Bashir and Esteve, 2000], descrit en el capítol anterior, per a avaluar el paper dels factors demogràfics i epidemiològics en la projecció.

4.4 Resultats

En aquest apartat presentarem els resultats generals obtinguts de l'anàlisi de les dades de mortalitat per CM a Europa. En primer lloc descriurem la mortalitat en termes absoluts, a partir de les taxes de mortalitat crues i ajustades. En segon lloc, avaluarem la tendència de la mortalitat per CM a Europa i les diferències segons els països i regions europees. En tercer lloc, complementarem l'anàlisi amb l'obtenció de la mortalitat per CM projectada per a l'any 2015 i, finalment, avaluarem quina part dels canvis és deguda a factors demogràfics i quina a factors epidemiològics relacionats amb el risc de morir per CM. Les taules i figures que contenen els següents apartats, seran comentades al llarg, i paral·lelament, en els diferents sub-apartats que a continuació presentem.

4.4.1 La mortalitat per càncer de mama a Europa els anys 2000 i 2005

Les taules 4.1, 4.2, 4.3, 4.4 i 4.5 recullen les taxes de mortalitat per CM crues i ajustades observades (i per alguns països estimades) per als anys 2000 i 2005 per als 32 països analitzats, segons la regió europea (Nord, Oest, Sud o Centre i Est). L'estudi de la taxa de mortalitat per CM crua ens permet avaluar la magnitud d'aquesta malaltia i l'estudi de la taxa ajustada, comparar els valors per a diferents països, analitzant així la seva heterogeneïtat a Europa. La figura 4.1 visualitza el rànking europeu en quant a la taxa ajustada de mortalitat per CM per a les dones majors de 35 anys i per als grups d'edat en el període 2000 i 2005.

L'any 2000 la taxa ajustada de mortalitat per CM a Europa va oscil·lar entre les 35 i 80 defuncions per 100.000 dones majors de 35 anys. Islàndia i Espanya van presentar les taxes més baixes (34,37 i 37,05 respectivament) mentre que Malta i Dinamarca les més altes (79,35 i 66,23 respectivament). Altres països que van presentar les taxes de mortalitat per CM en les 10 posicions més elevades foren Irlanda, Bèlgica, Hongria, els Països Baixos, Estònia, Luxemburg, el Regne Unit i Alemanya (oscil·lant la taxa entre 50,71 i 64,83), mentre que altres països amb taxes baixes foren Grècia, Polònia, Bielorrússia, Portugal, Bulgària, Finlàndia, Romania i Suècia (oscil·lant la taxa entre 37,97 i 43,03). L'any 2005, la taxa de mortalitat per CM a Europa va oscil·lar entre les 35 i 65 morts per 100.000 dones majors de 35 anys. Espanya i Islàndia van continuar presentant les taxes més baixes (33,38 i 35,77 respectivament) mentre que Dinamarca i Bèlgica presentaren les més altes (64,78 i 63,33 respectivament). Dins del conjunt de països europeus amb les 10 taxes de mortalitat més elevades cal afegir Irlanda, els Països Baixos, Malta, el Regne Unit, Hongria, Eslovènia, Alemanya i la República Txeca (oscil·lant la taxa entre 46,18 i 55,34), mentre que dels 10 països amb taxes baixes cal afegir Bulgària, Luxemburg, Grècia, Noruega, Portugal, Polònia, Finlàndia i Bielorrússia (oscil·lant la taxa entre 34,70 i 36,40).

Pel que fa a les taxes de mortalitat per CM ajustades segons els grups

d'edat s'observen grans diferències, presentant els grups més envellits una mortalitat per CM superior respecte les dones més joves. La taxa de mortalitat per CM ajustada per al grup de dones entre 35 i 49 anys es situa a Europa entre les 5 i 35 morts per 100.000 l'any 2000 i entre les 15 i 35 morts per 100.000 l'any 2005. Per a les dones entre 50 i 69 anys, la taxa ajustada oscil·la entre 45 i 100 l'any 2000 i entre 40 i 90 per a l'any 2005. Finalment, les dones majors de 70 anys presenten les taxes de mortalitat ajustades entre 65 i 230 defuncions l'any 2000 i entre 70 i 170 defuncions l'any 2005.

En general s'observa que el rànking de mortalitat per CM europeu descrit es manté en un cert grau, per als diferents grups d'edat. Per exemple, Bèlgica es situa dins dels grup dels 5 països europeus amb major mortalitat per CM, independentment del grup d'edat, o Espanya, que es situa dins dels països europeus amb menor mortalitat per CM per a tots els grups d'edat, tant en l'any 2000 com en l'any 2005. Tanmateix, s'observen països on aquesta circumstància no es dona. Per exemple, Grècia mostra unes taxes de mortalitat per CM de les més baixes a Europa, mentre que les dones majors de 70 anys mostren una mortalitat per CM semblant a l'observada en els països posicionats en el rànking mitjà-alt europeu. La Federació Russa també presenta una certa variabilitat, doncs tot i tenir en general una posició en el rànking europeu mitjana, les dones majors de 70 anys es situen clarament en els països europeus amb baixa mortalitat.

4.4.2 Anàlisi de la tendència de la mortalitat per càncer de mama a Europa durant els 10 darrers anys

Les taules 4.1, 4.2, 4.3, 4.4 i 4.5 recullen l'APC estimada en els models gaussians que avaluen la tendència de la mortalitat per CM en els darrers 10 anys disponibles, per cada país i, a més, contenen els intervals de confiança al 95% corresponents. Aquestes taules indiquen en negreta aquells intervals d'APC que no contenen la unitat i, per tant, són estadísticament significatius a un nivell del 5% ($\alpha = 0,05$). Paral·lelament, les figures 4.2, 4.3, 4.4 i 4.5 mostren les taxes ajustades de mortalitat per CM observades segons els grups

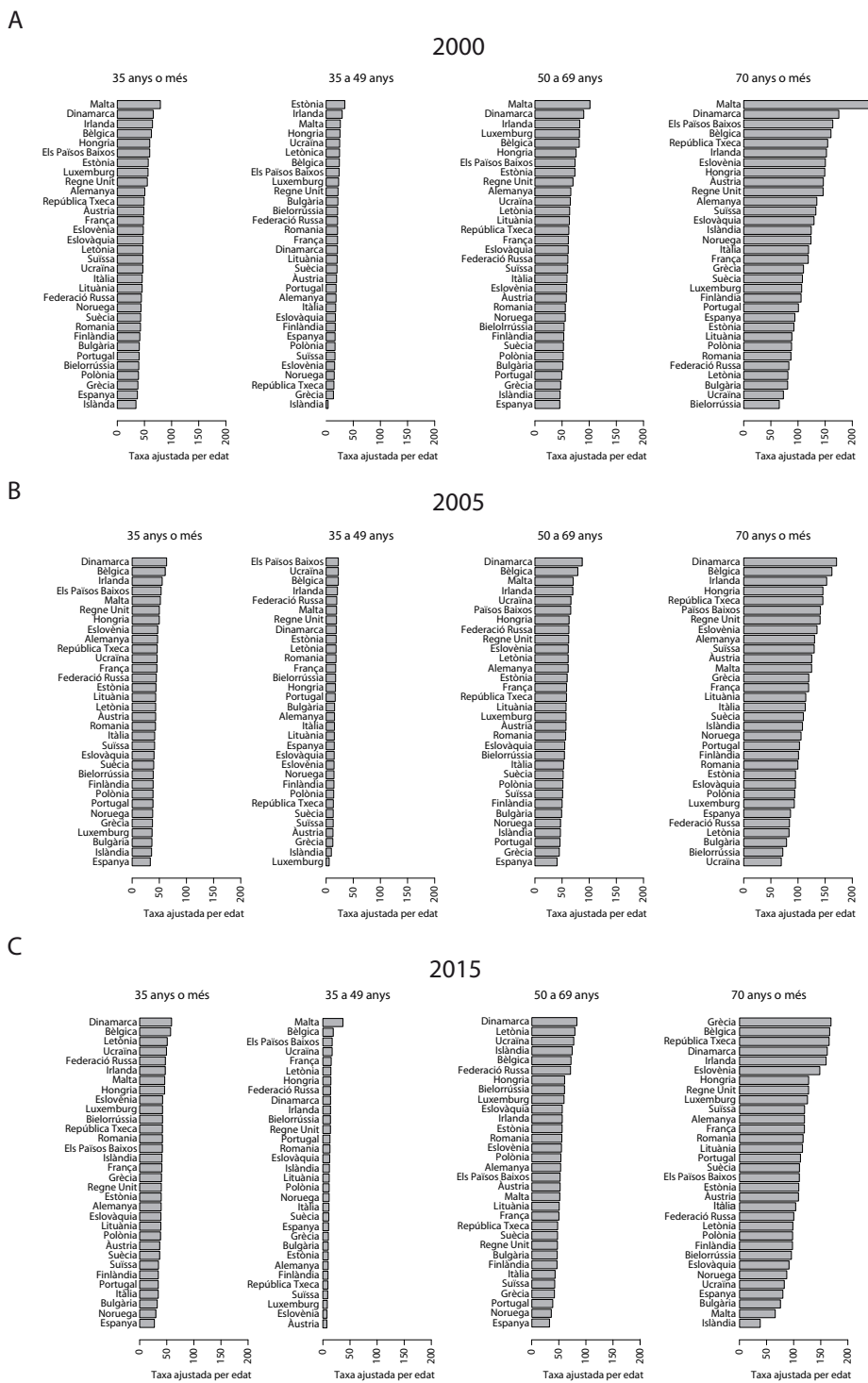


Figura 4.1: Rànkings de les taxes ajustades de mortalitat per CM per a les dones majors de 35 anys en els grups d'edat de majors de 35, 35-49, 50-69 i més de 70 anys, per als anys 2000 (A), 2005 (B) i 2015 (C).

Europa del Nord	País	Grup d'edat	APC (95% IC)	2000		2005		2015		Net Mida (%)	Canvi (N) 2015 vs 2005 Estructura (%)	Risc (%)	
				TA	N	TC	TA	N	TC				TA
Dinamarca (1992-2001)	35-49		-2,61 (-3,94; -1,26)	20,37	20,91	21,63	22,04	21,02	20,39	-6,35	-	-	
	50-69		-0,81 (-2,18; 0,57)	89,12	89,97	87,78	86,67	82,78	80,20	2,06	-	-	
	>70		-0,72 (-1,58; 0,14)	190,61	175,23	188,01	169,58	170,83	156,92	5,43	-	-	
	>35		-1,08 (-2,14; -0,02)	86,65	66,23	85,09	64,78	83,01	59,95	2,88	5,74	-8,17	
Finlàndia (1997-2006)	35-49		-3,94 (-4,82; -3,06)	17,07	16,86	14,59	14,42	14,83	14,67	-10,00	-	-	
	50-69		-0,82 (-1,93; 0,31)	52,44	53,15	50,19	50,13	46,96	45,55	4,09	-	-	
	>70		-0,43 (-1,28; 0,41)	111,82	105,71	107,77	100,92	100,11	91,00	4,99	-	-	
	>35		-1,33 (-1,94; -0,72)	52,87	41,51	51,38	38,56	50,99	35,71	3,16	8,27	-9,07	
Irlanda (1997-2006)	35-49		-3,93 (-7,11; 0,38)	28,84	29,61	20,38	20,92	17,80	18,62	9,09	-	-	
	50-69		-2,57 (-3,73; -1,41)	81,44	82,54	68,67	69,69	65,35	64,85	23,51	-	-	
	>70		0,61 (-0,51; 1,76)	155,88	152,45	167,61	152,85	135,63	123,47	-3,21	-	-	
	>35		-1,89 (-3,03; -0,76)	72,98	64,83	66,26	55,34	57,45	49,12	9,13	22,42	-13,16	
Islàndia (1997-2006)	35-49		-10,64 (-26,11; 1,82)	3,32	3,66	9,56	9,71	6,95	7,05	-33,33	-	-	
	50-69		1,29 (-4,75; 7,83)	46,12	46,42	47,11	46,98	29,87	29,76	-15,38	-	-	
	>70		-7,91 (-11,76; -4,16)	147,62	124,37	105,46	108,24	65,39	57,07	-33,33	-	-	
	>35		46,41	34,37	42,34	35,77	27,71	21,88	23	-25,81	7,79	1,12	-34,72
Noruega (1996-2005)	35-49		-3,39 (-6,57; -0,19)	14,76	15,01	14,30	14,55	12,66	12,32	-9,86	-	-	
	50-69		-2,81 (-3,67; -1,93)	55,56	55,84	46,63	47,07	39,13	38,31	-0,42	-	-	
	>70		-1,94 (-3,01; -0,88)	139,81	123,96	126,77	105,61	94,99	81,77	-20,42	-	-	
	>35		-2,67 (-3,45; -1,89)	61,32	43,67	52,83	37,83	42,21	30,57	-12,46	7,72	2,33	-22,51
Regne Unit (1997-2006)	35-49		-2,97 (-3,13; -2,81)	21,21	21,98	18,68	19,14	14,75	14,25	-25,18	-	-	
	50-69		-2,64 (-2,88; -2,39)	70,41	70,16	63,22	62,23	49,24	48,42	-10,07	-	-	
	>70		-1,15 (-1,49; -0,79)	164,26	146,26	163,49	140,71	151,88	128,33	0,87	-	-	
	>35		-2,27 (-2,41; -2,13)	74,78	55,35	69,75	50,08	62,11	40,65	-5,61	5,34	4,66	-15,61
Suècia (1996-2005)	35-49		-2,40 (-5,21; 0,44)	20,01	20,15	12,92	13,11	5,10	4,94	-60,17	-	-	
	50-69		-0,75 (-1,73; 0,25)	52,16	52,48	52,74	52,11	49,07	47,56	-0,52	-	-	
	>70		0,19 (-0,33; 0,71)	118,95	108,15	124,38	109,85	127,67	115,68	13,41	-	-	
	>35		-0,82 (-1,66; 0,02)	58,81	43,03	57,13	39,71	55,47	34,54	2,48	5,32	2,76	-5,60

Taula 4.1: Tendència de la mortalitat per CM segons país i grups d'edat a Europa del Nord. Projecció de la mortalitat per CM per l'any 2015 i contribució dels factors demogràfics. Taxa crua per 100.000 dones (TC), taxa ajustada a la població mundial per 100.000 dones (TA) i nombre de casos (N).

Europa de l'Oest País (Període)	Grup d'edat	APC (95% IC)	2000			2005			2015			Net Mida (%)	Canvi (N) 2015 vs 2005 Estructura (%)	Risc (%)	
			TC	TA	N	TC	TA	N	TC	TA	N				
Alemanya (1997-2006)	35-49	-4,20 (-4,59; -3,81)	17,62	18,35	1663	15,18	15,48	1515	10,68	10,09	885	-41,98	-	-	
	50-69	-1,39 (-1,76; -1,03)	68,41	66,12	6817	63,83	61,18	6671	54,62	54,56	6207	-6,96	-	-	
	>70	-0,79 (-1,12; -0,46)	147,95	134,71	9199	141,47	130,48	9191	138,50	126,55	10742	16,88	-	-	
Àustria (1997-2006)	>35	-1,71 (-1,98; -1,43)	69,02	50,71	17679	64,53	46,83	17377	65,07	41,07	17834	2,63	1,81	11,23	-10,41
	35-49	-5,47 (-7,45; -3,49)	18,18	19,27	165	12,33	12,68	124	7,75	7,46	71	-42,74	-	-	
	50-69	-0,92 (-1,26; -0,59)	58,68	57,85	530	58,53	57,21	560	51,73	52,26	584	4,29	-	-	
Bèlgica (1988-1997)	>70	-1,62 (-2,41; -0,84)	162,47	146,69	961	148,56	125,25	889	133,10	108,94	904	1,69	-	-	
	>35	-1,88 (-2,38; -1,38)	68,94	48,93	1656	61,43	43,28	1573	57,15	37,01	1559	-0,89	6,07	7,59	-14,55
	35-49	-1,59 (-2,72; -0,45)	26,33	26,91	300	26,48	26,41	308	26,06	25,41	279	-9,42	-	-	
França (1996-2005)	50-69	-0,85 (-1,70; -0,01)	86,18	84,51	988	83,58	82,91	998	81,00	79,80	1153	15,53	-	-	
	>70	0,23 (-0,59; 1,08)	165,61	154,06	1232	162,31	151,14	1286	165,72	145,47	1392	8,24	-	-	
	>35	-0,72 (-1,29; -0,15)	83,17	64,55	2520	82,21	63,33	2592	84,73	60,95	2824	8,95	5,84	7,31	-4,20
Luxemburg (1996-2005)	35-49	-1,47 (-2,78; -0,15)	20,99	21,1	1355	17,63	17,64	1149	17,06	16,80	1073	-6,61	-	-	
	50-69	-1,57 (-1,96; -1,18)	61,89	61,58	3924	57,97	58,33	4004	56,71	55,02	4725	18,01	-	-	
	>70	-0,14 (-0,63; 0,34)	131,01	118,91	5579	133,21	119,61	6055	130,29	110,79	6284	3,78	8,18	7,15	-7,53
Països Baixos (1996-2005)	>35	-1,18 (-1,51; -0,86)	63,67	48,44	10868	62,37	45,51	11208	62,13	42,78	12082	7,80	-	-	
	35-49	-4,56 (-9,52; 0,42)	21,81	23,04	11	5,35	5,66	3	15,83	15,47	9	200,00	-	-	
	50-69	-0,21 (-5,66; 5,54)	83,49	81,96	38	56,67	57,32	27	46,73	46,81	28	3,70	-	-	
Suïssa (1997-2006)	>70	-0,36 (-2,93; 3,34)	116,56	106,76	32	110,02	93,24	31	134,06	112,48	42	35,48	-	-	
	>35	-1,14 (-4,16; 1,93)	65,64	56,66	81	46,26	36,63	61	53,38	38,87	79	29,51	15,05	12,14	2,32
	35-49	-3,76 (-5,59; -1,92)	23,36	23,92	424	22,27	22,56	422	17,83	16,96	304	-27,96	-	-	
Suïssa (1997-2006)	50-69	-2,49 (-3,08; -1,91)	73,55	73,96	1256	65,79	65,98	1237	53,27	52,35	1209	-2,26	-	-	
	>70	-2,72 (-3,21; -2,24)	182,12	164,06	1712	164,77	141,34	1621	130,59	113,05	1469	-9,38	-	-	
	>35	-2,53 (-2,99; -2,08)	76,01	59,74	3392	68,93	53,34	3280	58,49	41,98	2982	-9,99	6,05	8,45	-23,59
Suïssa (1997-2006)	35-49	-2,38 (-3,22; -1,53)	15,67	16,36	132	12,81	13,07	115	12,34	11,71	100	-13,04	-	-	
	50-69	-2,56 (-3,78; -1,34)	60,49	60,39	505	51,74	50,99	463	45,02	44,38	474	2,38	-	-	
	>70	-1,03 (-1,96; -0,11)	149,72	132,91	691	143,36	129,88	708	117,89	104,26	701	-0,99	-	-	
		-2,28 (-3,22; -1,33)	62,11	47,17	1328	56,24	41,33	1286	51,87	35,21	1275	-0,86	6,85	9,31	-17,02

Taula 4.2: Tendència de la mortalitat per CM segons país i grups d'edat a Europa de l'Oest. Projectió de la mortalitat per CM per l'any 2015 i contribució dels factors demogràfics. Taxa crua per 100.000 dones (TC), taxa ajustada a la població mundial per 100.000 dones (TA) i nombre de casos (N).

País (Període)	Grup d'edat	APC (95% IC)			2000			2005			2015			Net (%)	Canvi (N) Mida (%)	Estructura (%)	Risc (%)
		APC (95% IC)	TC	TA	N	TC	TA	N	TC	TA	N	TC	TA				
Europa del Sud (1997-2006)	35-49	-6,26 (-8,81; -3,71)	15,66	15,73	36	14,92	14,69	34	7,77	7,77	17	7,77	17	-50,00	-	-	-
	50-69	-1,38 (-3,56; 0,82)	60,26	58,55	137	61,33	61,36	149	54,91	54,76	155	54,91	155	4,03	-	-	-
	>70	0,18 (-1,46; 1,84)	156,31	150,42	190	143,52	135,05	201	172,75	148,32	276	172,75	276	37,31	-	-	-
>35	-1,72 (-3,13; -0,29)	62,72	47,94	363	62,86	47,01	384	67,82	42,35	449	67,82	449	16,93	9,21	10,02	-2,30	
Espanya (1996-2005)	35-49	-3,24 (-4,09; -2,38)	16,03	16,58	688	14,66	15,06	727	10,54	10,64	593	10,54	593	-18,43	-	-	-
	50-69	-2,44 (-2,83; -2,06)	46,94	45,81	2068	41,22	40,79	1918	33,08	32,62	1899	33,08	1899	-0,99	-	-	-
	>70	-1,08 (-1,53; -0,63)	101,95	94,19	2850	94,36	86,22	3025	99,86	80,10	3623	99,86	3623	19,77	-	-	-
>35	-2,26 (-2,61; -1,92)	48,78	37,05	5606	44,24	33,38	5670	40,80	27,20	6114	40,80	6114	7,83	15,62	4,22	-12,01	
Grècia (1997-2006)	35-49	-2,28 (-4,15; -0,39)	13,18	13,51	151	12,20	12,41	150	9,63	9,59	125	9,63	125	-16,67	-	-	-
	50-69	-0,83 (-1,66; 0,01)	49,16	47,46	642	46,70	44,92	626	42,71	40,93	630	42,71	630	0,64	-	-	-
	>70	2,81 (1,65; 4,01)	114,91	110,13	798	126,51	119,85	999	193,37	155,58	1808	193,37	1808	80,98	-	-	-
>35	0,05 (-0,59; 0,69)	50,57	37,97	1591	52,84	37,41	1775	69,18	38,15	2563	69,18	2563	44,39	13,54	11,86	18,99	
Itàlia (1994-2003)	35-49	-3,13 (-3,93; -2,32)	17,04	17,67	1057	16,04	16,55	1085	15,56	15,00	1050	15,56	1050	-3,23	-	-	-
	50-69	-2,14 (-2,55; -1,73)	60,59	59,15	4366	56,55	55,20	4151	50,80	50,03	4098	50,80	4098	-1,28	-	-	-
	>70	-0,96 (-1,45; -0,47)	129,27	119,43	5814	125,43	113,76	6289	120,67	103,11	6816	120,67	6816	8,38	-	-	-
>35	-2,02 (-2,37; -1,65)	62,75	45,85	11237	60,28	43,06	11525	58,47	39,03	11964	58,47	11964	3,81	6,87	7,26	-10,32	
Malta (1996-2005)	35-49	-4,61 (-12,92; 3,92)	26,31	26,22	11	19,99	19,27	8	8,44	8,87	3	8,44	3	-62,50	-	-	-
	50-69	-3,58 (-7,34; 0,22)	101,33	101,44	45	67,79	70,51	33	37,93	36,26	21	37,93	21	-36,36	-	-	-
	>70	-6,02 (-10,53; -1,55)	228,63	230,06	44	135,88	125,12	30	83,62	75,88	22	83,62	22	-26,67	-	-	-
>35	-4,44 (-7,71; -1,16)	94,81	79,35	100	64,09	51,98	71	37,40	27,45	46	37,40	46	-35,21	7,08	6,13	-48,42	
Portugal (1994-2003)	35-49	-2,45 (-3,92; -0,98)	18,17	18,47	199	16,67	16,79	192	13,12	13,20	163	13,12	163	-15,10	-	-	-
	50-69	-1,91 (-2,75; -1,06)	49,68	49,39	599	47,02	46,26	591	39,45	38,64	556	39,45	556	-5,92	-	-	-
	>70	0,36 (-0,71; 1,43)	105,84	100,79	711	108,54	102,71	814	133,24	112,60	1157	133,24	1157	42,14	-	-	-
>35	-1,46 (-2,03; -0,89)	50,77	40,15	1509	50,55	38,24	1597	53,26	34,38	1878	53,26	1878	17,60	12,24	5,66	-0,30	

Taula 4.3: Tendència de la mortalitat per CM segons país i grups d'edat a Europa del Sud. Projecció de la mortalitat per CM per l'any 2015 i contribució dels factors demogràfics. Taxa crua per 100.000 dones (TC), taxa ajustada a la població mundial per 100.000 dones (TA) i nombre de casos (N).

País (Període)	Grup d'edat	APC (95% IC)	2000		2005		2015		Net Mida (%)	Canvi (N) Mida Estructura (%)	Risc (%)	
			TC	TA	TC	TA	TC	TA				N
Bielorrússia (1994-2003)	50-69	-2,01 (-3,98; 0,02)	20,72	21,25	16,14	15,59	188	7,24	76	-59,57	-	
	>70	0,82 (-0,88; 2,57)	54,61	53,72	48,34	48,96	558	37,69	486	-12,90	-	
	>35	2,77 (0,91; 4,71)	65,21	65,52	62,97	63,61	410	59,30	326	-20,49	-	
Bulgària (1995-2004)	50-69	0,41 (-0,73; 1,56)	43,11	39,58	38,94	34,70	1156	29,41	888	-23,18	1,41	
	>70	-4,22 (-6,09; -2,34)	21,97	21,73	15,78	15,60	126	9,88	79	-37,30	-	
	>35	-0,48 (-1,91; 0,98)	52,08	51,61	49,46	49,17	514	48,02	469	-2,33	-	
Eslovàquia (1996-2005)	50-69	-0,54 (-1,57; 0,48)	81,57	80,96	82,12	78,92	456	82,97	469	2,85	-	
	>70	-1,44 (-2,42; -0,46)	47,81	40,54	45,78	36,40	1097	43,60	1050	-4,26	1,82	
	>35	-1,73 (-4,21; 0,78)	17,33	17,07	15,27	14,88	88	11,84	74	-15,91	-	
Estònia (1996-2005)	50-69	-0,51 (-1,61; 0,58)	62,47	61,35	53,67	54,93	330	58,24	423	28,18	-	
	>70	-1,79 (-3,36; -0,21)	131,88	129,52	101,64	95,30	286	102,36	320	11,89	-	
	>35	-1,07 (-2,22; 0,09)	55,93	47,55	47,81	40,21	704	49,16	817	16,05	13,17	
Federació Rússa (1997-2006)	50-69	-5,03 (-10,04; -0,03)	34,66	34,38	19,22	18,56	28	19,14	19,33	26	-7,14	-
	>70	-1,15 (-2,48; 0,17)	73,11	72,64	61,39	59,29	110	58,22	57,61	106	-3,64	-
	>35	0,77 (-1,53; 3,14)	92,48	92,44	96,93	95,48	101	87,59	83,63	96	-4,95	-
Honrria (1996-2005)	50-69	-1,64 (-3,21; -0,07)	63,71	57,92	55,70	43,80	239	53,27	42,22	228	-4,60	-0,22
	>70	-2,66 (-3,09; -2,23)	21,13	21,17	20,68	19,65	3531	13,50	13,82	2046	-42,06	-
	>35	0,97 (0,44; 1,49)	61,35	60,73	62,93	62,37	11308	66,29	64,12	13696	21,12	-
Honrria (1996-2005)	50-69	1,41 (0,88; 1,92)	48,95	44,35	85,92	84,27	7479	96,28	86,31	8121	8,58	-
	>70	0,23 (-0,13; 0,62)	27,26	25,59	17,83	17,25	181	12,12	12,40	139	-23,30	-
	>35	-1,22 (-2,22; -0,21)	77,47	75,94	63,19	62,84	858	55,74	51,60	753	-12,24	-
Honrria (1996-2005)	50-69	-1,41 (-2,28; -0,54)	154,32	149,33	148,40	145,35	1036	142,73	145,15	1082	4,44	-
	>70	-3,16 (-4,94; -1,38)	76,09	59,81	65,57	49,99	2075	60,63	43,19	1974	-4,87	5,41
	>35	-1,64 (-2,28; -1,01)	2308	2308	2308	2075	2075	1974	1974	1974	2,23	-12,51

Taula 4.4: Tendència de la mortalitat per CM segons país i grups d'edat a Europa Central i de l'Est. Projectió de la mortalitat per CM per l'any 2015 i contribució dels factors demogràfics. Taxa crua per 100.000 dones (TC), taxa ajustada a la població mundial per 100.000 dones (TA) i nombre de casos (N).

Europa Central i de l'Est		2000		2005		2015		Canvi (N) 2015 vs 2005				
País (Període)	Grup d'edat	TC	TA	N	TC	TA	N	Net (%)	Mida (%)	Estructura (%)	Risc (%)	
Letònia (1997-2006)	35-49	24,47	24,78	65	18,88	18,43	48	11,24	11,43	26	-45,83	-
	50-69	64,97	64,02	203	62,25	61,18	191	78,73	77,05	242	26,70	-
	>70	83,13	81,42	144	85,91	83,46	156	91,83	81,65	181	16,03	-
	>35	54,83	47,27	412	53,19	43,28	395	60,92	46,45	449	13,67	3,27
Lituània (1997-2006)	35-49	19,07	20,43	74	15,29	15,16	42	6,81	6,72	24	-42,86	-
	50-69	65,09	63,52	272	58,44	57,37	227	38,10	37,11	172	-24,23	-
	>70	90,31	88,68	198	112,66	114,29	326	131,50	105,51	356	9,20	-
	>35	53,07	45,76	544	54,67	43,37	595	51,84	29,92	552	-7,23	1,43
Polònia (1997-2006)	35-49	17,07	16,48	759	14,87	14,01	595	10,64	11,18	422	-29,08	-
	50-69	52,61	52,15	2114	50,55	51,57	2283	55,20	52,17	2959	29,61	-
	>70	90,71	88,14	1806	96,88	93,96	2180	111,42	96,84	2731	25,28	-
	>35	44,74	39,04	4679	46,97	38,33	5058	51,88	37,46	6112	20,84	7,58
República Txeca (1996-2005)	35-49	15,16	14,41	163	13,42	13,35	136	8,77	9,46	104	-23,53	-
	50-69	62,41	62,47	766	57,91	58,21	774	52,12	47,63	730	-5,68	-
	>70	162,65	154,68	1005	154,34	145,81	995	186,66	165,55	1355	36,18	-
	>35	66,22	49,41	1934	63,62	46,18	1905	66,09	41,96	2189	14,90	4,05
Romania (1998-2007)	35-49	22,11	21,13	490	18,06	18,32	406	11,59	11,57	291	-28,33	-
	50-69	57,24	55,94	1435	57,28	56,62	1477	57,34	55,28	1506	1,96	-
	>70	87,48	86,97	991	99,14	99,83	1250	120,67	112,86	1645	31,60	-
	>35	49,81	42,72	2916	51,47	43,02	3133	52,91	40,64	3442	9,86	2,13
Ucraïna (1996-2005)	35-49	24,96	25,02	1422	23,41	22,54	1295	20,19	20,61	971	-25,02	-
	50-69	66,41	65,21	4071	67,28	66,24	4264	75,32	72,80	4914	15,24	-
	>70	73,24	73,09	2418	70,61	69,21	2248	62,05	58,40	2065	-8,14	-
	>35	52,29	47,01	7911	52,16	45,85	7771	54,21	46,56	7950	2,30	1,61

Taula 4.5: (continuació) Tendència de la mortalitat per CM segons país i grups d'edat a Europa Central i de l'Est. Projecció de la mortalitat per CM per l'any 2015 i contribució dels factors demogràfics. Taxa crua per 100.000 dones (TC), taxa ajustada a la població mundial per 100.000 dones (TA) i nombre de casos (N).

d'edat.

De l'anàlisi de les taules i gràfic citats anteriorment podem destacar que, en general, hi ha una certa homogeneïtat en la mortalitat per CM entre els països pertanyents a una mateixa regió europea. D'aquesta manera, podem observar com al Nord d'Europa (taula 4.1 i figura 4.2) la majoria de països presenten unes tendències de la mortalitat per CM decreixent, sobretot influïda pels grups de dones més joves. És aquest, per exemple, el cas de Finlàndia que presenta un ritme de decreixement anual del 1,3%, estadísticament significatiu, i que està principalment influït per la disminució significativa de la mortalitat en les dones més joves (de 35 a 49 anys), on el decreixement anual és del 4%. Dinamarca presenta també un perfil anàleg al de Finlàndia. En el cas de Noruega, el patró de la mortalitat per CM és similar a l'observat en el Regne Unit, on es pot constatar un descens estadísticament significatiu per a tots els grups d'edat, essent aquest major per les dones més joves (entre el 2% i el 3,4%). Pel que fa a Irlanda, el patró descrit és similar, encara que el decreixement és també significatiu en el grup d'edat intermig. A Suècia es pot assumir que les taxes de mortalitat per CM s'han estabilitzat en els darrers 10 anys (figura 4.2). Finalment, a Islàndia s'observa un patró de descens de la mortalitat consistent, en que els grups de dones més joves tenen un APC més esmorteït, encara que en aquest cas presenten uns valors, en termes absoluts, superiors, tot i que no són estadísticament significatius, degut a la mida petita de la població del país.

Pel que fa als països de l'Europa de l'Oest (taula 4.2 i figura 4.3) s'observa un patró de la mortalitat per CM en general similar al dels països nòrdics, essent el decreixement de la mortalitat major en les dones més joves. En aquests països també s'observa, de forma general, que la disminució de les taxes de mortalitat afecta sobretot als grups de dones amb edat intermèdia i alta, és a dir, no s'observa l'estabilització observada en els països nòrdics. Aquest és el cas d'Alemanya, que presenta un decreixement de la mortalitat per CM estadísticament significatiu per a tots els grups d'edat i que és alhora major per a les dones més joves. De forma similar es situen països com Suïssa,

Àustria i els Països Baixos. També s'observa aquest patró a Bèlgica i França, tot i que l'APC no és significatiu per al grup de dones majors de 70 anys, indicant una estabilització de la mortalitat per aquest grup. Observem també que els valors de l'APC són sensiblement majors que per als països nòrdics, com és el cas de les dones entre 35 i 49 anys austríaques en les que es pot constatar un decreixement anual del 5,5%. En general, s'observa que per a les dones de més de 35 anys la mortalitat per CM decreix entre el 1,18% de França i el 2,53% dels Països Baixos. El cas de Luxemburg presenta un patró de la mortalitat similar tot i que no és estadísticament significatiu.

En els països de l'Europa del Sud (taula 4.3 i figura 4.4) observem, en el cas d'Espanya, Itàlia i Portugal un patró similar al descrit pels països de l'Europa de l'Oest, constatant un descens de la mortalitat per CM estadísticament significatiu en tots els grups d'edat. Globalment, les dones italianes i espanyoles varen experimentar un descens anual de la mortalitat al voltant del 2%, essent aquest decrement superior en les dones majors. Així doncs, per a les dones italianes i espanyoles de 35 a 49 anys el decreixement fou del 1%, en les de 50 a 69 anys del 2-2,5% i per les de més de 70 anys superà el 3%, situant-se els darrers dos grups propers a països com Suïssa o els Països Baixos. Eslovènia mostra que hi ha un decrement de la mortalitat anual del 6,26% per a les dones entre 35 i 49 anys, influint en el descens global, que també és estadísticament significatiu. Cal remarcar Grècia pel comportament força particular i diferenciat de la resta de països de la seva àrea. Encara que hi ha una estabilització global de la mortalitat per CM, també es constata un decrement del 2,28% per les de dones més joves, alhora que s'observa un augment anual del 2,81% per a les dones majors de 70 anys, essent ambdues estadísticament significatives. Finalment, en el cas de Malta s'observa també un decrement general de la mortalitat per CM tot i que cal tenir present que es tracta d'un país amb molt pocs habitants comparat amb la resta i això pot influir en la variabilitat de les taxes.

En els països del'Europa Central i de l'Est (taules 4.4 i 4.5 i figura 4.5) s'observa un patró de la mortalitat per CM clarament diferenciat a la resta de

regions d'Europa. Tot i que hi ha un nivell de variabilitat important dins dels països d'aquesta zona es pot constatar que, en general, s'observa un decreixement global de les taxes de mortalitat per CM, essent més important en els grups d'edat més joves. Tanmateix però, és constatable que molts d'aquests països continuen presentant una tendència creixent de la mortalitat, específicament en les dones de major edat. A Ucraïna s'observa un descens anual de la mortalitat per CM del 2,49%, mentre que per al grup de dones entre 50 i 69 i per a les majors de 70 anys la mortalitat va augmentar a un ritme del 1,35% i 0,99% respectivament, essent estadísticament significatives. En la Federació Russa no observem cap decreixement de la mortalitat per CM en cap grup d'edat, constatant un augment significatiu de la mortalitat entre el 1 i el 1,5% per a les dones majors de 50 anys. Països com Romania i Letònia presenten un augment significatiu de la mortalitat en les dones majors de 50 anys, tot i que també s'observa un descens significatiu per les dones més joves. En els països on hi ha canvis de tendència de la mortalitat per CM segons els grups d'edat es constata una estabilització de la mortalitat. Hongria presenta un perfil similar al d'Espanya, on hi ha un descens general de la mortalitat, però essent més important en les dones més joves. Bulgària, Eslovàquia, Estònia, Lituània i la República Txeca també es situarien en un perfil similar tot i que per les dones més grans la mortalitat per CM s'estabilitza. Finalment, els resultats obtinguts per a Polònia indiquen que hi ha un decrement de la mortalitat únicament per al grup de dones més joves, mentre que per a Bielorrússia s'observa una tendència global a la disminució de la mortalitat tot i que per a cap grup d'edat s'observen diferències significatives, suggerint que al ser una població petita podria influir en aquests resultats.

4.4.3 La mortalitat per càncer de mama projectada a l'any 2015

La taula 4.6 conté els resultats, pel que fa a la bondat d'ajust, dels tres models (MAAP, AAAP i MAP) que hem ajustat per a cadascun dels 32

Europa del Nord

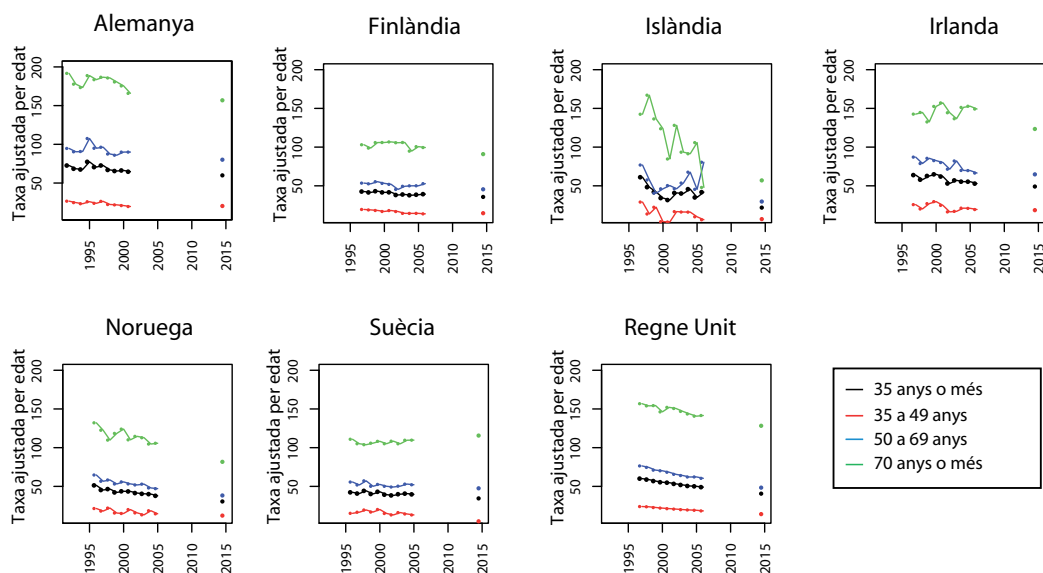


Figura 4.2: Taxes ajustades de mortalitat per càncer de mama a l'Europa del Nord i taxa projectada per l'any 2015, per tres grups d'edat).

Europa de l'Oest

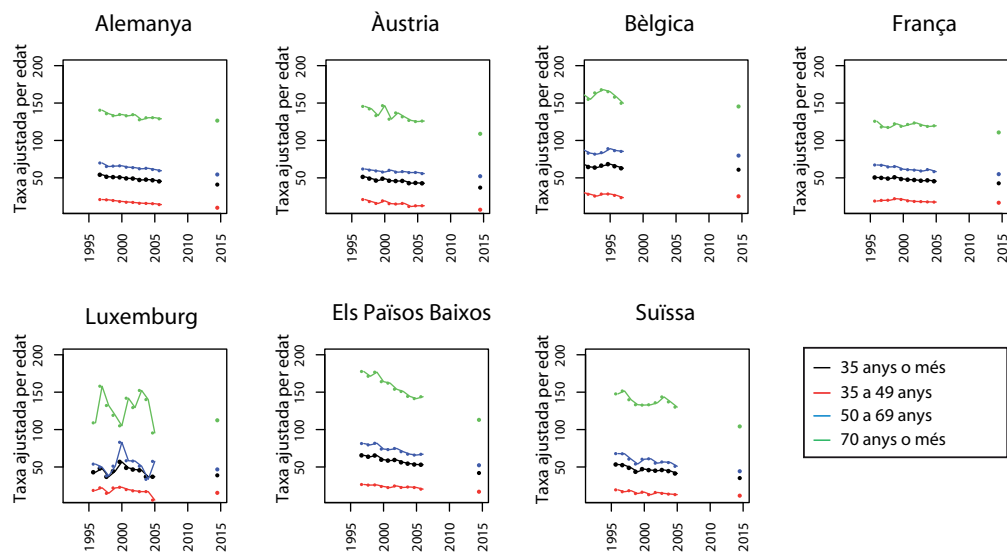


Figura 4.3: Taxes ajustades de mortalitat per càncer de mama a l'Europa de l'Oest i taxa projectada per l'any 2015, per tres grups d'edat).

Europa del Sud

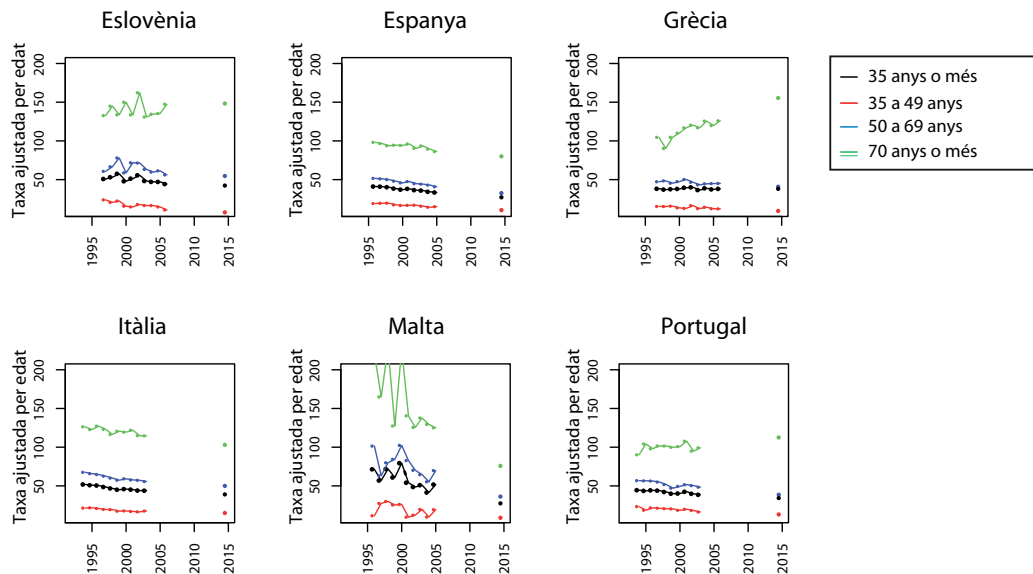


Figura 4.4: Taxes ajustades de mortalitat per càncer de mama a l'Europa del Sud i taxa projectada per l'any 2015, per tres grups d'edat).

països. En aquesta taula es mostra l'estadístic de la χ^2 de Pearson, els graus de llibertat corresponents i el p-valor, assumint la hipòtesi nul·la que el model ajusta bé a les dades. Seguint els criteris exposats en la secció 4.3, la taula indica quin ha estat el model escollit per a cada país. D'aquesta forma, per a 21 dels 32 països es va escollir un model amb un ajust acceptable, ja que presentaven un p-valor superior a 0,05, que permetia assumir com a certa la hipòtesi nul·la que el model ajusta suficientment bé les dades. En canvi, per a 11 països no es va trobar cap model que presentés un ajust acceptable. D'aquesta forma, en aquests 11 països restants, es va reduir el nombre d'anys del període d'ajust fins obtenir un p-valor superior a 0,05 (taula 4.7). Aquesta metodologia va permetre establir un model acceptable per a 10 dels 11 països restants. A Lituània no es va establir un model que ajustés, fet atribuïble a la qualitat de les dades i la mida petita de la població d'aquest país, optant pel model ajustat pels darrers 5 anys. Per la resta de països, el període d'ajust va oscil·lar des de els 9, 8 ó 7 anys respectivament

Europa del Centre i de l'Est

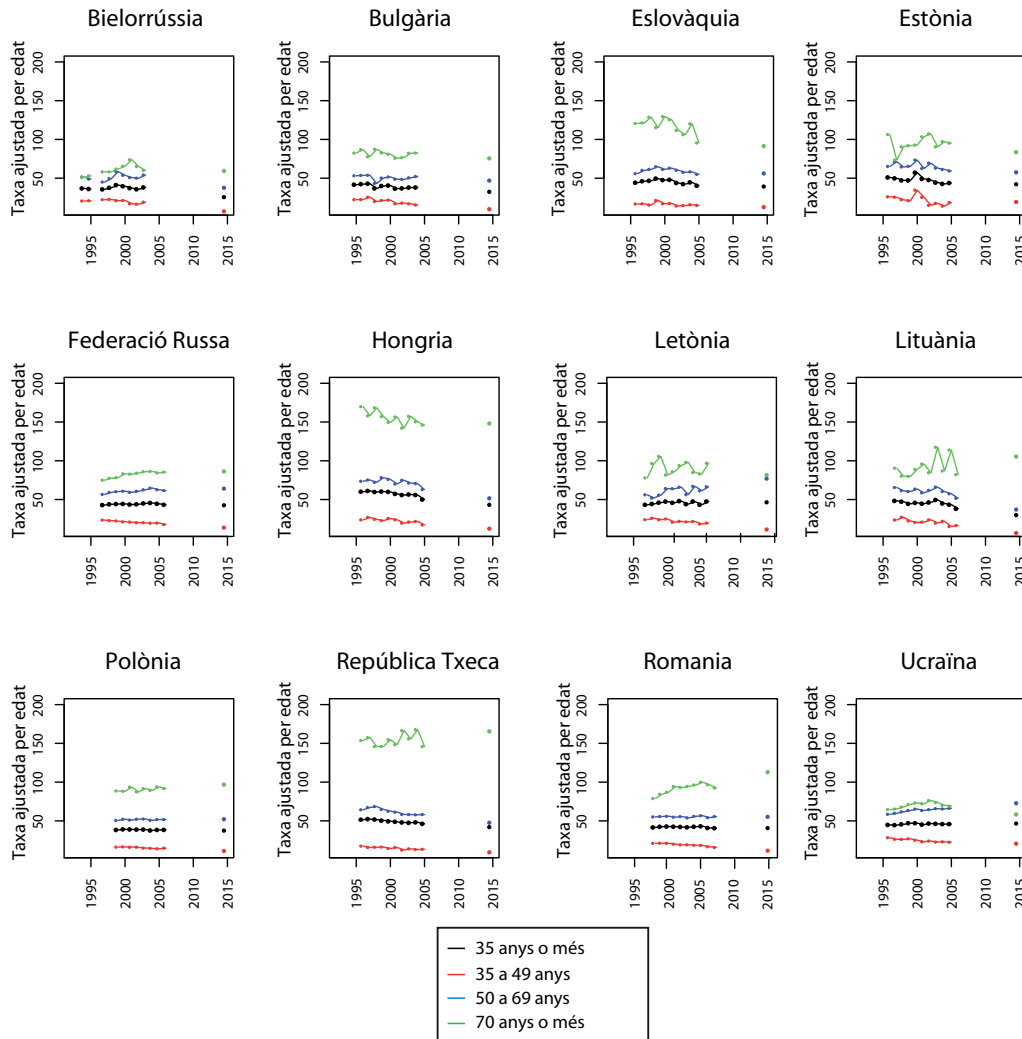


Figura 4.5: Taxes ajustades de mortalitat per càncer de mama a l'Europa del Centre i de l'Est i taxa projectada per l'any 2015, per tres grups d'edat).

per a Letònia, Suècia i el Regne Unit, 5 anys per a França, els 4 darrers anys per a Ucraïna i Itàlia i, els 3 darrers anys per a Alemanya, Bielorrússia, Federació Russa i Hongria. El fet que s'hagi d'escurçar el període d'ajust per tal que el model sigui acceptable suggereix que es tracta de països on, per algun grup d'edat o per a tots ells, hi ha hagut un canvi de tendència de la mortalitat per CM en els darrers anys (figures 4.2, 4.3, 4.4 i 4.5). Per exemple, a França s'observa que hi ha un cert canvi de tendència de la mortalitat per CM més marcat, en els darrers 5 anys, específicament per al grup de major edat.

D'aquesta forma, disposem per cadascun dels 32 països d'un model ajustat sobre el qual efectuar la projecció de la mortalitat per CM l'any 2015, partint dels darrers 10 anys disponibles en la majoria de casos o bé d'un període menor fins als tres darrers anys. El model edat-període complet, que contempla efectes diferents per a grups d'edat (MAAP) ha estat el model escollit per a 11 països (Alemanya, Àustria, Bulgària, Eslovàquia, Eslovènia, Espanya, Hongria, Lituània, Portugal, República Txeca i Regne Unit). Un model, també multiplicatiu, però que assumeix la mateixa tendència de mortalitat per CM per a tots els grups d'edat (MAP) ha estat l'escollit per a 13 països (Bèlgica, Dinamarca, Estònia, Finlàndia, França, Irlanda, Islàndia, Itàlia, Luxemburg, Malta, Noruega, Països Baixos i Suïssa). Finalment, el model additiu que assumeix tendències de mortalitat per CM diferents segons els grups d'edat (AAAP) ha permès model·lar les dades per a 8 països (Bielorrússia, Federació Russa, Grècia, Letònia, Polònia, Romania, Suècia i Ucraïna). Hem emprat els paràmetres estimats per aquests models per a projectar les defuncions per CM que caldria esperar per a l'any 2015. Per realitzar aquesta projecció hem utilitzat la població europea predita per a aquest any i que està descrita en la secció 4.1.

L'obtenció de les projeccions de mortalitat per CM segons grups d'edats ens ha permès obtenir una estimació per a la taxa bruta i ajustada de mortalitat per CM, a més del nombre de defuncions per a les dones majors de 35 anys i per grups d'edat. Les taules 4.1, 4.2, 4.3, 4.4 i 4.5 mostren les

Període ajust Inici Fi	País	Model.MAAP			Model.AAAP			Model.MAP			Model escollit
		X^2	g.l.l.	p-valor	X^2	g.l.l.	p-valor	X^2	g.l.l.	p-valor	
1997 2006	Alemanya	320,71	88	0,00	316,88	88	0,00	590,48	98	0,00	-
1997 2006	Àustria	89,52	88	0,44	90,86	88	0,40	135,41	98	0,01	MAAP
1994 2003	Bielorrússia	115,12	77	0,00	113,90	77	0,00	149,17	87	0,00	-
1988 1997	Bèlgica	101,47	88	0,15	101,36	88	0,16	118,77	98	0,08	MAP
1995 2004	Bulgària	103,19	88	0,13	102,10	88	0,14	130,97	98	0,01	MAAP
1992 2001	Dinamarca	90,86	88	0,40	90,02	88	0,42	116,03	98	0,10	MAP
1996 2005	Esllovàquia	105,85	88	0,09	105,38	88	0,10	124,12	98	0,04	MAAP
1997 2006	Eslovènia	97,80	88	0,22	97,16	88	0,24	123,81	98	0,04	MAAP
1996 2005	Espanya	97,56	88	0,23	97,10	88	0,24	175,49	98	0,00	MAAP
1996 2005	Estònia	102,03	88	0,15	101,84	88	0,15	113,73	98	0,13	MAP
1997 2006	Federació Russa	592,73	88	0,00	586,58	88	0,00	1277,35	98	0,00	-
1997 2006	Finlàndia	97,40	88	0,23	97,12	88	0,24	115,53	98	0,11	MAP
1996 2005	França	172,77	88	0,00	173,60	88	0,00	318,89	98	0,00	-
1997 2006	Grècia	105,12	88	0,10	105,74	88	0,10	193,21	98	0,00	AAAP
1996 2005	Hongria	135,38	88	0,00	133,90	88	0,00	168,49	98	0,00	-
1997 2006	Irlanda	80,10	88	0,71	79,50	88	0,73	113,29	98	0,14	MAP
1997 2006	Islàndia	86,49	88	0,53	85,12	88	0,57	101,10	98	0,39	MAP
1994 2003	Itàlia	158,10	88	0,00	160,82	88	0,00	284,00	98	0,00	-
1997 2006	Letònia	116,08	88	0,02	115,50	88	0,03	141,40	98	0,00	-
1997 2006	Lituània	116,43	88	0,02	116,20	88	0,02	144,64	98	0,00	-
1996 2005	Luxemburg	89,16	88	0,45	88,63	88	0,46	95,36	98	0,56	MAP
1996 2005	Malta	103,11	88	0,13	102,07	88	0,14	113,67	98	0,13	MAP
1996 2005	Noruega	91,71	88	0,37	92,56	88	0,35	103,15	98	0,34	MAP
1997 2006	Països Baixos	88,38	88	0,47	88,02	88	0,48	117,66	98	0,09	MAP
1997 2006	Polònia	75,04	66	0,21	74,12	66	0,23	136,78	76	0,00	AAAP
1994 2003	Portugal	105,12	88	0,10	103,90	88	0,12	154,21	98	0,00	MAAP
1997 2006	Regne Unit	132,01	88	0,00	135,43	88	0,00	282,03	98	0,00	-
1996 2005	República Txeca	102,92	88	0,13	102,51	88	0,14	148,40	98	0,00	MAAP
1998 2007	Romania	112,78	88	0,04	109,35	88	0,06	206,32	98	0,00	AAAP
1996 2005	Suècia	115,54	88	0,03	115,13	88	0,03	128,54	98	0,02	-
1996 2005	Suïssa	86,03	88	0,54	86,76	88	0,52	114,32	98	0,12	MAP
1996 2005	Ucraïna	241,27	88	0,00	239,15	88	0,00	431,17	98	0,00	-

Taula 4.6: Resultats dels models escollits per al càlcul de les projeccions de mortalitat per CM, emprant un període de 10 anys. Estadístic de Pearson (X^2), graus de llibertat (g.l.l.) i p-valor associat al model.

País	Període ajust		Nombre d'anys	Ajust model			Model escollit
	Inici	Fi		X^2	g.ll.	p-valor	
Alemanya	2004	2006	3 anys	17,00	11	0,11	MAAP
Bielorrússia	2001	2003	3 anys	19,23	11	0,06	AAAP
Federació Russa	2004	2006	3 anys	15,94	11	0,14	AAAP
França	2001	2005	5 anys	58,42	43	0,06	MAP
Hongria	2003	2005	3 anys	16,19	11	0,13	MAAP
Itàlia	2000	2003	4 anys	45,51	32	0,06	MAP
Letònia	1998	2006	9 anys	96,59	77	0,06	AAAP
Lituània	2002	2006	5 anys	100,48	66	0,00	MAAP
Regne Unit	2000	2006	7 anys	67,26	55	0,12	MAAP
Suècia	1998	2005	8 anys	85,68	66	0,05	AAAP
Ucraïna	2002	2005	4 anys	24,35	22	0,33	AAAP

Taulla 4.7: Resultats dels models escollits per al càlcul de les projeccions de mortalitat per CM, emprant un període inferior a 10 anys. Estadístic de Pearson (X^2), graus de llibertat (g.ll.) i p-valor associat al model.

projeccions per a l'any 2015. A més, els gràfics 4.2, 4.3, 4.4 i 4.5 mostren la projecció estimada, com un punt a l'any 2015, constatant-se que la projecció obtinguda és raonablement satisfactòria per a tots els països i grups d'edat inclosos en l'estudi. D'aquesta forma, en els casos en els que la mortalitat per CM augmenta, el creixement és sostingut, essent en cap cas exponencial. Per l'altra banda, en el cas en el que la mortalitat per CM s'estabilitza o decreix es constata una taxa de mortalitat projectada en el mateix nivell, o bé seguint un cert esmorteïment, cosa que té força sentit des del punt de vista epidemiològic. Per exemple, a Grècia podem observar a la figura 4.4 com la projecció de la mortalitat per CM en les dones majors de 70 anys és la corresponent a un creixement additiu i que, per tant, sosté que el model AAAP és el més vàlid en aquest cas. En els casos de decreixements importants de la mortalitat per CM, com és el cas de Noruega per a tots els grups d'edat, observem l'efecte dels models exponencials que ofereixen una projecció en una tendència decreixent que s'esmorteïx.

Finalment, la taulla 4.8 mostra les taxes de mortalitat per CM brutes i ajustades, així com el nombre de morts estimats per a l'any 2005 i els projectats a l'any 2015, agregades segons les regions d'Europa prèviament definides. Així doncs, a Europa s'estima que l'any 2015, hi hauran 127.983 defuncions per CM en les dones majors de 35 anys, específicament 10.288 casos (8,04%) en les dones entre 35 i 49 anys, 54.226 casos (42,37%) en les de

50 i 69 anys i 63.469 casos (49,59%) en les de més de 70 anys. És remarcable que, el nombre de dones que varen morir per aquesta causa l'any 2005 va ser de 122.264, molt similar al nombre de defuncions projectades per a l'any 2015. En termes absoluts, la majoria de defuncions per CM s'observaran a l'Europa del Centre i de l'Est (49.513, 38,69%), seguit de l'Europa de l'Oest (38.636, 30,19%) i l'Europa del Sud i Nord (23.014 i 16.820, 17,98% i 13,14% respectivament). En termes de la taxa de mortalitat per CM projectada, a Europa s'estima que l'any 2015 hi hauran 56,91 morts per cada 100.000 dones majors de 35 anys. Per al grup de dones entre 35-49 anys aquesta taxa es redueix fins a 13,60 casos per 100.000, en els dones entre 50-69 anys augmenta a 56,28 casos per 100.000 i, finalment, en les dones de més de 70 anys la taxa crua s'estima de 120,07 casos per cada 100.000. En la taula 4.8 s'observa com les taxes de mortalitat per CM projectades també són força diferents segons la regió d'Europa. Els Països del Nord i Oest presenten les taxes més altes, amb un valor al voltant de 60 defuncions per 100.000 dones, arribant als 135 i 145 casos per 100.000 en el grup de més de 70 anys. Els països del Sud, Centre i Est d'Europa presenten unes taxes de mortalitat per CM sensiblement menors, al voltant de 50 casos per 100.000 en les dones majors de 35 anys, observant-se però com aquesta taxa augmenta amb l'edat, tal i com s'ha observat en altres àrees, tot i que de forma menys pronunciada.

4.4.4 Anàlisi dels factors demogràfics i de risc en les diferències entre la mortalitat de l'any 2015 respecte l'any 2005

La taula 4.8 conté els resultats d'aplicar el mètode descrit en els apartats 3.5 i 4.3, que avalua quina part de l'increment o descens de la mortalitat per CM projectada a l'any 2015 respecte l'observada l'any 2005 és atribuïble a canvis demogràfics o a factors epidemiològics associats al risc del morir. La figura 4.6 visualitza els percentatges de canvi observats en la mortalitat nets i separatament segons els factors demogràfics i de risc.

Europa				Canvi (N) 2015 vs 2005							
Regió	Grup d'edat	2005			2015			Net Mida Estructura Risc (%)			
		TC	TA	N	TC	TA	N				
Nord	35-49	17,91	18,28	1741	14,22	13,84	1338	-23,15	-	-	-
	50-69	62,16	61,37	6348	51,25	50,15	5996	-5,55	-	-	-
	>70	155,21	135,00	9353	143,52	123,86	9486	1,42	-	-	-
	>35	67,18	48,71	17442	60,68	40,71	16820	-3,56	6,11	4,33	-14,00
Oest	35-49	16,90	17,11	3637	14,22	13,64	2720	-25,21	-	-	-
	50-69	62,50	61,30	13960	56,11	55,05	14381	3,02	-	-	-
	>70	141,92	128,44	19776	135,87	120,55	21534	8,89	-	-	-
	>35	64,67	47,43	37374	63,74	42,33	38636	3,38	4,83	9,16	-10,61
Sud	35-49	15,28	15,68	2196	12,87	12,68	1951	-11,16	-	-	-
	50-69	50,18	49,20	7468	43,21	42,48	7362	-1,42	-	-	-
	>70	114,47	104,54	11354	121,63	101,04	13701	20,67	-	-	-
	>35	53,66	39,14	21018	52,95	34,53	23014	9,49	10,82	6,04	-7,37
Centre i Est	35-49	19,51	18,72	6647	13,39	13,75	4278	-35,64	-	-	-
	50-69	60,60	60,22	22909	63,08	60,71	26488	15,62	-	-	-
	>70	89,95	88,32	16874	97,96	90,13	18747	11,1	-	-	-
	>35	51,23	43,52	46430	53,19	41,56	49513	6,64	2,81	3,34	0,49
Europa	35-49	17,85	17,74	14221	13,60	13,51	10288	-27,66	-	-	-
	50-69	59,46	58,74	50685	56,28	54,67	54226	6,99	-	-	-
	>70	117,93	110,27	57357	120,07	106,08	63469	10,66	-	-	-
	>35	57,25	44,74	122264	56,91	40,58	127983	4,68	5,28	6,20	-6,80

Taula 4.8: Mortalitat per CM segons regió i grups d'edat a Europa l'any 2005. Projectió de la mortalitat per CM per l'any 2015 i contribució dels factors demogràfics. Taxa crua per 100.000 dones (TC), taxa ajustada a la població mundial per 100.000 dones (TA) i nombre de casos (N).

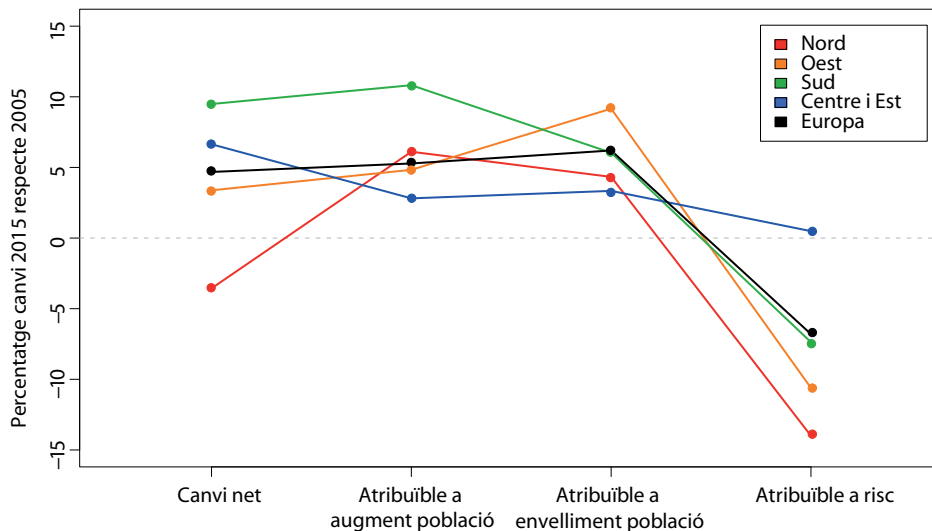


Figura 4.6: Percentatges de canvi observats en la mortalitat per CM a Europa l'any 2015 respecte l'any 2005, nets i separadament segons els factors de risc i demogràfics

En general, de l'anàlisi de la taula 4.8 i de la figura 4.6, es pot destacar que a Europa hi haurà un augment del 4,68%, en el nombre de defuncions per CM en dones majors de 35 anys (122.264 casos l'any 2005 respecte 127.983 casos l'any 2015). Tanmateix però, aquest increment de la mortalitat és degut en un 5,28% a l'augment previst de la població europea, i en un 6,20% al seu futur envelliment. En canvi, es pot constatar que hi haurà una disminució del nombre de morts per CM en un 6,8%, atribuïble a una disminució del risc de morir per aquesta causa. Així doncs, aquests resultats ens permeten afirmar que el risc de morir per CM disminuirà sensiblement tot i que, degut a qüestions demogràfiques hi haurà un augment molt important del nombre de morts. Pel que fa a les 4 regions d'Europa definides, podem observar com la disminució del nombre de defuncions per CM atribuïbles a una disminució del risc de morir serà marcadament superior als països nòrdics (14%), seguit dels països de l'Oest (10,61%) i dels països del Sud (7,37%). En canvi, pels països del Centre i Est d'Europa es preveu que el risc de morir continuarà

augmentant, tot i que molt lleugerament (0,49%). En els països del Nord d'Europa es preveu per l'any 2015 una disminució en termes absoluts del nombre de morts per CM (3,56%) respecte l'any 2005 tot i que augmentarà la població i l'envelliment. La disminució del nombre de defuncions per CM serà major en el grup d'edat més jove (23,15%). En els països de l'Oest d'Europa, qüestions demogràfiques semblen jugar un paper més rellevant i, en general, s'espera un augment del nombre de defuncions per CM per l'any 2015 d'un 3,38%. Malgrat això, per a les dones entre 35 i 49 anys, s'estima una reducció de la mortalitat del 25,21%, similar a la dels països nòrdics. Els països del Sud d'Europa tindran un augment important del nombre de defuncions per CM, del 9,49%, atribuïble al major augment demogràfic (10,82%), afegit a que la disminució del nombre de defuncions a causa que el risc és menor a la dels països de l'Oest i nòrdics. Finalment, pel que fa als països del Centre i Est d'Europa s'observa que el nombre de defuncions per CM augmentarà en un 6,64%, atribuïble a un creixement moderat de la població (2,81%), un envelliment també moderat (3,34%) i, alhora, a un augment del risc de morir per CM (0,49%).

Les taules 4.1, 4.2, 4.3, 4.4 i 4.5 mostren també els resultats obtinguts en avaluar els efectes dels factors demogràfics i epidemiològics, separadament per a cadascun dels 32 països. Per tal de facilitar la interpretació dels resultats, la figura 4.7 mostra el rànking de països segons la importància en l'augment del nombre de defuncions per CM (percentual) net i separadament segons la mida i estructura de la població o el risc de morir per CM. D'aquesta forma podem observar com Grècia serà el país europeu que experimentarà un augment percentual de les defuncions per CM major (44,39%), atribuïble sobretot a l'augment de defuncions en les dones de més de 70 anys (80,98%). Així mateix, Grècia es manté en les primeres posicions del rànking europeu pel que fa als augments percentuals del risc de morir, mida o estructura de la població. L'increment del nombre de defuncions per CM atribuïble a canvis en el risc de morir es constata en Letònia, Luxemburg, Bielorrússia, Federació Russa, Ucraïna, Polònia, Romania i Estònia, la majore-

ria de l'Europa de l'Est i Central. En canvi, els Països Baixos presenten una disminució menor en el nombre de defuncions per CM atribuïbles a canvis en el risc (23,59%) seguit de Noruega (22,51%), el Regne Unit (15,61%), Suïssa (17,02%), Àustria (14,55%), Alemanya (10,41%), Espanya (12,01%), Itàlia (10,32%) i Irlanda (13,16%), entre altres. Finalment, la figura 4.7 posa de manifest diferències importants a Europa pel que fa al nombre de defuncions per CM que hi hauran atribuïbles a creixements demogràfics i a l'envelliment de la població. En general, s'observa que hi haurà un augment important de la població a Irlanda, Luxemburg, Espanya, Grècia o Eslovàquia, que comportarà un augment del nombre de defuncions per CM. Aquest augment serà inferior en Alemanya, Lituània, Finlàndia, el Regne Unit i Suècia i, fins i tot, hi haurà un descens demogràfic en Ucraïna i Letònia el que incidirà en una disminució de la mortalitat per aquest tumor. Finalment, l'augment del nombre de defuncions per CM degut a l'envelliment de la població europea serà similar en els 32 països d'Europa oscil·lant entre el -0,13% de Irlanda (l'únic negatiu) i el 12,14% de Luxemburg. Altres països que presentaran un augment important del nombre de defuncions per CM degut a un envelliment de la població seran Grècia, Alemanya, Eslovènia, Suïssa, els Països Baixos, Finlàndia, Lituània i Àustria.

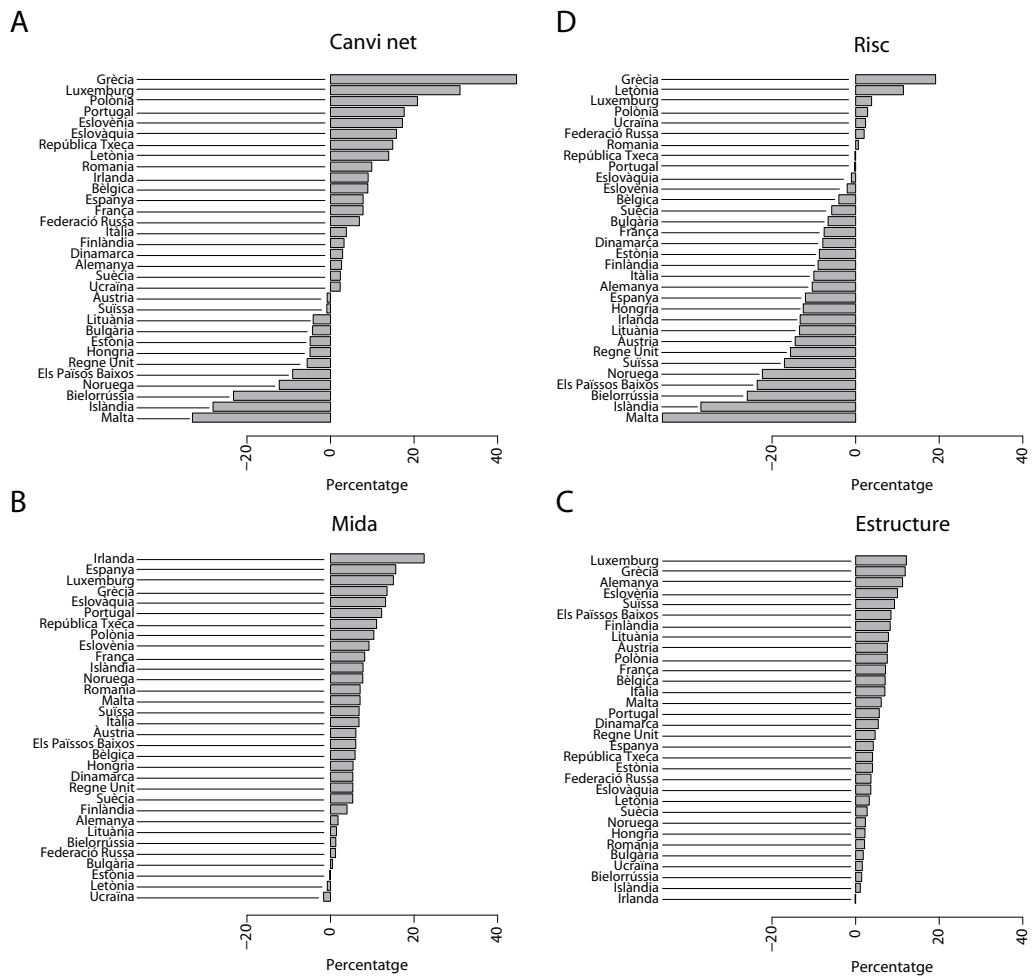


Figura 4.7: Rànking del canvi net (A) en el nombre de morts per càncer de mama de l'any 2015 respecte l'any 2005, en percentatges. Partició de la diferència neta en la part atribuïble a mida de la població (B), a l'estructura de la població (C) i part deguda a risc (D).

Capítol 5

Eines computacionals desenvolupades

Els resultats principals d'aquesta tesi han estat en part possibles gràcies al desenvolupament d'un conjunt d'eines que ens han permès executar els càlculs pertinents amb les dades, ajustar els models, avaluar la bondat de l'ajust i, finalment, obtindre la projecció de la mortalitat per CM. A més, també hem desenvolupat les eines computacionals necessàries per a produir les diferents visualitzacions dels resultats.

Totes les anàlisis presentades en aquesta tesi han estat programades mitjançant el paquet estadístic **R** [R, 2007], el que constitueix una part important de les contribucions en el conjunt del treball realitzat. En primer lloc, hem desenvolupat un conjunt de funcions, en codi **R**, per tal de calcular els APCs i ajustar els models per a obtindre les posteriors projeccions. Aquest conjunt de funcions, per a l'anàlisi de tendències i projeccions, es poden descarregar de la pàgina web <http://sites.google.com/site/joanvallstthesis> i són lliurement accessibles. A més, també es poden obtindre les dades analitzades en aquesta tesi amb el codi d'anàlisi emprat, fet que permet reproduir els resultats o bé modificar les anàlisis sobre les mateixes dades. En segon lloc, hem desenvolupat un conjunt de funcions per a implementar el mètode per a avaluar l'efecte dels factors demogràfics respecte els factors

epidemiològics associats al risc en les diferències observades en la mortalitat [Bashir and Esteve, 2000], que hem anomenat *RiskDiff*. Finalment, hem implementat una interfície web, més amigable, que permet, lliurement, realitzar les anàlisis proposades. Aquesta, es troba disponible a l'adreça <http://rht.iconcologia.net/riskdiff.htm>, on també es poden descarregar les funcions en codi R originals.

En aquest capítol descrivim els dos conjunts de funcions desenvolupats i, també, el funcionament general de la interfície web *RiskDiff*. Per il·lustrar l'ús de les funcions mostrem els resultats obtinguts amb l'anàlisi de la mortalitat per CM a Espanya. Finalment, fem referència a l'arxiu que conté el codi de les anàlisis realitzades i que també és accessible lliurement.

5.1 Eines per a l'anàlisi de tendències i projeccions de taxes d'incidència o mortalitat

5.1.1 Funcions implementades

El fitxer `CancerStats.R` conté el codi font de les dues funcions principals que hem emprat per al càlcul de l'APC i per a l'ajust de les dades de mortalitat i posterior projecció. El codi d'aquestes funcions, que hem anomenat `compute.apc()` i `project.mortality()`, es troba a l'annex A i també es pot descarregar a <http://sites.google.com/site/joanvallsthesis/>.

Ambdues funcions tenen com a paràmetres d'entrada el període d'anàlisi (any inicial i any final), el país seleccionat i els grups d'edats quinquennals seleccionats. Per a la segona funció, addicionalment, cal indicar l'any (o anys) de projecció així com el model escollit per model·lar les dades. Totes dues funcions retornen un objecte tipus llista, que conté els resultats estadístics. En l'apartat següent mostrem un exemple d'ús d'aquestes funcions.

5.1.2 Mortalitat a Espanya els anys 1996-2005 i projecció l'any 2015

Per tal d'il·lustrar l'ús de les funcions implementades, exposem a continuació els resultats obtinguts en l'anàlisi de la mortalitat per CM a Espanya en els darrers 10 anys disponibles, per a les dones majors de 35 anys. En primer lloc, per tal d'avaluar la tendència de la mortalitat per CM al llarg d'aquest període, podem emprar la funció `compute.apc()` indicant el període, el país i els grups d'edats seleccionats. La funció retorna un objecte tipus llista que conté les dades agregades per any, amb el nombre de morts i la taxa ajustada per edat i, finalment, l'APC estimada amb l'interval de confiança corresponent (95%), segons el model gaussià, i el p-valor associat que avalua l'hipòtesi que l'APC és nul·la. Mostrem a continuació la crida de la funció i els resultats obtinguts a Espanya, que ens permeten destacar com ha disminuït la taxa considerablement els darrers anys, a un nivell de decreixement anual del 2,26% (I.C. 1,9%- 2,6%), essent aquesta disminució molt significativa (p-valor=0,0000013).

```
> source("CancerStats.R")
> older.35 <- c("35-39", "40-44", "45-49", "50-54", "55-59", "60-64", "65-69",
               "70-74", "75-79", "80-84", "85+")
> compute.APC(start=1996, end=2005, country="spain", ages=older.35)
Waiting for profiling to be done...
$data
  year population cases AS.rate country
1 1996  10793200  5683 41.09494  spain
2 1997  10968400  5707 40.89694  spain
3 1998  11143600  5724 40.21559  spain
4 1999  11318800  5634 38.44766  spain
5 2000  11494000  5607 37.05481  spain
6 2001  11758400  5864 38.08383  spain
7 2002  12022800  5738 36.41593  spain
8 2003  12287200  5868 35.85783  spain
9 2004  12551600  5766 34.29083  spain
10 2005  12816000  5671 33.38408  spain

$ages
[1] "35-39" "40-44" "45-49" "50-54" "55-59"
     "60-64" "65-69" "70-74" "75-79" "80-84" "85+"

```

```

$results
country start end APC APC.LCI APC.UCI APC.pvalue
spain 1996 2005 -2.264793 -2.608947 -1.919987 1.379064e-06

```

En segon lloc, per tal d'obtindre la projecció de la mortalitat per CM a Espanya l'any 2015, plantejarem dos models possibles: MAP i MAAP, corresponents als dos models multiplicatius (descrits a la secció 4.3), ja que la tendència és decreixent i, per tant un model additiu (AAP) podria donar lloc a projeccions fins i tot negatives. Per a aquesta anàlisi, cal especificar a la funció `project.mortality()` el tipus de model que es vol emprar. La funció retorna un objecte amb estadístiques sobre el model ajustat i un segon objecte que conté, pròpiament, les dades projectades. Mostrem a continuació els resultats obtinguts a Espanya. Observem, primerament, que l'estadístic de bondat d'ajust de Pearson revela una clara falta d'ajust per al model MAP (p -valor = 0,0000025) el que ens indica que pot haver-hi efectes diferenciats segons els grups d'edat quinquennals. En canvi, el model MAAP presenta un bon ajust ($X^2 = 97,55$ amb p -valor associat 0,22) i que ens permetrà projectar la mortalitat. D'aquesta forma, obtenim el nombre de casos projectats per a cadascun dels grups d'edat quinquennals, que oscil·len entre les 75,99 morts per al grup de menor edat fins a les 1751,52 per les dones majors de 85 anys previst per l'any 2015. Convé remarcar que les projeccions obtingudes a partir dels dos models són substancialment diferents, el que recolza la importància de la tria del model.

```

> project.mortality(start=1996,end=2005,up.to=2015,country="spain"
,ages=older.35,type="map")
$model
  start end up.to country Pearson.chisq df.residual pvalue type
1 1996 2005 2015 spain 175.4890 98 2.545717e-06 map

$projected.data
  country year age.group population projected.cases
1 spain 2015 35-39 1886000 124.8080
2 spain 2015 40-44 1915000 243.5901
3 spain 2015 45-49 1819000 359.0589
4 spain 2015 50-54 1705000 470.1715
5 spain 2015 55-59 1518000 510.2012

```

```

6   spain 2015    60-64   1306000    519.6660
7   spain 2015    65-69   1210000    559.9347
8   spain 2015    70-74   1022000    553.1715
9   spain 2015    75-79    883000    586.3005
10  spain 2015    80-84    831000    701.9084
11  spain 2015     85+    892000    1225.0879

> project.mortality(start=1996,end=2005,up.to=2015,country="spain"
                    ,ages=older.35,type="maap")

$model
  start end up.to country Pearson.chisq df.residual  pvalue type
1 1996 2005 2015  spain      97.55946           88 0.2278769 maap

$projected.data
  country year age.group population projected.cases
1   spain 2015  35-39   1886000      75.99359
2   spain 2015  40-44   1915000     187.67418
3   spain 2015  45-49   1819000     328.88780
4   spain 2015  50-54   1705000     390.14607
5   spain 2015  55-59   1518000     515.29556
6   spain 2015  60-64   1306000     435.94567
7   spain 2015  65-69   1210000     557.19536
8   spain 2015  70-74   1022000     585.23160
9   spain 2015  75-79    883000     571.47515
10  spain 2015  80-84    831000     714.54343
11  spain 2015   85+    892000    1751.52931

```

5.2 Riskdiff: anàlisi de factors de risc i demogràfics en l'evolució de la incidència o mortalitat d'una malaltia

5.2.1 Funcions implementades

Hem desenvolupat les funcions `risk.diff()` i `plot.risk.diff()`, que permeten realitzar les anàlisis que avaluen el paper dels factors demogràfics i epidemiològics (el risc de morir o desenvolupar càncer) en dades de mortalitat o incidència [Bashir and Esteve, 2000]. El codi d'aquestes dues funcions es troba disponible en l'arxiu `riskdiff.R` que es pot descarregar des de la pàgina web <http://sites.google.com/site/joanvallstthesis/>, on també es

pot accedir a material complementari, tal com arxius de dades i exemples d'ús d'aquestes funcions. En l'article, ja publicat, sobre aquesta implementació [Valls et al., 2009] estan descrits àmpliament aquestes dues funcions, amb exemples il·lustratius. En l'annex A presentem el codi d'aquestes funcions i en l'annex B reproduïm l'article esmentat.

La primera funció té com a paràmetres d'entrada el nombre de casos i població a risc per a les dues poblacions, la de referència o basal respecte la de comparació, i retorna un objecte amb els resultats obtinguts. La segona funció proporciona una sortida gràfica i té com a paràmetre d'entrada alguns dels resultats obtinguts amb la primera la primera funció. En l'apartat següent mostrem un exemple d'ús d'aquestes funcions.

5.2.2 Comparació de la mortalitat a Espanya l'any 2015 respecte el 2005 i avaluació de factors demogràfics

Els objectes `data.2005` i `data.2015`, presentats a continuació, contenen el nombre de morts per càncer de mama que es van registrar a Espanya l'any 2005, per als grups d'edat quinquennals corresponents a dones majors de 35 anys, i el nombre de defuncions projectat per a l'any 2015 (emprant el model MAAP), descrits en l'apartat anterior.

```
> data.2005
  country year age.group population cases
1  spain 2005 35-39      1765000    128
2  spain 2005 40-44      1679000    254
3  spain 2005 45-49      1513000    345
4  spain 2005 50-54      1319000    399
5  spain 2005 55-59      1247000    495
6  spain 2005 60-64      1086000    519
7  spain 2005 65-69      1001000    505
8  spain 2005 70-74      1071000    723
9  spain 2005 75-79       912000    663
10 spain 2005 80-84       669000    694
11 spain 2005 85+         554000    946

> data.2015
  country year age.group population projected.cases
1  spain 2015 35-39      1886000          75.99359
2  spain 2015 40-44      1915000         187.67418
```

3	spain 2015	45-49	1819000	328.88780
4	spain 2015	50-54	1705000	390.14607
5	spain 2015	55-59	1518000	515.29556
6	spain 2015	60-64	1306000	435.94567
7	spain 2015	65-69	1210000	557.19536
8	spain 2015	70-74	1022000	585.23160
9	spain 2015	75-79	883000	571.47515
10	spain 2015	80-84	831000	714.54343
11	spain 2015	85+	892000	1751.52931

Mostrem a continuació els resultats obtinguts amb la funció `risk.diff()` per avaluar quina part de les diferències observades es atribuïble a risc de morir respecte de la que ho és a factors demogràfics.

```
> source("RiskDiff.R")
> results.risk.diff <- risk.diff(cases.init=data.2005$cases,
+                               cases.end=data.2015$projected.cases,
+                               pop.init=data.2005$population,
+                               pop.end=data.2015$population)
> results.risk.diff
[[1]]
      group  cases population crude.rate
1 baseline group 5671.000  12816000  44.24938
2 comparison group 6113.918  14987000  40.79481

[[2]]
      change.in crude.rate Perc.crude.rate  Number Perc.Number
1      Risk    -5.319997    -12.022763 -681.8109  -12.022763
2 Structure    1.865429     4.215718  239.0733    4.215718
3      Size           NA           NA  885.6553   15.617268
4 Net change  -3.454569    -7.807045  442.9177    7.810223

[[3]]
[1] "5671 total cases observed in baseline group and 6113.91772238813 in
comparison group, representing a crude rate of 44 and 41 cases per 100000
persons respectively."
[2] "The net change in terms of crude rate is -3.45 cases per 100000
persons (-7.81%), which can be partitioned into -5.32 cases (-12.02%) due
to a decrement of the risk and 1.87 cases (4.22%) due to changes in the
population structure."
[3] "The net change in terms of total number of cases is 442.92 cases
(7.81%), which can be partitioned into -681.81 cases (-12.02%) due to
a decrement of the risk, 239.07 (4.22%) due to changes in the population
structure and 885.66 cases (15.62%) due to an increment of the population
size."
```

```
> plot.risk.diff(results.risk.diff[[2]]$crude.rate  
                ,results.risk.diff[[2]]$Number)
```

Mostrem la sortida gràfica produïda per la funció `plot.risk.diff()` a la figura 5.1. Així doncs, podem constatar que el nombre de defuncions per CM a Espanya l'any 2005 va ser de 5.671, representant una taxa crua de mortalitat de 44,24 defuncions per 100.000 dones majors de 35 anys, mentre que per l'any 2015 hem projectat un nombre de 6.113,91 defuncions que representaria una taxa bruta de 40,79. La segona taula, analitza quina part de l'augment en el nombre de defuncions per CM (respecte el canvi net) és atribuïble a canvis en factors demogràfics (mida o estructura) i quina part és atribuïble a canvis en el risc de morir per aquest tumor. El nombre de defuncions augmentarà l'any 2015 en un 7,81% (442,91 casos) respecte l'any 2005. Tanmateix, es pot obtenir una partició d'aquest percentatge que indica que, de fet, el nombre de morts per càncer de mama disminuirà en un 12% (681,81 casos) degut a una disminució del risc de morir, mentre que, el nombre de defuncions augmentarà en 15,6% (885,65 casos) degut a un increment de la mida de la població espanyola i també en un 4,2% (239,07 casos) degut a un envelliment progressiu de la població. El gràfic mostrat a la figura 5.1 aporta una representació d'aquesta partició en termes absoluts (a partir del nombre de casos) i en termes relatius (a partir de les unitats de la taxa de mortalitat). Remarquem, finalment, que la utilitat d'aquest mètode és identificar aquella malaltia en la que la variació en el nombre de defuncions o casos és atribuïble a canvis en el risc de morir respecte a canvis demogràfics. Això comporta que en aquestes malalties s'ha d'intervenir amb la finalitat de reduir-ne el risc, bé millorant l'accés sanitari, o altres aspectes que hi estiguin relacionats.

5.2.3 Interfície web *RiskDiff*

Les anàlisis estadístiques exposades en l'apartat anterior poden igualment ser dutes a terme mitjançant la interfície web, que hem anomenat *RiskDiff*, i que es troba disponible a <http://rht.iconcologia.net/riskdiff.htm>.

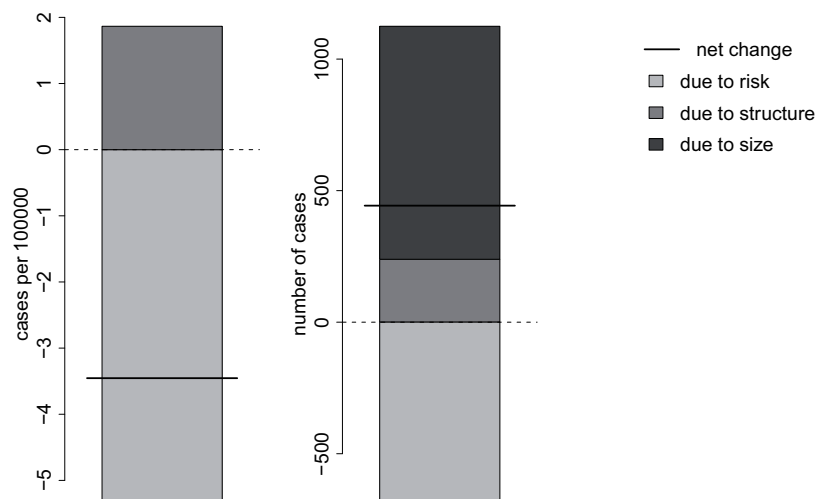


Figura 5.1: Representació de la part atribuïble a risc respecte a la part atribuïble a factors demogràfics (mida i població) en l'augment de la mortalitat per càncer de mama a Espanya l'any 2015 respecte l'any 2005

La implementació d'aquesta eina s'ha realitzat mitjançant el llenguatge de programació PHP i, de fet, s'utilitza un servidor en Linux remot per tal d'executar les funcions en R descrites prèviament. La figura 5.2 mostra l'aparença general d'aquesta eina web. Bàsicament, les dades poden introduir-se de dues maneres, bé emprant un fitxer de text o bé introduint directament les dades. En qualsevol cas, cal indicar el format del fitxer (pel que fa a separadors i si conté o no una capçalera) i, en una segona finestra indicar les etiquetes dels dos grups (referència i comparació). Un cop superada aquesta fase, *RiskDiff* retorna una pàgina web amb les dades originals, els resultats de l'anàlisi i també una representació gràfica i, addicionalment, s'ofereix l'opció de descarregar els resultats sobre un fitxer de tipus *pdf*.



RiskDiff

A web tool for the analysis of the difference due to risk and demographic factors for incidence or mortality data

This tool can be useful to study the differences in the incidence or mortality observed in two given situations (such as time points, geographical areas, or males versus females). The method performed splits the observed differences into three components: (1) the risk itself, (2) difference attributed to changes in the population size and (3) difference attributed to changes in population structure. For its use, you may provide the number of cases (or deaths), for separate age-groups (e.g. 5 years) and the respective reference population for both groups.

Upload a Data Set into RiskDiff

A dataset can be uploaded into RiskDiff in two ways: (1) *using a tab-separated text file with 4 columns*, one for each vector or (2) *directly typing the data into the web interface separately for each vector*.

The dataset should have 2 columns with the number of incident/death cases (one column for the reference group and another for the comparison one) as well as 2 columns for the person-years at risk for each age group considered. Once the dataset has been uploaded, RiskDiff will ask for the group labels to identify each column. Finally, RiskDiff produces a web page with summary tables, graphical representations and a short paragraph of text to facilitate the interpretation of the results. The RiskDiff user can find in this web page two example datasets with the required format. The source of data from which these example datasets have been created can be also found in this web page.

* We recommend to be at least 2 age groups to perform RiskDiff, otherwise there is no evolution of the age structure and changes can be only attributed to population size and risk.

Data files:

[Example dataset 1:Mortality.cancer.Catalonia.men.1985.2004.txt](#)
[Example dataset 2:Mortality.cancer.Catalonia.women.1985.2004.txt](#)

Source data of the example datasets (not to be run with RiskDiff): [CancerMortalityCataloniaData.txt](#)

Using data from [Example dataset 1](#), let us assume that we want to compare the evolution of mortality in Catalonia comparing year 2004 with 1985. Assume that reference group could be year 1985 data whereas the comparison one is year 2004 data.
 In this dataset, columns 1 and 2 are the Baseline and Comparison cases respectively, whereas columns 3 and 4 are the Baseline and Comparison populations at risk. You can copy this example and paste it into the box below in order to see how RiskDiff works.

Files to download to perform the analysis with R

[Send us an e-mail to get functions source code](#)

Source R file containing some examples using R functions: [ExampleRiskDiff.R](#)

WARNING: THIS PROCESS MAY TAKE SOME TIME

Enter the text file with your data

or enter data here

Read Labels from file header

Finally, select delimiter between columns:

©2003-2009 Institut Català d'Oncologia. Tots els drets reservats.

Figura 5.2: Interfície web RiskDiff.

5.3 Codi de les anàlisis presentades en el capítol 4

En la pàgina web <http://sites.google.com/site/joanvallsthesis/> es pot descarregar un arxiu comprimit, anomenat `vallsthesis.zip` que conté, a més dels arxius de dades necessaris, un arxiu amb el codi emprat per a obtenir els resultats presentats en aquesta tesi. Les anàlisis que defineix estan estructurades de forma que es pot seguir ordenadament els resultats obtinguts i també s'hi han afegit comentaris per a la seva millor comprensió. D'aquesta forma, es poden replicar les anàlisis presentades en aquest document i facilitar, per exemple, la seva discussió crítica o també l'ús del codi per a altres conjunts de dades.

Capítol 6

Discussió

6.1 Metodologia estadística

El treball realitzat en aquesta tesi aporta una estratègia global per a la modelització de les taxes de mortalitat o d'incidència d'una malaltia determinada, obtenir la seva projecció en un temps futur i, finalment, avaluar el paper dels factors demogràfics i del risc de morir o desenvolupar la malaltia, en les diferències detectades entre la mortalitat o incidència observada respecte de la projectada. Hem aplicat aquesta estratègia d'anàlisi a les dades de mortalitat per càncer de mama observades per a 32 països d'Europa en els darrers 10 anys disponibles i, la metodologia emprada, ens ha permès aportar una estimació de la mortalitat que hi haurà a Europa per aquesta malaltia l'any 2015. Els resultats obtinguts amb aquest conjunt de dades ens permeten discutir la seva aplicabilitat des del punt de vista empíric i metodològic, i per tant avaluar l'objectiu estadístic principal, perseguit per aquesta tesi, en un context heterogeni.

El marc de modelització estadística proposat en aquesta tesi, assumeix que les dades històriques de mortalitat es poden modelar eficientment a partir de tres models edat-període que hem anomenat MAAP, MAP i AAAP. Els models proposats assumeixen que el nombre de defuncions es distribueix segons una llei de Poisson i que la població a risc està fixada. La diferen-

cia entre aquests models fa referència a si, aquest és multiplicatiu (MAAP i MAP) o si és additiu (AAAP). Els primers models assumeixen un creixement (o decreixement) exponencial, mentre que el segon assumeix un creixement (o decreixement) constant. En aquest punt, cal remarcar que el nostre objectiu persegueix la cerca d'un model que ajusti raonablement bé les dades en el període observat però, que alhora, permeti obtenir una predicció en el futur, és a dir fora del rang observat en la variable explicativa temps. Així doncs, tenint en compte aquesta perspectiva, els models multiplicatius poden tenir sentit, sempre i quan la tendència de les taxes no sigui creixent, doncs podria donar lloc a projeccions futures extremadament elevades. Per l'altre costat, el model additiu podria no tenir sentit quan les taxes de mortalitat històriques tenen una tendència decreixent, ja que podria fins i tot conduir a una projecció negativa. Aquesta, constitueix doncs la primera consideració que cal fer alhora de modelar unes dades d'incidència o mortalitat per a obtenir una projecció en el futur. Tanmateix, un altre aspecte ineludible que emergeix quan s'analitzen aquest tipus de dades històriques és el període de temps analitzat. És evident que l'evolució de la incidència o mortalitat d'una malaltia, a la realitat, és un fet molt variable i molt sovint es poden donar canvis de tendència. En el nostre treball, hem considerat oportú treballar amb els 10 darrers anys disponibles per cadascun dels països. Aquest aspecte, tot i que en part subjectiu, ens ha permès avaluar les dades més recents assumint que, a priori, 10 anys són suficients per avaluar la tendència *recent* de l'evolució de la mortalitat per càncer de mama en un país.

En primer lloc, per tal d'avaluar si la tendència de la mortalitat per CM dels països és creixent, decreixent o bé estabilitzada, hem emprat un model alternatiu que assumeix que les taxes de mortalitat es distribueixen segons una llei Normal, que ens ha permès estimar el canvi percentual anual (APC) i fer-ne inferència, conjuntament amb l'anàlisi descriptiva, mitjançant visualitzacions de l'evolució de la mortalitat per cada país. Aquest primer pas ens ha permès descartar models per a països concrets. En segon lloc, un cop descartats alguns models per alguns països, hem escollit aquell model que

ofereix millor ajust segons un criteri de bondat d'ajust basat en l'estadístic de Pearson. Per avaluar aquest estadístic, hem introduït també el principi de parsimònia. D'aquesta forma, en els casos que el model ajusta raonablement bé tant pel model MAAP com pel model MAP, hem optat per aquest segon, ja que assumeix la mateixa tendència per a tots els grups d'edat quinquennals i, per tant, el nombre de paràmetres és menor. D'altra banda, en el cas d'obtenir un ajust raonable tant pel model MAAP com AAAP, específicament pels països amb una tendència de mortalitat estable, hem optat per avaluar l'estimació puntual de l'APC, independentment de la seva significació estadística, i optar pel model MAAP en cas de APC negativa i pel model MAP en cas contrari. Finalment, en aquells casos on cap dels tres models presentava un ajust raonable, hem optat per reduir el període de temps ajustat, doncs principalment cal assumir que es tracta de països en els que la tendència de mortalitat ha sofert canvis abruptes darrerament.

Els resultats obtinguts permeten, en primer lloc, constatar que el marc de modelització proposat ha proporcionat uns resultats acceptables. D'aquesta forma, és destacable que per a tots els països, amb l'excepció de Lituània, s'ha pogut obtenir un model amb una bona qualitat d'ajust. En la majoria de països (21) aquest model emprà el període de 10 anys complet, el que suggereix que la seva tendència és consistent en el període més recent. Per als països restants (11) el període s'ha reduït de forma variable, des dels nou fins als darrers tres anys. A Lituània no ha estat possible trobar un model acceptable, atribuint-se en part a la baixa qualitat de les dades, juntament amb una mida de la població baixa, que podria haver provocat una major variabilitat de les taxes. En segon lloc, la modelització ha permès l'estimació puntual del nombre de morts que hi hauran un un moment concret del futur. L'avaluació, estrictament visual, de les taxes projectades, juntament amb les taxes observades, semblen indicar que aquesta projecció ofereix un escenari, a deu anys vista, força plausible donades les dades històriques observades.

Per un altre costat, l'aplicació del mètode que avalua quina part dels canvis en la mortalitat projectada l'any 2015 respecte l'observada l'any 2005,

és deguda a factors de risc de la que és deguda a canvis en factors demogràfics (augment i envelliment de la població) ha esdevingut un complement adequat i necessari per a la correcta interpretació de les dades. Els resultats obtinguts als països d'Europa ens han permès també constatar la idoneïtat del seu ús, degut a que l'augment de la població i el seu envelliment són aspectes importants en molts països d'Europa i, a més, també ens ha permès copsar una certa variabilitat de patrons, pel que fa aquests aspectes.

El treball presentat en aquesta tesi presenta un conjunt de limitacions important, que detallem a continuació. En primer lloc, pel que fa als models proposats, en el nostre treball hem obviat una família de models que s'utilitza també sovint en aquest context: els models edat-període-cohort. L'avantatge principal d'aquest tipus de models és que poden ajustar un efecte diferenciat segons la cohort de naixement de la població i, per tant, ajustar diferents efectes que podrien adquirir sentit a la realitat, doncs pretenen introduir canvis generacionals en una població, com podria ser hàbits de tabaquisme o, per exemple, en el cas del càncer de mama, les polítiques de cribatge. Tanmateix, un dels inconvenients d'aquests models és l'augment de la complexitat en la seva especificació, pel que fa al nombre de paràmetres, així com la seva interpretació, ja que hi ha una confusió entre les tres components: edat, període i cohort que, a la pràctica, pot afectar l'explicabilitat del model. Per una altra banda, en el nostre cas, el fet que el període de temps analitzat sigui de 10 anys, abona l'assumpció en favor dels models edat-període, ja que els models edat-període-cohort s'utilitzen habitualment quan es vol descriure la tendència d'una malaltia al llarg d'un període llarg de temps i que, per tant, puguin emergir efectes destacable atribuïbles a cohorts específiques. Cal doncs remarcar que el treball realitzat en aquesta tesi es centra en la modelització de dades de mortalitat, en un període curt, amb l'objectiu immediat d'obtenir una posterior projecció i, per tant, simplement es desitja que aquesta projecció estigui en concordança amb la tendència més recent observada per un país. En segon lloc, hem emprat en les nostres anàlisis un esquema de màxima versemblança per ajustar els models, utilitzant la teoria dels models

lineals generalitzats. Aquest punt també presenta unes limitacions importants, doncs haguéssim pogut plantejar anàlisis bayesianes o també models no paramètrics, mitjançant la introducció de mètodes de remostreig tipus *bootstrap*. Tanmateix, els models emprats han demostrat oferir un bon ajust de les dades, fet que és sense dubte en part atribuïble al període curt per a el que s'ha observat la mortalitat. Finalment, una tercera limitació important ha estat que no hem aportat cap estimació per a la variabilitat de la projecció, per exemple, en forma d'un interval de confiança. Com hem indicat, aquesta qüestió està resolta per als models edat-període [Hakulinen and Dyba, 1994], tot i que no està implementada en cap programari. Pel que fa a aquest punt, podem realitzar dues remarques. En primer lloc, les dades que analitzem són d'abast poblacional i per tant, censal. D'aquesta forma, no es tracta doncs estrictament d'una mostra aleatòria simple d'una població i per tant, podem considerar que una estimació puntual podria ser ja, per si mateixa, satisfactòria a la pràctica, en termes de salut pública. En segon lloc, tot i que creiem que, des del punt de vista estadístic, és molt important disposar d'aquesta estimació per interval, hem optat per excloure aquest treball d'aquesta tesi i creiem que pot constituir una línia futura de treball. Més concretament, en un futur creiem molt interessant programar en `R` l'estimació per interval de les projeccions segons el mètode proposat per Dyba i Hakulinen, fet que ens permetria avaluar la precisió d'aquestes. També estem interessats en desenvolupar un mètode tipus *bootstrap* que, mitjançant remostreigs de les dades, permeti obtenir, de forma no paramètrica, un interval de confiança per a la projecció. Pel que fa a aquest darrer punt, disposem ja en l'actualitat d'alguns resultats preliminars tot i que hem optat per no presentar-los en aquesta tesi.

Pel que fa a les funcions que hem programat en `R`, aquestes ens han permès realitzar les anàlisis proposades de manera eficient, sense dubte gràcies a les prestacions i versatilitat del programa `R`. A més, el fet de programar les funcions en aquest llenguatge ha aportat dos aspectes remarcables. En primer lloc, els resultats obtinguts en aquesta tesis són fàcilment reproduïbles

per qualsevol investigador amb coneixements de R, el que creiem que facilita la seva discussió crítica. En segon lloc, les funcions són lliurement accessibles per a qualsevol investigador amb problemes de recerca similars i, a més el codi d'aquestes funcions és també obert. En el cas de l'eina *RiskDiff*, la implementació en una pàgina web amplia el context d'ús pel que fa a potencials usuaris. Tanmateix, creiem que hagués estat molt positiu haver realitzat una llibreria amb R que contingués les funcions aportades. Aquesta és una limitació del treball presentat, tot i que estem treballant en l'actualitat en la creació d'aquesta llibreria i esperem que formi part d'una contribució futura arran del treball presentat en aquesta tesi.

6.2 Anàlisis epidemiològiques

Els resultats presentats en aquesta tesi aporten una visió exhaustiva i actualitzada respecte la situació de la mortalitat del càncer de mama a Europa i la seva projecció a l'any 2015. En el treball exposat, s'han confirmat alguns dels aspectes epidemiològics ja coneguts respecte a la mortalitat per càncer de mama, però també s'han posat a la llum alguns resultats que, fins al nostre coneixement, són nous. El fet d'avaluar com els canvis en factors demogràfics (augment i envelliment de la població) i en el risc de morir afecten al futur nombre de morts per càncer de mama a Europa l'any 2015 i, en general, l'abordatge de les anàlisis, aporta una nova mirada sobre aquesta patologia i ofereix, en conjunt, uns resultats originals que passarem a discutir a continuació.

En primer lloc, pel que fa a la interpretació epidemiològica dels resultats obtinguts, resulta convenient distingir entre la mortalitat en termes absoluts (i.e. països que presenten majors o menors taxes) respecte de la tendència de la mortalitat (i.e. creixent, decreixent, estabilitzada o bé variable segons el grup d'edat). Així doncs, respecte a la magnitud de la taxa de mortalitat, els nostres resultats han permès constatar la gran variabilitat en la mortalitat per CM a Europa, tant per a l'any 2000 com per a l'any 2005, oscil·lant

entre les 35 i 85 morts per cada 100.000 dones majors de 35 anys. No hem detectat un patró geogràfic clar que pugui explicar, de forma substancial, aquesta heterogeneïtat. En aquest sentit, països com Dinamarca o Hongria, per exemple, presenten taxes molt elevades, mentre que països com Espanya i Finlàndia mostren taxes menors. Respecte a la tendència de la mortalitat estimada, mitjançant el canvi anual percentual (APC) observem que si que sembla existir, en aquest cas, un cert patró geogràfic. En línies generals, en els països nòrdics s'observa una tendència de la mortalitat decreixent, al voltant del 2%, accentuada en els grups de dones més joves però, en canvi, aquest descens és més esmorteït en els grups d'edat més avançats. Els països de l'Europa de l'Oest mostren una tendència similar a l'anterior tot i que, en general, el decreixement és de major grau, especialment en el grup de dones més joves (oscil·lant entre el 2% i el 6%) i en les dones de 50-64 anys el descens és moderat. Els països del Sud d'Europa presenten un patró similar als països del nord-oest d'Europa, disminuint la mortalitat en tots els grups d'edat, encara que amb un grau de decreixement menor. Finalment, els països de l'Est i Centre d'Europa, presenten un patró marcadament diferenciat a la resta de països. Malgrat observar-se, en general, un cert decreixement global de les taxes de mortalitat, també es pot constatar com les dones de més de 50 anys presenten una estabilització de la mortalitat per CM, fins i tot evidenciant-se per alguns països un creixement significatiu de la mortalitat en les dones majors de 70 anys.

Una limitació d'aquest estudi emergeix del fet que no hem emprat dades d'incidència del CM, que ens permetria predir millor si en un futur els casos nous d'aquesta malaltia augmentarien o no. D'altra banda, tampoc hem inclòs en l'estudi l'heterogeneïtat existent en la implantació de sistemes de cribatge de CM a Europa. En qualsevol cas, el treball presentat obre noves possibilitats per a determinar el paper d'alguns aspectes com l'implementació de cribatges de mama, que podrien explicar parcialment diferents patrons de mortalitat observats a Europa. Tanmateix, el decreixement general de les taxes de mortalitat que hem pogut constatar en els nostres resultats està en

la línia de les tendències prèviament publicades (un exemple recent es troba a [Hery et al., 2008b]). De fet, es coneix que el decreixement de mortalitat per càncer de mama ocorre en paral·lel a la implantació de sistemes de cribatge de CM mitjançant la mamografia en la majoria de països des dels inicis dels anys 1990 i també degut a un millor accés a tractaments efectius [Hery et al., 2008a,b, Hermon and Beral, 1996, Botha et al., 2003, IARC, 2002b].

Pel que fa a les projeccions de la taxa de mortalitat presentades en aquesta tesi, ens han permès estimar un escenari versemblant per a la situació de la mortalitat per càncer de mama a Europa a l'any 2015. Així doncs, hem pogut estimar que l'any 2015 hi hauran a Europa 127.983 morts degudes a aquesta patologia en dones majors de 35 anys, que es reparteixen en 10.288 defuncions (8,04 %) per a dones entre 35 i 49 anys, 54.226 (42,37%) per a dones entre 50 i 69 anys i 63.469 (49,59%), per a dones de més de 70 anys. És remarcable que el major nombre de defuncions es donarà a l'Europa del Centre i de l'Est (38,69%), seguit per l'Europa de l'Oest (30,19%) i, en una magnitud menor, Europa del Sud i del Nord (17,98% i 13,14% respectivament). La taxa bruta de mortalitat per CM a Europa serà l'any 2015 de 56,91 defuncions per cada 100.000 dones majors de 35 anys, observant-se marcades diferències segons els grups d'edat, essent aquesta taxa de 13,60, 56,28 i 120,07 per als grups 35-49, 50-69 i majors de 70 anys respectivament.

Una de les originalitats del treball presentat en aquesta tesi ha estat la comparació del nombre de morts per càncer de mama l'any 2005 amb el nombre de defuncions projectat l'any 2015, tenint en compte els canvis en els factors demogràfics de la població. Així doncs, hem estimat que a Europa a l'any 2015 hi haurà un augment del nombre de defuncions respecte l'any 2005 del 4,68%. Tanmateix, hi haurà un decrement del 6,8% en el nombre de morts per càncer de mama degut a una reducció del risc de morir. Així doncs, els resultats ens permeten remarcar que el fet que el nombre de casos augmenti és degut principalment a un augment de la població europea (que contribuirà en un augment del 5,28%, en nombre casos) i un envelliment d'aquesta (que contribuirà en un augment del 6,20%, en nombre de casos). D'aquesta forma,

l'augment del nombre de morts en un 4,68% equival a la suma de les tres components (5,28%+ 6,20% -6,8%) . La descomposició de l'augment del nombre de defuncions per CM en l'augment atribuïble a factors demogràfics, ens permet quantificar la part de variació atribuïble al risc de morir per la malaltia. Un risc negatiu, ens suggereix que en un país determinat, l'accés sanitari és adequat i per tan el cribatge i el tractament és adequat. En canvi, una variació positiva en el risc de morir per CM ens suggereix que l'accés sanitari podria no ésser el més adequat. Per tant, en planificació sanitària, el fet de detectar països amb risc de morir positius hauria de ser motiu de revisió i, en cas de confirmar-se l'augment, hauria de motivar millores a l'accés sanitari. En el nostre estudi observem que, en general, la disminució deguda a risc serà molt major en els països nòrdics, seguida pels països de l'Oest i del Sud (14%, 10,61% i 7,37% respectivament). En canvi pels països de l'Est es preveu que el risc augmenti lleugerament, en un 0,49%. Pel que fa als resultats obtinguts individualment per cadascun dels 32 països, és molt destacable el cas de Grècia on estimem el major augment de defuncions degut a risc (en un 44,39%), influït principalment per l'augment que experimentarà el nombre de defuncions en les dones majors de 70 anys (80,98%). S'observen també augments rellevants en un bon nombre de països de l'Europa de l'Est i Central, específicament per a les dones d'edat més avançada. Per l'altre costat, disminucions importants del risc de morir s'han constatat en països com els Països Baixos, Noruega, el Regne Unit, Suïssa, Àustria, Alemanya, Espanya, Itàlia o Irlanda. D'altra banda, l'augment de la població en els països d'Europa és força heterogeni i provocarà augments remarcables en el nombre de morts en països com Irlanda, Luxemburg, Espanya, Grècia o Eslovàquia i, aquest augment serà inferior en països com Alemanya, Lituània, Finlàndia, el Regne Unit i Suècia, observant-se fins i tot disminucions en el cas de Ucraïna i Letònia. En canvi, pel que fa a l'envelliment, aquest serà més homogeni i provocarà també un augment en el nombre de defuncions arreu d'Europa, essent el màxim augment del 12,14% en Luxemburg.

La implantació de sistemes de cribatge de càncer de mama té un efecte

sobre la mortalitat, de tal forma que s'ha constatat una disminució de la mortalitat deguda a la possibilitat de detectar el càncer de mama en un estadi inicial. S'ha pogut estimar que la reducció de la mortalitat per CM atribuïble al cribatge és del 16% en general per a totes les edats, del 15% per a dones entre 40 i 49 anys, encara que en aquest darrer existeix un cert dubte sobre la seva efectivitat [IARC, 2002b], i del 22% per a les dones majors de 50 anys. Tanmateix, es requereix un cert lapse de temps entre la implementació d'un programa de cribatge i la detecció del seu efecte sobre la mortalitat, estimant-se un període mínim de 10 anys per a detectar una disminució de la mortalitat [Jatoi and Miller, 2003]. Per altra banda, el nostre estudi ha posat de manifest que l'augment en el nombre de defuncions es pot atribuir principalment a l'augment de mida de la població i al seu envelliment, així com també per l'augment del risc de morir en els països de l'Europa Central i de l'Est i Grècia. En canvi, per la resta de països el risc de morir disminueix, essent atribuïble a la implementació del cribatge de mama a partir de la dècada de 1990 [Shapiro et al., 1998, Svendsen et al., 2006, Tornberg et al., 2006, Schulz-Wendtland et al., 2007, Kovacs et al., 2007, Parkin and Fernandez, 2006, Payne, 1999, Autier et al., 2002, Giordano et al., 2007, Ernst et al., 2007, Hofvind et al., 2007]. Així mateix, en els anys 1990 es van implementar millores en els tractament (principalment per tamoxifè) produint també una reducció en la mortalitat per aquest tumor [Hery et al., 2008a,b, Botha et al., 2003, Cleries et al., 2006, Otten et al., 2008]. En els països de l'Europa Central i de l'Est, només a Hongria es va implementar un sistema de cribatge en els anys 1990 [Kovacs et al., 2007], mostrant els nostres resultats un decreixement significatiu de la mortalitat per aquest país.

En diversos països, la disminució de la mortalitat per càncer de mama també s'ha associat al paper del cribatge oportunístic lligat a una major consciència entre la població femenina respecte aquesta malaltia [Vutuc et al., 2006, Hertl et al., 2006]. A més de l'efecte del cribatge, la mortalitat per càncer de mama reflecteix l'efecte de l'activitat diagnòstica i del tractament sobre el període de temps transcorregut entre els símptomes i el tractament

[Berry et al., 2005, Jatoi and Miller, 2003]. Així doncs, un augment de la mortalitat per càncer de mama podria estar associat a un augment de la incidència de la malaltia, conjuntament amb una manca de consciència sobre el càncer de mama, i de limitacions en l'accés a teràpies eficients [Schopper and de Wolf, 2009].

Cal remarcar que a Grècia només un 53,8% de les dones entre 50 i 69 anys van rebre serveis de criatge de càncer de mama, percentatge inferior que el recomanat pel consell d'Europa [Dimitrakaki et al., 2009]. D'aquesta forma, la manca de consciència entre la població grega i l'accés limitat a cirurgia i tractaments de càncer [Petrakis and Paraskakis, 2009] podrien estar associats al creixement en la mortalitat per aquest tumor observat en el nostre estudi, específicament per al grup de dones de major edat. Així doncs, aquest patró hauria de ser estudiat en detall per a definir els factors socio-sanitaris implicats i implementar millores.

En el cas d'Hongria i la República Txeca, s'han constatat decreixements significatius de la mortalitat, específicament en les dones entre 35 i 69 anys. En aquests dos països, aquest descens ja s'havia prèviament descrit en els anys 1990 [Hery et al., 2008a,b, Tyczynski et al., 2004]. Els nostres resultats suggereixen que les dones entre 50 i 69 anys, en el conjunt de països format per Bielorrússia, Bulgària, Estònia, Letònia, Polònia, Federació Russa i Lituània, tindrien la major contribució en la càrrega de la mortalitat per càncer de mama l'any 2015, el que es podria explicar per la poca implantació dels sistemes de criatge de CM [Szynglarewicz et al., 2009, Tekkel et al., 2007, Boguseviciute et al., 2003]. És remarcable que les nostres prediccions per a l'any 2015 situïn a Letònia, Ucraïna i la Federació Russa en les primeres posicions pel que fa a la mortalitat per CM, suggerint que aquesta tendència pot estar associada a un augment de la incidència del càncer de mama [Hery et al., 2008a] i a les dificultats en l'accés a les proves de criatge o tractaments eficients [Tyczynski et al., 2004]. Per una altra banda, la millora en la supervivència relativa a 5 anys del CM, observada per alguns països del centre-est europeu [Berrino et al., 2007] podria explicar les diferències entre els països

d'aquesta regió, malgrat que la manca de registres de càncer no fa possible una millor avaluació de la heterogeneïtat. També cal tenir present que, tal i com s'ha suggerit [Hery et al., 2008a, Tyczynski et al., 2004], la transició que han experimentat aquests països, d'una economia de mercat centralitzada a sistemes més capitalistes, podria explicar diferències en alguns factors de risc associats al càncer de mama, com són els factors reproductius i els canvis en els estils de vida.

Els resultats presentats en aquest treball parteixen de l'anàlisi d'un conjunt ben definit, però fitat, de dades, que no estan exemptes de limitacions, que exposem a continuació.

1. Una primera limitació emergeix del fet d'emprar un període de 10 anys (o menys per alguns països) per a obtenir una projecció al cap de 10 anys, aproximadament. Aquest és un punt discutible i, òbviament, altres esquemes aportarien resultats que no serien, com a mínim, exactament iguals. Tanmateix, existeix a la pràctica un cert consens en emprar les dades més recents per a projectar la incidència o mortalitat d'una malaltia i, sembla que un període de 10 anys pot recollir la tendència més recent [Dyba and Hakulinen, 2008].
2. Pel que fa a les limitacions de les pròpies dades també cal remarcar que, en el cas de Bèlgica, les dades més recents foren les de l'any 1997, i per tant, vam haver d'estimar la mortalitat de l'any 2000 i 2005, a més de la projecció per l'any 2015. En el cas de Bulgària, Itàlia, Portugal, Bielorrússia i Dinamarca, també vam haver d'estimar les dades per l'any 2005, tot i que el període analitzat es més recent que en el cas crític de Bèlgica.
3. En un altre ordre de qüestions cal situar el cas de països com Islàndia, Malta i Luxemburg, en els que la mida petita de la població constitueix una limitació important a l'hora d'analitzar els resultats, ja que canvis en relativament pocs nombres de casos podrien donar lloc a canvis de tendència espúries.

4. Els resultats discutits en aquesta tesi depenen directament de l'estimació de la població futura dels països europeus analitzats. Aquestes estimacions depenen de factors com les taxes de natalitat i migració, entre altres [Bray and Moller, 2006]. D'aquesta forma, canvis en aquestes prediccions podrien afectar a l'evolució de la càrrega de la mortalitat per càncer de mama en termes de la mida de la població, envelliment d'aquesta i el risc de morir per la pròpia malaltia [Bashir and Esteve, 2000].
5. Les dades que hem analitzat estan subjectes a la qualitat dels certificats de defunció. És conegut que aquest és un aspecte complex que pot generar en alguns casos una infra-estimació del nombre de defuncions i per tant, introduir una nova font d'heterogeneïtat entre països, que el nostre estudi no ha tingut en compte.
6. Finalment, no hem inclòs en l'anàlisi les variacions en la implantació de sistemes de cribatge a Europa així com les dades d'incidència en els països analitzats, fet que podria explicar una bona part dels resultats obtinguts, tal i com hem discutit prèviament.

Tanmateix també volem remarcar que el nostre estudi presenta un conjunt de punts forts, que exposem tot seguit.

1. L'estudi aporta una imatge actual i exhaustiva de la situació de la mortalitat per CM a Europa.
2. Presentem un estudi de les tendències recents i avaluem posteriorment possibles paral·lelismes amb la implementació de sistemes de cribatge i tractament, així com pel que fa a la incidència del CM.
3. Les projeccions obtingudes encenen una llum d'alerta en alguns països on es preveuen augments del nombre de defuncions deguts a un augment del risc de morir per CM. Per tant, els avaluadors sanitaris haurien de considerar aquests resultats per confirmar-los, en primer lloc, i per millorar la situació, en segon lloc.

Finalment, en termes més propis de salut pública, és prou conegut que l'èxit del desenvolupament dels serveis de salut pública involucrats en la detecció, diagnòstic i tractament del CM està intrínsecament lligat a la disminució del risc de mortalitat per CM en la majoria de països. Així doncs, la implementació d'un programa de cribatge de CM en les àrees on aquest no hi és present, i el conseqüent augment en les taxes de participació de la població així com una promoció de la conscienciació envers aquesta patologia, resulta cabdal en les estratègies a seguir en els països on no hem previst un augment important de la mortalitat per CM l'any 2015. També cal ser conscients que, per a que aquesta projecció no sigui assolida cal que els països involucrats es dotin del suport financer necessari.

Capítol 7

Conclusions

7.1 Metodologia estadística

1. Hem desenvolupat un marc de modelització estadística que permet analitzar eficientment les dades de mortalitat o incidència d'una malaltia determinada durant un període de temps, a partir de models edat-període, i projectar la mortalitat o incidència en el futur.
2. La estratègia de modelització proposada ha donat lloc a uns resultats raonablement acceptables per analitzar la mortalitat per càncer de mama a 32 països d'Europa els darrers 10 anys disponibles i obtindre una projecció de la mortalitat per a l'any 2015.
3. Hem desenvolupat un conjunt de funcions en R, lliurement accessibles, que permeten reproduir els resultats obtinguts i, alhora, faciliten l'anàlisi de nous conjunts de dades.
4. Hem desenvolupat una interfície web, anomenada *RiskDiff* per analitzar la diferència deguda a factors de risc respecte els factors demogràfics per a dades d'incidència o mortalitat d'una malaltia determinada, amb l'objectiu de promoure i divulgar l'ús del mètode proposat per Bashir i Estève.

7.2 Anàlisis epidemiològiques

1. Hem aportat una estimació de la mortalitat per càncer de mama l'any 2005 i la seva projecció l'any 2015 a 32 països europeus, i hem identificat diferents patrons de mortalitat.
2. En general, els països del Nord, Oest i Sud d'Europa presenten una tendència decreixent de la mortalitat per càncer de mama, al voltant del 2%. Els països de l'Europa de l'Oest presenten un decreixement de major grau, especialment en el grup de dones més joves, mentre que pels països nòrdics s'observa un cert esmorteïment en les dones de major edat. En els països de l'Europa del Centre i de l'Est, tot i observar-se en general un cert decreixement global, més evident en les dones joves, es constata per les dones amb edat mitjana o elevada una estabilització de la mortalitat o, fins i tot un creixement.
3. Hem estimat que l'any 2015 a Europa hi hauran 127.983 morts per càncer de mama en dones majors de 35 anys, el que representa un augment dels casos del 4,68% respecte l'any 2005. 10.288 defuncions (8,04%) correspondran a dones entre 35 i 49 anys, 54.226 defuncions (42,37%) a dones entre 50 i 69 anys i, la majoria de defuncions, 63.469 (49,59%) a dones majors de 70 anys. El major nombre de defuncions correspondrà a dones de països de l'Europa del Centre i Est (38,69%), seguit per dones de l'Europa de l'Oest (30,19%) i, en una magnitud menor, dones de l'Europa del Sud i del Nord (17,98% i 13,14% respectivament). La taxa bruta de mortalitat per càncer de mama a Europa l'any 2015 serà de 56,91 defuncions per cada 100.000 dones majors de 35 anys, observant-se diferències marcades segons els grups d'edat, essent aquesta taxa de 13,60, 56,28 i 120,07 per als grups 35-49, 50-69 i majors de 70 anys respectivament.
4. Pel que fa al augment previst del nombre de morts per càncer de mama l'any 2015 respecte l'any 2005 del 4.68%, aquest es pot descomposar en

tres components: augment de mida de la població (5,28%), augment per envelliment de la població (6,20%) i disminució deguda a risc de morir per la malaltia (-6,80%). D'aquesta forma hem pogut constatar que l'augment del nombre de defuncions és degut principalment a factors demogràfics (mida i envelliment) mentre que el risc de morir per aquesta patologia disminuirà de forma important.

5. La disminució del risc de morir serà superior en l'Europa del Nord, seguit de l'Oest i del Sud (14,0%, 10,61% i 7,37% respectivament). S'espera que hi hagi un augment moderat del nombre de defuncions deguts a un augment del risc de morir en l'Europa del Centre i est (0,49%).
6. Per tal de disminuir el nombre de defuncions previstos per l'any 2015, específicament a l'Europa del Centre i de l'Est i també a Grècia, seria recomanable, en termes de salut pública, aprofundir en la implantació de sistemes de cribatge i teràpies més eficients, així com una promoció de la conscienciació pública envers aquesta patologia.

Apèndix A

Codi de les funcions

A.1 Funcions per a l'anàlisi de tendències i projeccions

Funció `compute.APC()`

```
compute.APC <- function(start,end,country,ages)
  #start;end : integers with the starting and ending year for computing APC
  #country: country name
  #ages: array containing age-groups for being included.
  #E.g. ages=c("35-39","40-44","45-49") for 35-49 group
{
  data <- breast.mortality[breast.mortality$country==country
    & breast.mortality$age.group %in% ages
    & breast.mortality$year %in% c(start:end),]
  data$crude.rate <- data$cases / data$population * 100000
  data$WSP <- WSP.pyramid$weight[as.integer(data$age.group)]
    /sum(WSP.pyramid[ages,2])
  data$AS.rate <- data$crude.rate*data$WSP
  data <- aggregate(data[,c(4,5,8)],by=list(data$year),FUN=sum)
  data$country <- rep(country,nrow(data))
  names(data)[1] <- "year"
  data[,1] <- as.numeric(as.character(data[,1]))
  m <- glm(AS.rate ~ year, family=gaussian(link="log"),data=data)
  m.APC <- (exp( m$coeff["year"])-1)*100
  m.APC.CI <- (exp(confint(m)["year",])-1)*100
  m.APC.pvalue <- summary(m)$coefficients[2,4]

  result <- list(data=data,ages=ages
    ,results=data.frame(country=country,start=start,end=end
```

```

        ,APC=m.APC,APC.LCI=m.APC.CI[1],APC.UCI=m.APC.CI[2]
        ,APC.pvalue=m.APC.pvalue))
row.names(result$results) <- ""
result
}

```

Funció project.mortality()

```

project.mortality <-
function(start,end,up.to,country,ages,type="log")
  #start;end: integers with the starting and ending year
  #country: country name
  #ages: array containing age-groups for being included.
  #E.g. ages=c("35-39","40-44","45-49") for 35-49 group
  #type: "log" or "additive" or "log-age-drift"
{
  data <- breast.mortality[breast.mortality$country==country
                           & breast.mortality$age.group %in% ages
                           & breast.mortality$year %in% c(start:end),]

  if(type=="log" | type=="maap")
  {
    m <- glm(cases ~ offset(log(population))+age.group+age.group*year-1
             ,family=poisson(link="log"),data=data)
  }
  if(type=="log-age-drift" | type=="map")
  {
    m <- glm(cases ~ offset(log(population))+age.group+year - 1
             ,family=poisson(link="log"),data=data)
  }
  if(type=="additive" | type=="aaap")
  {
    m <- glm(cases/population ~ age.group + age.group*year - 1
             ,family=poisson(link="identity"),weights=population,data=data)
  }

  m.Pearson.chisq <- sum(resid(m,type="pearson")^2)
  m.df.residual <- m$df.residual
  m.pvalue <- 1-pchisq(m.Pearson.chisq,m$df.residual)

  newdata <- europe.population[europe.population$country==country
                               & europe.population$year==up.to
                               & europe.population$age.group %in% ages,]

  m.predict <- predict(m,newdata=newdata,type="response")
  if(type=="additive" | type=="aaap")
  {

```

```

    m.predict <- m.predict*newdata$population
  }

  newdata <- data.frame(newdata,projected.cases=m.predict)
  rownames(newdata) <- 1:nrow(newdata)
  result <- list(data.frame(start=start,end=end,up.to=up.to,country=country
    ,Pearson.chisq=m.Pearson.chisq,df.residual=m.df.residual
    ,pvalue=m.pvalue,type=type),projected.data=newdata)
  names(result)[1] <- "model"
  result
}

```

A.2 Funcions per a l'anàlisi les diferències degudes a risc i a factors demogràfics en dades d'incidència o mortalitat

Funció risk.diff()

```

risk.diff <- function(cases.init,cases.end,pop.init,pop.end
  ,labels=c("baseline group","comparison group"))
#cases.init and cases.end are vectors containing age-group cases for
the baseline(reference) and comparison groups respectively
#(e.g. init and end years)
#pop.init and pop.end are vectors containing population pyramids
#for both groups
#labels is a vector containing text labels for baseline and
#comparison groups respectively
{
  adjusted.population.init <- pop.init/sum(pop.init)* 100000
  adjusted.population.end <- pop.end/sum(pop.end)* 100000
  R1 <- cases.init / pop.init * adjusted.population.init
  R2 <- cases.end / pop.end * adjusted.population.end
  R3 <- cases.init / pop.init * adjusted.population.end
  S1 <- sum(R1)
  S2 <- sum(R2)
  S3 <- sum(R3)
  table.out <- data.frame(change.in=c("Risk","Structure","Size","Net change")
    ,crude.rate=rep(NA,4),Perc.crude.rate=rep(NA,4)
    ,Number=rep(NA,4),Perc.Number=rep(NA,4))

  table.out[4,2] <- S2-S1
  table.out[1,2] <- S2-S3
  table.out[2,2] <- S3-S1
  table.out[,3] <- table.out[,2] / S1 * 100

  table.out[4,4] <- sum(cases.end) - sum(cases.init)
  table.out[1:2,4] <- sum(cases.init) * table.out[1:2,3] /100

```

```

table.out[3,4] <- table.out[4,4] - table.out[1,4] - table.out[2,4]

table.out[1:2,5] <- table.out[1:2,3]
table.out[3,5] <- table.out[3,4]/sum(cases.init) * 100
table.out[4,5] <- (sum(cases.end)-sum(cases.init))/sum(cases.init) * 100

summary.table <- data.frame(group=c(labels[1],labels[2]),cases=rep(NA,2)
                           ,population=rep(NA,2),crude.rate=rep(NA,2))
summary.table[1,2] <- sum(cases.init)
summary.table[2,2] <- sum(cases.end)
summary.table[1,3] <- sum(pop.init)
summary.table[2,3] <- sum(pop.end)
summary.table$crude.rate<-summary.table$cases/summary.table$population * 100000

summary.text.1 <- paste(summary.table[1,2]," total cases observed in "
                       ,labels[1]," and ",summary.table[2,2]," in "
                       ,labels[2]," , representing a crude rate of "
                       ,round(summary.table[1,4])," and ",round(summary.table[2,4])
                       ," cases per 100000 persons respectively.",collapse="",sep="")
if(table.out[1,2]>=0) risk <- "increment"
else risk <- "decrement"
if(table.out[3,4]>=0) size <- "increment"
else size <- "decrement"
summary.text.2 <- paste("The net change in terms of crude rate is "
                       ,round(table.out[4,2],2)," cases per 100000 persons
                       (" ,round(table.out[4,3],2),"%), which can be
                       partitioned into " ,round(table.out[1,2],2)
                       ," cases (" ,round(table.out[1,3],2),"% due to
                       a " ,risk," of the risk and " ,round(table.out[2,2],2)
                       ," cases (" ,round(table.out[2,3],2),"% due to changes
                       in the population structure.",collapse="",sep="")
summary.text.3 <- paste("The net change in terms of total number
                       of cases is " ,round(table.out[4,4],2)
                       ," cases (" ,round(table.out[4,5],2),"%
                       , which can be partitioned into " ,round(table.out[1,4],2)
                       ," cases (" ,round(table.out[1,5],2),"%
                       due to a " ,risk," of the risk, " ,round(table.out[2,4],2)
                       ," (" ,round(table.out[2,5],2),"% due to
                       changes in the population structure and "
                       ,round(table.out[3,4],2)," cases (" ,round(table.out[3,5],2)
                       ,"% due to an " ,size," of the population size."
                       ,collapse="",sep="")

list(summary.table,table.out,c(summary.text.1,summary.text.2,summary.text.3))
}

```

Funció plot.risk.diff()

```
plot.risk.diff <-
```

```

function(crude.rate,number,file=NULL,option="pdf",cex.plot=2)
#crude.rate and number are vectors containing the difference
#observed in incidence or mortality between two groups, as
#provided by function risk.diff()
#file can contain a path with a pdf file to be created
#with the graphical representation
#option indicates the type of file to be created "pdf" or "jpg"
{
  if(!is.null(file))
  {
    if(option=="pdf")
    {
      pdf(file=file,width=6,height=4)
    }
    if(option=="jpg")
    {
      jpeg(filename = file, width = 600, height = 400)
    }
  }
  par(mfrow=c(1,3))
  if(crude.rate[1]*crude.rate[2]>0)
  {
    plot(1,type="n",ylim=range(c(crude.rate,0),na.rm=T),
          xlim=c(-0.6,0.6),xlab="",ylab="cases per 100000"
          ,xaxt="n",bty="n",cex.lab=cex.plot,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,crude.rate[1]
          ,crude.rate[1]),col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(crude.rate[1],crude.rate[1]
          ,crude.rate[1]+crude.rate[2],crude.rate[1]+crude.rate[2])
          ,col="gray50")
  }
  if(crude.rate[1]*crude.rate[2]<0)
  {
    plot(1,type="n",ylim=range(crude.rate,na.rm=T),xlim=c(-0.6,0.6)
          ,xlab="",ylab="cases per 100000",xaxt="n",bty="n",cex.lab=cex.plot,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,crude.rate[1]
          ,crude.rate[1]),col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,crude.rate[2]
          ,crude.rate[2]),col="gray50")
  }
  lines(c(-0.5,0.5),rep(crude.rate[4],2),lwd=2)
  abline(h=0,lty=2)

  if(number[1]<0 & number[2]<0 & number[3] <0)
  {
    plot(1,type="n",ylim=range(c(number[1:3],0)),xlim=c(-0.6,0.6)
          ,xlab="",ylab="number of cases",xaxt="n",bty="n"
          ,cex.lab=cex.plot,cex.axis=cex.plot)
  }
}

```

```

    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[1],number[1])
           ,col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[1],number[1]
           ,number[1]+number[2],number[1]+number[2]),col="gray50")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[1]+number[2]
           ,number[1]+number[2],number[1]+number[2]+number[3]
           ,number[1]+number[2]+number[3]),col="gray25")
}
if(number[1]<0 & number[2]<0 & number[3] >0)
{
    plot(1,type="n",ylim=c(sum(number[1:3][number[1:3]<0])
           ,sum(number[1:3][number[1:3]>0])),xlim=c(-0.6,0.6),xlab=""
           ,ylab="number of cases",xaxt="n",bty="n",cex.lab=cex.plot
           ,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[1],number[1]),col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[1],number[1]
           ,number[1]+number[2],number[1]+number[2]),col="gray50")
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[3],number[3]),col="gray25")
}
if(number[1]<0 & number[2]>0 & number[3] <0)
{
    plot(1,type="n",ylim=c(sum(number[1:3][number[1:3]<0])
           ,sum(number[1:3][number[1:3]>0])),xlim=c(-0.6,0.6)
           ,xlab="" ,ylab="number of cases",xaxt="n",bty="n"
           ,cex.lab=cex.plot,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[1],number[1]),col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[1],number[1],number[1]+number[3]
           ,number[1]+number[3]),col="gray25")
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[2],number[2]),col="gray25")
}
if(number[1]<0 & number[2]>0 & number[3] >0)
{
    plot(1,type="n",ylim=c(sum(number[1:3][number[1:3]<0])
           ,sum(number[1:3][number[1:3]>0])),xlim=c(-0.6,0.6)
           ,xlab="" ,ylab="number of cases",xaxt="n",bty="n"
           ,cex.lab=cex.plot,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[2],number[2]),col="gray50")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[2],number[2]
           ,number[2]+number[3],number[2]+number[3]),col="gray25")
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[1],number[1]),col="gray75")
}
if(number[1]>0 & number[2]<0 & number[3] <0)
{
    plot(1,type="n",ylim=c(sum(number[1:3][number[1:3]<0])
           ,sum(number[1:3][number[1:3]>0])),xlim=c(-0.6,0.6),xlab=""
           ,ylab="number of cases",xaxt="n",bty="n",cex.lab=cex.plot
           ,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[1],number[1]),col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[2],number[2]),col="gray50")
}

```

```

    polygon(c(-0.4,0.4,0.4,-0.4),c(number[2],number[2]
      ,number[2]+number[3],number[2]+number[3]),col="gray25")
  }
  if(number[1]>0 & number[2]<0 & number[3] >0)
  {
    plot(1,type="n",ylim=c(sum(number[1:3][number[1:3]<0])
      ,sum(number[1:3][number[1:3]>0])),xlim=c(-0.6,0.6),xlab=""
      ,ylab="number of cases",xaxt="n",bty="n",cex.lab=cex.plot
      ,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[1],number[1]),col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[2],number[2]),col="gray50")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[1],number[1]
      ,number[1]+number[3],number[1]+number[3]),col="gray25")
  }
  if(number[1]>0 & number[2]>0 & number[3] <0)
  {
    plot(1,type="n",ylim=c(sum(number[1:3][number[1:3]<0])
      ,sum(number[1:3][number[1:3]>0])),xlim=c(-0.6,0.6)
      ,xlab="" ,ylab="number of cases",xaxt="n",bty="n"
      ,cex.lab=cex.plot,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[1],number[1]),col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[1],number[1]
      ,number[1]+number[2],number[1]+number[2]),col="gray50")
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[3],number[3]),col="gray25")
  }
  if(number[1]>0 & number[2]>0 & number[3]>0)
  {
    plot(1,type="n",ylim=range(c(number[1:3],0)),xlim=c(-0.6,0.6)
      ,xlab="" ,ylab="number of cases",xaxt="n",bty="n"
      ,cex.lab=cex.plot,cex.axis=cex.plot)
    polygon(c(-0.4,0.4,0.4,-0.4),c(0,0,number[1],number[1]),col="gray75")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[1],number[1]
      ,number[1]+number[2],number[1]+number[2]),col="gray50")
    polygon(c(-0.4,0.4,0.4,-0.4),c(number[1]+number[2]
      ,number[1]+number[2],number[1]+number[2]+number[3]
      ,number[1]+number[2]+number[3]),col="gray25")
  }
  lines(c(-0.5,0.5),rep(number[4],2),lwd=2)
  abline(h=0,lty=2)

  plot(1,type="n",ylim=c(1,4),xlim=c(0,4),xlab="" ,ylab="" ,bty="n",xaxt="n",yaxt="n")
  legend(0,4,"net change",lwd=2,bty="n",cex=cex.plot)
  legend(0.1,3.8,"due to risk",fill="gray75",bty="n",cex=cex.plot)
  legend(0.1,3.6,"due to structure",fill="gray50",bty="n",cex=cex.plot)
  legend(0.1,3.4,"due to size",fill="gray25",bty="n",cex=cex.plot)
  if(!is.null(file))
  {
    dev.off()
  }
}

```


}

Apèndix B

RiskDiff: a web tool for the analysis of the difference due to risk and demographic factors for incidence or mortality data

Reproduïm a continuació una còpia de l'article que presenta les funcions i interfície web implementades i, alhora, conté una aplicació a dades de mortalitat per càncer a Catalunya, que mostren particularment l'interès d'aquests mètodes.

Software

Open Access

RiskDiff: a web tool for the analysis of the difference due to risk and demographic factors for incidence or mortality dataJoan Valls*^{1,2,3}, Ramon Clèries^{1,4}, Jordi Gálvez¹, Victor Moreno^{4,5}, Rosa Gispert⁶, Josep M Borràs¹ and Josepa Ribes^{1,4}

Address: ¹Catalan Cancer Registry, Catalan Institute of Oncology, Barcelona, Catalonia, ²Biomedical Research Institute, Lleida, Catalonia, Spain, ³Department of Mathematics, Autonomous University of Barcelona, Barcelona, Catalonia, Spain, ⁴Department of Clinical Sciences, University of Barcelona, Barcelona, Catalonia, Spain, ⁵Bioinformatics and Biostatistics Unit, Catalan Institute of Oncology, Barcelona, Catalonia, Spain and ⁶Catalan Mortality Registry, Barcelona, Catalonia, Spain

Email: Joan Valls* - joan.valls@iconcologia.net; Ramon Clèries - r.cleries@iconcologia.net; Jordi Gálvez - j.galvez@iconcologia.net; Victor Moreno - v.moreno@iconcologia.net; Rosa Gispert - rosa.gispert@gencat.cat; Josep M Borràs - jmborras@iconcologia.net; Josepa Ribes - j.ribes@iconcologia.net

* Corresponding author

Published: 18 December 2009

Received: 16 July 2009

BMC Public Health 2009, 9:473 doi:10.1186/1471-2458-9-473

Accepted: 18 December 2009

This article is available from: <http://www.biomedcentral.com/1471-2458/9/473>

© 2009 Valls et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Analysing the observed differences for incidence or mortality of a particular disease between two different situations (such as time points, geographical areas, gender or other social characteristics) can be useful both for scientific or administrative purposes. From an epidemiological and public health point of view, it is of great interest to assess the effect of demographic factors in these observed differences in order to elucidate the effect of the risk of developing a disease or dying from it. The method proposed by Bashir and Estève, which splits the observed variation into three components: risk, population structure and population size is a common choice at practice.

Results: A web-based application, called RiskDiff has been implemented (available at <http://rht.iconcologia.net/riskdiff.htm>), to perform this kind of statistical analyses, providing text and graphical summaries. Code from the implemented functions in R is also provided. An application to cancer mortality data from Catalonia is used for illustration.

Conclusions: Combining epidemiological with demographical factors is crucial for analysing incidence or mortality from a disease, especially if the population pyramids show substantial differences. The tool implemented may serve to promote and divulgate the use of this method to give advice for epidemiologic interpretation and decision making in public health.

Background

The analysis of the observed differences in the incidence or mortality of a given disease can be of great interest both for scientific and administrative purposes [1]. Studies frequently focus on comparing the number of incident or deceased cases in two given situations, with the aim of

quantifying the differences observed, for further epidemiological interpretations and to give advice for decision making in public health. In this situation, time trends are usually performed to study the historical evolution of risk and to assess the occurrence of a disease in a certain period of time [2], such as comparing two different time points,

in a simple trend analysis. In the same line, geographical variation of risk of a disease can be evaluated by comparing incidence or mortality rates between two areas. These comparisons are usually reported with the absolute difference in the observed number of cases (incidents or deaths) or by using the difference in the crude rates (usually per 100 000 persons), and sometimes the percentage of change is also computed [3]. Although crude rates can be used to compare different diseases in the same population, they are not useful for comparing rates of the same disease in different populations or over time [4]. To overcome this, standardized measures of risk are used to compare the evolution of risk [2], using a common reference population (the world standard population is a common choice [5,6]), and, the percentage of change of the disease is then computed [1]. However, these changes could be partially attributed to the effect of demographic factors and not only to risk, especially if the population pyramids involved in the two situations show substantial differences. A variation of the population size over time could explain variation of the number of cases, due to the consequent increment (or decrement) of persons at risk to develop or die from a certain disease. In addition, changes in the age structure between the populations involved could also lead to substantial changes in the number of cases. Regarding to this, in a number of diseases such as cancer, ageing is known to be clearly associated with molecular, cellular and physiological changes that influence carcinogenesis and subsequent cancer growth [7], and, therefore, an increase of cases among the oldest age-groups is expected [8]. In addition, another situation can arise when migration flows lead to changes in the population structure. For example, recently an increase of measles cases in Catalonia was reported, which has been partially attributed to immigration coming from undeveloped countries with poor measles vaccination coverage [9].

Bashir and Estève developed a method for partitioning the variation in the incidence or mortality from a disease between two groups, quantifying the percentage of change attributable to demographic factors (population size and structure) with respect to that which could be attributed to changes in the risk of developing or dying from a particular disease [10]. The method is based on the idea of first computing the incidence or mortality that one would have observed if the population size and structure were the same for both groups, and secondly attributing this difference with respect to the net change to demographic factors. In addition, the change attributed to demographic factors can then itself be split into that due to variation in population size and that due to changes in the population structure [10]. Thus, this method can evaluate differences in mortality (or incidence) data due to risk and demographic factors, which is not possible directly using stand-

ardized mortality (or incidence) data, since the reference population is a common standard and differences could only be attributed to risk.

The main aim of this paper is to present a set of functions in R code [11], that we have implemented, based on the method proposed by Bashir and Estève. These functions also provide convenient tables and graphical representations. In order to make these functions more widely available, we have implemented a web tool, called RiskDiff (publicly available at <http://rht.iconcologia.net/riskdiff.htm>) where the users can easily perform their analysis. Code for R functions is also freely available on the same web page.

Finally, to illustrate the use of this web tool, we analyse the differences in the number of deceased individuals from cancer in Catalonia in 1985 with respect to 2004, through a long period of 20 years, which is quite relevant from an epidemiological point of view.

Implementation

The functions implemented in R are based on the method presented in the article by Bashir and Estève [10]. This method assumes two groups: the baseline (or reference) group and the comparison group. Number of incident or death cases are given for both groups, aggregated by age-groups (usually 5-years groups). The observed difference in the total number of cases or deaths between both groups can then be split into three components: cases due to changes in population size, cases due to changes in population structure (age distribution) and, for last, cases attributable to changes on the risk to develop or die due to the disease itself. Crude rates (per 100 000 people) have to be first computed, and then, the difference in the crude rates from both groups have to be partitioned in those due to risk and those due to population structure, using the following formula [10]:

$$\frac{S_2 - S_1}{S_1} = \frac{S_3 - S_1}{S_1} + \frac{S_2 - S_3}{S_1}$$

Where S_1 and S_2 are the crude rates (per 100 000 people) for the baseline and comparison group respectively and S_1 is an intermediate rate obtained for the baseline group but using the comparison group as reference population.

Thus, $\frac{S_2 - S_1}{S_1}$ represents the proportional change between the observed rates in two groups, which is then partitioned in the proportional change due to population structure $\left(\frac{S_3 - S_1}{S_1}\right)$ and the proportional change due to differences in risk $\left(\frac{S_2 - S_3}{S_1}\right)$.

Two functions have been implemented in R [11]. The first one, `risk.diff()` needs four parameters called `cases.init`, `cases.end`, `pop.init` and `pop.end` which are vectors of the same length that contain the number of cases (or deaths) and the population for the two groups, for each age-group. As a result this function provides two tables that summarize the difference observed between the groups involved and a short text to facilitate interpretation. The second one, `plot.risk.diff()`, generates a graphical representation from the obtained results. These functions are available as a source text file and some examples of use are also provided. The implementation of these functions in a web interface has been made using PHP programming language [12]. Functions are executed on a remote Linux server, and results are provided on-line.

For the example illustrated in this paper, we have used cancer mortality data for the period 1985-2004 provided from the Catalan Mortality Registry. In 1985, the Catalan population was about 6 million people and near to 7 million in 2004. Population pyramids have been provided by the Catalan Statistical Institute [13]. The number of cancer deaths and the population at risk have been grouped in 5-year age bands. Registered deaths from all cancer location sites are included except those from non-skin melanoma (C44 as coded by ICD-10 [14]).

Results

The number of cancer deaths observed for both sexes in Catalonia in 1985 and 2004, and the respective Catalan population pyramids for these years are shown in tables 1 and 2. To perform the analyses with RiskDiff the user must provide four vectors with the same size containing the number of observed cases or deaths and the population in the both situations, i.e. baseline and comparison groups, for each age group. For our example mortality data from years 1985 and 2004 will be the baseline and comparison groups, respectively. Data can be plugged into RiskDiff in two ways: (1) using a tab-separated text file with 4 rows, one for each vector, with a similar structure as the one shown in table 1 and 2 or (2) directly typing the data into the web interface separately for each vector. Group labels can also be introduced in order to identify the groups. RiskDiff then produces a web page with summary tables, graphical representations and a short paragraph of text to facilitate the interpretation of the results. The results obtained when analysing mortality data from tables 1 and 2 are shown in figures 1 and 2 respectively.

Regarding the changes in the observed cancer mortality in Catalonia, a relatively high increment of both the number of deaths and crude rate is observed through the period 1985 to 2004. However, more thorough analysis reveals

that the risk of dying from cancer has experienced a clear decline. More precisely, for women, the net change in the crude rate was 17 deaths per 100 000 person-years (from 151 to 168), representing an increment of 11.02%. However, our results indicate a decrease of 31.55 deaths per 100 000 person-years (21%) attributable to changes in risk while an increment of 48.19 deaths per 100 000 person-years (32%) was due to changes in population structure, i.e. ageing of the Catalan population. In terms of the absolute number of deaths, the net change was of 1088 deaths (from 4629 to 5717), representing an increment of 23.5%. In the same line, this can be partitioned into that due to an increase of the population size (577.67 deaths, 12%), that due to the ageing of the population (1477.41 deaths, 32%) and that due to risk, which represent a decrement of 967.08 deaths (21%). Analogously, for men the net change in the crude rate was 63 deaths per 100 000 person-years (from 225 to 288), representing an increment of 27.8%. Similar to that of women, a decrement of 8.77 deaths per 100 000 person-years (4%) was attributable to changes in risk while an increment of 71.44 deaths per 100 000 person-years (32%) was due to changes in population structure. In terms of the absolute number of deaths, the net change was 2918 deaths (from 6632 to 9550), representing an increment of 44%. Once again, this can be partitioned into that due to an increase of the population size (1073.32 deaths, 16%), that due to the ageing of the population (2102.76 deaths, 32%) and that due to risk, which represents a decrement of 258.09 deaths (4%).

As stated by the authors of the method, sometimes looking at two points may not be useful or may not paint a clear picture of what is actually happening [10], and they suggest that multiple comparisons can be adequate for this purpose. In these cases, one baseline group could be used, along with a set of comparison groups. This would be the case of analysing the incidence or mortality of a disease for a number of consecutive years. Even though this is not explicitly considered in RiskDiff, this analysis could actually be done by directly using the functions in R code. To illustrate this, Figure 3 shows the evolution of the percent change in terms of crude rate for the years 1986 to 2004, with respect to the baseline year 1985 (R code from this analysis is available at the web page). These results show that the real decline of the risk of dying from cancer in Catalonia started in mid 80s for women and early 90s for men. However, analysing the evolution of the net change in the crude rate, which is analogous to analyzing the evolution of the crude rate itself, indicates erroneously that mortality started to decline in posterior years and, in addition, this decline is much more clear than the one observed for the decline in the risk itself. In addition to this, these results could also give some clues for quantifying the effect of migration flows and ageing in Catalonia

RiskDiff Results

Summary table

group	cases	population	crude.rate
CancerWomenDeaths1985	4629	3065733	150.99
CancerWomenDeaths2004	5717	3410329	167.64

Results table

Change in	Crude rate	%	Number	%
Risk	-31.55	-20.89	-967.09	-20.89
Structure	48.19	31.92	1477.41	31.92
Size	NA	NA	577.67	12.48
Net change	16.65	11.02	1088.00	23.50

4629 total cases observed in CancerWomenDeaths1985 and 5717 in CancerWomenDeaths2004, representing a crude rate of 151 and 168 cases per 100000 persons respectively. The net change in terms of crude rate is 16.65 cases per 100000 persons (11.02%), which can be partitioned into -31.55 cases (-20.89%) due to a decrement of the risk and 48.19 cases (31.92%) due to changes in the population structure. The net change in terms of total number of cases is 1088 cases (23.5%), which can be partitioned into -967.09 cases (-20.89%) due to a decrement of the risk, 1477.41 (31.92%) due to changes in the population structure and 577.67 cases (12.48%) due to an increment of the population size.

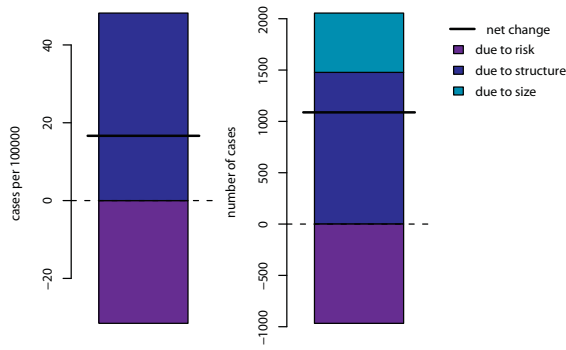


Figure 1
Results obtained with RiskDiff to evaluate the change in the observed mortality for years 1985 respect to 2004, in women from Catalonia.

RiskDiff Results

Summary table

group	cases	population	crude.rate
CancerMenDeaths1985	6632	2943372	225.32
CancerMenDeaths2004	9550	3316066	287.99

Results table

Change in	Crude rate	%	Number	%
Risk	-8.77	-3.89	-258.09	-3.89
Structure	71.44	31.71	2102.77	31.71
Size	NA	NA	1073.33	16.18
Net change	62.67	27.81	2918.00	44.00

6632 total cases observed in CancerMenDeaths1985 and 9550 in CancerMenDeaths2004, representing a crude rate of 225 and 288 cases per 100000 persons respectively. The net change in terms of crude rate is 62.67 cases per 100000 persons (27.81%), which can be partitioned into -8.77 cases (-3.89%) due to a decrement of the risk and 71.44 cases (31.71%) due to changes in the population structure. The net change in terms of total number of cases is 2918 cases (44%), which can be partitioned into -258.09 cases (-3.89%) due to a decrement of the risk, 2102.77 (31.71%) due to changes in the population structure and 1073.33 cases (16.18%) due to an increment of the population size.

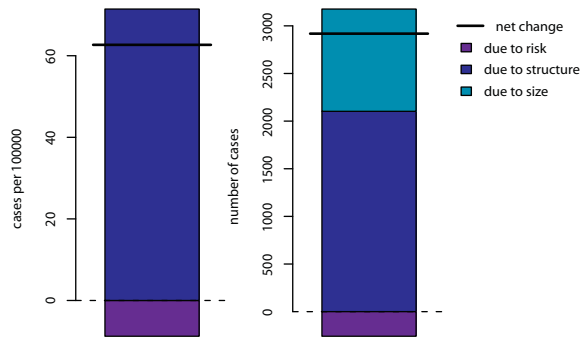


Figure 2
Results obtained with RiskDiff to evaluate the change in the observed mortality for years 1985 respect to 2004, in men from Catalonia.

Table 1: Number of observed deaths from cancer (all sites except non-skin melanoma) and population for years 1985 and 2004, for women in Catalonia.

Age in years	Deaths		Population (%)	
	1985	2004	1985	2004
0-4	6	2	163597 (5.3%)	171047 (5.0%)
5-9	7	3	230080 (7.5%)	146822 (4.3%)
10-14	6	2	247955 (8.1%)	145507 (4.3%)
15-19	10	3	238448 (7.8%)	158541 (4.6%)
20-24	16	8	228928 (7.5%)	215665 (6.3%)
25-29	13	18	220139 (7.2%)	291501 (8.5%)
30-34	49	39	207732 (6.8%)	287951 (8.4%)
35-39	96	80	207676 (6.8%)	269903 (7.9%)
40-44	94	99	181344 (5.9%)	252005 (7.4%)
45-49	147	178	167197 (5.5%)	232065 (6.8%)
50-54	261	260	189080 (6.2%)	210241 (6.2%)
55-59	363	335	180805 (5.9%)	205309 (6.0%)
60-64	451	399	164187 (5.4%)	166926 (4.9%)
65-69	567	500	133730 (4.4%)	156719 (4.6%)
70-74	685	768	117116 (3.8%)	166723 (4.9%)
75-79	722	927	92075 (3.0%)	141981 (4.2%)
80-84	667	932	59506 (1.9%)	105100 (3.1%)
85+	469	1164	36138 (1.2%)	86323 (2.5%)
Total	4629	5717	3065733 (100%)	3410329 (100%)

Table 2: Number of observed deaths from cancer (all sites except non-skin melanoma) and population for years 1985 and 2004, for men in Catalonia.

Age in years	Deaths		Population (%)	
	1985	2004	1985	2004
0-4	13	3	175499 (6.0%)	180877 (5.5%)
5-9	9	5	244280 (8.3%)	156506 (4.7%)
10-14	8	4	263451 (9.0%)	153960 (4.6%)
15-19	17	10	251663 (8.6%)	167438 (5.0%)
20-24	25	17	235813 (8.0%)	226394 (6.8%)
25-29	23	25	216962 (7.4%)	314418 (9.5%)
30-34	32	28	206881 (7.0%)	314187 (9.5%)
35-39	79	85	207612 (7.1%)	287527 (8.7%)
40-44	106	109	183591 (6.2%)	256270 (7.7%)
45-49	240	296	167220 (5.7%)	228945 (6.9%)
50-54	443	493	181749 (6.2%)	203690 (6.1%)
55-59	636	724	168626 (5.7%)	195976 (5.9%)
60-64	927	896	146463 (5.0%)	157911 (4.8%)
65-69	909	1204	103788 (3.5%)	138642 (4.2%)
70-74	1068	1527	82637 (2.8%)	134291 (4.0%)
75-79	1015	1670	58032 (2.0%)	100964 (3.0%)
80-84	706	1419	32430 (1.1%)	62208 (1.9%)
85+	376	1035	16675 (0.6%)	35862 (1.1%)
Total	6632	9550	2943372 (100%)	3316066 (100%)

on the future mortality of cancer [15], which can be useful for decision making in public health.

Discussion and conclusions

Evaluating the differences in the incidence or mortality of a disease in two given situations (such as time points, geographical areas or males versus females) without adjusting by the populations at risk involved, could lead to incorrect results [10]. Thus, it is necessary to take into account demographic factors, i.e. population size and population structure, in order to more precisely attribute which part of the observed changes is due to risk. The method presented by Bashir < Estève [10] is a good solution and a common choice at practice. This work presents a publicly available web tool that performs this analysis and provides graphical summaries and tables, with the intention of contributing to the divulgation of the method and to promote its use in epidemiology and public health sciences, which may contribute to its use at an applied level.

The results obtained from the analysis of the Catalan cancer mortality are useful to illustrate the method and its application. Thus, these results give an appropriate example that supports the importance of taking into account the changes in the population, since a simple analysis would have concluded that there was an increment in the mortality for cancer in Catalonia of 23% for women and 44% for men throughout the 20 year period analysed (1985 to 2004), however by using this method it can be stated that, actually the risk of dying from cancer has been reduced by 20% for women and 4% for men, and the major reasons for the apparent net increment was the increase of the population size (12% for women and 16% for men) and ageing of society (32% for both women and men). Thus, the use of this method is highly recommended when comparing data from heterogeneous populations, which is translated into large variability between them. The effect of immigration on the assessment of risk when comparing two time periods could be ascertained through this method, as it is the case of Catalonia [13,16,17]. Similar to other regions in Europe, the stated decline of the mortality from cancer in Catalonia in this period may be due to a number of factors such as advances in cancer treatment and diagnostic techniques as well as the decrease in the prevalence of smoking habits [18], which is somewhat similar to other regions in Europe[19].

Regarding statistical issues, the method developed by Bashir < Estève [10] does not consider specific methods for assessing whether the observed differences are significant or not, so that it is unclear how this type of hypothesis could be tested. Thus, RiskDiff has to be considered just as a tool for describing mortality or incidence data. In the case of a population-based register the differences

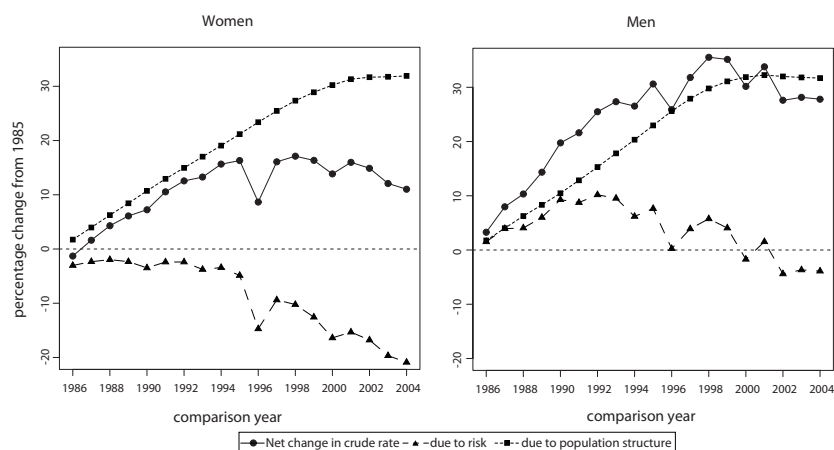


Figure 3
Evolution of the differences in cancer mortality in Catalonia from 1986 to 2004 respect the baseline year 1985.

observed can be considered as the true ones and, therefore, the differences described directly refer to the differences in the population. However, in the case of sampling a general population, these differences have to be taken with caution. In a future, a non parametric procedure, such a bootstrap one, could be implemented to RiskDiff, so that a confidence interval for the observed differences could then be provided.

In conclusion, analysing incidence or mortality data without taking into account demographic effects, can lead to results that are not easily usable for policy making. In this situation, data on the absolute number of cases and demographic determinants is highly relevant for planning purposes and for assessing future needs. This work supports the idea of combining epidemiology with demography when performing statistical analysis on the incidence or mortality from a disease, especially in dynamic populations that are affected also by other risk factors as well, that may also vary across time, gender or geographic regions.

Availability and requirements

Project name: RiskDiff

Project home page: The webtool can be used through the following website, <http://rht.iconcologia.net/risk>

[diff.htm](#). In addition, files for the R functions and examples of use can be available as supplementary material (Additional file 1) and can also be downloaded from the web site.

Operating system: Platform independent for accessing the public web server

Programming language: R and PHP

Requirement: R statistical software available at <http://www.r-project.org/> is required for the functions implemented.

License: None

Any restriction to use by non-academics: None

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JV, JR and RC initially conceived the tool and were involved in its design. JV did the statistical analysis and the implementation in R. JG implemented the web interface. All authors have been involved in drafting the man-

uscript and revising it critically. All authors approved the final version.

Additional material

Additional file 1

R code for the functions and examples of use. A compressed (zip) file which contains: one file with the R code for the functions that perform riskdiff analysis (RiskDiff.R); one file with R code with examples of use of the functions (ExampleRiskDiff.R); three text files (tab-separated) that contain the data from cancer mortality in Catalonia used in the paper and that can be used as example for the R functions (Mortality.cancer.catalonia.men.1985.2004.txt, Mortality.cancer.catalonia.women.1985.2004.txt and CancerMortalityCataloniaData.txt). Click here for file [http://www.biomedcentral.com/content/supplementary/1471-2458-9-473-S1.ZIP]

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2458/9/473/prepub>

References

- Hakulinen T, Hakama M: **Predictions of epidemiology and the evaluation of cancer control measures and the setting of policy priorities.** *Soc Sci Med* 1991, **33**(12):1379-1383.
- Esteve J, Benhamou E, Raymond L: **Statistical methods in cancer research. Volume IV. Descriptive epidemiology.** *IARC Sci Publ* 1994:1-302.
- Breslow NE, Day NE: **Statistical methods in cancer research. Volume II--The design and analysis of cohort studies.** *IARC Sci Publ* 1987:1-406.
- Sasieni PD, Adams J: **Standardized lifetime risk.** *Am J Epidemiol* 1999, **149**(9):869-875.
- Segi M: **Cancer mortality for selected sites in 24 countries (1950-57)** Sendai: Tohoku University School of Public Health; 1960.
- Doll R, Payne P, Waterhouse J: **Cancer incidence in five continents: a technical report** Berlin: Springer-Verlag; 1966.
- Balducci L, Ersler WB: **Cancer and ageing: a nexus at several levels.** *Nat Rev Cancer* 2005, **5**(8):655-662.
- Bray F, Moller B: **Predicting the future burden of cancer.** *Nat Rev Cancer* 2006, **6**(1):63-74.
- Dominguez A, Torner N, Barrabeig I, Rovira A, Rius C, Cayla J, Plasencia E, Minguell S, Sala MR, Martinez A, et al.: **Large outbreak of measles in a community with high vaccination coverage: implications for the vaccination schedule.** *Clin Infect Dis* 2008, **47**(9):1143-1149.
- Bashir S, Esteve J: **Analysing the difference due to risk and demographic factors for incidence or mortality.** *Int J Epidemiol* 2000, **29**(5):878-884.
- R Development Core Team: **R: a language and environment for statistical computing.** Vienna, Austria: R Foundation for Statistical Computing; 2007.
- PHP Hypertext preprocessor [http://www.php.net]
- Catalan Statistical Institute [http://www.idescat.cat]
- World Health Organization: **International Statistical Classification of Diseases and Related Health Problems 10th Revision, Geneva 2007.**
- Ribes J, Cléries R, Buxo M, Ameijide A, Valls J, Gispert R: **Predictions of cancer incidence and mortality in Catalonia to 2015 by means of Bayesian models.** *Med Clin* 2008, **131**(Suppl):32-41.
- Cabré A, Domingo A: **Demografia i immigració, 1991-2005.** *Papers de demografia* 2007, **32**:1-32.
- Non-national populations in the EU Member States [http://epo.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NK-06-008/EN/KS-NK-06-008-EN.PDF]
- Gispert R, Cléries R, Puigdefàbregas A, Freitas A, Esteban L, Ribes J: **Cancer mortality trends in Catalonia, 1985-2004.** *Med Clin* 2008, **131**(Suppl):25-31.
- Karim-Kos HE, de Vries E, Soerjomataram I, Lemmens V, Siesling S, Coebergh JW: **Recent trends of cancer in Europe: a combined approach of incidence, survival and mortality for 17 cancer sites since the 1990s.** *Eur J Cancer* 2008, **44**(10):1345-1389.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



Apèndix C

Differences due to risk and demographic factors for the projections and burden of breast cancer mortality in Europe from 2005 to 2015

Reproduïm a continuació una còpia del text l'article (sotmès i en revisió) que presenta els resultats principals de la tesi pel que fa a la mortalitat per càncer de mama a Europa i projeccions. No hem inclòs aquí les figures ni les taules, així com tampoc la llista de referències bibliogràfiques, doncs coincideixen substancialment amb les figures, taules i referències en aquesta tesi.

Differences due to risk and demographic factors for the projections and burden of breast cancer mortality in Europe from 2005 to 2015

Ramon Clèries^{*1,2}, Joan Valls^{3,4}, Josepa Ribes¹, Josep Maria Borràs^{1,2}, Tadeusz Dyba⁵

1. Cancer Registry of Catalonia. Generalitat de Catalunya and IDIBELL, Barcelona, Spain. Av. Gran Via Km 2,7 08907 Hospitalet de Llobregat. Catalonia. Spain.
2. Dept. of Clinical Sciences. University of Barcelona. Campus de Bellvitge, Edifici Pavelló de Govern, Feixa Llarga s/n, 08907 L'Hospitalet de Llobregat, Barcelona, Catalonia. Spain.
3. Mathematics Department, Autonomous University of Barcelona. 08193 Cerdanyola del Vallés. Catalonia. Spain.
4. Institut de Recerca Biomèdica de Lleida. Av. Alcalde Rovira Roure nº80 25198 Lleida. Catalonia. Spain.
5. Finnish Cancer Registry, Institute for Statistical and Epidemiological Cancer Research, Pieni Roobertinkatu 9, FI-00130 Helsinki, Finland

Key words: Mortality, Breast Cancer, burden, prediction

Address of the corresponding author:

Dr Ramon Clèries

Cancer Registry of Catalonia. Generalitat de Catalunya and IDIBELL, Barcelona, Spain.

Av. Gran Via Km 2,7 08907 Hospitalet de Llobregat. Spain

Tel.:+34 932607417

e-mail: r.cleries@iconcologia.net

Abstract

The goal of this study is to assess the burden of breast cancer (BC) mortality in women older than 35 years of age between 2005 and 2015 in 32 European countries. Using the last 10 years of mortality data available in each country, short-term prediction methods have been used to compute the estimated annual percent change (EAPC) and the predictions of BC mortality rates. The difference in the number of deaths comparing 2005 with 2015 (net change) has been partitioned into three additive percentages: 1) population size, 2) age distribution (structure) and 3) BC risk of death by BC. A net change of 4.68% on the number of BC deaths in Europe comparing year 2005 (N=122,264) with year 2015 (N=127,983) is expected. This net change is partitioned into 5.28% due to population increase, 6.20% due to population ageing and -6.80% due to decrease of BC risk of death. BC mortality rates decreased in most European countries, although a significant rise was detected among women older than 70 years in Greece (2.81%), Belarus (2.77%), Romania (1.81%), Russian Federation (1.31%) and Ukraine (0.99%) and among women aged 50-69 years in Russian Federation (0.97%), Ukraine (1.25%) and Latvia (1.92%). Risk of death from BC comparing years 2005 with 2015 is expected to rise only in Latvia (11.1%), Poland (2.89%), Ukraine (2.34%), Russian Federation (2.04%), Romania (0.64%) and Greece (18.99%). The rising BC risk of death in some Eastern European countries and Greece requires specific assessment of BC trends.

Introduction

During 2006, breast cancer (BC) was the most common tumour related-death among European women. There were estimated 131,900 women who died from BC in Europe during this year⁽¹⁾. In 2002 the highest BC mortality rates were found in Western and Northern Europe whereas Eastern and Central European (ECE) countries showed the lowest mortality rates⁽²⁻⁴⁾.

Changes in BC clinical management and implementation of screening occurred during mid 80s⁽⁵⁾ have led to stage shift at presentation and towards a decreasing trend of BC mortality rates in Europe since the mid 90s⁽⁶⁾. However, these BC mortality trends have not been homogeneous across European areas and they have been stable or upwards up to 2004 in most of the ECE countries^{(2, 3),(6)}. Since the interventions for the reduction of BC mortality has not been homogeneous between countries⁽⁷⁾ is therefore of interest to consider the burden of BC mortality in Europe in the years to come.

The aim of the present study is to quantify the burden of BC mortality in Europe for 2015 assessing the role of demographic factors for each country considered in the study.

Methods

Data

Mortality data has been obtained from the World Health Organization database⁽⁸⁾, available at the International Agency for Research on Cancer⁽⁹⁾ web site (<http://www-dep.iarc.fr/>, last accessed November 2009), for European countries which had mortality data since 1990 onwards. Population data for the years analysed and for projections have been extracted from the Population Division web site of the United Nations, through the World Population Prospects: the 2006 Revision⁽¹⁰⁾ (<http://esa.un.org/unpp/> last accessed November 2009). A total of 32 European countries were included at the time when the study was carried out. The last year with data available was 2007 for Romania; 2006 for Finland, Iceland, Ireland, United Kingdom (UK: England, Wales and Scotland), Austria, Germany, Switzerland, Greece, Slovenia, Latvia, Lithuania, Poland and Russian Federation; 2005 for Norway, Sweden, France, Luxembourg, The

Netherlands, Malta, Spain, Czech Republic, Estonia, Hungary, Slovakia and Ukraine; 2004 for Bulgaria; 2003 for Bulgaria, Italy and Portugal; 2001 for Denmark and 1997 for Belgium. European countries have been divided into four areas: Northern (NE), Western (WE), Southern (SE) and Eastern-Central (ECE). Data for 1996 in Belarus was estimated through interpolation. Countries with small populations, and therefore, low number of BC deaths such as Iceland, Malta and Luxembourg, have been included in the analysis even that caution should be taken when interpreting their data due to estimates are based on small size populations. Both mortality and population data have been collected in a 5-year age group structure, for women aged 35 or older.

Statistical analyses

Age-adjusted rates (AAR) to the World Standard Population⁽¹¹⁾ and crude rates per 100,000 person-years for all ages and for three different age-groups: 35-49, 50-69 and ≥ 70 years of age were obtained. We have calculated the Estimated Annual Percent Change (EAPC) of these rates for the last 10 years of data available. Short-term prediction methods based on log-linear models^{(12),(13),(14)} have been used to obtain predictions for year 2015.

A method described by Bashir an Estève⁽¹⁵⁾ has been used to assess differences between the number of BC deceased cases projected for 2015 respect to those observed in 2005. These differences could be attributable to changes in population size, changes in population structure (ageing), and/or changes of BC risk itself. This is achieved by partitioning the global net change between the crude BC mortality rates of the two time points into three percent changes: population size and structure, and risk of death from BC itself, therefore the sum of these changes is the percent net change in the number of cases comparing BC deaths in 2015 with those in 2005. Note that these changes should be interpreted with caution for Malta, Iceland and Luxembourg because instability in trends due to small population sizes.

Equations related with these statistical methods are described in detail in the Appendix. There has been developed a set of macros on R (www.cran-r.org) and Stata (www.stata.com) format which can be obtained through previous e-mail to the corresponding author.

Results

The burden of BC mortality by European regions in 2015 (Table 1)

A net change of 4.68% in the number of BC deaths in Europe comparing year 2005 (N=122,264) with year 2015 (N=127,983) is expected based on projecting the most recent time trends of BC mortality. This net change is partitioned into 5.28% due to population size, 6.20% due to population ageing and -6.80% due to risk of death by BC. This pattern could be also observed in the WE region (net change 3.38%: 4.83% size, 9.16% ageing and -10.61% due to BC risk of death, respectively) and in the SE region (net change 6.64%: 10.82% size, 6.04% ageing and -7.37% due to BC risk of death, respectively). The decrease expected in the number of BC deceased cases in NE (-3.56%) could be due to a large decrease in 14% on the BC risk of death. An increase of BC deaths among women older than 50 years is predicted in the ECE and WE regions (Women 50-69: 15.62% in ECE and 3.02 in the WE countries; Women older than 70: 11.10% and 8.89%, respectively). Among women younger than 70, the number of BC deaths is expected to drop in NE (Women 35-49: -23.15%; Women 50-69: -5.55%) and SE (Women 35-49: -11.16%; Women 50-69: -1.42%) countries, whereas a rise in the number of BC deaths is expected in women aged 70 and older in all areas.

Trends assessment: EAPC based on the last 10 years with available data (Figure 1 and Table 2)

A significant decrease of BC mortality rates (EAPC<0) was found among women older than 35 with the exception of Iceland, Sweden, Luxembourg, Greece, Belarus, Latvia, Poland, Romania, Russian Federation, Slovakia and Ukraine. BC mortality declined significantly in all age-groups separately in Norway, UK, Austria, Germany, Switzerland, The Netherlands, Italy and Spain, whereas in Denmark, Finland, Greece and Slovenia, these rates dropped only in women aged 35 to 49. BC mortality rates significantly rose among women aged 50 to 69 in Latvia (EAPC: 1.92%), Russian Federation (0.97%) and Ukraine (1.25%), whereas decreased or remained stable in the other countries. BC mortality rates also significantly rose among women older than 70 in Greece

(2.81%), Belarus (2.77%), Romania (1.81%), Russian Federation (1.31%) and Ukraine (0.99%).

*Changes of BC mortality in 2015 according to BC mortality rates in 2005
(Table 2 and Figure 2)*

An increase in the number of BC deceased cases (positive percent net change) is expected for all European countries except Switzerland, Austria, Bulgaria, Estonia, Hungary, UK, Lithuania, The Netherlands, Norway, Belarus, Iceland and Malta. Greece showed the highest positive net change in Europe (44.39%) in terms of absolute number of cases (N 2005: 1,775 and N 2015: 2,563). In this line, comparing the number of BC deaths in Greece during 2015 according to BC deaths in 2005, BC mortality could increase 80.98% among women older than 70 years.

An increase in population size as well as ageing of the population may lead to an expected increase in the number of BC deaths. In all countries, the expected increase in the number of cases could be partially due to i) increase in population size except for Estonia, Latvia and Ukraine, and to ii) ageing of the population except for Ireland. Germany and France, the countries that most contribute to the burden of BC mortality in Western Europe, showed a positive percent net change, 2.63% and 7.80%, respectively, mainly attributed to the effect of increase in population size (1.81% and 8.18%, respectively) and to the effect of population ageing (11.23% and 7.15%, respectively). In UK, the country which contributes the most to the NE burden of BC mortality, the dramatic drop in BC mortality observed during the last 10 years is expected to continue during period 2005-2015 leading to -5.61% net change in the number of BC deaths comparing years 2015 with 2005.

Although BC risk of death decreased in most European countries, a rise of this risk is expected in some ECE countries and in Greece and Luxembourg. Greece showed the highest percent change due to increase in the BC risk of death itself (18.99%). The remaining countries with increase of BC risk of death itself were Latvia (11.17%), Poland (2.89%), Ukraine (2.34%), Luxembourg (2.32%), Russian Federation (2.04%) and Romania (0.64%).

Discussion

Our study depicts a pattern of BC mortality rates for 2015, if most recent time trends remain unchanged since 2005 among the European countries considered, assessing how changes in demographic factors (population size or ageing) and in the risk of death from BC could affect to the future burden of BC mortality in Europe. This study is based on projecting future trends assuming that nothing drastic happens in factors such as occurrence or treatment which can be useful as future reference scenario for BC mortality. This is a major difference with a statistical modelling study, where variations of these and other factors are assessed through different statistical models in order to obtain different scenarios for the disease in study^{(16),(17)}.

BC mortality burden in Europe points out an expected increase of 4.68% in the number of BC deaths in Europe when comparing predicted BC mortality in 2015 with that of year 2005. In NE countries, the number of BC deaths is expected to decrease by 3.56% during the study period in parallel with the higher decreasing risk of death from BC attributable to factors other than demographic ones. However, in the remaining European areas the number of BC deaths is expected to increase. All countries showed a decrease in the number of cases among women younger than 50, whereas among women older than 50 years of age an increase is expected in ECE and in WE countries. Among women older than 70 the highest increase is expected in SE (20.67%) and the lowest in NE (1.42%) countries. Although BC death risk is decreasing in most European countries, the increase in the number of cases might be due to the increase in population size and population ageing^{(15),(17)} as it has been shown in Figure 2.

The most striking EAPC decrease has been found among women between 35 to 69 years of age, continuing the trend previously detected during 1990-2002⁽³⁾. In these age groups, the decreasing BC mortality trend occurs at the same time as the increasing access to 1) mammography screening in some countries since the beginning of the 90s⁽⁷⁾ and to 2) access effective treatments^{(2, 3, 18-21),(22)}. Detection of BC in early stage could partially explain the decrease in BC deceased cases and could reflect an impact on the decrease of BC risk of death. Its effect has been estimated to be 16% decrease of BC mortality

among women in all ages, 15% in women aged 40-49 years and 22% in women older than 50⁽²³⁾.

On the contrary, an increase of risk of death from BC can be associated to a rise of BC incidence and to low survival. The lack of population BC awareness and access to specific therapies may affect BC survival⁽²⁴⁾. Changing exposure factors that may affect the risk of developing BC must be considered such as childbearing, breast-feeding, type of diet and obesity, and alcohol and tobacco consumption^{(25),(26)}. Risk factors can be also related to female sex hormones with oestrogenic and progestagenic activity either produced within the body or given as hormonal contraceptives or hormone-replacement therapy⁽²⁷⁾. However, the decline in the use of hormone-replacement therapy has been suggested to have an impact on decreasing BC incidence, but its impact on mortality is not clear⁽²⁸⁾.

In our study, an increase in the risk of death from BC has been found in some ECE countries which could partially be related to a recent increase of BC incidence⁽²⁹⁾. After the transition from centralized to market economy it has been suggested that rising BC incidence could be explained with changes in some risk factors associated to BC such as reproductive factors, differences in childbearing and lifestyles^(3,4). In addition, the increase in BC mortality may also be partially related to the health care system development during this economical transition since 1990s, which could not be a major benefactor of the changes of the market-oriented reform in these countries at early 90s⁽⁴⁾. However, decreasing BC mortality among women aged 35-69 years has been found to be statistically significant in Czech Republic and Hungary, continuing the mortality trend previously observed at the since the end of the 90s^{(2-4),(6)}. During 1990-2005 BC mortality in the ECE has been lower than in the remaining European areas with the exception of these two countries. Women aged 50-69 in Belarus, Bulgaria, Estonia, Latvia, Poland, Russian Federation and Lithuania, could have the most important contribution to the burden of BC mortality in ECE for 2015.

Information about BC screening activities has been reported in a region of Poland among women aged 50-69, with a participation rate of only 41%⁽³⁰⁾. There are also two studies, one of them carried out in Estonia⁽³¹⁾ among women 45-64 and the other one carried out in Lithuania⁽³²⁾, which reported a

participation rate of 50% and 31.8%. These figures could show evidence of the low BC awareness among these populations or also reflect that women were more likely to have undergone screening in countries with nationwide breast cancer screening programs than in those with opportunistic screening⁽³³⁾. The 5-year relative survival improvement of BC in some ECE observed during 1990-94 might be partially associated with their substantial improvements in cancer care⁽³⁴⁾, although the lack of population-based cancer registries and therefore survival data do not make possible a better assessment of differences in survival among them. A study conducted in Estonia during the transition to open-market economy, suggested that difference in BC survival between Estonia and other European countries could be attributable to a late diagnosis, particularly in elderly women⁽³⁵⁾. These results suggest that in some ECE countries only young and middle-aged women could be beneficiaries of the improvements in treatment and detection of BC.

None of the countries with rising risk of BC death started a BC screening program before year 2005 with the exception of Greece, which started a regional program in 1989⁽³⁶⁾. In that region of Greece was estimated that only 53.8% of women aged 50-69 years received BC screening services, being lower than European council recommendations⁽³⁷⁾. Greece has a major contribution to the expected increase of BC deaths in Southern Europe between years 2005 and 2015. In this country, women older than 70 years contribute 56.28% to the BC deaths in 2005 whereas their contribution can be 70.54% in 2015. The lack of BC awareness among the Greek population and the low access to cancer treatment could be related to the increasing mortality from this cause among oldest women⁽³⁸⁾, despite the adjuvant therapy available. The BC mortality trend in these age groups since the beginning of the 60s (AAR: 26.15 per 100,000 women-years for 1960; AAR: 126.04 for 2006; EAPC 3.35%; <http://www-dep.iarc.fr/>) confirms that Greece could be at the top of the ranking of European BC mortality among women older than 70 in 2015, almost doubling their number of BC deaths for year 2015 (N: 1808) compared to those observed for 2005 (N: 999). Unfortunately, information about BC incidence data in this country is not available and, therefore, this continuous rising trend should be further investigated as well as the access to treatments by age groups among women.

Some limitations are necessary to take into account in this study, in order to interpret these results: i) Mortality trends used for projections are based on the last 10-years where data was available leading to varying time periods in each country. The burden of BC mortality for 2005 in Belgium was not observed and it was based on projections. In the same line, we used data from projections in Bulgaria, Italy, Portugal, Belarus and Denmark to quantify the burden of BC for 2005; ii) The variability of data for Iceland, Malta and Luxembourg -countries with low number of BC cases- is also a limitation when assessing changes in the risk of death or demographic factors for BC mortality. Even changes in few deceased BC cases lead to abrupt changes in BC mortality rates and in modelling trends but not in statistically significance of EAPC and iii) projections are based on models fitted to past BC mortality trends and those depend on estimates of future European population, for which precision may vary due to changes in migration, birth rates and other factors⁽¹⁷⁾. These estimates could affect the interpretation of results related with the evolution of the burden of BC mortality in terms of population size, ageing and risk of death⁽¹⁵⁾.

The non-availability of BC incidence data at a national level for most of the countries included in the study poses difficult assessment for future time trends of BC mortality. In this line, the projected increases in the number of BC deaths for Greece and some ECE countries could have several possible causes as it has been discussed, but the lack of BC incidence data makes difficult to ascertain a complete picture of the future BC mortality burden in these countries. On the other hand, BC death certification is sufficiently reliable to permit meaningful inference in most of the countries, even in some ECE countries underrecording of cancer deaths may be found due to a fall in precision of coding causes of death⁽³⁹⁾.

To sum up, the successful development of health care services involved in early detection, diagnosis and management of BC are associated to the decrease in BC mortality risk in most European countries. The implementation of BC cancer screening programs and specific treatments in those areas with lack of them and their assessment of BC incidence cannot be expected without financial support. Targeting public health messages to specifically the segments of the population that could not have access to health care knowledge may also lead

to BC awareness. These efforts should be taken into account in order to assess a future decrease of the burden of BC mortality for the coming years.

Acknowledgements:

Dr Ramon Clèries and Joan Valls were funded by AGAUR grants of the Catalan Government with references 2006 BE 00614 and 2007 BE-1 00046, respectively. Also, the support of the Spanish Network of cancer research (RD06/0020/0089) is acknowledged. The authors want to acknowledge Meritxell Nomen for her secretarial support.

Contributors

All authors participated in the development, design, analysis, and interpretation of this work and in the writing of this paper.

Conflict of interest statement

The named authors declare that they have no conflict of interest.

Role of the funding sources

All funding sources are Dr Ramon Clèries and Joan Valls research awards funded by the Catalan Government. Funding sources had no decision in study design, in the collection, analysis and interpretation of data; in the writing of the manuscript; and in the decision to submit the manuscript for publication.

Statistical Methods Appendix

1. Models for projections

Following the method described by Dyba and Hakulinen^{(12),(13),(14)}, three models were used for the statistical analysis:

$$E(c_{it} / n_{it}) = \alpha_i + \beta_i t \quad (1.1),$$

$$\ln(E(c_{it} / n_{it})) = \exp\{\alpha_i + \beta_i t\} \quad (1.2),$$

$$\ln(E(c_{it} / n_{it})) = \exp\{\alpha_i + \beta_i t\} \quad (1.3),$$

where $E(c_{it})$ and n_{it} are, respectively, the expected number of deaths and population counts for the age-group i at time t , and α_i, β_i are model parameters to be estimated. The first model allows for different age-specific changes of mortality rates on an additive scale over time and it is applied for cancer sites with increasing or stable mortality rates. The next two models are applied for cancer sites with stable or decreasing mortality trends as they use the logarithmic scale. Model (1.2) assumes the same constant change of mortality rate over all age groups whereas model (1.3) allows for different age-specific changes. The choice of a model for projection was based on its goodness of fit to the data using Pearson's X^2 statistic⁽¹³⁾. The most parsimonious model was chosen in case of competing models. Each model was fitted to the most recent period of available data, starting with the last 10 years. When model did not fit, data was sequentially restricted to 9, 8, 7, 6 and 5 last years, respectively, until model had adequate goodness of fit. The predicted number of deceased cases was obtained projecting mortality time trends of the selected model into the future and using the projected population. The following age groups were used in the process: group 35-49 (35-39, 40-44, 45-49), group 50-69 (50-54, 55-59, 60-64, 65-69), group ≥ 70 (70-74, 75-79, 80-84, ≥ 85) and group ≥ 35 (all age groups).

2. Assessment of the differences in the total number of cases between year 2005 and 2015

The method for splitting the differences due to risk and demographic factors (population size and structure) assumes that there are two groups¹⁵: the baseline (or reference) group and the comparison group. The number of incident or death cases has been observed for both groups, and it is aggregated by age-groups (usually 5-year band). The observed difference in the total number of cases or deaths between both groups can then be split into three components: cases due to changes in population size, cases due to changes in population structure (age distribution) and, lastly, the rest of cases attributable to changes on the risk to develop or die due to the disease itself. Let us assume that we have observed k age groups for each one of the two populations to compare which are indexed by $i \in \{1, \dots, k\}$ and $j \in \{1, 2\}$, respectively. In each age group of these populations we have C_{ij} cases and P_{ij} person-years at risk, where $C_j = \sum_i C_{ij}$ and $P_j = \sum_i P_{ij}$. Assume that we have C_1 and C_2 cases and a total population of P_1 and P_2 , respectively. The relative difference between C_1 and C_2 can be expressed as

$$\frac{C_2 - C_1}{C_1} = \frac{C_2}{C_1} - 1 \quad (2.1),$$

which is the quantity to be analysed. Note that to eliminate the effect of the population size on the differences between the number of cases, crude rates (per 100,000 people) have to be first calculated.

Let us define the rates,

$$S_2 = \frac{C_2}{P_2} \quad (2.2),$$

and

$$S_1 = \frac{C_1}{P_1} \quad (2.3),$$

then using (2.1-2.3) we have

$$\frac{C_{.2}}{C_{.1}} - 1 = \frac{S_2 P_2}{S_1 P_1} - 1 = \left[\frac{S_2}{S_1} - 1 \right] + \left[\frac{P_2}{P_1} - 1 \right] \frac{S_2}{S_1} \quad (2.4).$$

The quantity

$$\left[\frac{S_2}{S_1} - 1 \right] = \frac{S_2 - S_1}{S_1} \quad (2.5)$$

will be split into a component due to differences in risk and another component due to differences in population structure, whereas

$$\left[\frac{P_2}{P_1} - 1 \right] \frac{S_2}{S_1} \quad (2.6)$$

reflects the effect of the change in the population size on the crude rates.

Let us define $\lambda_{ij} = \frac{C_{ij}}{P_{ij}}$ as the rate in the age group i in the population j , and

$w_{ij} = \frac{P_{ij}}{P_j}$ as the proportion (weight) of population in the age group i and the population j . Let us also consider

$$S_3 = \sum_i w_{i2} \lambda_{i1} \quad (2.7)$$

to be an intermediate rate, which is a standardized rate in the population 1 with weights of population proportion in population 2. Using (2.5) and (2.7) and

$\frac{S_2 - S_1}{S_1} = \frac{S_3 - S_1}{S_1} + \frac{S_2 - S_3}{S_1}$ we have the following relation

$$\frac{\sum_i w_{i2} \lambda_{i2} - \sum_i w_{i1} \lambda_{i1}}{\sum_i w_{i1} \lambda_{i1}} = \frac{\sum_i w_{i2} \lambda_{i1} - \sum_i w_{i1} \lambda_{i1}}{\sum_i w_{i1} \lambda_{i1}} + \frac{\sum_i w_{i2} \lambda_{i2} - \sum_i w_{i2} \lambda_{i1}}{\sum_i w_{i1} \lambda_{i1}} = U + R \quad (2.8).$$

The first component of the right hand-side of (2.8) represents the proportion of the difference in the crude rate between populations 1 and 2 due to differences of population structure and the second component represents the proportion due to differences in risk.

Using (2.8) and (2.6) we could decompose (2.4) into

$$\frac{C_{.2}}{C_{.1}} - 1 = \left[\frac{S_2 - S_1}{S_1} \right] + \left[\frac{P_2}{P_1} - 1 \right] \frac{S_2}{S_1} \Leftrightarrow Z = U + R + Q \quad (2.9),$$

where Z is the percent net change, U is the proportion of Z due to population structure, R is the proportion of Z due to risk itself and Q is the proportion of Z due to population size.

The basis of the method to obtain (2.9) requires a two-step procedure based on the crude rates:

- 1) Using (2.6) we have the effect of changes in population size
- 2) Using (2.8), we compute the crude rates and estimate differences on the net change of these crude rates due to population structure and differences in risk

Bibliografia

- M. D. Althuis, J. M. Dozier, W. F. Anderson, S. S. Devesa, and L. A. Brinton. Global trends in breast cancer incidence and mortality 1973-1997. *Int J Epidemiol*, 34(2): 405-12, 2005.
- P. Autier, F. Shannoun, A. Scharpantgen, C. Lux, C. Back, G. Severi, S. Steil, and D. Hansen-Koenig. A breast cancer screening programme operating in a liberal health care system: the luxembourg mammography programme, 1992-1997. *Int J Cancer*, 97(6):828-32, 2002.
- L. Balducci and W. B. Ershler. Cancer and ageing: a nexus at several levels. *Nat Rev Cancer*, 5(8):655-62, 2005.
- J. A. Baron, P. A. Newcomb, M. P. Longnecker, R. Mittendorf, B. E. Storer, R. W. Clapp, G. Bogdan, and J. Yuen. Cigarette smoking and breast cancer. *Cancer Epidemiol Biomarkers Prev*, 5(5):399-403, 1996.
- S. Bashir and J. Esteve. Analysing the difference due to risk and demographic factors for incidence or mortality. *Int J Epidemiol*, 29(5):878-84, 2000.
- I. C. Bennett, M. Gattas, and B. T. Teh. The genetic basis of breast cancer and its clinical implications. *Aust N Z J Surg*, 69(2):95-105, 1999.
- C. S. Berkey, H. R. Rockett, M. W. Gillman, and G. A. Colditz. One-year changes in activity and in inactivity among 10- to 15-year-old boys and girls: relationship to change in body mass index. *Pediatrics*, 111(4 Pt 1):836-43, 2003.
- L. Bernstein. Epidemiology of endocrine-related risk factors for breast cancer. *J Mammary Gland Biol Neoplasia*, 7(1):3-15, 2002.
- F. Berrino. The eurocare study: strengths, limitations and perspectives of population-based, comparative survival studies. *Ann Oncol*, 14 Suppl 5:v9-13, 2003.
- F. Berrino, R. De Angelis, M. Sant, S. Rosso, M. Bielska-Lasota, J. W. Coebergh, and M. Santaquilani. Survival for eight major cancers and all cancers combined for european adults diagnosed in 1995-99: results of the eurocare-4 study. *Lancet Oncol*, 8(9):773-83, 2007.
- D. A. Berry, K. A. Cronin, S. K. Plevritis, D. G. Fryback, L. Clarke, M. Zelen, J. S. Mandelblatt, A. Y. Yakovlev, J. D. Habbema, and E. J. Feuer. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med*, 353(17):1784-92, 2005.

- A. Boguseviciute, A. Dirzyte, and A. Bogusevicius. [motivation to participate in mammography screening program for breast cancer]. *Medicina (Kaunas)*, 39(10):1007–15, 2003.
- J.M. Borrás, A. Ameijide, L. Vilardell, J. Valls, R. Marcos-Gragera, and A. Izquierdo. Trends in cancer incidence in catalonia, 1985-2002. *Med Clin*, 131(Suppl):11–18, 2008.
- J. L. Botha, F. Bray, R. Sankila, and D. M. Parkin. Breast cancer incidence and mortality trends in 16 european countries. *Eur J Cancer*, 39(12):1718–29, 2003.
- P. Boyle and D. M. Parkin. Cancer registration: principles and methods. statistical methods for registries. *IARC Sci Publ*, (95):126–58, 1991.
- F. Bray and B. Moller. Predicting the future burden of cancer. *Nat Rev Cancer*, 6(1):63–74, 2006.
- F. Bray, A. Guilloux, R. Sankila, and D. M. Parkin. Practical implications of imposing a new world standard population. *Cancer Causes Control*, 13(2):175–82, 2002.
- N. E. Breslow and N. E. Day. Statistical methods in cancer research. volume ii—the design and analysis of cohort studies. *IARC Sci Publ*, (82):1–406, 1987.
- L. A. Brinton, L. Bernstein, and G. A. Colditz. Summary of the workshop: Workshop on physical activity and breast cancer, november 13-14, 1997. *Cancer*, 83(3 Suppl):595–9, 1998.
- P. Bruzzi, E. Negri, C. La Vecchia, A. Decarli, D. Palli, F. Parazzini, and M. R. Del Turco. Short term increase in risk of breast cancer after full term pregnancy. *BMJ*, 297(6656):1096–8, 1988.
- H. J. Burstein, K. Polyak, J. S. Wong, S. C. Lester, and C. M. Kaelin. Ductal carcinoma in situ of the breast. *N Engl J Med*, 350(14):1430–41, 2004.
- C.L. Chiang. *Introduction to Stochastic Processes in Biostatistics*. Wiley, New York, 1968.
- R. Cleries, J. Ribes, L. Esteban, J. M. Martínez, and J. M. Borrás. Time trends of breast cancer mortality in spain during the period 1977-2001 and bayesian approach for projections during 2002-2016. *Ann Oncol*, 17(12):1783–91, 2006.
- R. Cleries, J. Valls, J. Ribes, J.M. Borrás, T. Hakulinen, and T. Dyba. Differences due to risk and demographic factors for the projections and burden of breast cancer mortality in europe from 2005 to 2015. 2010, submitted.
- G. A. Colditz and B. Rosner. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the nurses’ health study. *Am J Epidemiol*, 152(10):950–64, 2000.
- G. A. Colditz, H. J. Baer, and R. M. Tamimi. *Breast cancer*. Cancer Epidemiology and prevention. Oxford, New York, 2006a.
- G. A. Colditz, T. A. Sellers, and E. Trapido. Epidemiology - identifying the causes and preventability of cancer? *Nat Rev Cancer*, 6(1):75–83, 2006b.

- M. P. Coleman, J. Esteve, P. Damiecki, A. Arslan, and H. Renard. Trends in cancer incidence and mortality. *IARC Sci Publ*, (121):1–806, 1993.
1996. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. collaborative group on hormonal factors in breast cancer. *Lancet*, 347(9017):1713–27, 1996.
1997. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. collaborative group on hormonal factors in breast cancer. *Lancet*, 350(9084):1047–59, 1997.
2002. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet*, 360(9328):187–95, 2002.
- C. Dimitrakaki, D. Boulamatsis, A. Mariolis, N. Kontodimopoulos, D. Niakas, and Y. Tountas. Use of cancer screening services in greece and associated social factors: results from the nation-wide hellas health i survey. *Eur J Cancer Prev*, 18(3):248–57, 2009.
- T. Dyba. Correction. *Stat Med*, 19(9):1251, 2000.
- T. Dyba and T. Hakulinen. Comparison of different approaches to incidence prediction based on simple interpolation techniques. *Stat Med*, 19(13):1741–52, 2000.
- T. Dyba and T. Hakulinen. Do cancer predictions work? *Eur J Cancer*, 44(3):448–53, 2008.
- T. Dyba, T. Hakulinen, and L. Paivarinta. A simple non-linear model in incidence prediction. *Stat Med*, 16(20):2297–309, 1997.
- D. Easton, D. Ford, and J. Peto. Inherited susceptibility to breast cancer. *Cancer Surv*, 18:95–113, 1993.
- J. Einhorn. Cancer by the year 2000. educational requirements for future oncologists. *Acta Oncol*, 28(5):723–8, 1989.
- M. F. Ernst, L. V. van de Poll-Franse, J. A. Roukema, J. W. Coebergh, C. M. van Gestel, G. Vreugdenhil, M. J. Louwman, and A. C. Voogd. Trends in the prognosis of patients with primary metastatic breast cancer diagnosed between 1975 and 2002. *Breast*, 16(4):344–51, 2007.
- J. Esteve, E. Benhamou, and L. Raymond. Statistical methods in cancer research. volume iv. descriptive epidemiology. *IARC Sci Publ*, (128):1–302, 1994.
- J. Ferlay, F. Bray, P. Pisani, and D.M. Parkin. Globocan 2002: Cancer incidence, mortality and prevalence worldwide. *IARC CancerBase No. 5. version 2.0*, page IARCPress, 2004.

- J. Ferlay, P. Autier, M. Boniol, M. Heanue, M. Colombet, and P. Boyle. Estimates of the cancer incidence and mortality in europe in 2006. *Ann Oncol*, 18(3):581–92, 2007.
- B. Fisher, J. P. Costantino, D. L. Wickerham, C. K. Redmond, M. Kavanah, W. M. Cronin, V. Vogel, A. Robidoux, N. Dimitrov, J. Atkins, M. Daly, S. Wieand, E. Tan-Chiu, L. Ford, and N. Wolmark. Tamoxifen for prevention of breast cancer: report of the national surgical adjuvant breast and bowel project p-1 study. *J Natl Cancer Inst*, 90(18):1371–88, 1998.
- C. M. Friedenreich. Review of anthropometric factors and breast cancer risk. *Eur J Cancer Prev*, 10(1):15–32, 2001.
- Jr. Gambrell, R. D., R. C. Maier, and B. I. Sanders. Decreased incidence of breast cancer in postmenopausal estrogen-progestogen users. *Obstet Gynecol*, 62(4):435–43, 1983.
- J. Garcia, A. Jemal, EM. Ward, MM. Center, Y. Hao, RL. Siegel, and MJ. Thun. Global cancer facts and figures 2007. *American Cancer Society*, 2007.
- L. Giordano, D. Giorgi, P. Piccini, L. Ventura, V. Stefanini, C. Senore, E. Paci, and N. Segnan. Time trends of process and impact indicators in italian mammography screening programmes–1996-2004. *Epidemiol Prev*, 31(2-3 Suppl 2):21–32, 2007.
- B. R. Goldin, H. Adlercreutz, S. L. Gorbach, J. H. Warram, J. T. Dwyer, L. Swenson, and M. N. Woods. Estrogen excretion patterns and plasma levels in vegetarian and omnivorous women. *N Engl J Med*, 307(25):1542–7, 1982.
- G. P. Gupta and J. Massague. Cancer metastasis: building a framework. *Cell*, 127(4):679–95, 2006.
- T. Hakulinen and T. Dyba. Precision of incidence predictions based on poisson distributed observations. *Stat Med*, 13(15):1513–23, 1994.
- T. Hakulinen and M. Hakama. Predictions of epidemiology and the evaluation of cancer control measures and the setting of policy priorities. *Soc Sci Med*, 33(12):1379–83, 1991.
- J. R. Harris, M. E. Lippman, U. Veronesi, and W. Willett. Breast cancer. *N Engl J Med*, 327(5):319–28, 1992.
- R. E. Hendrick, C. Klabunde, A. Grivegne, G. Pou, and R. Ballard-Barbash. Technical quality control practices in mammography screening programs in 22 countries. *Int J Qual Health Care*, 14(3):219–26, 2002.
- C. Hermon and V. Beral. Breast cancer mortality rates are levelling off or beginning to decline in many western countries: analysis of time trends, age-cohort and age-period models of breast cancer mortality in 20 countries. *Br J Cancer*, 73(7):955–60, 1996.
- K. Hertl, M. Primic-Zakelj, J. Zgajnar, and I. Kocijancic. Performance of opportunistic breast cancer screening in slovenia. *Neoplasma*, 53(3):237–41, 2006.
- C. Hery, J. Ferlay, M. Boniol, and P. Autier. Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with caucasian-majority populations. *Ann Oncol*, 19(6):1187–94, 2008a.

- C. Hery, J. Ferlay, M. Boniol, and P. Autier. Changes in breast cancer incidence and mortality in middle-aged and elderly women in 28 countries with caucasian majority populations. *Ann Oncol*, 19(5):1009–18, 2008b.
- S. Hofvind, B. Geller, P. M. Vacek, S. Thoresen, and P. Skaane. Using the european guidelines to evaluate the norwegian breast cancer screening program. *Eur J Epidemiol*, 22(7):447–55, 2007.
- T. R. Holford. The estimation of age, period and cohort effects for vital rates. *Biometrics*, 39(2):311–24, 1983.
- L. Holmberg and H. Anderson. Habits (hormonal replacement therapy after breast cancer—is it safe?), a randomised comparison: trial stopped. *Lancet*, 363(9407):453–5, 2004.
- L. Holmberg, O. E. Iversen, C. M. Rudenstam, M. Hammar, E. Kumpulainen, J. Jaskiewicz, J. Jassem, D. Dobaczewska, H. E. Fjosne, O. Peralta, R. Arriagada, M. Holmqvist, and J. Maenpaa. Increased risk of recurrence after hormone replacement therapy in breast cancer survivors. *J Natl Cancer Inst*, 100(7):475–82, 2008.
- G. R. Howe, T. Hirohata, T. G. Hislop, J. M. Iscovich, J. M. Yuan, K. Katsouyanni, F. Lubin, E. Marubini, B. Modan, T. Rohan, and et al. Dietary factors and risk of breast cancer: combined analysis of 12 case-control studies. *J Natl Cancer Inst*, 82(7):561–9, 1990.
- M. Hu, J. Yao, D. K. Carroll, S. Weremowicz, H. Chen, D. Carrasco, A. Richardson, S. Violette, T. Nikolskaya, Y. Nikolsky, E. L. Bauerlein, W. C. Hahn, R. S. Gelman, C. Allred, M. J. Bissell, S. Schnitt, and K. Polyak. Regulation of in situ to invasive breast carcinoma transition. *Cancer Cell*, 13(5):394–406, 2008.
- D. J. Hunter, D. Spiegelman, H. O. Adami, L. Beeson, P. A. van den Brandt, A. R. Folsom, G. E. Fraser, R. A. Goldbohm, S. Graham, G. R. Howe, and et al. Cohort studies of fat intake and the risk of breast cancer—a pooled analysis. *N Engl J Med*, 334(6):356–61, 1996.
- International Agency for Research on Cancer IARC. Cancer incidence in five continents. volume viii. *IARC Sci Publ*, (155):1–781, 2002a.
- International Agency for Research on Cancer IARC. *IARC Handbooks of Cancer Prevention. Breast Cancer Screening*. Lyon, 2002b.
- I. Jatoi and A. B. Miller. Why is breast-cancer mortality declining? *Lancet Oncol*, 4(4):251–4, 2003.
- A. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, C. Smigal, and M. J. Thun. Cancer statistics, 2006. *CA Cancer J Clin*, 56(2):106–30, 2006.
- J. L. Kelsey, M. D. Gammon, and E. M. John. Reproductive factors and breast cancer. *Epidemiol Rev*, 15(1):36–47, 1993.
- A. Kovacs, L. Dobrossy, A. Budai, I. Boncz, and A. Cornides. [the state of the organised mammography screening in hungary in 2006]. *Orv Hetil*, 148(19):879–85, 2007.

- C. Lassarre, S. Hardouin, F. Daffos, F. Forestier, F. Frankenne, and M. Binoux. Serum insulin-like growth factors and insulin-like growth factor binding proteins in the human fetus. relationships with growth in normal subjects and in subjects with intrauterine growth retardation. *Pediatr Res*, 29(3):219–25, 1991.
- M. P. Longnecker. Alcoholic beverage consumption in relation to risk of breast cancer: meta-analysis and review. *Cancer Causes Control*, 5(1):73–82, 1994.
- N. Mantel and W. Haenszel. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*, 22(4):719–48, 1959.
- P. McCullagh and J.A. Nelder. *Generalized Linear Models*. Chapman and Hall, London, 1989.
- S. H. Moolgavkar, N. E. Day, and R. G. Stevens. Two-stage model for carcinogenesis: Epidemiology of breast cancer in females. *J Natl Cancer Inst*, 65(3):559–69, 1980.
- J. D. Otten, M. J. Broeders, J. Fracheboud, S. J. Otto, H. J. de Koning, and A. L. Verbeek. Impressive time-related influence of the dutch screening programme on breast cancer incidence and mortality, 1975-2006. *Int J Cancer*, 123(8):1929–34, 2008.
- D. M. Parkin and L. M. Fernandez. Use of statistics to assess the global burden of breast cancer. *Breast J*, 12 Suppl 1:S70–80, 2006.
- D. Payne. Ireland announces national breast screening programme. *BMJ*, 318(7190):1025, 1999.
- R. Peto, J. Boreham, M. Clarke, C. Davies, and V. Beral. Uk and usa breast cancer deaths down 25years. *Lancet*, 355(9217):1822, 2000.
- I. E. Petrakis and S. Paraskakis. Breast cancer in the elderly. *Arch Gerontol Geriatr*, 2009.
- D. L. Preston, H. Kato, K. Kopecky, and S. Fujita. Studies of the mortality of a-bomb survivors. 8. cancer mortality, 1950-1982. *Radiat Res*, 111(1):151–78, 1987.
- P. Prior, C. B. Woodman, S. Wilson, and A. G. Threlfall. Reliability of underlying incidence rates for estimating the effect and efficiency of screening for breast cancer. *J Med Screen*, 3(3):119–22, 1996.
- Development Core Team R. R: a language and environment for statistical computing, 2007.
- J. Ribes, R. Cleries, M. Buxo, A. Ameijide, J. Valls, and R. Gispert. Predictions of cancer incidence and mortality in catalonia to 2015 by means of bayesian models. *Med Clin*, 131(Suppl):32–41, 2008.
- L.A.G Ries, M.P. Eisner, C.L. Kosary, and Et al. *Seer Cancer Statistics Review 1975-2002*. National Cancer Insitute, Bethesda, 2005.
- I. Romieu, M. Hernandez-Avila, E. Lazcano, L. Lopez, and R. Romero-Jaime. Breast cancer and lactation history in mexican women. *Am J Epidemiol*, 143(6):543–52, 1996.

- B. Rosner, G. A. Colditz, and W. C. Willett. Reproductive risk factors in a prospective study of breast cancer: the nurses' health study. *Am J Epidemiol*, 139(8):819–35, 1994.
- M. Sant, C. Allemani, M. Santaquilani, A. Knijn, F. Marchesi, and R. Capocaccia. Eurocare-4. survival of cancer patients diagnosed in 1995-1999. results and commentary. *Eur J Cancer*, 2009.
- P. D. Sasieni and J. Adams. Standardized lifetime risk. *Am J Epidemiol*, 149(9):869–75, 1999.
- C. Schairer, J. Lubin, R. Troisi, S. Sturgeon, L. Brinton, and R. Hoover. Menopausal estrogen and estrogen-progestin replacement therapy and breast cancer risk. *JAMA*, 283(4):485–91, 2000.
- D. E. Schaubel, H. I. Morrison, M. Desmeules, D. Parsons, and S. S. Fenton. End-stage renal disease projections for canada to 2005 using poisson and markov models. *Int J Epidemiol*, 27(2):274–81, 1998.
- D. Schopper and C. de Wolf. How effective are breast cancer screening programmes by mammography? review of the current evidence. *Eur J Cancer*, 45(11):1916–23, 2009.
- R. Schulz-Wendtland, N. Becker, K. Bock, K. Anders, and W. Bautz. [mammography screening]. *Radiologe*, 47(4):359–69; quiz 370, 2007.
- M. Segi. *Cancer mortality fort selected sites in 24 countries (1950-57)*. Tohoku University School of Public Health, Sendai, 1960.
- S. Shapiro, E. A. Coleman, M. Broeders, M. Codd, H. de Koning, J. Fracheboud, S. Moss, E. Paci, S. Stachenko, and R. Ballard-Barbash. Breast cancer screening programmes in 22 countries: current policies, administration and guidelines. international breast cancer screening network (ibsn) and the european network of pilot projects for breast cancer screening. *Int J Epidemiol*, 27(5):735–42, 1998.
- R. A. Smith, D. Saslow, K. A. Sawyer, W. Burke, M. E. Costanza, 3rd Evans, W. P., Jr. Foster, R. S., E. Hendrick, H. J. Eyre, and S. Sener. American cancer society guidelines for breast cancer screening: update 2003. *CA Cancer J Clin*, 53(3):141–69, 2003.
- S. A. Smith-Warner, D. Spiegelman, S. S. Yaun, H. O. Adami, W. L. Beeson, P. A. van den Brandt, A. R. Folsom, G. E. Fraser, J. L. Freudenheim, R. A. Goldbohm, S. Graham, A. B. Miller, J. D. Potter, T. E. Rohan, F. E. Speizer, P. Toniolo, W. C. Willett, A. Wolk, A. Zeleniuch-Jacquotte, and D. J. Hunter. Intake of fruits and vegetables and risk of breast cancer: a pooled analysis of cohort studies. *JAMA*, 285(6):769–76, 2001.
- A. L. Svendsen, A. H. Olsen, M. von Euler-Chelpin, and E. Lynge. Breast cancer incidence after the introduction of mammography screening: what should be expected? *Cancer*, 106(9):1883–90, 2006.
- B. Szynglarewicz, R. Matkowski, P. Kasprzak, J. Kotowska, J. Forgacz, M. Pudelko, and J. Kornafel. [the effectiveness of population-based breast cancer screening programme]. *Pol Merkur Lekarski*, 26(152):117–20, 2009.

- M. Tekkel, T. Veideman, A. Baburin, and M. Rahu. Use of mammography and pap smear in estonia, a country without organized cancer screening. *Int J Public Health*, 52(2): 109–16, 2007.
- S. Tornberg, L. Kemetli, E. Lynge, A. Helene Olsen, S. Hofvind, H. Wang, A. Anttila, M. Hakama, and L. Nystrom. Breast cancer incidence and mortality in the nordic capitals, 1970-1998. trends related to mammography screening programmes. *Acta Oncol*, 45(5):528–35, 2006.
- D. Trichopoulos, B. MacMahon, and P. Cole. Menopause and breast cancer risk. *J Natl Cancer Inst*, 48(3):605–13, 1972.
- J. E. Tyczynski, I. Plesko, T. Aareleid, M. Primic-Zakelj, M. Dalmas, J. Kurtinaitis, A. Stengrevics, and D. M. Parkin. Breast cancer mortality patterns and time trends in 10 new eu member states: mortality declining in young women, but still increasing in the elderly. *Int J Cancer*, 112(6):1056–64, 2004.
- M. A. Unger and B. L. Weber. Recent advances in breast cancer biology. *Curr Opin Oncol*, 12(6):521–5, 2000.
- G. Ursin, M. P. Longnecker, R. W. Haile, and S. Greenland. A meta-analysis of body mass index and risk of premenopausal breast cancer. *Epidemiology*, 6(2):137–41, 1995.
- J. Valls, R. Cleries, J. Galvez, V. Moreno, R. Gispert, J. M. Borrás, and J. Ribes. Riskdiff: a web tool for the analysis of the difference due to risk and demographic factors for incidence or mortality data. *BMC Public Health*, 9:473, 2009.
- P. A. van den Brandt, D. Spiegelman, S. S. Yaun, H. O. Adami, L. Beeson, A. R. Folsom, G. Fraser, R. A. Goldbohm, S. Graham, L. Kushi, J. R. Marshall, A. B. Miller, T. Rohan, S. A. Smith-Warner, F. E. Speizer, W. C. Willett, A. Wolk, and D. J. Hunter. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol*, 152(6):514–27, 2000.
- U. Veronesi, P. Boyle, A. Goldhirsch, R. Orecchia, and G. Viale. Breast cancer. *Lancet*, 365(9472):1727–41, 2005.
- C. Vutuc, T. Waldhoer, and G. Haidinger. Breast cancer trends: opportunistic screening in austria versus controlled screening in finland and sweden. *Eur J Cancer Prev*, 15(4): 343–6, 2006.
- S Weisberg. *Applied Linear Regression*. Wiley, New York, 1985.
- W. C. Willett. Diet and breast cancer. *J Intern Med*, 249(5):395–411, 2001.
- X. F. Yang, W. G. Beamer, H. Huynh, and M. Pollak. Reduced growth of human breast cancer xenografts in hosts homozygous for the lit mutation. *Cancer Res*, 56(7):1509–11, 1996.
- S. M. Zhang, W. C. Willett, J. Selhub, D. J. Hunter, E. L. Giovannucci, M. D. Holmes, G. A. Colditz, and S. E. Hankinson. Plasma folate, vitamin b6, vitamin b12, homocysteine, and risk of breast cancer. *J Natl Cancer Inst*, 95(5):373–80, 2003.