

Grau en Estadística

Títol: Estimació de la incidència del càncer a totes les províncies d'Espanya

Autor: Andrea Elías Vidal

Codirectors: Alberto Ameijide i Xavier Puig

Departament: Estadística i Investigació Operativa (UPC)

Convocatòria: 29 de juny de 2018



UNIVERSITAT DE
BARCELONA



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Facultat de Matemàtiques i Estadística

RESUM

El càncer és una de les malalties que més afecta actualment tant a la població catalana com espanyola, sent aquesta la segona causa de mort més comuna. En l'àmbit estadístic, l'estudi d'aquesta malaltia parteix d'una sèrie d'indicadors epidemiològics que permeten mesurar-ne l'impacte que genera el càncer en la societat, així com observar quina va sent la seva evolució al llarg del temps.

D'aquests indicadors epidemiològics, el present treball se centra en l'anàlisi de la incidència, indicador que permet mesurar el nombre de nous casos que es registren sobre una població en una determinada zona i moment.

La incidència del càncer en un territori és una dada desconeguda a excepció d'alguns territoris on existeixen Registres de Càncer de Base Poblacional (RCBP). L'objectiu d'aquest estudi és estimar la incidència del càncer a Espanya a partir de les dades d'incidència i mortalitat conegudes i validar la idoneïtat d'aquest estimador.

A partir de les dades d'incidència d'aquelles províncies i períodes en els que es disposen de dades dels registres de càncer de base poblacional i dades de mortalitat per aquelles mateixes zones i períodes, s'ha fet l'estimació de la raó entre la incidència i mortalitat (RIM) pel període 1994-2007, mitjançant models lineals generalitzats mixtes bayesians per a cada tipus de tumor i sexe.

S'ha estimat al nombre de casos incidents durant el període 2009-2013 a les províncies de Tarragona i Girona en tres escenaris diferents que representaven diferents hipòtesis de la evolució del RIM, constant lineal i quadràtica, i s'han comparat i validat aquestes estimacions amb les dades reals de Tarragona i Girona per al mateix període seleccionant l'escenari que millor ajustava.

A partir del escenari triats, s'ha estimat la incidència per cada una de les províncies espanyoles pel període 2009-2013. En el cas dels homes dels 22 tipus de tumors estudiats, 8 han presentat estimacions no acceptables mentre que en el cas de les dones dels 24 tipus estudiats, 5 han presentat estimacions no acceptables.

La conclusió final és que el model estudiat produeix estimacions acceptables en la majoria dels tumors, però té certes limitacions que poden produir estimacions no acceptables en determinades situacions, on seria necessari seguir estudiant nous mètodes i/o models d'estimacions que presentin uns valors de casos incidents més acurats.

Paraules clau: epidemiologia, *incidència*, *registres de càncer de base poblacional*, *control*, *prevenció*, *model jeràrquic bayesià*

Classificació AMS: 60J22, 62F15, 62-07, 62P10, 62Q99

ABSTRACT

Cancer is one of the diseases that most affects both the Catalan and Spanish population, and this is the second most common cause of death. In the statistical field, the study of this disease presents a set of epidemiological indicators that allow us to measure the impact that cancer causes in society, as well as to observe the evolution of the disease over time.

Of these epidemiological indicators, the present project focuses on the analysis of incidence, an indicator that allows to measure the number of new cases that are recorded on a population in a specific area and time.

The incidence of cancer in one area is unknown, with the exception of some territories where there are Population-Based Cancer Registries (RCBP). The objective of this study is to estimate the incidence of cancer in Spain based on known incidence and mortality data and validate the suitability of this estimator.

From the incidence data of those provinces and periods in which data on population-based cancer registers and mortality data are available for those same zones and periods, the estimation of the ratio between incidence and mortality (RIM) for the period 1994-2007 it's done using Bayesian mixed generalized linear models for each type of tumor and sex.

The number of incidents in the period 2009-2013 has been estimated in the provinces of Tarragona and Girona in three different scenarios that represented different hypotheses of the evolution of RIM, linear and quadratic constant, and these estimates have been compared and validated with the real data of Tarragona and Girona for the same period, selecting the setting that best adjusted.

Based on the scenarios chosen, the incidence was estimated for each of the Spanish provinces during the period 2009-2013. In the case of men from the 22 types of tumors studied, 8 have presented unacceptable estimates while in the case of women of the 24 types studied, 5 have presented non-acceptable estimates.

The final conclusion is that the model studied produces acceptable estimates in most tumors, but it has certain limitations that can produce estimates not acceptable in certain situations, where it would be necessary to continue studying new methods and / or estimation models that present a more accurate incident case values.

Keywords: *epidemiology, incidence, population-based cancer registries, control, prevention, hierarchical bayesian model*

ÍNDEX

1	INTRODUCCIÓ	1
1.1	CÀNCER I SALUT PÚBLICA	1
1.2	INDICADORS PRINCIPALS	2
1.3	REGISTRES DE CàNCER DE BASE POBLACIONAL (RCBP)	2
1.4	REGISTRES DE CàNCER A ESPANYA	4
1.5	REGISTRES DE CàNCER A CATALUNYA	5
1.6	DESCRIPCIÓ DEL PROJECTE	6
1.7	OBJECTIUS	8
1.8	PROGRAMARI	8
2	DESCRIPCIÓ DE LES DADES	9
2.1	INDICADORS PER AL CONTROL DEL CàNCER	9
2.1.1	<i>Incidència</i>	9
2.1.2	<i>Mortalitat</i>	9
2.1.3	<i>Supervivència</i>	10
2.1.4	<i>Prevalença</i>	10
2.2	METODOLOGIA DELS REGISTRES DE CàNCER DE CATALUNYA	10
2.2.1	<i>Aspectes geogràfics, demogràfics i temporals</i>	11
2.2.2	<i>Definició de cas</i>	12
2.2.3	<i>Variables que monitoritzen els Registre de Càncer de Tarragona i Girona</i>	13
2.3	FONT I VARIABLES DE LES BASES DE DADES	13
2.3.1	<i>Població</i>	14
2.3.2	<i>Mortalitat</i>	14
2.3.3	<i>Incidència</i>	14
2.3.4	<i>Codis identificadors per als tumors</i>	15
3	ANÀLISI DESCRIPTIVA DE LES DADES	17
3.1	POBLACIÓ	18
3.2	INCIDÈNCIA	19
3.3	MORTALITAT	22
4	MODELITZACIÓ ESTADÍSTICA BAYESIANA	25
4.1	INTRODUCCIÓ	25
4.2	INFERÈNCIA BAYESIANA	26
4.3	COMPUTACIÓ BAYESIANA	27
5	PRESENTACIÓ DEL MODEL	28
5.1	DESCRIPCIÓ DEL MODEL	28
5.2	OBSERVACIONS DEL MODEL	31
5.3	ESTIMACIÓ DE LA RIM PER AL PERÍODE 1994-2007	32
6	DEFINICIÓ I SELECCIÓ DEL MILLOR ESCENARI	35
6.1	DIFERENTS ESCENARIS PER L'EVOLUCIÓ TEMPORAL DE LA RIM	35
6.2	SELECCIÓ DEL MILLOR ESCENARI	37
7	ESTIMACIÓ DE LA INCIDÈNCIA PER A TOTES LES PROVÍNCIES	42
8	CONCLUSIONS	47
	BIBLIOGRAFIA	49
	ANNEX 1. TAXES BRUTES D'INCIDÈNCIA	53
	ANNEX 2. TAXES BRUTES DE MORTALITAT	54
	ANNEX 3. EFECTE DE L'ANY SOBRE LA RIM	55

ANNEX 4. EFECTE DE L'EDAT SOBRE LA RIM	58
ANNEX 5. TAXES BRUTES I MAPES D'ESPANYA	61
ANNEX 6. CODI D'R	71

1 INTRODUCCIÓ

1.1 CÀNCER I SALUT PÚBLICA

El càncer és un dels principals problemes de salut tant a Catalunya com a Espanya. S'estima que, a Catalunya, la mitjana anual de nous casos de càncer en el període 1998-2002 fou de 29.062 [1] i de 36.048 casos a l'any 2010 [2]. Paral·lelament, el càncer va ser responsable de 16.888 defuncions a l'any 2015 [2], essent la segona causa de mort en la població catalana després de les malalties de l'aparell circulatori.

Es tracta de xifres que malgrat el pas dels anys i els avenços en la tecnologia no s'han aconseguit de reduir. En el futur milers de pacients a Catalunya i a Espanya continuaran amb la necessitat de tractaments pels nous càncers diagnosticats, o bé per recurrències de càncers ja diagnosticats i tractats prèviament, i alhora, caldran recursos tant per a la recerca en general com per al tractament i suport d'aquests mateixos pacients.

Tenint en compte que molts càncers es desenvolupen durant anys o dècades, qualsevol efecte de les mesures de control que s'apliquin a dia d'avui, reflectiran reduccions en la incidència i/o mortalitat d'aquí a cinc, deu o més anys a partir d'ara. En aquest aspecte, és molt important les mesures de prevenció i control que s'estableixen sobre una determinada malaltia.

A nivell individual, el càncer és una malaltia crònica i progressiva que molts cops requereix un control a llarg termini per avaluar l'efectivitat dels esforços per controlar-lo. A nivell comunitari, el principal objectiu és disminuir els nous casos diagnosticats i la mortalitat, i l'avaluació d'aquests paràmetres s'ha de realitzar sempre amb una perspectiva a llarg termini.

Tot i que el risc de patir un càncer és estable, l'envelliment de la població provoca que el nombre de nous casos de càncer augmenti tant a Catalunya com a Espanya, però gràcies als avenços en un diagnòstic precoç i els tractaments que es donen, s'ha aconseguit reduir la mortalitat en aquells tipus de tumors que se'ls ha destinat la suficient investigació. El cost econòmic per al diagnòstic, el tractament, el suport i l'atenció a llarg termini del càncer és immens i el cost emocional incalculable.

Per a l'estudi d'aquesta malaltia, hi ha una sèrie d'indicadors principals que permeten mesurar la seva magnitud així com l'evolució al llarg del temps. Aquests indicadors són la incidència, la mortalitat, la supervivència i la prevalença. La combinació de tots quatre permet establir certes mesures de control dirigides a reduir la magnitud del càncer, a millorar la supervivència i també la qualitat de vida dels pacients amb aquest diagnòstic. L'eina fonamental que permet disposar d'aquests indicadors són els registres de càncer de base poblacional.

1.2 INDICADORS PRINCIPALS

La incidència fa referència al nombre de nous casos de càncer durant un període especificat en una població determinada. Es tracta d'una mesura del grau d'exposició d'una població als factors de risc d'aquesta malaltia. La incidència és la millor mesura per monitoritzar les tendències del risc de càncer.

La mortalitat es defineix com el nombre de defuncions per càncer durant un període especificat de temps en una població determinada. És una mesura fonamental d'impacte del càncer en la comunitat, ja que representa la quantificació dels casos que han tingut més mal pronòstic i han provocat la mort del pacient.

La supervivència es defineix com la proporció de malalts amb aquest diagnòstic que sobreviuen més d'un temps determinat. Les taxes de supervivència són els indicadors més directes de la gravetat del càncer i de l'impacte que té el tractament.

Per últim, la prevalença és la proporció d'una població que està afecte de càncer en un moment donat del temps. La prevalença inclou totes les persones que, havent estat diagnosticades en qualsevol moment del passat continuen vivint, estiguin curades o no. Proporciona una informació molt útil per a la quantificació i l'estimació de la càrrega assistencial actual que impliquen els pacients, curats o no, ja sigui en forma de tractament inicial com en forma de seguiment i/o rehabilitació.

1.3 REGISTRES DE CÀNCER DE BASE POBLACIONAL (RCBP)

Els registres de càncer de base poblacional són sistemes d'informació que recullen, tracten i emmagatzemen de forma continuada i sistemàtica dades de tots els casos nous de càncer d'una població ben definida. Aquests registres introdueixen i conserven informació de forma homogènia sobre els nous casos de càncer, l'actualitzen i estan en condicions de reutilitzar-la [3].

La utilització d'aquestes dades mitjançant la producció d'estadístiques útils per conèixer les característiques dels càncers i la seva evolució en el temps en una àrea determinada és un dels principals objectius d'aquests sistemes.

Els registres de càncer de base poblacional de diferents llocs de la geografia mundial segueixen una àmplia llista de principis, criteris i normes de funcionament que permeten obtenir uns resultats comparables [4, 5].

La gran majoria de registres poblacionals de càncer són generals, és a dir, que recullen informació de totes les localitzacions tumorals, però alguns estan limitats solament a alguna característica en funció de l'edat (registres pediàtrics) o el tipus tumoral (registres monogràfics). Exemples d'això són els registres de càncer infantil, els registres de tumors digestius i els de tumors hematològics.

Des d'un punt de vista històric el primer registre de càncer com a tal va ser el d'Hamburg (1929), tot i que la seva finalitat era més de tipus clínic que no pas epidemiològic. Poc temps després van aparèixer els primers registres amb finalitats purament epidemiològiques: New York (1940) i Connecticut (1941) als Estats Units d'Amèrica, Dinamarca (1942), Saskatchewan a Canadà (1944), i Bèlgica i Anglaterra a Europa també van iniciar les seves tasques ben aviat [6].

Amb això, al 1950 l'Organització Mundial de la Salut va establir un comitè sobre l'enregistrament dels casos de càncer i aquest va publicar les primeres guies metodològiques per a l'enregistrament del càncer [7].

A partir d'aquest moment els fets van anar succeint ràpidament, es va organitzar per part de la Unió Internacional contra el Càncer (UICC), el "Simposi Internacional sobre Patologia Geogràfica i Demografia del Càncer" [8], que va representar una altra fita, destacant la necessitat d'enumerar tots els nous casos de càncer en una àrea definida. A partir de la base de recomanacions del Simposi, la UICC va establir un Comitè de Patologia Geogràfica, sorgeix el primer volum de la sèrie *Cancer Incidence in Five Continents* [9].

El 1965 es va establir l'Agència Internacional per a la Recerca en Càncer (IARC) com un centre d'investigació especialitzat en càncer de l'Organització Mundial de la Salut i es van implantar nous registres, fenomen que es va estendre molt a partir dels anys 70 del segle XX. Així el nombre de registres membre de l'IACR era de 87 l'any 1979, de 449 al 2006 [10] i de 535 a dia d'avui.

Actualment, la població coberta per registres de càncer varia molt en funció de la regió del món i, en general, els països en vies de desenvolupament estan molt menys coberts que els països amb més riquesa. Tanmateix, els registres de càncer també són necessaris en aquells països menys desenvolupats [11]. Paral·lelament a la seva expansió numèrica, també es van anar ampliant les seves funcions. Inicialment es concentraven en definir els patrons i les tendències de la incidència. Posteriorment, molts van entrar en l'estudi de la supervivència i finalment han abastat un ampli ventall de funcions dirigides a la planificació i l'avaluació de les activitats de control del càncer.

Exemples d'això són el seu ampli ús en l'avaluació i monitoratge de programes de cribratge [12] o els estudis poblacionals que avaluen no solament la supervivència sinó també la qualitat de vida [13].

1.4 REGISTRES DE CÀNCER A ESPANYA

La *Red Española de Registros de Cáncer* (REDECAN), és l'òrgan cooperatiu dels registres de càncer de base poblacional (RCBP) arreu d'Espanya. Comprèn les províncies d'Albacete, Astúries, Canàries, Comunitat Valenciana, Cuenca, Ciutat Real, Girona, Granada, La Rioja, Mallorca, Múrcia, Navarra, País Basc (Alava, Guipúscoa i Vizcaya) Saragossa i Tarragona. La figura 1.1 mostra un mapa que situa les províncies enumerades.

La REDECAN es va crear a finals de l'any 2010 [14] amb l'objectiu d'obtenir i proporcionar a les autoritats sanitàries i a la comunitat científica els resultats sobre incidència, supervivència i prevalença del càncer a Espanya, i també la creació d'una base de dades conjunta que assegurí la qualitat de les dades i la màxima actualització possible de les mateixes.

A Espanya, els dos registres de càncer poblacional més antics són el Registre de Càncer de Saragossa, creat l'any 1960, i el de Navarra, creat el 1970. El 1976 es va crear el Pla Nacional de Registres de Càncer de Població, afegint-se nous registres: d'Astúries, Tenerife, Sevilla i Valladolid, tot i que no tots van mantenir la seva continuïtat al llarg dels anys.

En l'actualitat, existeixen 14 RCBP on les seves dades estan incloses a la publicació de referència *Cancer Incidence in Five Continents* editada per l'IARC, *International Agency for Research on Cancer*, mencionada amb anterioritat. A més, existeix un RCBP monogràfic de mama a Castelló i un registre monogràfic de Càncer Infantil a la Comunitat Valenciana.

En el seu conjunt, es cobreix una població aproximada de 10 milions, xifra que representa el 27% de la població espanyola. Pel que fa a països del Sud d'Europa, propers en geografia i cultura, com per exemple França o Itàlia, els seus registres cobreixen aproximadament la meitat de la població total.

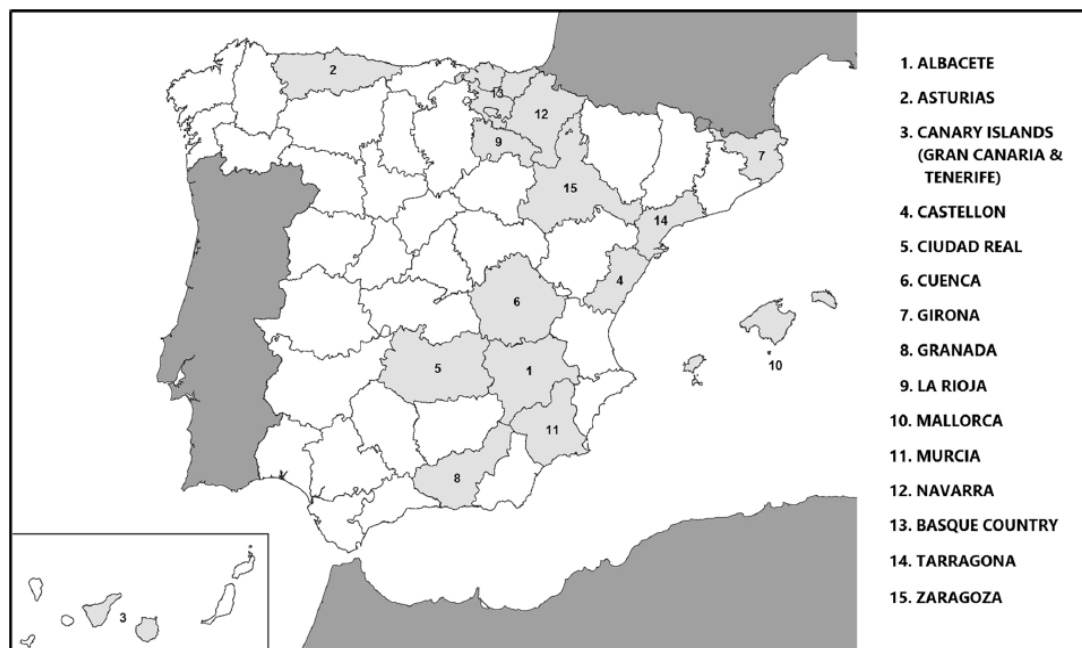


Figura 1.1 Províncies amb RCBP a Espanya. Font: Galceran J, et. al. *Cancer incidence in Spain, 2015*.

1.5 REGISTRES DE CÀNCER A CATALUNYA

A Catalunya, com es pot observar a la Figura 1.1, existeixen dos Registres de Càncer de Base Poblacional, el de Tarragona i el de Girona.

El Registre de Càncer de Tarragona [15] és un registre gestionat per la Fundació Lliga per a la Investigació i Prevenció del Càncer (FUNCA) i va iniciar la seva activitat l'any 1979 de forma que conté informació de base poblacional des de l'any 1980.

El Registre de Càncer de Girona [16], va ser establert l'any 1994 com ampliació del Registre Monogràfic de Càncer de Mama i Genital Femení, que va estar en funcionament des de l'any 1980 fins l'any 1989, recull informació sobre tots els tipus de càncer de la província de Girona des de l'any 1994. La Unitat d'Epidemiologia i Registre de Càncer del Pla director d'Oncologia és l'encarregada de la seva gestió.

La fiabilitat de les dades d'un Registre es basa sobretot en l'exhaustivitat de la casuística, i en la precisió i la comparabilitat de les dades. Per fer possible això, s'han emprat els criteris i les normes internacionals que van des de la definició de cas fins a l'elaboració de resultats [3,4,17,18].

Ambdós registres col·laboren habitualment en les principals publicacions internacionals sobre incidència i supervivència.

El Registre de Càncer de Tarragona va publicar per primera vegada els seus resultats al cinquè volum de la sèrie "*Cancer Incidence in Five Continents*" de l'Agència Internacional d'Investigació del Càncer (IARC) de l'Organització Mundial de la Salut (OMS), editada a l'any 1987 [19] i que abastava el període 1980-1983, mentre que el Registre de Càncer de Girona va publicar per primera vegada al volum vuit [22] d'aquesta sèrie i que abastava el període 1994-1997. Posteriorment, ambdós registres han publicats els resultats de períodes més recents en els volums sisè al onzè de la mateixa sèrie [19,20,21,22,23].

En quan a les publicacions internacionals de supervivència, l'any 1995 es va publicar la primera monografia del projecte EURO CARE que mesurava la supervivència poblacional dels malalts diagnosticats de càncer a Europa entre 1978 i 1984 [24], posteriorment s'han publicat els volums II, III IV i V d'aquest estudi [25,26,27] corresponents als períodes d'anys de 1985-1989, 1990-1994, 1995-1999, 2000-2007. El Registre de Càncer de Tarragona ha participat en aquests quatre últims volums i el Registre de Càncer de Girona en els dos últims.

En relació al càncer infantil, el Registres de Tarragona i Girona han participat en el segon volum de la sèrie "*International Incidence of Childhood Cancer*" [28] de l'Agència Internacional d'Investigació del Càncer, així com en el projecte ACCIS (*Automated Childhood Cancer Information System*) [29].

1.6 DESCRIPCIÓ DEL PROJECTE

La incidència d'una malaltia en una població sol ser una dada generalment desconeguda en el cas del càncer, a excepció d'aquells territoris on existeixen registres de càncer de base poblacional (RCBP).

A Espanya el conjunt de RCBP contenia, el gener de 2013, una població equivalent al 28,7% de la població espanyola (13.521.802 habitants d'un total de 47.129.783).

No seria correcte estimar les taxes d'incidència de càncer a Espanya a partir de les taxes de tot el conjunt dels RCBP, sense considerar que pugui haver diferències d'incidència entre les províncies amb registre i les províncies sense registre de càncer de base poblacional i per això, per a corregir el biaix entre la incidència de les províncies amb registre de càncer de base poblacional i les províncies sense registre s'utilitza habitualment com a variable correctora, les dades de mortalitat per càncer.

En primer lloc, es plantejarà un model per a l'estimació de dades conjuntes d'incidència i mortalitat, seguint les pautes del mètode REDECAN [30], el qual parteix sota la hipòtesi que per a cada tipus tumoral i sexe, la raó entre la incidència i la mortalitat (RIM), és similar en totes les províncies.

Amb això, els factors que puguin explicar la variació que hi ha en el RIM dins d'un mateix tipus de tumor i sexe són l'edat al diagnòstic i l'any de diagnòstic, per això s'inclouen aquestes variables com a predictores del model.

L'estimació per a cada RIM es farà per al període 1994-2007 a partir de les dades d'incidència d'aquelles províncies i períodes en els que es disposen de dades dels registres de càncer de base poblacional i dades de mortalitat per aquelles mateixes zones i períodes.

Aquesta anàlisi es durà a terme mitjançant models lineals generalitzats mixtes bayesians per a cada tipus de tumor i sexe, utilitzant mètodes de simulació del tipus Monte Carlo i Cadenes de Markov (Markov Chain Montecarlo Methods, MCMC).

Seguidament, es realitzarà una estimació per al nombre de casos incidents durant el període 2009-2013 a les províncies de Tarragona i Girona.

El capítol 6 es centrarà en el procés de comparació i validació per diferents escenaris de les estimacions de la incidència a Tarragona i Girona per al període 2009-2013, per així poder concloure en quines situacions aquest mètode funciona bé i en quines seria adequat pensar un altre tipus de mètode alternatiu.

Un cop feta aquesta comparació, al capítol 7 es presentaran les estimacions de la incidència per totes les províncies, amb i sense Registre de Càncer de Base Poblacional, amb la finalitat de tenir la incidència calculada arreu d'Espanya.

Aquest projecte està directament relacionat en un conveni de pràctiques amb la Fundació Lliga en la Investigació i Prevenció del Càncer (FUNCA) que es va començar al mes de març. Així doncs, al haver estat aquests mesos en contacte amb el Registre de Càncer de Tarragona, es decideix que els resultats més detallats es mostren sobre aquesta província, i conseqüentment, també amb la de Girona, ja que són les dues províncies amb RCBP de Catalunya.

1.7 OBJECTIUS

En el present treball, els principals objectius són:

- Estimar la RIM (raó entre incidència i mortalitat) per les províncies amb RCBP en el període 1994-2007 mitjançant el mètode REDECAN.
- En base a aquestes estimacions, fer una projecció del nombre de casos incidents per grup d'edat, sexe i tipus tumoral per les províncies de Tarragona i Girona durant el període 2009-2013, per a diferents futurs escenaris.
- Validar aquestes estimacions amb les dades reals de Tarragona i Girona per aquest mateix període.
- Estimar la incidència per tots els tipus de tumors i en totes les províncies d'acord amb els escenaris escollits.

1.8 PROGRAMARI

Per al càlcul de taxes i l'elaboració de gràfics s'utilitzarà el programa R, disponible a: <http://www.r-project.org>. Per a l'estimació dels paràmetres dels models bayesians implementats s'utilitzarà el programa JAGS, el qual està disponible a la pàgina web: <http://mcmc-jags.sourceforge.net>. Un avantatge de JAGS és que es pot executar des de l'R i així fer la manipulació de dades i la presentació de resultats des de l'R.

2 DESCRIPCIÓ DE LES DADES

En aquest capítol s'explicaran els indicadors utilitzats en aquest projecte, així com la metodologia del Registre de Càncer de Tarragona, tot explicant quins són els seus aspectes geogràfics, demogràfics i temporals, la definició d'allò que es considera cas i també les variables que utilitza en la recollida de dades. Per una altra banda, es farà una breu presentació de les bases de dades utilitzades.

2.1 INDICADORS PER AL CONTROL DEL CÀNCER

Els principals indicadors per a la magnitud del càncer i el progrés en el seu control són la incidència, la mortalitat, la supervivència i la prevalença. Aquest projecte però, es centra en dos d'aquests indicadors, que són la incidència i la mortalitat.

2.1.1 *Incidència*

Les dades brutes d'incidència de càncer en una població són recollides directament pels registres de càncer, de les històries clíniques i d'altres fonts similars. Com a conseqüència, la precisió diagnòstica és generalment elevada i els controls de qualitat dels registres eviten errors que són corregits a les pròpies fonts d'informació.

Les tendències de la incidència poden veure's afectades per canvis en l'exhaustivitat de l'enregistrament, per canvis en els mètodes diagnòstics, en les definicions patològiques o en la codificació dels tumors i per la introducció de programes de cribratge. Donat que no és possible assolir una prevenció completa dels càncers, és necessari monitoritzar també altres mesures com són la supervivència i la mortalitat.

2.1.2 *Mortalitat*

La reducció a llarg termini de la mortalitat per càncer és una mesura clau del progrés contra el càncer, però les taxes de mortalitat tenen algunes limitacions. Les taxes de mortalitat en un any donat depenen principalment del nombre de pacients diagnosticats recentment i de la seva supervivència però també del nombre de pacients diagnosticats en els deu anys previs o més i de l'abast dels seus tractaments que els van portar a la curació o al retard de la mort per càncer.

Aproximadament una quarta part dels pacients amb càncer moren per altres causes i les taxes de mortalitat per càncer no reflecteixen la seva experiència de la malaltia. Les diferències geogràfiques i les tendències temporals de la mortalitat per càncer són difícils d'interpretar sense informació sobre les diferències i les tendències en la incidència. La mortalitat per càncer és, per tant, un indicador incomplet del control del càncer i la interpretació de les tendències de mortalitat està difuminada per la seva complexa dependència de les tendències de la incidència i de la supervivència dels deu

o més anys anteriors.

2.1.3 Supervivència

La supervivència es defineix com la proporció de malalts diagnosticats de càncer que sobreviuen més d'un temps determinat. Les taxes de supervivència són els indicadors més directes de la gravetat del càncer i de l'impacte del tractament.

2.1.4 Prevalença

La prevalença que es defineix com la proporció d'una població que està afectada per càncer en un moment donat del temps. Inclou totes les persones que, havent estat diagnosticades en qualsevol moment del passat, encara estan vives, estiguin curades o no.

2.2 METODOLOGIA DELS REGISTRES DE CÀNCER DE CATALUNYA

Els Registres de Càncer de Tarragona (RCT) i Girona (RCG) són registres de càncer de base poblacional que tenen com a finalitat identificar i enregistrar, de forma exhaustiva i continuada, tots els casos incidents, els nous casos diagnosticats de càncer, entre els habitants residents a les demarcacions provincials de Tarragona i Girona, independentment del lloc on s'hagin diagnosticat, ja sigui dins de les províncies de Tarragona i Girona o fora d'aquestes.

El seu objectiu fonamental és determinar la incidència del càncer a les poblacions de les demarcacions provincials de Tarragona i Girona, la seva distribució segons múltiples variables com el sexe, l'edat, la localització tumoral, o el tipus histològic i la seva evolució en el temps. També són objectius majors analitzar la mortalitat, la supervivència i la prevalença; estimar la incidència futura; contribuir a la planificació de serveis sanitaris i avaluar les intervencions sanitàries tant a nivell assistencial com preventiu. Altres objectius són col·laborar en estudis etiològics i analítics, promoure la formació dels professionals sanitaris en temes d'epidemiologia del càncer, entre altres.

Per tal de garantir la fiabilitat i la comparabilitat de les seves dades amb la d'altres registres de càncer de base poblacional, utilitzen criteris internacionals que abasten des de la definició de cas fins els seus sistemes d'operació i l'elaboració de resultats. Els Registres de Càncer de Tarragona i Girona segueixen, com la resta de registres de càncer de base poblacional de l'Estat espanyol, d'Europa i del món, procediments estrictes per tal d'assegurar l'absoluta confidencialitat i protecció de les dades.

2.2.1 Aspectes geogràfics, demogràfics i temporals

L'àrea que cobreix el Registre de Càncer de Tarragona és la demarcació provincial de Tarragona, que engloba les regions sanitàries Camp de Tarragona (comarques de l'Alt Camp, Baix Camp, Baix Penedès, Conca de Barberà, Priorat i Tarragonès) i Terres de l'Ebre (comarques del Baix Ebre, Montsià, Ribera d'Ebre i Terra Alta) i que està localitzada al sud de Catalunya i al Nord-Est d'Espanya a la península ibèrica, estenent-se entre les latituds 40°32' i 41°35'. La seva superfície és de 6.303 *km*², i la longitud del seu perímetre és de 586 km, dels quals 216 limiten amb la mar Mediterrània. Els seus municipis es situen en altures que oscil·len entre el nivell basal del mar i els 823 metres per sobre d'aquest.

La població que abasta l'àrea que cobreix el Registre de Càncer de Tarragona ha augmentat de manera constant des de 1980, primer any amb dades completes de Registre. Segons el cens de 1981, aquell any hi havia 513.050 habitants, segons el de 2001 els habitants n'eren 609.673, i segons el de 2011, 807.044 habitants. En aquest últim cens, les ciutats de Tarragona (133.223 hab.) i Reus (106.849 hab.), representen gairebé un terç de la població provincial per aquest any.

La base de dades del Registre de Càncer de Tarragona té una cobertura de base poblacional provincial des de gener de 1980. Per tal de facilitar la no inclusió de casos prevalents en els primers anys del Registre, es van recollir la majoria dels casos incidents del període 1977-1979. Quant al seguiment de l'estat vital, els pacients tenen seguiment fins el 31 de desembre de 2015.

En quant a Girona l'àrea que cobreix el seu Registre de Càncer fa referència a la demarcació provincial de Girona, la qual engloba les comarques de l'Alt Empordà, Baix Empordà, La Garrotxa, Gironès, Pla de l'Estany, La Selva i el Ripollès, amb 210 municipis. La seva superfície és de 5569,5 *km*², amb una línia de costa de 246,8 km. La seva orografia és molt variada i comprèn zones d'alta muntanya als Pirineus i la zona costera Mediterrània.

La població que abasta l'àrea que cobreix el Registre ha augmentat de manera constant des dels noranta anys de Registre. Segons el cens de 1991, aquell any hi havia 509.628 habitants, segons el de 2001 els habitants n'eren 609.673, i segons el de 2011, 751.806 habitants. En aquest últim cens, la ciutat de Girona (96.113 hab.) esdevé la més poblada de la província.

La base de dades del Registre de Càncer de Girona té una cobertura de base poblacional provincial des de gener de 1994 i en quant al seguiment de l'estat vital, els pacients estan seguits fins el 31 de desembre de 2015.

2.2.2 Definició de cas

Els Registres de Càncer de Tarragona i Girona consideren cas qualsevol tumor maligne invasiu o no (in situ), així com tots els tumors del sistema nerviós central de comportament no especificat de qualsevol edat i els benignes del sistema nerviós central diagnosticats en nens i nenes fins 14 anys, diagnosticats en persones residents a les regions sanitàries Camp de Tarragona i Terres de l'Ebre en el moment del diagnòstic o en el moment de la defunció, si el certificat de defunció és l'única font d'informació.

El coneixement de les malalties neoplàsiques¹ ha anat evolucionant amb el temps i, com a conseqüència, han anat apareixent noves edicions de la classificació internacional de les malalties oncològiques, els Registres de Càncer de Tarragona i Girona han anat adaptant-se a aquestes classificacions.

Quant al diagnòstic, a més dels casos que tenen confirmació microscòpica, s'inclouen els tumors malignes amb un diagnòstic basat en diferents tècniques exploratòries no anatomo-patològiques com signes, proves d'imatge, cirurgia exploratòria, tests bioquímics i/o immunològics específics.

En aquestes circumstàncies, la inclusió final d'un cas utilitza tota la informació existent a la història clínica per determinar el grau de seguretat de que la patologia en qüestió sigui realment un càncer.

També s'enregistren els càncers procedents de les Butlletes Estadístiques de Defunció, quan aquesta és l'única font d'informació del cas. Aquests casos s'inclouen en la base de dades com a tumors malignes sempre que el certificat ho consideri així. En aquest sentit, el terme neoplàsia és considerat sinònim de neoplàsia maligna.

Els Registres de Càncer de Tarragona i Girona, com la resta de registres internacionals, enregistra tumors i no individus, assumint que un mateix individu pot presentar dos o més tumors. A nivell mundial existeixen criteris estàndard internacionals bàsics per la inclusió de càncers múltiples definits en les recomanacions conjuntes de la *European Network of Cancer Registries* (ENCR), la *International Agency for Research on Cancer* (IARC) i la *International Association of Cancer Registries* (IACR).

¹ En llenguatge mèdic és freqüent utilitzar el terme de neoplàsia maligna com a substitut de la paraula càncer.

2.2.3 Variables que monitoritzen els Registre de Càncer de Tarragona i Girona

Els principals ítems d'interès epidemiològic que s'enregistren de forma sistemàtica són:

Relatius al pacient:

- Sexe.
- Data de naixement.
- Província i país de naixement.
- Municipi i comarca de residència en el moment del diagnòstic.
- Edat en el moment del diagnòstic.

Relatius al tumor:

- Data del diagnòstic.
- Localització i sublocalització del tumor primari.
- Tipus Histològic.
- Comportament tumoral.
- Mètode més vàlid de diagnòstic.
- Mètode de detecció en relació al cribratge.
- Multiplicitat tumoral.
- TNM i estadi² des de l'any 2000.

Relatius a la defunció:

- Data de defunció.
- Causa de defunció (Càncer versus No càncer).
- Codi de defunció.

2.3 FONT I VARIABLES DE LES BASES DE DADES

Per a la classificació i codificació de les diferents variables de la base de dades s'utilitzen normes consensuades internacionalment i que, per una part d'elles en el cas d'Europa, estan reunides en una publicació conjunta de la *World Health Organization* (WHO), la *European Commission* (EU) i la *European Network Cancer Registries* (ENCR) de l'any 2003.

² L'estadificació és una forma de descriure quant càncer té un pacient en el seu cos i en quines parts d'aquest està localitzat. Això ajuda a determinar on està el tumor original, quant gran és i, en cas de que s'hagi propagat, en quin lloc ho ha fet. En la nomenclatura TNM, T descriu la mesura del tumor i la dispersió del càncer cap al teixit proper, la N fa referència a la dispersió del càncer fins als ganglis limfàtics propers i, per últim, la M descriu les metàstasis (disseminació del càncer cap a altres parts del cos).

2.3.1 Població

En quant a la base de dades de la població, conté les dades anuals d'individus que hi ha per província i per al conjunt d'Espanya durant el període 1994-2013, a partir de les estimacions intercensals anuals i les estimacions nacionals de població a curt termini de l'INE (www.ine.com).

Les variables que s'utilitzaran en aquest cas són:

- Província
- Any
- Sexe
- Edat

2.3.2 Mortalitat

Presenta les defuncions segons el tipus de càncer produïdes a cada província per al període 1994-2013. Aquestes dades són de disposició pública a través del Institut Nacional d'Estadística (www.ine.com).

En aquest cas hi ha les següents variables:

- Província
- Any
- Sexe
- ICDO (Codi identificador del tipus de tumor)
- Edat

2.3.3 Incidència

La base de dades de la incidència inclou, per una part, les dades d'incidència de càncer en el període 1994-2007, proporcionades pels RCBP espanyols (Albacete, Astúries, Canàries, Castelló, Ciutat Real, Cuenca, Euskadi, Girona, Granada, La Rioja, Mallorca, Múrcia, Navarra, Tarragona). Cada RCBP aporta la incidència disponible dins del període mencionat. Aquestes dades són de disposició pública, i es poden trobar per mig del portal ECIS (<https://ecis.jrc.ec.europa.eu>). Per una altra banda, inclou les dades d'incidència de càncer en el període 2009-2013, proporcionades directament pels RCBP de Tarragona i Girona.

Per a la incidència, s'utilitzaran les variables:

- Província
- Any
- Sexe
- ICDO (Codi identificador del tipus de tumor)
- Edat

2.3.4 Codis identificadors per als tumors

Les classificacions utilitzades en la presentació de resultats es mostren en base a la desena revisió de la Classificació Estadística Internacional de Malalties i Problemes relacionats amb la Salut (ICD-10). Per fer possible això, s'aplica una sèrie de conversions sobre les variables "Localització i sublocalització", "Histologia" i "Comportament" de cadascun dels casos enregistrats de la base de dades.

La presentació dels diferents tipus tumorals, normalment en funció de la localització del tumor primari, a les taules i figures dels resultats es realitza amb el mateix ordre que la ICD-10 amb alguns agrupaments respecte a la classificació original.

El total de càncers es comptabilitza fent dues agrupacions diferents. En la primera hi consten els tumors invasius sense excepció ("Totes" C00-96) i en la segona, tots els tumors invasius menys els de la pell que no són melanomes ("Totes excepte pell no mel." C00-96 (-C44)).

A la taula 2.1 es presenten els tipus de tumors que s'utilitzaran en aquest projecte, juntament amb el seu codi.

Codi	Sexe	Tipus tumoral
C01-06, C09-13	Home i Dona	Cavitat Oral i Faringe
C15	Home i Dona	Esòfag
C16	Home i Dona	Estómac
C18	Home i Dona	Còlon
C19	Home i Dona	Recte
C21	Home i Dona	Còlon i Recte
C22	Home i Dona	Fetge
C23	Home i Dona	Vies biliars
C25	Home i Dona	Pàncrees
C32	Home i Dona	Laringe
C34	Home i Dona	Pulmó
C43	Home i Dona	Melanoma de pell
C50	Dona	Mama
C53	Dona	Coll Uterí
C54	Dona	Cos Uterí
C56	Dona	Ovari
C61	Home	Pròstata
C62	Home	Testicles
C64	Home i Dona	Ronyó
C67	Home i Dona	Bufeta urinària
C70-72	Home i Dona	Encèfal i sistema nerviós
C73	Home i Dona	Tiroides
C81	Home i Dona	Limfoma Hodgkinià
C82-85, C96	Home i Dona	Limfoma No Hodgkinià
C90	Home i Dona	Mieloma
C91-95	Home i Dona	Leucèmia
C99	Home i Dona	Altres
C00-96 (-C44)	Home i Dona	Tots, excepte pell no melanoma

Taula 2.1 Codis identificadors per als tumors

3 ANÀLISI DESCRIPTIVA DE LES DADES

En aquest apartat es realitzarà una breu anàlisi exploratòria en quant a les diferents bases de dades: població, incidència i mortalitat.

Degut a que els rangs comuns en les tres bases de dades són des de 1994 fins a 2007, s'utilitzaran aquests resultats tant en els gràfics com en les taules presentades. Posteriorment, un cop realitzades les estimacions per un període més actual i s'hagi fet la validació d'aquestes, se'n mostrarà una anàlisi amb els resultats més actualitzats.

Per a les bases de dades d'incidència i mortalitat es calcularà la taxa bruta d'incidència (o mortalitat), i també les taxes específiques segons el sexe i l'edat.

La taxa bruta d' incidència (o mortalitat) (TB) mesura la freqüència de la incidència (o mortalitat) en una població durant un període determinat. La TB és el quocient entre el nombre de casos incidents (o morts) en un període de temps i la població en aquell mateix període de temps. Com que la població pot ser canviant, el denominador es considera persones-temps a risc. Així, el denominador recull la suma dels temps a risc de tots els individus de la població.

La taxa bruta expressada per 100.000 persones-any, és:

$$TB = \frac{D}{P} \cdot 10^5$$

on, D és el nombre total de casos incidents (o morts) durant el període de temps determinat, i P són les persones-any del període.

Al estar incloent la dimensió del temps en el càlcul d'aquesta taxa, fa que aquesta s'interpreti com la velocitat a la que incideixen o moren les persones que formen part de la població d'estudi. A més, en aquest cas, en lloc d'utilitzar el terme persones-any s'utilitzarà el terme d'habitants, que sol ser més habitual i entenedor.

Com que la taxa bruta ignora la distribució de la població per sexe i per edat, no resulta del tot útil a l'hora de fer comparacions, és per això que també es faran els càlculs de les taxes específiques tant per sexe com per edat.

Les taxes específiques d'incidència i mortalitat, es caracteritzen perquè en lloc d'agafar tot el conjunt de casos incidents (o defuncions) i població, només consideren una part d'aquests. En aquest cas es mostraran primer els resultats de les taxes específiques per sexe, i posteriorment, per edats.

Les taxes específiques segons el sexe expressades per 100.000 persones-any són:

$$TE_x = \frac{d_x}{P_x} \cdot 10^5$$

on x fa referència al sexe, home o dona, de forma que d_x és el nombre total de casos incidents (o morts) per a cada valor de x , i P_x són les persones-any també per a cada valor de x , és a dir, del sexe.

De la mateixa forma, les taxes específiques segons el grup d'edat expressades per 100.000 persones-any són:

$$TE_y = \frac{d_y}{P_y} \cdot 10^5$$

on y fa referència al grup d'edat, i d_y és el nombre total de casos incidents (o morts) en interval d'edat y , i P_y són les persones-any de l'interval d'edat y .

3.1 POBLACIÓ

Arran d'aquesta base de dades es pot veure el desenvolupament que hi ha hagut en la població durant el període d'anys 1994-2007. A continuació es fa una representació de quatre piràmides que fan referència a Tarragona i Girona, pels anys 1994 i 2007.

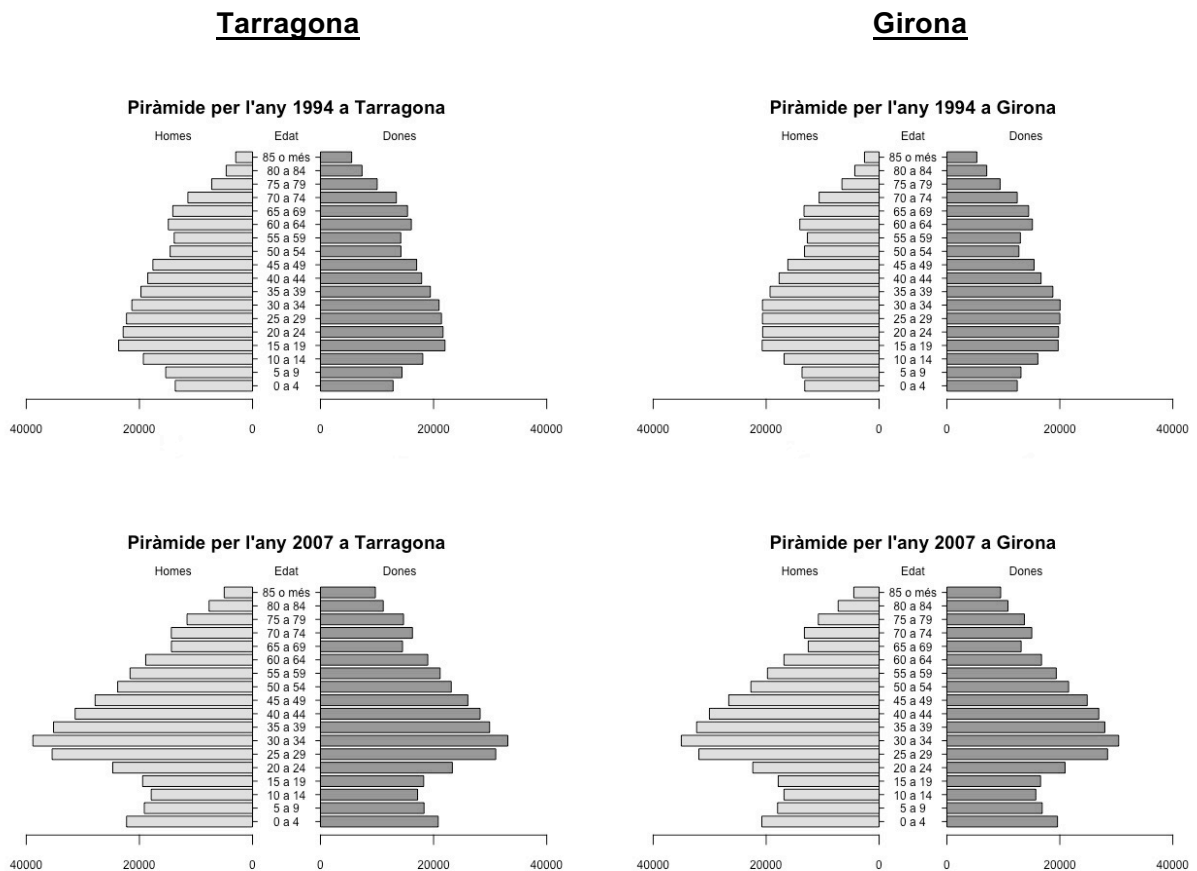


Figura 3.1 Piràmides de població per sexe per Tarragona i Girona pels anys 1994 i 2007.

Així a la figura 3.1 es poden veure els canvis que han hagut en la població en el transcurs d'aquests anys, de forma que hi ha una distribució força semblant en els dos àmbits territorials de Tarragona i Girona.

3.2 INCIDÈNCIA

A continuació es mostren una sèrie de gràfics que permeten observar quina ha sigut l'evolució al llarg del temps d'alguns dels tipus de tumor més recurrents en les províncies de Tarragona i Girona, i per a cada tipus de sexe. La representació d'aquests gràfics fa referència al nombre total de casos incidents que hi ha hagut a cada any sobre cada tipus de tumor.

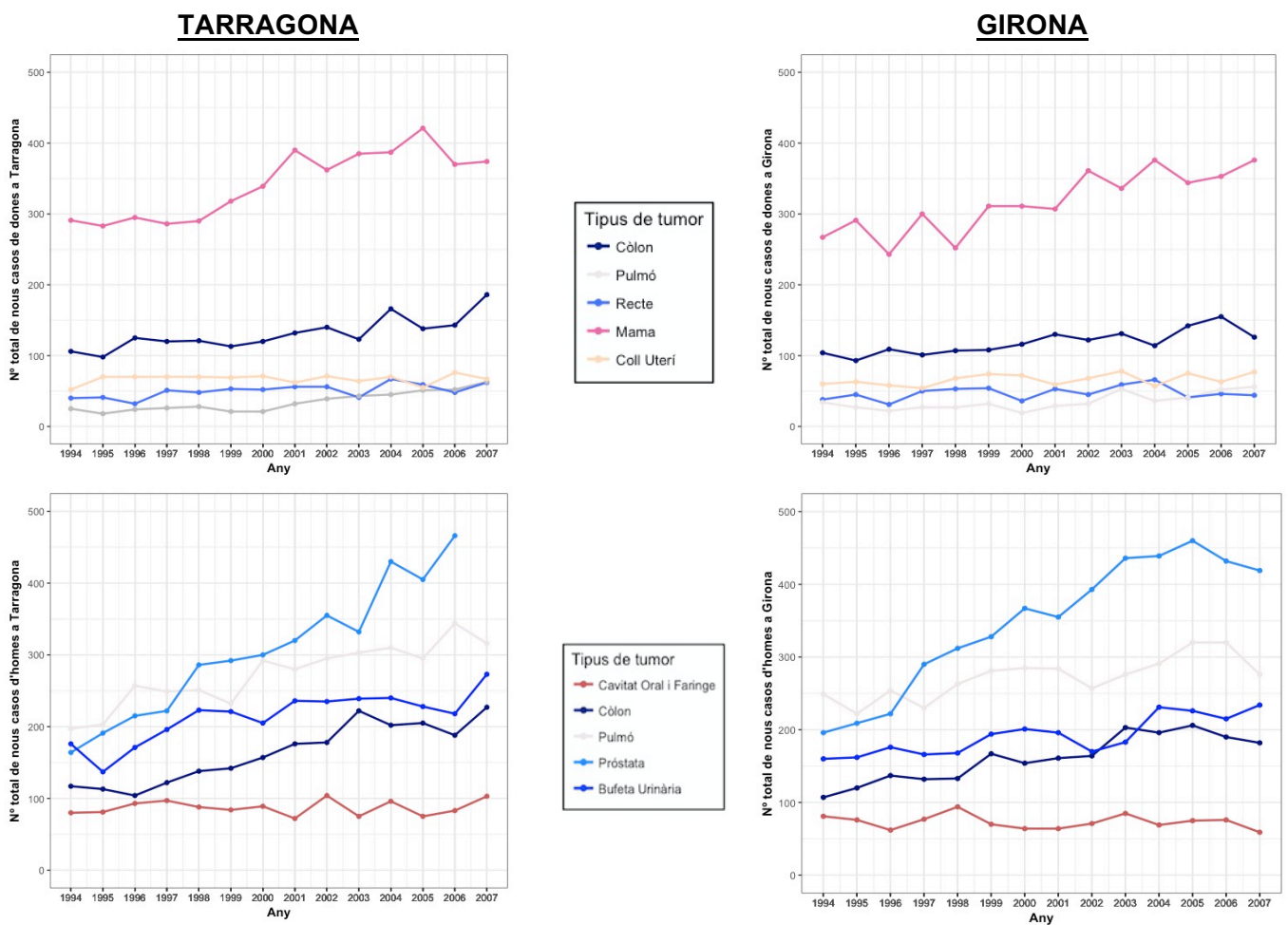


Figura 3.2 Representació de l'evolució d'alguns dels càncers més freqüents per sexe a Tarragona i Girona pel període d'anys 1994-2007.

A la figura 3.2 es pot veure que en el cas de les dones el càncer de Mama ha estat sempre per sobre dels altres, tant a Tarragona com a Girona, mentre que la resta es troben més o menys, més igualats, tot i poder establir que l'ordre és primer càncer de Còlon, seguit del Coll Uterí, i quedant gairebé igualats el càncer de Recte i el de Pulmó.

Pel que fa als homes, es pot observar com les distribucions també són força similars entre Tarragona o Girona. En el cas de Tarragona, es veu com el càncer de Pròstata és el que més incidència té amb el pas del temps, seguit del càncer de Pulmó, el qual en els quatre primers anys estava per sobre. A continuació el tipus de tumor que rep més nous casos incidents és el de Bufeta Urinària, tot seguit del càncer de Còlon. Ja per últim, té menys incidència el càncer que engloba la Cavitat Oral i Faringe.

Com s'ha dit anteriorment, aquest projecte farà càlculs per tots els tipus de tumor que s'han presentat, tot i que els resultats es mostraran més concretament per al Còlon, de forma que en aquest apartat, les taxes que es mostren tot seguit fan referència a només aquest tumor.

Tarragona			
	1994	2000	2007
Nº casos incidents	223	277	413
Població	560014	595598	765178
Taxa Bruta	39.82	46.51	53.97
Girona			
	1994	2000	2007
Nº casos incidents	211	270	308
Població	518546	550302	707567
Taxa Bruta	40.69	49.06	43.53

Taula 3.1 Nombre de casos incidents i taxes brutes per 100.000 habitants de càncer de Còlon a Tarragona i Girona durant el període 1994-2007.

Visualitzant la taula 3.1 es pot veure que a Tarragona la taxa bruta per al càncer de Còlon ha variat de forma ascendent en el transcurs del temps. El 1994 es detectaren aproximadament 40 casos cada 100.000 habitants, mentre que al 2000 se'n van detectar uns 47, i al 2007 gairebé 54. Aquest fet difereix en el cas de Girona, de forma que per al 1994 el nombre de nous casos detectats van ser-ne uns 41, l'any 2000, 49, i al 2007 aquest nombre va baixar a 44. A l'Annex 1 es mostra el resultat d'aquestes taxes brutes per a tots els tipus de càncer comuns en homes i dones.

Per veure aquests resultats contrastats d'una forma més específica, a continuació es fa el càlcul de les taxes específiques, les quals en lloc d'agafar tot el conjunt de defuncions i població, només consideren una part d'aquestes. En aquest cas es mostraran primer els resultats de les taxa específiques per sexe, i posteriorment, per edats.

Tarragona			
	1994	2000	2007
TB Dones	37.59	40.1	49.5
TB Homes	42.08	52.98	58.29

Girona			
	1994	2000	2007
TB Dones	39.73	41.7	36.18
TB Homes	41.67	56.59	50.65

Taula 3.2 Taxes específiques d'incidència del càncer de Còlon segons el sexe per 100.000 persones-any a Tarragona i Girona durant el període 1994-2007.

	Tarragona			Girona		
	1994	2000	2007	1994	2000	2007
0-4 anys	0	0	0	0	0	0
5-9 anys	0	0	0	0	0	0
10-14 anys	0	0	0	0	0	0
15-19 anys	0	0	0	0	0	0
20-24 anys	0	0	0	0	0	0
25-29 anys	0	2.08	3.01	0	0	3.31
30-34 anys	2.36	0	0	2.46	4.54	4.58
35-39 anys	12.77	6.56	6.15	7.88	2.26	1.66
40-44 anys	13.73	11.71	8.39	0	21.8	12.29
45-49 anys	17.33	15.44	20.41	41.14	38.17	11.66
50-54 anys	20.83	29.73	55.36	23.11	41.7	29.37
55-59 anys	53.47	60.89	74.75	42.79	43.5	71.55
60-64 anys	67.83	78.58	87.15	82.23	96.04	86.3
65-69 anys	128.85	122.96	162.87	115.19	159.42	140.09
70-74 anys	136.77	181.41	189.39	186.3	208.13	123.82
75-79 anys	202.83	247.35	281.7	231.15	183.86	228.72
80-84 anys	266.4	301.66	324.68	246.46	230.72	265.69
85-99 anys	294.29	228.06	407.66	164.7	251.99	314.06

Taula 3.3 Taxes específiques d'incidència del càncer de Còlon segons el grup d'edat per 100.000 persones-any a Tarragona i Girona durant el període 1994-2007.

A la taula 3.2 es pot observar com les taxes específiques per al sexe mostren valors més elevats en el cas dels homes que de les dones, tant per a Tarragona com Girona, de la mateixa forma que a la taula 3.3, els nombres més elevats de nous casos de càncer de Còlon corresponent a les edats més avançades.

3.3 MORTALITAT

Com en el cas de la incidència, es fa una representació gràfica del nombre total de morts que s'han donat durant el període d'anys marcat, per a cada sexe i per a cada tipus dels tumors més recurrents en les societats de Tarragona i Girona.

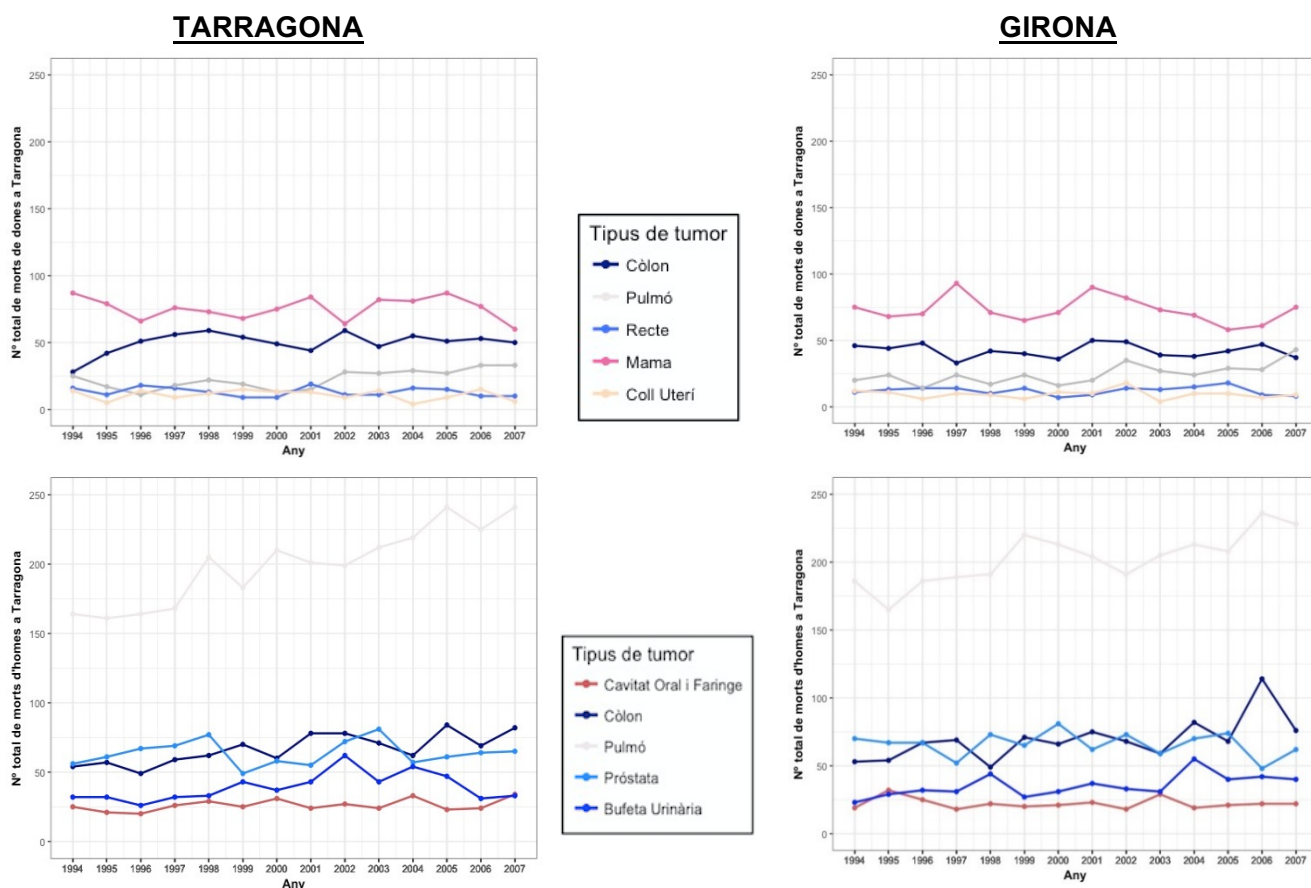


Figura 3.3 Representació de l'evolució d'alguns dels càncers més freqüents per sexe a Tarragona i Girona pel període d'anys 1994-2007.

Es pot observar a la figura 3.3 que en el cas de les dones el càncer de Mama és el que més mortalitat té, tot i que en els tres últims anys experimenta una lleugera baixada i s'igualava bastant amb el de Còlon. Mentrestant, els tumors de Pulmó, Recte i Coll Uterí mostren uns valors de mortalitat menys elevats, respectivament. Pel que fa als homes, es veu amb molta diferència com el càncer de pulmó presenta uns índexs de mortalitat molt elevats, mentre que la resta es troben per sota. Els tumors de Còlon i Pròstata tenen uns resultats prou semblants que van variant segons els anys, de forma que en els primers anys la Pròstata té una mortalitat més elevada mentre que cap als darrers anys, hi ha més morts per càncer de Còlon. Els tumors a la Bufeta Urinària i a la Cavitat Oral i Faringe són els que presenten menys morts dels representats en els gràfics.

Seguidament es presenta la taxa bruta de mortalitat (TB), que calcula la freqüència de la mortalitat en una població durant un període determinat. S'entén que en aquest cas les taxes calculades també fan referència al càncer de Còlon.

Tarragona			
	1994	2000	2007
Nº defuncions	105	135	183
Població	560014	595598	765178
Taxa Bruta	18.75	22.67	23.92
Girona			
	1994	2000	2007
Nº defuncions	119	125	150
Població	518546	550302	707567
Taxa Bruta	22.95	22.71	21.19

Taula 3.4 Nombre de defuncions i taxes brutes per 100.000 habitants del càncer de Còlon a Tarragona i Girona durant el període 1994-2007.

A la taula 3.4 se'n dedueix que el nombre de defuncions va en augment durant el transcurs dels anys per a Tarragona, mentre que a Girona es manté molt igualat entre els anys 1994 i 2000, fent un lleuger descens (de gairebé 23 defuncions a 21) en el darrer any que és el 2007. A l'Annex 2 es mostra el resultat d'aquestes taxes brutes per a tots els tipus de càncer comuns en homes i dones.

Tarragona			
	1994	2000	2007
TB Dones	14.9	21.39	19.42
TB Homes	22.66	23.96	28.24
Girona			
	1994	2000	2007
TB Dones	24.07	17.97	15.79
TB Homes	21.81	27.56	26.44

Taula 3.5 Taxes específiques de defuncions per càncer de Còlon segons el sexe per 100.000 persones-any a Tarragona i Girona durant el període 1994-2007.

	Tarragona			Girona		
	1994	2000	2007	1994	2000	2007
0-4 anys	0	0	0	0	0	0
5-9 anys	0	0	0	0	0	0
10-14 anys	0	0	0	0	0	0
15-19 anys	0	0	0	0	0	0
20-24 anys	0	0	0	0	0	0
25-29 anys	0	2.08	0	0	2.3	0
30-34 anys	0	0	0	0	0	1.53
35-39 anys	2.55	2.19	1.54	2.63	0	0
40-44 anys	5.49	7.03	0	2.91	7.27	1.76
45-49 anys	5.78	12.87	3.71	9.49	5.45	1.94
50-54 anys	6.94	2.70	14.9	11.55	14.89	11.3
55-59 anys	14.26	16.03	9.34	23.34	7.25	17.89
60-64 anys	29.07	21.43	26.41	27.41	36.02	38.69
65-69 anys	50.86	48.54	62.38	46.8	69.31	31.13
70-74 anys	68.38	78.26	81.64	108.31	57.81	60.14
75-79 anys	98.52	132.84	133.24	112.45	106.44	85.77
80-84 anys	108.23	150.83	159.68	184.84	176.88	215.88
85-99 anys	270.75	237.18	346.51	253.39	222.91	264.1

Taula 3.6 Taxes específiques de defuncions per càncer de Còlon segons el grup d'edat per 100.000 persones-any a Tarragona i Girona durant el període 1994-2007.

Es pot veure a la taula 3.5 que el nombre de defuncions per cada 100.000 habitants és més elevat en el cas dels homes que de les dones per a Tarragona i Girona, a excepció de l'any 1994 a Girona. Amb això, es podria dir que tot i haver un registre de nous casos més elevat en les dones (taula 3.2), aquest tipus de càncer afecta més greument als homes, ja que la seva mortalitat és més elevada.

Altrament, a la taula 3.6 es pot visualitzar com el nombre de defuncions per càncer de Còlon va en augment a mesura que es tenen més anys.

4 MODELITZACIÓ ESTADÍSTICA BAYESIANA

En aquest capítol es presenta la modelització estadística des d'un punt de vista bayesià. Per una banda, es farà una introducció a l'Estadística Bayesiana per poder entendre què és i com funciona, així com quins són els elements que hi intervenen. I per una altra banda, es farà referència a la inferència bayesiana i la forma de computar-la. Els següents punts presentats es basen en un curs d'Estadística Bayesiana [31, 32, 33, 34].

4.1 INTRODUCCIÓ

Un Model Estadístic, M , és una llista de models de probabilitat indexada per un paràmetre θ , on es sap que θ pertany a un espai de paràmetres Ω , $\theta \in \Omega$, i es pot escriure com $M = \{P(y|\theta); \theta \in \Omega\}$. El Model Estadístic és l'objecte central de l'estadística, és el punt de partida comú de la estadística Freqüentista i la Bayesiana abans de bifurcar-se.

Un model bayesià parteix d'un model estadístic, $M = \{P(y|\theta), \theta \in \Omega\}$, però tracta el paràmetre θ com una variable aleatòria, i a més, està disposat a triar una distribució de probabilitat sobre l'espai de paràmetres, Ω , que reflecteixi el coneixement que es té a priori sobre el paràmetre abans d'observar les dades. Aquesta nova distribució de probabilitat s'escriu com $\pi(\theta)$, i s'anomena distribució a priori.

En un model bayesià, $P(y|\theta), \pi(\theta), \theta \in \Omega$ és una llista de distribucions de probabilitat ordenada. És a dir, $\pi(\theta)$ ordena la llista de distribucions del Model Estadístic de més a menys creïble. En l'estadística Bayesiana es fan dos tries subjectives: el model estadístic, que reflecteix el coneixement a priori sobre l'espai mostral i la distribució a priori, que reflecteix el coneixement a priori sobre l'espai paramètric.

Per entendre-ho millor, la diferència entre un freqüentista i un bayesià, és que el punt de partida abans d'observar les dades, el freqüentista té una llista de distribucions de probabilitat no ordenada (Model Estadístic), mentre que el bayesià té una llista de distribucions de probabilitat ordenada de més a menys creïble en funció de la distribució a priori.

Abans d'observar les dades es parteix del model Bayesià $M = \{P(y|\theta), \pi(\theta), \theta \in \Omega\}$, on $\pi(\theta)$ reflecteix el coneixement o incertesa a priori sobre θ i, per especificar el model Bayesià cal triar els paràmetres de la distribució a priori. De la mateixa forma la distribució marginal de les dades, que s'anomena predictiva a priori, $P(\tilde{y})$, i es calcula com $P(\tilde{y}) = \int P(\tilde{y})\pi(\theta)d\theta$ on \tilde{y} representa una dada futura, reflectirà el coneixement o incertesa sobre l'espai mostral abans d'observar les dades.

Un cop observades les dades l'únic que canvia del model es que s'actualitza la $\pi(\theta)$ per la $\pi(\theta|y)$, anomenada distribució a posteriori, on es reflecteix tot el coneixement sobre el paràmetre un cop tinguda en compte la informació provinent de les dades, i de la mateixa manera la predictiva a posteriori reflecteix tot el coneixement de les dades futures. La predictiva a posteriori s'utilitza per fer prediccions. En molts casos és més rellevant fer inferència sobre l'espai mostral que no sobre l'espai de paràmetres. Si es vol fer inferència sobre l'espai mostral s'utilitzarà la predictiva a posteriori i si es vol fer inferència sobre l'espai de paràmetres s'utilitzarà la distribució a posteriori.

Per tal de calcular la distribució a posteriori s'utilitza el Teorema de Bayes de la següent forma:

$$\pi(\theta|y) = \frac{P(y|\theta)}{P(y)} = \frac{P(y|\theta)\pi(\theta)}{P(y)}$$

El denominador, $P(y)$ és la predictiva a priori avaluada a les dades, és a dir, $P(\tilde{y} = y)$, i és el que pot ser més difícil de calcular ja que s'ha d'integrar. Llavors es té:

$$\pi(\theta|y) = \frac{P(y|\theta)\pi(\theta)}{P(y)} \propto P(y|\theta)\pi(\theta) = L_y(\theta)\pi(\theta)$$

on $L_y(\theta)$ és la funció de versemblança.

Molts cops no serà necessari calcular $P(y)$, és a dir, no caldrà integrar el denominador ja que pot ser que l'expressió del numerador permeti reconèixer quina distribució segueix $\pi(\theta|y)$.

La distribució a posteriori sempre és un compromís entre la distribució a priori i la versemblança, és a dir, entre el que sabíem abans d'observar les dades i la informació que han aportat les dades.

4.2 INFERÈNCIA BAYESIANA

Fer inferència vol dir que a partir dels resultats obtinguts mitjançant la mostra estudiada, es vol saber les característiques que tindrà la població general. De la mateixa manera que la distribució a priori reflexa el coneixement que es té del paràmetre abans d'observar les dades, la distribució a posteriori és la que reflexa tot el què es sap del paràmetre després de recollir les dades.

A l'hora de fer inferència es pot obtenir una estimació puntual d'aquesta distribució o bé, una estimació per interval d'aquesta.

Com a estimador puntual del paràmetre θ , que és el paràmetre del model estadístic, s'utilitza qualsevol mesura de localització de la distribució a posteriori, per exemple:

- L'esperança: $\hat{\theta}_{ep} = E(\theta|y) = \int_{\Omega} \theta \pi(\theta|y) d\theta$

- La moda: $\hat{\theta}_{modp} = Moda(\theta|y)$

Es defineix un interval de credibilitat a posteriori amb probabilitat p, C_p , de θ , a qualsevol conjunt de Ω tal que $P(\theta \in C_p|y) = p$, és a dir, $P(\theta \in C_p|y) = \int_{C_p} \pi(\theta|y)d\theta = p$.

Els intervals centrals de credibilitat més utilitzats es basen en els percentils. Són invariants davant de reparametrizacions i fàcils de calcular a partir de simulacions de la distribució a posteriori.

El què s'ha de fer per calcular-los és ordenar les simulacions per tal d'obtenir els percentils $\frac{1-p}{2}$ i $1 - \frac{1-p}{2}$.

Es poden calcular intervals de credibilitat de qualsevol distribució, tant de les distribucions a priori i posteriori com de les distribucions predictives.

4.3 COMPUTACIÓ BAYESIANA

La computació Bayesiana fa referència a tots els procediments que aproximen la distribució a posteriori a través de la simulació.

Tot tipus d'inferència Bayesiana es basa en obtenir la distribució a posteriori, així com les predictives a posteriori. Moltes vegades aquestes distribucions no es poden obtenir de forma tancada, és a dir, tenir de forma explícita la seva funció de densitat, degut a que les integrals són difícils de calcular o simplement no tenen primitiva. Per exemple, per calcular $\pi(\theta = (\theta_1, \dots, \theta_p)|y)$ s'ha de resoldre com:

$$\pi(\theta = (\theta_1, \dots, \theta_p)|y) = \frac{L_y(\theta_1, \dots, \theta_p)\pi(\theta_1, \dots, \theta_p)}{\int_{\Omega_1} \dots \int_{\Omega_p} L_y(\theta_1, \dots, \theta_p)\pi(\theta_1, \dots, \theta_p)d\theta_p \dots d\theta_1}$$

Llavors, si a partir del numerador es pot identificar de quina distribució es tracta, aleshores no és necessari calcular les integrals.

Els mètodes de MCMC, Markov Chain Monte Carlo permeten simular la distribució a posteriori sense necessitat de resoldre la integral. Són mètodes que requereixen unes simulacions d'escalfament per tal que les cadenes convergeixen cap a la distribució a posteriori.

En l'apartat 5.2 (Estimació de la RIM per al període 1994-2007) s'explica amb més deteniment els criteris que s'han utilitzat per a dur a terme aquestes simulacions.

5 PRESENTACIÓ DEL MODEL

5.1 DESCRIPCIÓ DEL MODEL

En aquest apartat procedeix a la presentació de com es farà l'estimació de les raons d'incidència i mortalitat (RIM) per al període 1994-2007 mitjançant el mètode de REDECAN [30].

En primer lloc, es parteix de la hipòtesi que per a cada tipus de tumor i sexe, el valor de la RIM és similar en totes les províncies:

$$\frac{\text{Incidència}_{Edat,Any,Espanya}}{\text{Mortalitat}_{Edat,Any,Espanya}} \approx \frac{\text{Incidència}_{Edat,Any,RCBP}}{\text{Mortalitat}_{Edat,Any,RCBP}}, \text{ per tot RCBP.}$$

Un cop feta aquesta reflexió, l'objectiu en aquest cas és estimar la RIM per a cada combinació de tipus de tumor i sexe a partir de les dades d'incidència per a aquelles províncies i períodes dels que se'n disposen dades de la incidència aportades per Registres de Càncer de Base Poblacional, i dades de la mortalitat per a aquelles mateixes províncies i períodes.

Les anàlisi, com ja s'ha mencionat, es realitzaran mitjançant models lineals generalitzats mixtes per a cada tipus de tumor i sexe. En aquests models, es suposa que el nombre de casos incidents a cada província segueix una distribució de Poisson.

Pel que fa a l'any de diagnòstic, el fet de tenir un efecte més aviat quadràtic que no pas lineal, mostra la necessitat d'utilitzar una funció polinòmica de segon grau. A més, aquesta mateixa variable s'utilitza de forma centrada de manera que $Any = Any \text{ de diagnòstic} - 1994$.

L'efecte de l'edat es suavitza mitjançant un spline lineal amb 4 nodes, tècnica que consisteix en recategoritzar la variable en 5 grups segons 4 punts de tall o nodes. Per a cada combinació de tipus de tumor i sexe, s'escullen 4 nodes en funció del pool de defuncions (valors al percentil 10 (P10), tercil 1 (T1), tercil 2 (T2) i percentil 90 (P90)). De forma que els valors de la variable edat a cada node quedarien definits com:

$$Edat_1 = \begin{cases} Edat - P_{10} & \text{si } edat \geq P_{10} \\ 0 & \text{si } edat < P_{10} \end{cases} \quad Edat_2 = \begin{cases} Edat - T_1 & \text{si } edat \geq T_1 \\ 0 & \text{si } edat < T_1 \end{cases}$$
$$Edat_3 = \begin{cases} Edat - T_2 & \text{si } edat \geq T_2 \\ 0 & \text{si } edat < T_2 \end{cases} \quad Edat_4 = \begin{cases} Edat - P_{90} & \text{si } edat \geq P_{90} \\ 0 & \text{si } edat < P_{90} \end{cases}$$

Es crea un efecte aleatori relacionat amb la província de residència per tenir en compte la heterogeneïtat que hi ha entre la raó d'incidència i mortalitat de les diferents províncies.

Amb això, es considera el nombre de casos incidents com a terme de dependent, mentre que el nombre de defuncions és l'offset del model, i els termes independents són el sexe, la funció spline relacionada amb l'edat, la funció polinòmica de l'any de diagnòstic i l'efecte aleatori de la variable província.

La variable resposta que recull el nombre de casos incidents al grup d'edat e , l'any de diagnòstic a i la província p , s'escriu com $I_{e,a,p}$, mentre que el nombre de defuncions també per a cada grup d'edat, any de diagnòstic i província, s'escriu com $M_{e,a,p}$. A continuació es mostra el plantejament general que s'ha seguit per a la descripció del model:

$$I_{e,a,p} \sim \text{Poisson}(\lambda_{e,a,p})$$

on, $e = 1, 2, \dots, 18$, representa el grup d'edat
 $a = 1994, 1995, \dots, 2007$, representa l'any de diagnòstic
 $p = 1, 2, \dots, 15$, que fa referència a les províncies amb RCBP

Sent $\lambda_{e,a,p}$ el nombre de casos incidents esperats al grup d'edat e , any de diagnòstic a i província p , que es pot expressar com:

$$\lambda_{e,a,p} = \theta_{e,a,p} \cdot M_{e,a,p}$$

on $\theta_{e,a,p}$ representa el valor de la RIM per a cada grup d'edat, any de diagnòstic i província. Utilitzant el link canònic del model de Poisson, es pot escriure el logaritme del valor esperat com:

$$\log(\theta_{e,a,p} \cdot M_{e,a,p}) = \log(\theta_{e,a,p}) + \log(M_{e,a,p})$$

on $\log(M_{e,a,p})$ és l'offset. I el predictor lineal serà:

$$\log(\theta_{e,a,p}) = \alpha + \delta_p + (\beta_1 \text{edat}_1 + \beta_2 \text{edat}_2 + \beta_3 \text{edat}_3 + \beta_4 \text{edat}_4) + (\gamma_1 \text{any}^2 + \gamma_2 \text{any}) + \ln(M_{e,a,p})$$

on, α és la constant del model,
 δ_p és l'efecte aleatori de la variable província,
 β_1 és el coeficient del primer node de l'edat,
 β_2 és el coeficient del segon node de l'edat,
 β_3 és el coeficient del tercer node de l'edat,
 β_4 és el coeficient del quart node de l'edat,
 γ_1 és el coeficient del terme quadràtic de l'any de diagnòstic,
 γ_2 és el coeficient de l'any de diagnòstic

L'efecte aleatori δ_p es suposa que segueix una distribució $Normal(0, \sigma^2)$. Per a tots els paràmetres del model s'escullen distribucions a priori no informatives. Al Quadre 5.1 es troba el Model Bayesià complet.

$$I_{e,a,p} | \theta_{e,a,p} \cdot M_{e,a,p} \sim \text{Poisson}(\theta_{e,a,p} \cdot M_{e,a,p})$$

on, $e = 1, 2, \dots, 18$, representa el grup d'edat

$a = 1994, 1995, \dots, 2007$, representa l'any de diagnòstic

$p = 1, 2, \dots, 15$, que fa referència a les províncies amb RCBP

$$\log(\theta_{e,a,p}) = \alpha + \delta_p + (\beta_1 \text{edat}_1 + \beta_2 \text{edat}_2 + \beta_3 \text{edat}_3 + \beta_4 \text{edat}_4) + (\gamma_1 \text{any}^2 + \gamma_2 \text{any}) + \ln(M_{e,a,p})$$

$$\text{on } \delta_p \sim \text{Normal}(0, \sigma^2).$$

Les distribucions a priori són:

$$\alpha \sim \text{Normal}(0, 1000)$$

$$\beta_1 \sim \text{Normal}(0, 1000)$$

$$\beta_2 \sim \text{Normal}(0, 1000)$$

$$\beta_3 \sim \text{Normal}(0, 1000)$$

$$\beta_4 \sim \text{Normal}(0, 1000)$$

$$\gamma_1 \sim \text{Normal}(0, 1000)$$

$$\gamma_2 \sim \text{Normal}(0, 1000)$$

$$\frac{1}{\sigma^2} \sim \text{Gamma}(0.01, 0.01)$$

Quadre 5.1 Model per a l'estimació de la RIM en cada grup d'edat e , any a i província p

5.2 OBSERVACIONS DEL MODEL

Amb la informació aportada a la Secció 5.1 sobre el model, es presenta alguna observació en aquest apartat per tal de poder estudiar amb una mica més de deteniment els factors que hi influeixen en aquest.

S'observa que $\theta_{e,a,p}$ es pot escriure de la següent forma:

$$\theta_{e,a,p} = e^{\alpha + \delta p} \cdot e^{\beta_1 edat_1 + \beta_2 edat_2 + \beta_3 edat_3 + \beta_4 edat_4} \cdot e^{\gamma_1 any^2 + \gamma_2 any} \cdot M_{e,a,p}$$

on es poden desglossar els factors $e^{\beta_1 edat_1 + \beta_2 edat_2 + \beta_3 edat_3 + \beta_4 edat_4}$ i $e^{\gamma_1 any^2 + \gamma_2 any}$ per veure l'efecte que aquestes variables tenen sobre el model.

Sent l'Any una variable introduïda com $Any = Any \text{ de diagnòstic} - 1994$ es sap que per l'any 1994 el valor serà 0, per tant el valor de la seva exponencial 1. Quan aquest factor prengui un valor superior a 1 estarà dient que el valor de la RIM ($\theta_{e,a,p}$) serà més gran que la calculada per l'any 1994.

Amb això, resultarà interessant el fet de visualitzar l'efecte temporal que es produeix. Anàlogament, també serà interessant visualitzar l'efecte de l'edat.

5.3 ESTIMACIÓ DE LA RIM PER AL PERÍODE 1994-2007

L'estimació dels paràmetres dels models es realitza en base a una aproximació estadística bayesiana utilitzant mètodes de simulació del tipus Monte Carlo i Cadenes de Markov (Markov Chain Montecarlo Methods, MCMC).

Es calcularan tres cadenes de simulacions dels paràmetres segons l'algorisme MCMC. Cadascuna d'aquestes cadenes precisa d'un període inicial d'entrenament el qual es fixarà inicialment en 10.000 iteracions, mentre que per al període posterior d'anàlisi, s'utilitzaran 40.000 iteracions, de les quals se'n seleccionaran una de cada cinc per tal d'obtenir un total de 24.000 simulacions de la distribució a posteriori de cadascun dels paràmetres. En el cas del càncer de mama, el període inicial d'entrenament necessari ha estat de 60.000 iteracions abans no s'ha assolit la convergència.

Per fer aquestes simulacions s'utilitza el software R, que permet executar els models amb el programa JAGS, mitjançant els paquets "R2jags" i "Rjags".

Amb les simulacions obtingudes a continuació es mostra la representació dels efectes que tenen les variables *Any* i *Edat* tal com s'ha explicat a la Secció 5.2. S'explica amb més deteniment per al càncer de Còlon, i es presenten els diferents gràfics per a tots els tumors als Annexos 3 i 4.

Per interpretar els gràfics que mostren l'efecte de l'any sobre la RIM (Figura 5.1, Annex 3) cal tenir present que l'any 1994 és l'any basal i l'efecte serà sempre 1, mentre que per la resta d'anys els valors majors que 1 indiquen una major incidència que mortalitat, fet que es pot associar a una major supervivència i, contràriament, valors inferiors a 1 a una menor supervivència, sempre respecte de l'any de referència.

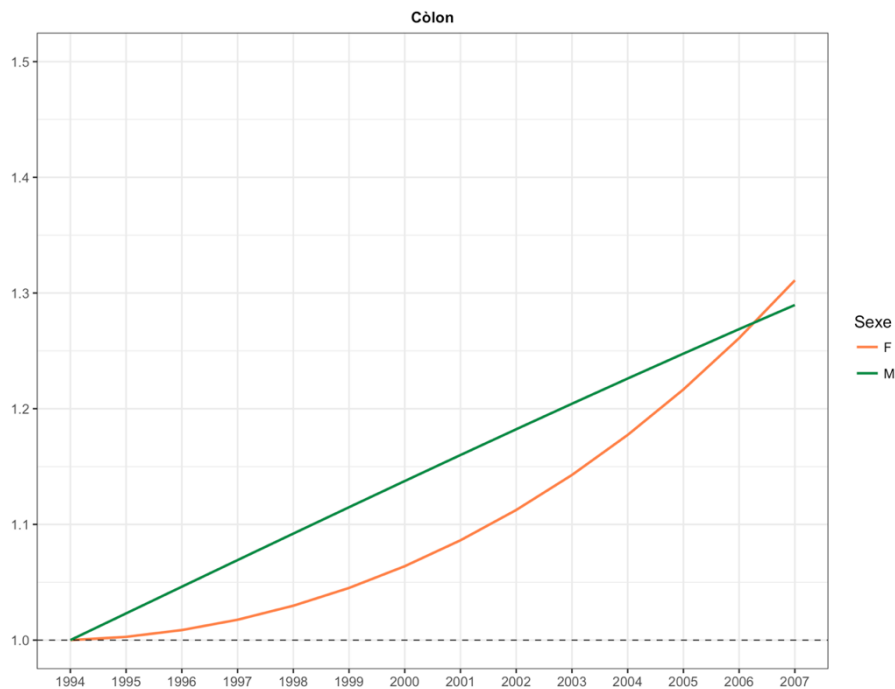


Figura 5.1 Efecte de l'any per al càncer de Còlon

A la Figura 5.1 es mostra que per al càncer de Còlon tant les dones com els homes tenen un comportament al llarg del temps amb una influència molt similar, tot i que les dones mostren uns valors lleugerament per sota dels homes excepte pel darrer any.

El fet de que la raó entre la incidència i mortalitat augmenti amb el pas del temps, portaria a deduir que el nombre de nous casos diagnosticats amb càncer de Còlon és cada cop més superior al nombre de morts per aquest mateix tumor. En termes de supervivència, s'estaria dient que aquesta ha anat augmentant amb el pas dels anys observant les dades des de l'any 1994.

A l'Annex 3 hi ha la presentació completa per a l'efecte dels anys en la resta de tipus de tumors. Es pot observar que dels tumors que són comuns per a homes i dones, el d'Estómac, Recte, Fetge, Pàncrees, Pulmó, Ronyó, Bufeta Urinària, Encèfal i sistema nerviós, Leucèmies, Mieloma i Altres presenten unes corbes molt similars entre si, per la qual es podria dir que no hi ha gaires diferències segons el sexe, a més, a la major part d'aquests els caracteritza una corba lleugerament ascendent, fet que indica que amb el pas dels anys el nombre de casos incidents ha sigut superior sobre el nombre de morts en aquell determinat tumor. Tot i això, per als tumors pertanyents als grups d'Encèfal i sistema nerviós i Altres, les corbes són lleugerament descendents, per tant, aquest factor indica que per als anys on el valor del seu efecte és inferior a 1, el nombre de nous casos incidents va ser inferior al nombre de casos morts. Contràriament, els altres tumors, com són Cavitat Oral i Faringe, Esòfag, Vies biliars, Laringe, Melanoma de pell, Tiroides, Limfoma de Hodgkin i Limfoma no hodgkinià tenen diferents efectes al

llarg del temps i entre sexes. Per una altra banda, els darrers gràfics fan referència als tipus de tumor que només són per a dones, o homes.

Les primeres edats (en concret, fins al percentil 10) corresponen a la categoria basal on l'efecte de l'edat (veure secció 5.2) és igual a 1. Per a la resta d'edats els valors poden ser majors o menors que 1, depenent si la incidència és major o menor que la mortalitat en comparació de les edats amb l'edat basal.

L'efecte de la variable edat és més complex d'analitzar ja que amb la finalitat de suavitzar el seu efecte s'ha estimat com un spline lineal.

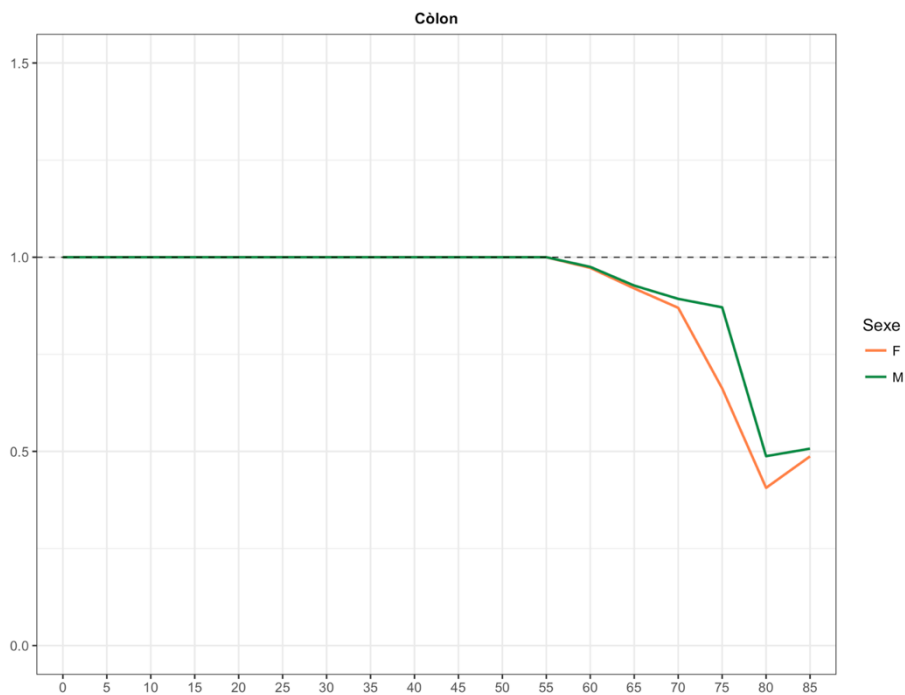


Figura 5.2 Efecte de l'edat per al càncer de Còlon

S'observa a la Figura 5.2 que tant per a les dones com per als homes, el comportament en els diferents grups d'edat és força similar, tot i que les dones mostren uns valors lleugerament inferiors als homes.

El gràfic estaria indicant una suposició de què a edats més avançades el risc per sobreviure és va fent menor. A l'Annex 4 es poden observar els gràfics per a tots els tumors, on per la majoria de tumors es repeteix un patró similar.

6 DEFINICIÓ I SELECCIÓ DEL MILLOR ESCENARI

6.1 DIFERENTS ESCENARIS PER L'EVOLUCIÓ TEMPORAL DE LA RIM

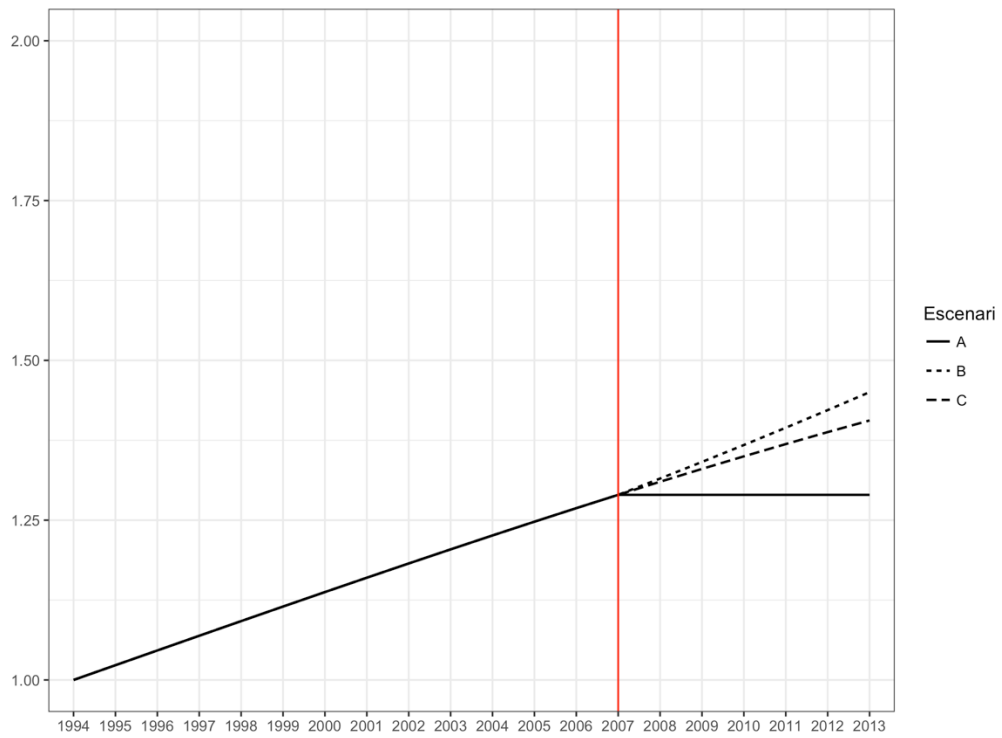
Per estimar el nombre de casos incidents en el període 2009-2013 el que es fa és aplicar a la mortalitat real de Tarragona i Girona per a aquest mateix interval d'anys les raons estimades d'incidència/mortalitat (RIM) a partir del model descrit a la secció 5.3 per a aquestes províncies en funció de diferents escenaris.

Els diferents escenaris que es plantegen fan referència a una hipòtesi diferent sobre l'evolució de l'estimació de la RIM per als anys posteriors, de forma que s'estableix que:

- Escenari A (constant): es considera que la RIM és estable des del 2009 fins al 2013 i igual a la RIM de l'any 2007.
- Escenari B (lineal): tendència lineal del logaritme de la RIM en el període 1994-2007 que és prolongada fins l'any 2013.
- Escenari C (quadràtic): tendència quadràtica del logaritme de la RIM en el període 1994-2007 que és prolongada fins l'any 2013.

A la figura 6.1 es troba representat l'efecte de l'any segons l'evolució de la RIM per al càncer de Còlon en homes i dones tenint en compte cada tipus d'escenari.

Homes



Dones

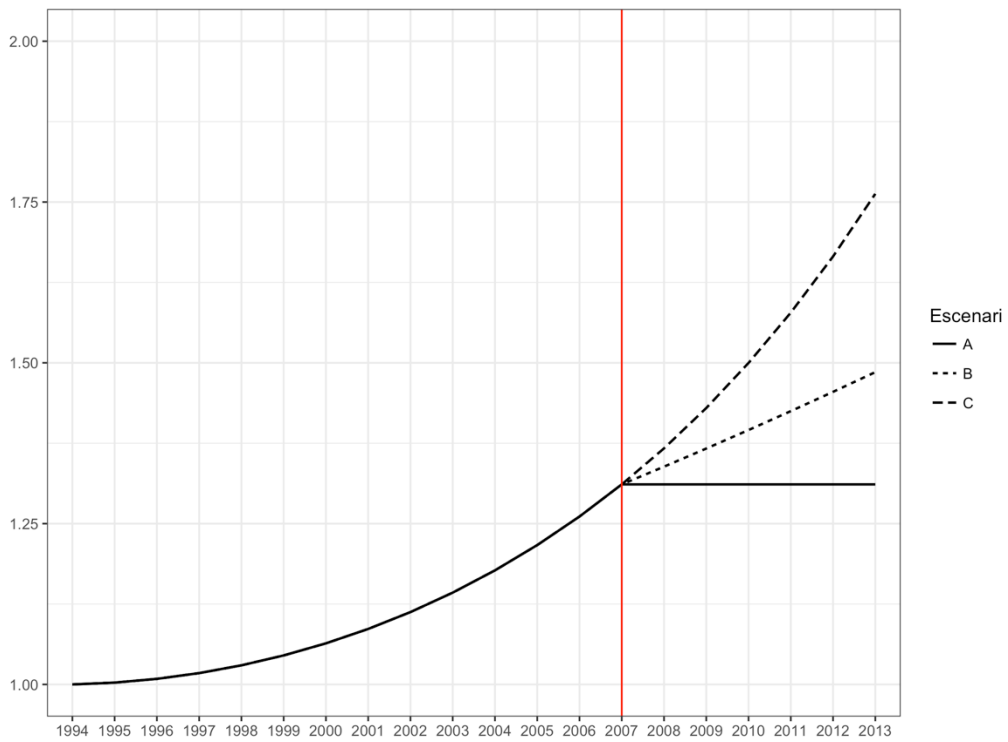


Figura 6.1 Efecte de l'any sobre l'evolució de la RIM en el càncer de Còlon i segons tipus d'escenari.

El càlcul de casos incidents en aquest apartat es fa únicament per les províncies de Tarragona i Girona perquè són les que s'utilitzaran per dur a terme el procés de validació, el qual servirà per poder escollir quin escenari ajusta millor les dades segons cada tipus de tumor.

6.2 SELECCIÓ DEL MILLOR ESCENARI

En l'aplicació de mètodes estadístics es poden considerar quatre fases generals:

1. Anàlisi descriptiva de les dades recollides, que permet veure quina informació aporten les dades i quin és el seu comportament.
2. Creació i estimació d'un model que permeti veure la relació que hi pot haver entre la variable d'interès i les variables explicatives.
3. Validació del model mitjançant la recollida de noves dades més actualitzades. Serveix per veure si el model utilitzat realment s'ajusta en bona mesura a les dades observades.
4. Prediccions futures en base al model creat sempre i quan aquest sigui representatiu.

Havent dut a terme les dues primeres fases, aquesta secció tracta la validació del model. El basar la validació amb el fet que les prediccions siguin consistents amb els valors observats, permetrà al mateix temps seleccionar aquell escenari per al qual les prediccions hagin estat més acurades.

Per a cadascun dels escenaris estudiats es presenta un indicador, que servirà per definir quin és el millor escenari possible per a cada tipus de tumor i també, per a determinar la precisió de les estimacions d'incidència a Tarragona i Girona pel període 2009-2013. Així, per a cada combinació tipus tumoral i sexe, es proposa el càlcul del següent indicador de bondat d'ajust per cada escenari:

$$\begin{aligned} \text{Estimador bondat d'ajust}(\text{Escenari}) &= \\ &= \sum_{PROV} \sum_{ANY=2009}^{2013} \sum_{GR.EDAT} \frac{(\text{Casos Esp.}_{PROV,ANY,GR.EDAT,ESCENARI} - \text{Casos Obs}_{PROV,ANY,GR.EDAT})^2}{\text{Casos Esp.}_{PROV,ANY,GR.EDAT,ESCENARI}} \end{aligned}$$

Donat que s'ha utilitzat l'algorisme MCMC, per a cada tipus tumoral s'han utilitzat les 24.000 simulacions de la distribució conjunta a posteriori dels paràmetres per tal de simular la distribució predictiva a posteriori. Per a la determinació del millor escenari per cada tipus tumoral i sexe es busca per a cadascuna de les simulacions quin és l'escenari que obté l'estimador de bondat de l'ajust més petit.

L'escenari que en les 24000 simulacions d'un determinat tipus tumoral i sexe, hagi sigut triat més vegades com a "escenari que millor ajusta" serà considerat l'escenari més plausible per aquest tipus de tumor.

Per validar si l'estimació que s'ha realitzat de la incidència és considerada acceptable o no acceptable es mirarà que el nombre de casos observats per a cada tipus de tumor i sexe estigui dintre de l'interval de credibilitat (IC) del 95% dels casos esperats.

Per aquells tipus de tumors on el nombre de casos observats no s'ha trobat dins d'aquest interval de credibilitat serà convenient explorar un altre mètode alternatiu al que s'ha realitzat en aquest projecte.

Les taules 6.1 i 6.2 mostren els resultats d'aquests càlculs, el percentatge de vegades que s'escull cada escenari així com l'IC per a cada tipus de tumor i sexe, tot indicant-se en color vermell aquells tumors on els resultats de l'estimació que s'ha fet no es pot considerar vàlida, i que per tant, s'hauria de buscar un mètode alternatiu.

HOMES	DETERMINACIÓ ESCENARI QUE MILLOR AJUSTA EN CADA SIMULACIÓ						MILLOR ESCENARI			
	A		B		C		Escenari	Casos	Casos	IC 95% (C. Esp)
TIPUS TUMORAL	N.sim	%	N.sim	%	N.sim	%	Triat	Observats	Esperats	
Cavitat oral i faringe	11316	47.15%	8806	36.69%	3878	16.16%	A	813	740	(560 - 930)
Esòfag	7364	30.68%	8145	33.94%	8491	35.38%	C	324	335	(255 - 431)
Estómac	6871	28.63%	8194	34.14%	8935	37.23%	C	690	768	(643 - 910)
Colon	5070	21.12%	10433	43.47%	8497	35.40%	B	2418	2665	(2330 - 2992)
Recte	8189	34.12%	8071	33.63%	7740	32.25%	A	1182	1470	(1294 - 1648)
Fetge	7027	29.28%	11275	46.98%	5698	23.74%	B	710	758	(620 - 904)
Vesícula biliar	10715	44.65%	5245	21.85%	8040	33.50%	A	163	160	(109 - 216)
Pàncrees	7837	32.65%	9132	38.05%	7031	29.30%	B	575	539	(454 - 632)
Laringe	7081	29.50%	8088	33.70%	8831	36.80%	C	494	523	(417 - 644)
Pulmó	7338	30.58%	10828	45.12%	5834	24.31%	B	3377	3297	(2986,98 - 3613)
Melanoma de pell	22763	94.85%	1001	4.17%	236	0.98%	A	427	614	(476 - 762)
Pròstata	7304	30.43%	4719	19.66%	11977	49.90%	C	4716	5241.5	(4159 - 6378,02)
Testicle	3	0.01%	0	0.00%	23997	99.99%	C	225	50	(35 - 68)
Ronyó	7159	29.83%	8868	36.95%	7973	33.22%	B	600	658	(557 - 767)
Bufeta urinària	13197	54.99%	10364	43.18%	439	1.83%	A	2694	3183	(2822 - 3546)
Encèfal i sistema nerviós	8668	36.12%	9015	37.56%	6317	26.32%	B	359	396	(327 - 476)
Tiroides	2	0.01%	0	0.00%	23998	99.99%	C	120	18	(9 - 31)
Limfoma de Hodgkin	12	0.05%	10	0.04%	23978	99.91%	C	134	24	(10 - 45)
Limfoma no hodgkinià	15118	62.99%	8849	36.87%	33	0.14%	A	604	777	(672 - 892)
Mieloma	8713	36.30%	6306	26.28%	8981	37.42%	C	246	227	(161 - 311)
Leucèmia	9937	41.40%	8839	36.83%	5224	21.77%	A	564	596	(483 - 719)
Altres	12428	51.78%	5862	24.42%	5710	23.79%	A	1324	1047	(936 - 1161)

Taula 6.1 Selecció del millor escenari així com l'IC al 95% per la predicció total de casos per cada tipus tumoral en homes durant el període 2009-2013.

DONES	DETERMINACIÓ ESCENARI QUE MILLOR AJUSTA EN CADA SIMULACIÓ						MILLOR ESCENARI			
	A		B		C		Escenari	Casos Observats	Casos Esperats	IC 95% (C. Esp)
TIPUS TUMORAL	N.sim	%	N.sim	%	N.sim	%	Triat			
Cavitat oral i faringe	10030	41.79%	3225	13.44%	10745	44.77%	C	264	278	(204 - 370)
Esòfag	4258	17.74%	5026	20.94%	14716	61.32%	C	53	6	(0 - 20)
Estómac	8219	34.25%	8230	34.29%	7551	31.46%	B	414	456	(379 - 541)
Colon	6776	28.23%	9595	39.98%	7629	31.79%	B	1715	1880	(1625 - 2148)
Recte	8772	36.55%	6574	27.39%	8654	36.06%	A	661	765	(669 - 875)
Fetge	8720	36.33%	9936	41.40%	5344	22.27%	B	248	240	(183 - 308)
Vesícula biliar	7790	32.46%	7848	32.70%	8362	34.84%	C	190	193	(133 - 273)
Pàncrees	7164	29.85%	8319	34.66%	8517	35.49%	C	479	504	(390 - 641,02)
Laringe	301	1.25%	1340	5.58%	22359	93.16%	C	38	7	(1 - 21)
Pulmó	8118	33.82%	8978	37.41%	6904	28.77%	B	719	753	(623 - 888)
Melanoma de pell	5371	22.38%	1042	4.34%	17587	73.28%	C	400	300	(219 - 395)
Mama	2261	9.42%	16172	67.38%	5567	23.20%	B	4682	4902	(4306 - 5507)
Coll uterí	9512	39.63%	1820	7.58%	12668	52.78%	C	304	536	(243 - 978)
Cos uterí	8854	36.89%	4626	19.28%	10520	43.83%	C	829	765	(326 - 1385)
Ovari	8796	36.65%	8550	35.62%	6654	27.72%	A	445	397	(328 - 474)
Ronyó	16034	66.81%	7126	29.69%	840	3.50%	A	290	328	(270 - 393)
Bufeta urinària	15211	63.38%	7809	32.54%	980	4.08%	A	453	531	(430 - 639)
Encèfal i sistema nerviós	7723	32.18%	8450	35.21%	7827	32.61%	B	270	242	(185 - 310)
Tiroides	0	0.00%	0	0.00%	24000	100.00%	C	400	22	(11 - 36)
Limfoma de Hodgkin	15	0.06%	0	0.00%	23985	99.94%	C	95	6	(1 - 15)
Limfoma no hodgkinià	18570	77.38%	5335	22.23%	95	0.40%	A	483	730	(638 - 832)
Mieloma	8770	36.54%	7083	29.51%	8147	33.95%	A	220	205	(146 - 271)
Leucèmia	8276	34.48%	7897	32.90%	7827	32.61%	A	428	398	(304 - 500)
Altres	8987	37.45%	8998	37.49%	6015	25.06%	B	1105	872	(375 - 1575)

Taula 6.2 Selecció del millor escenari així com l'IC al 95% per la predicció total de casos per cada tipus tumoral en dones durant el període 2009-2013.

Un cop es té quin és el millor escenari per a l'ajust de les dades, es pot observar per al període 2009-2013 la diferència que ha hagut entre les dades reals i les observades a la província de Tarragona. A les Figures 6.2 i 6.3 hi ha representats el nombre total de nous casos observats i predits de càncer de còlon d'homes i dones respectivament, on l'escenari triat ha sigut el B. A més, es representen també les dades de períodes anteriors, utilitzades per a fer la descriptiva del Capítol 2, i que correspon també al període en què s'han ajustat els models, per tal de poder veure amb més claredat quina ha estat l'evolució al llarg del temps.

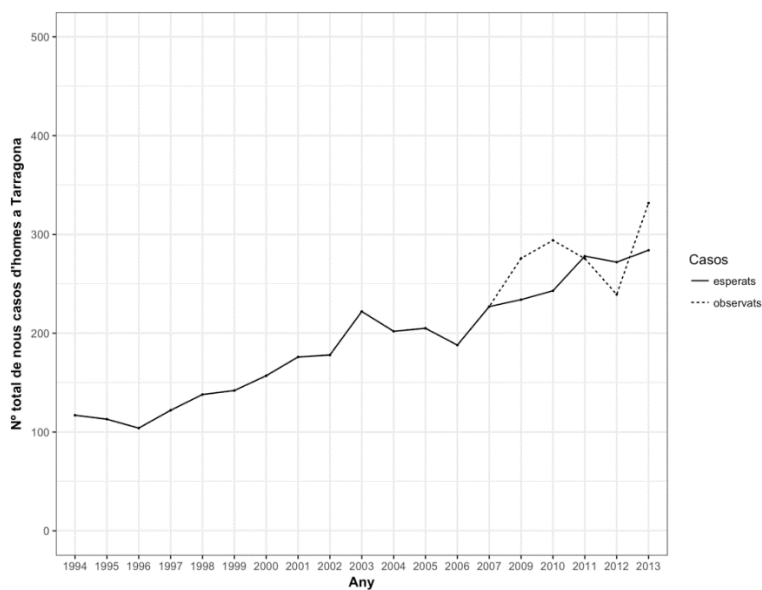


Figura 6.2 Evolució casos incidents observats i predits d'homes amb càncer de còlon a Tarragona, sota l'escenari B.

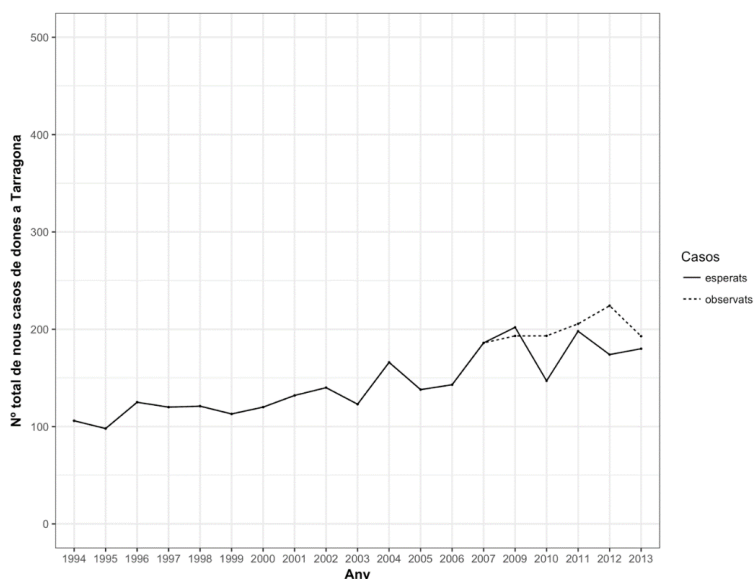


Figura 6.3 Evolució casos incidents observats i predits de dones amb càncer de còlon a Tarragona, sota l'escenari B.

7 ESTIMACIÓ DE LA INCIDÈNCIA PER A TOTES LES PROVÍNCIES

Un cop es tenen els resultats de la validació es procedeix a l'estimació de la incidència de cada tipus de tumor a totes les províncies d'Espanya.

Els resultats per a cada tumor es calculen tenint en compte quin dels tres escenaris és el que millor ajusta segons les dades dels Registres de Catalunya (Taules 6.1 i 6.2). A més, aquesta estimació es farà per al període 2009-2013 ja que d'aquesta forma es continua dins del mateix rang d'anys portat a terme en l'estudi.

Val a dir que aquest període d'anys seleccionat és també degut a què entre l'últim any que un Registre té dades registrades i l'any en curs hi ha un cert retard, i aquest fet succeeix en tots els RCBP. Això és necessari per tal de poder assegurar que queden registrats tots els casos incidents d'un període, de forma que l'adequat és que passin almenys dos o tres anys des de la fi del període, ja que hi haurà casos incidents que no es troben per les fonts d'informació habituals (hospitals), sinó que acaben arribant al cap d'un temps per altres fonts com el Registre de Mortalitat o altres.

El procediment es du a terme mitjançant els resultats extrets del model, de forma que per a les províncies amb RCBP es té un determinat valor que forma part de l'efecte aleatori d'aquella província, mentre que per a les províncies que no en tenen, a aquest factor aleatori se li assigna una $Normal(0, \sigma^2)$ (veure Quadre 5.1).

Un cop obtingudes aquestes estimacions es presentaran detalladament per al càncer de còlon.

En primer lloc, es calcularan les taxes brutes per a cada província. Com ja s'ha definit al Capítol 3, la taxa bruta es calcula com $TB = \frac{D}{P} \cdot 10^5$,

on, D és el nombre total de casos incidents en una província per un determinat any
 P és el nombre total d'habitants en aquella província

Caldrà tenir present que hi ha dues poblacions, la femenina i la masculina, les quals es tracten per separat. Amb això el resultat de la taxa bruta expressa el nombre de nous casos detectats per cada 100.000 habitants en l'any i població d'estudi.

La figura 7.1 és una representació de la intensitat de la taxa bruta per als homes diagnosticats amb càncer de còlon a Espanya. Es pot veure com ha canviat en cada província la intensitat de nous casos detectats al llarg dels quatre anys, així com l'existència d'un patró espacial on les taxes en general són més altes al nord-oest i disminueixen en la direcció sud-est. Aquest fet pot suggerir hipòtesis sobre els factors de risc associats al càncer de còlon. Cal tenir present, però, que es presenten les taxes

brutes i que per tant, part de les diferències entre províncies poden ser degudes a canvis en l'estructura de l'edat de les piràmides poblacionals.

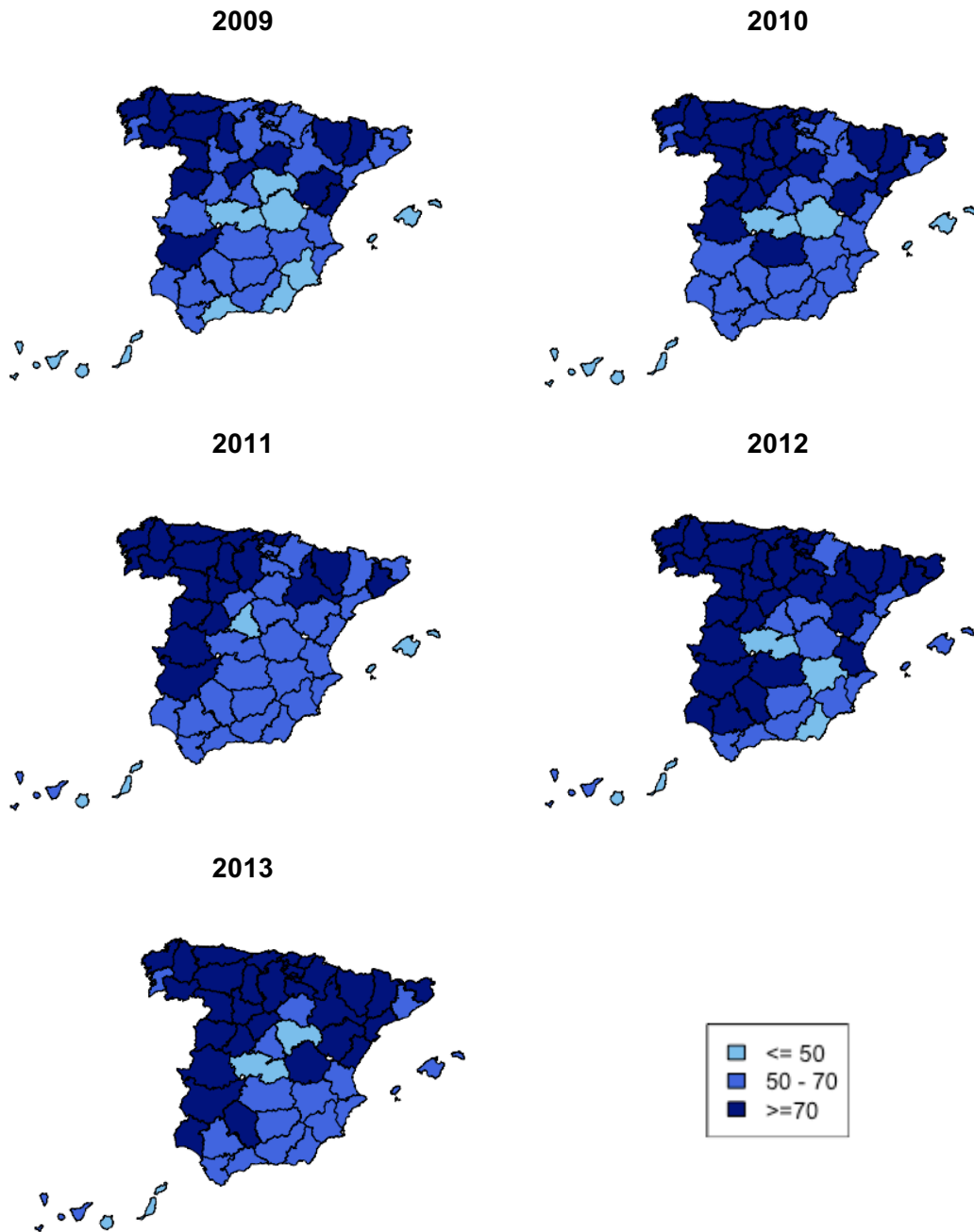


Figura 7.1 Representació taxes brutes per als casos incidents estimats en els anys 2009, 2010, 2011, 2012 i 2013 en homes amb càncer de còlon.

De la mateixa forma, s'observa a la Figura 7.2 que en el cas de les dones el nombre de nous casos diagnosticats amb càncer de còlon també ha anat variant en les diferents províncies de forma similar als homes, les taxes brutes més elevades s'observen sobretot al nord-oest.

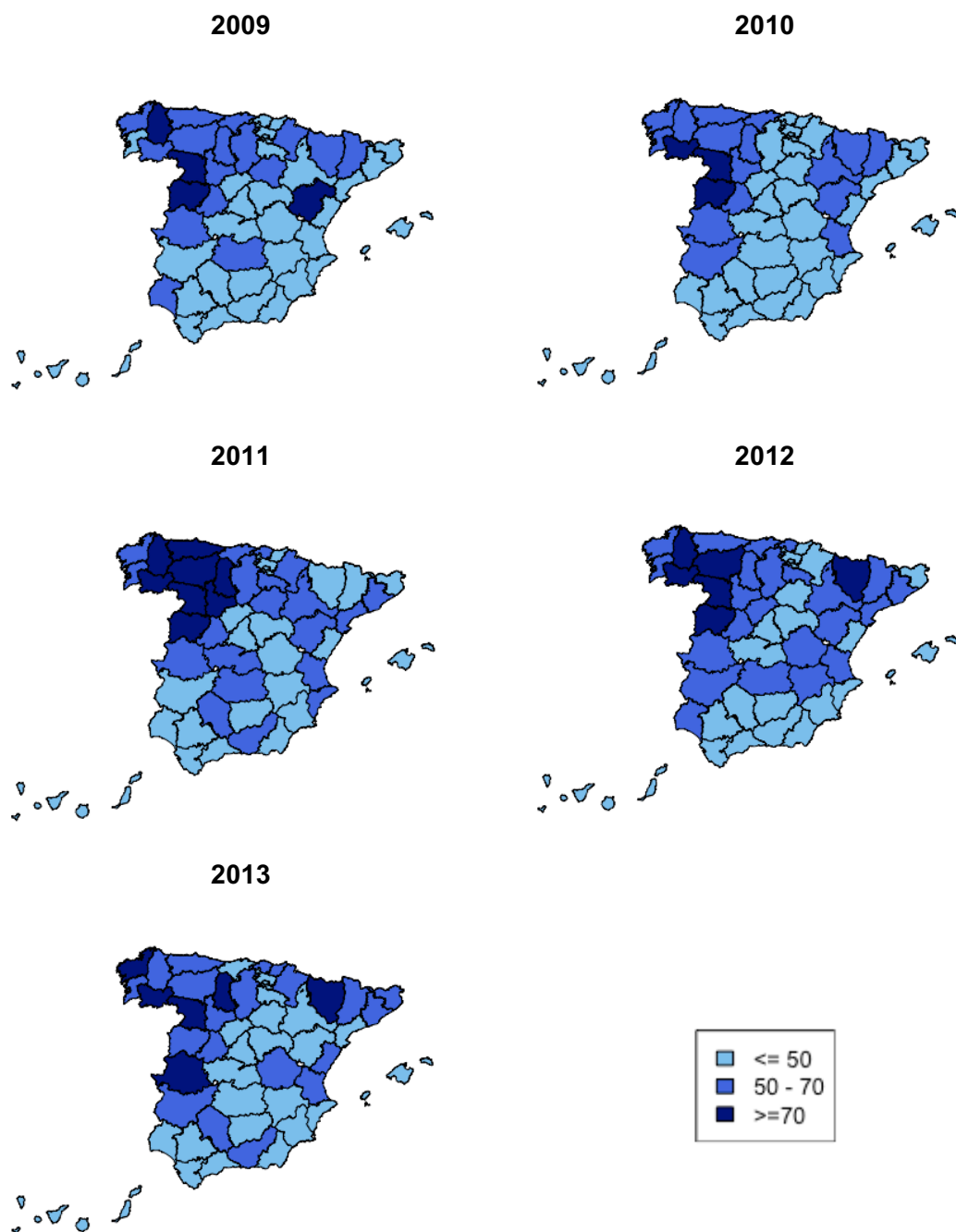


Figura 7.2 Representació taxes brutes per als casos incidents estimats en els anys 2009, 2010, 2011, 2012 i 2013 en dones amb càncer de còlon.

La taula 7.1 mostra el resultat de les taxes brutes per càncer de còlon per a cadascuna de les províncies per homes i dones a l'any 2013. La columna de la N fa referència al nombre total de casos incidents estimats per a l'any 2013 amb càncer de còlon. I seguidament, es presenta l'interval de credibilitat al 95% per aquesta estimació. Finalment la columna TB mostra el resultat de la taxa bruta que serien el nombre de nous casos per cada 100.000 habitants dones (o homes) en aquella mateixa província.

Província	Dones			Homes		
	N	IC	TB	N	IC	TB
Álava/Araba	78	(49 - 116)	48.45	135	(91 - 190)	85.53
Albacete	83	(52 - 124)	41.86	113	(76 - 159)	56.96
Alicante/Alacant	442	(297 - 630)	47.58	609	(427 - 836)	66.07
Almería	123	(79 - 179)	36.01	222	(153 - 306)	63.38
Ávila	52	(31 - 80)	62.97	64	(41 - 93)	75.58
Badajoz	221	(146 - 317)	63.53	297	(206 - 409.02)	86.77
Illes Balears	249	(165 - 357)	44.86	285	(197 - 392.02)	51.14
Barcelona	1561	(1065 - 2214)	55.75	1829	(1302 - 2483)	68.61
Burgos	120	(77 - 176)	65.88	156	(106 - 220)	84.63
Cáceres	157	(101 - 228)	76.54	216	(148 - 299)	105.79
Cádiz	267	(178 - 380)	42.56	400	(280 - 548)	64.38
Castellón/Castelló	176	(116 - 255)	60.50	211	(145 - 294)	72.85
Ciudad Real	126	(81 - 183)	48.04	146	(99 - 205)	56.47
Córdoba	228	(150 - 327)	56.02	316	(221 - 436)	80.56
A Coruña	419	(281 - 596)	71.11	528	(370 - 722)	96.66
Cuenca	70	(43 - 105)	67.91	87	(57 - 124)	82.05
Girona	224	(183 - 268)	60.19	282	(235 - 330)	75.28
Granada	241	(160 - 346)	51.93	282	(196 - 387)	61.64
Guadalajara	61	(37 - 92)	49.08	59	(37 - 87)	45.46
Guipúzcoa/Gipuzkoa	189	(123 - 273)	52.24	265	(183 - 367)	76.81
Huelva	109	(69 - 160)	41.34	201	(139 - 279)	77.92
Huesca	82	(51 - 122)	74.73	100	(66 - 143)	88.49
Jaén	141	(92 - 205)	42.73	215	(148 - 299)	65.68
León	174	(114 - 251)	69.96	257	(177 - 355)	107.95
Lleida	113	(71 - 166)	52.80	159	(107 - 222)	72.19
La Rioja	54	(32 - 82)	34.15	122	(82 - 172)	77.71
Lugo	103	(64 - 152)	58.34	185	(126 - 258)	110.76
Madrid	1411	(960 - 2000)	42.52	1844	(1310 - 2509)	59.99
Málaga, Ceuta y Melilla	388	(261 - 552)	43.16	530	(373 - 725.02)	60.04
Murcia	273	(181 - 389)	37.53	378	(265 - 520)	51.46
Navarra	196	(129 - 281)	61.14	241	(167 - 334)	76.26

Ourense	130	(84 - 189)	78.01	224	(155 - 312)	143.69
Asturias	362	(240 - 522)	65.35	553	(389 - 758)	108.63
Palencia	61	(37 - 91)	71.86	80	(52 - 114)	96.12
Las Palmas	245	(163 - 350)	44.71	252	(175 - 348)	45.81
Pontevedra	280	(187 - 402)	57.31	308	(214 - 424)	66.63
Salamanca	123	(78 - 181)	69.51	165	(112 - 231)	97.61
Santa Cruz de Tenerife	200	(132 - 289)	39.37	296	(206 - 409)	59.17
Cantabria	150	(97 - 218)	49.80	239	(166 - 330)	83.11
Segovia	30	(16 - 48)	38.03	61	(38 - 89)	75.44
Sevilla	439	(295 - 624)	44.51	625	(441 - 854)	65.90
Soria	23	(11 - 37)	49.25	29	(16 - 44)	60.52
Tarragona	193	(155 - 232)	48.55	332	(279 - 386)	82.68
Teruel	27	(14 - 44)	39.75	81	(53 - 116)	113.70
Toledo	167	(109 - 242)	48.53	176	(120 - 244)	49.59
Valencia/València	768	(522 - 1089)	59.66	859	(606 - 1173)	68.81
Valladolid	169	(111 - 244)	62.65	221	(152 - 306)	85.13
Vizcaya/Bizkaia	367	(248 - 524)	62.09	494	(347 - 677)	89.40
Zamora	86	(54 - 127)	90.74	142	(96 - 200)	152.79
Zaragoza	244	(161 - 352)	49.63	365	(255 - 501)	76.26

Taula 7.1 Nombre total de casos incidents estimats, interval de credibilitat del 95% per al nombre total de casos incidents i taxa bruta per a cada província i sexe, durant el període 2009-2013.

Per fer els mapes mostrats en aquest apartat, així com la darrera la taula, s'han generat dues funcions amb el software R. D'aquesta forma, introduint el codi d'un altre tipus de tumor es poden generar els mateixos resultats que els que s'han mostrat en aquesta Secció 7.

A l'Annex 5 es mostren els mapes per l'any 2013 per a alguns dels tumors més coneguts. S'ha utilitzat el mateix criteri de separació per a les taxes brutes que en el cas del Càncer de Còlon, és a dir, el mapa ha quedat representat segons si la taxa bruta és menor de 50, entre 50 i 70 o major de 70, de forma que es pugui veure de forma més comparativa com són els impactes segons els diferents tumors sobre les províncies d'Espanya. A més, al final de l'Annex 6 es poden veure les funcions utilitzades per generar-ho.

8 CONCLUSIONS

Aquest projecte ha presentat un mètode per estimar les dades d'incidència en la situació on només es coneix la incidència d'una part del territori, però es disposa de la mortalitat de tot el territori, així com la validació d'aquest mateix mètode.

Més concretament, el mètode es basa en que a partir de les dades d'incidència i mortalitat d'aquelles províncies i períodes (del 1994 al 2007) en els que es disposen de dades dels registres de càncer, es pugui modelar la raó entre la incidència i la mortalitat, incloent l'edat i l'any de diagnòstic com a variables predictores del model i la província com a un efecte aleatori.

Tot seguit, per validar aquest model s'han comparat les dades reals d'incidència a Tarragona i Girona per al període 2009-2013 amb una estimació del nombre de casos incidents durant aquest mateix període en aquestes províncies segons tres possibles escenaris d'evolució de la RIM (raó entre incidència i mortalitat): constant, lineal o quadràtica.

Al capítol 6, es presenten les taules on es determina l'escenari més plausible per a cada tipus de tumor, així com el nombre de casos estimats en aquest escenari i el seu interval de credibilitat al 95%. Per validar si l'estimació que s'ha realitzat de la incidència és considerada acceptable o no acceptable s'ha mirat que el nombre de casos observats per a cada tipus de tumor i sexe estigui dintre de l'interval de credibilitat dels casos esperats.

En el cas dels homes dels 22 tipus de tumors estudiats, 8 han presentat estimacions no acceptables mentre que en el cas de les dones dels 24 tipus estudiats, 5 han presentat estimacions no acceptables.

Dintre dels tipus de tumors amb estimacions que no es poden considerar com a acceptables, una bona part d'ells (Esòfag, Melanoma de Pell, Laringe, Testicle, Tiroides, Limfoma de Hodgkin, Limfoma no Hodgkin) corresponen amb baixa letalitat³ i en alguns casos, poca mortalitat, la qual cosa implica que els valors de la RIM són alts i inestables a diferència dels tumors letals amb RIMs altes i estables.

En el cas dels tipus tumorals Recte i Bufeta Urinària s'ha produït una sobreestimació. En el cas del càncer de Recte aquesta sobreestimació podria ser deguda a un increment artificial de la incidència en els últims anys del període d'ajust del model motivat per l'avançament del diagnòstic que produeix en un territori la implantació del cribratge, mentre que en el cas del càncer de Bufeta Urinària aquestes diferències podrien estar

³ La taxa de letalitat fa referència a la proporció de persones que moren per una malaltia entre els afectats per la mateixa en un període i àrea determinats.

influïdes pels canvis en la definició de càncer invasiu de Bufeta Urinària al llarg del temps.

També és interessant ressaltar que una gran part dels tumors amb estimacions no acceptables tenien com a escenari triat el C, que representa una hipòtesi de tendència quadràtica de l'evolució temporal del logaritme de la RIM. La conclusió que resulta d'aquesta dada és que s'ha d'anar amb molta cura amb les estimacions fetes en les hipòtesis quadràtiques sobre el logaritme de la RIM, sobretot si aquestes són molt diferents dels escenaris A (evolució de la RIM constant) i B (evolució de la RIM log-lineal).

La conclusió final sobre els mètodes que utilitzen models RIM per estimar la incidència en la situació on només es coneix la incidència d'una part del territori, però es disposa de la mortalitat del territori, és que aquests mètodes produeixen estimacions acceptables en la majoria dels tipus de tumors, però que tenen certes limitacions i poden produir estimacions no acceptables en les següents situacions:

- Quan el nombre de morts és petit
- Quan les raons entre incidència i mortalitat (RIM) són altes degut a la baixa letalitat del tipus tumoral
- Quan hi ha canvis bruscos de la RIM en relació amb el període o l'edat al diagnòstic
- Quan el cribratge produeix canvis forts en la RIM en relació amb l'edat, el període de diagnòstic o la província
- Quan ràpids increments de la incidència poden produir un augment sobtat de la RIM, inclús mantenint-se la proporció de casos fatals.

En totes aquestes situacions, seria necessari seguir estudiant nous mètodes i/o models d'estimacions que presentin uns valors de casos incidents més acurats.

BIBLIOGRAFIA

1. Marcos-Gragera R, Cardó X, Galceran J, Ribes J, Izquierdo A, Borràs J. Incidencia del càncer en Catalunya, 1998-2002. *Med Clin (Barc)*. 2008 Oct ;131 Suppl 1:4-10.
2. Clèries R, Ameijide A, Marcos-Gragera R, Pareja L, Carulla M, Vilardell ML, Esteban L, Buxó M, Espinàs JA, Puigdefàbregas A, Ribes J, Izquierdo A, Galceran J, Borràs JM. Predicting the cancer burden in Catalonia between 2015 and 2025: the challenge of cancer management in the elderly. *Clin Transl Oncol*. 2018 May;20(5):647-657.
3. MacLennan R, Muir CS, Steinitz R, Winkler A, editors. *Cancer registration and its techniques*. IARC Scientific Publication No. 21. Lyon: International Agency for Research on Cancer; 1978.
4. Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG. *Cancer registration. Principles and methods*. IARC Scientific Publication No. 95. Lyon: International Agency for Research on Cancer; 1991.
5. Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL. *Comparability and quality control in Cancer registration*. IARC Technical report No. 19. Lyon: International Agency for Research on Cancer; 1994.
6. Wagner G. *Cancer registration: Historical aspects*. A: Parkin DM, Wagner G, Muir CS. *The Role of the Registry in Cancer Control*. IARC Scientific Publication No. 66. Lyon: International Agency for Research on Cancer; 1985.
7. Stocks P. *Cancer registration and studies of incidence by surveys*. *Bull World Health Organ*. 1959;20:697-715.
8. Clemmesen J. *Symposium on the geographic pathology and demography of cancer*. Paris: Council for the Coordination of International Congresses of Medical Statistics; 1951.
9. Doll R, Payne P, Waterhouse JAH, editors. *Cancer Incidence in Five Continents, Vol I*. Geneva; Union Internationale Contre le Cancer; 1966.
10. Parkin DM. The evolution of the population-based cancer registry. *Nature reviews* 2006;6:603-612.

11. Valsecchi MG, Sterialova-Foucher E. Cancer registration in developing countries: luxury or necessity?. *Lancet Oncology* 2008;9:159-67.
12. Sankila R, Démaret E, Hakama M, Lynge E, Schouten L, Parkin DM, editors. Evaluation and monitoring of cancer screening Programmes. Brussels, Luxembourg: European Commission; 2001.
13. Sankila R, Black R, Coebergh JW, Démarte E, Forman D, Gatta G, Parkin DM. Evaluation of Clinical Care by Cancer Registries. IARC Technical Publication No. 37. Lyon, France, 2003.
14. REDECAN: Red Española de Registros de Cáncer: <http://redecana.org> (visitada el maig de 2018).
15. FUNCA: Fundació Lliga per a la Investigació i Prevenció del càncer: <http://funca.cat/registre> (visitada el maig de 2018).
16. ICO: Institut Català d'Oncologia, Registre del càncer: http://ico.gencat.cat/ca/professionals/serveis_i_programes/registre_del_cancer/ (visitada el maig de 2018).
17. Muir CS, Waterhouse J. Comparability and quality of data: reliability of registration. En: Muir CS, Waterhouse J, Mack T, Powell J, Whelan S, editors. Cancer incidence in five continents, Vol V. IARC Scientific Publication No. 88. Lyon: International Agency for Research on Cancer; 1987.
18. Tyczynski JE, Démaret E, Parkin DM, editors. Standards and guidelines for cancer registration in Europe. The ENCR recommendations. IARC Technical Publication No. 40. Lyon: International Agency for Research on Cancer; 2003.
19. Borràs J, Galceran J, Creus J, Bosch FX, Viladiu P, Campillo M. Cancer incidence in Tarragona, Catalonia, Spain. 1980-1983. En: Muir C., Waterhouse J, Mack T, Powell J, Whelan S, editors. Cancer incidence in five continents. Volume V. IARC Scientific Publications N° 88. Lyon: International Agency for Research on Cancer; 1987.
20. Borràs J, Galceran J, Moreno V, Anglada LI, Bosch FX, Viladiu P. Cancer incidence in Tarragona, Catalonia, Spain. 1984-1987. En: Parkin DM, Muir CS, Whelan SL, Gao YT, Ferlay J, Powell J. Cancer incidence in five continents. Volume VI. IARC Scientific Publications N° 120. Lyon: International Agency for Research on Cancer; 1992.
21. Galceran J, Borràs J, Moreno V, Anglada LI, Bosch FX, Viladiu P.. Cancer incidence in Tarragona, Catalonia, Spain. 1988 1992. En: Parkin DM, Whelan

- SL, Ferlay J , Raymond L , Young J. Cancer incidence in five continents. Volume VII. IARC Scientific Publications N°143. Lyon: International Agency for Research on Cancer; 1997.
22. Borràs J, Galceran J, Piñol JLI, Rubió J, Moreno V, Muñoz J, Lafuerza A, Borràs JM, Viladiu P, Bosch FX. Cancer Incidence in Tarragona, Spain, 1993-1998. En: Parkin DM., Whelan SL., Ferlay J., Teppo L., Thomas DB, editors. Cancer incidence in five continents. Volume VIII. IARC Scientific Publications N° 155. Lyon: International Agency for Research on Cancer; 2002.
23. Borràs J, Galceran J, Piñol JLI, Cardó X, Ameijide A, Izquierdo A, Marcos R, Moreno V, Viladiu P, Bosch FX. Cancer incidence in Tarragona, Spain, 1998-2001. En: Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P, editors. Cancer incidence in five continents, Vol. IX. IARC Scientific Publications No. 160, Lyon: International Agency for Research on Cancer; 2007.
24. Berrino F, Sant M, Verdecchia A, Capocaccia R, Hakulinen T, Estève J, editors. Survival of cancer patients in Europe. The EURO CARE Study. IARC Scientific Publications No. 132. Lyon: International Agency for Research on Cancer; 1995.
25. Berrino F, Capocaccia R, Estève J, Gatta G, Hakulinen T, Micheli A, Sant M, Verdecchia A, editors. Survival of cancer patients in Europe. The EURO CARE-2 Study. IARC Scientific Publications No. 151. Lyon: International Agency for Research on Cancer; 1999.
26. Berrino F, Capocaccia R, Coleman MP, Estève J, Gatta G, Hakulinen T, Micheli A, Sant M, Verdecchia A, editors. Survival of cancer patients in Europe: The EURO CARE-3 Study. *Ann Oncol.* 2003;14 Suppl 5:v1-155.
27. Capocaccia R, Gavin A, Hakulinen T, Lutz JM, Sant M, editors. Survival of cancer patients in Europe, 1995-2002: The EURO CARE-4 Study. *Eur J Cancer*, 2009 Apr; 45(6):901-1094.
28. Galceran J. Childhood Cancer incidence in Tarragona, Spain. En: Parkin DM, Kramárová E, Draper GJ, Masuyer E, Michaelis J, Neglia J, Qureshi S, Stiller CA, editors. International Childhood Cancer Incidence. Vol. II. IARC, Scientific Publications N° 144. Lyon: International Agency for Research on Cancer; 1998.
29. Steliarova-Foucher E, Coebergh JW, Kaatsch P, Pritchard-Jones K, Stiller C, editors. Cancer in Children and Adolescents in Europe. *Eur J Cancer* 2006 Sep;42(13):1915-2190.
30. Galceran J, Ameijide A, Carulla M, Mateos A, Quirós JR, Rojas D, Alemán A, Torrella A, Chico M, Vicente M, Díaz JM, Larrañaga N, Marcos-Gragera R,

Sánchez MJ, Perucha J, Franch P, Navarro C, Ardanaz E, Bigorra J, Rodrigo P, Bonet RP and REDECAN Working Group. Cancer incidence in Spain, 2015. *Clin Transl Oncol.* 2017 Jul;19(7):799-825.

31. Bolstad, William M. *Introduction to Bayesian Statistics*. 2nd ed. Hoboken, N.J.: John Wiley, 2007.
32. Gelman, Andrew. *Bayesian data analysis*. 3rd ed. London: Chapman & Hall, 2014.
33. Congdon, Peter. *Applied Bayesian Modelling*. Wiley, 2014.
34. Kruschke, John K. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Amsterdam: Academic Press, 2015.

ANNEX 1. TAXES BRUTES D'INCIDÈNCIA

La següent taula mostra els resultats de les taxes brutes de casos incidents per 100.000 habitants a les províncies de Tarragona i Girona per aquells tumors que són comuns en homes i dones i pels anys 1994, 2000 i 2007.

Tipus tumoral	Tarragona			Girona		
	1994	2000	2007	1994	2000	2007
Cavitat Oral i Faringe	16.96	17.29	16.99	17.93	15.08	11.73
Esòfag	4.64	5.20	4.18	6.94	6.36	5.51
Estómac	21.43	19.98	10.59	27.96	21.81	17.67
Còlon	39.82	46.51	53.97	40.69	49.06	43.53
Recte	17.86	22.33	24.83	21.02	18.72	17.38
Fetge	6.79	12.59	9.41	13.31	11.45	13.43
Vies biliars	5.00	4.87	3.92	6.75	4.18	5.37
Pàncrees	11.07	11.75	10.72	8.68	12.90	11.17
Laringe	11.61	9.23	7.19	10.41	10.72	5.94
Pulmó	39.64	52.55	49.53	54.58	55.24	46.92
Melanoma de pell	8.39	10.24	10.59	5.01	8.18	10.18
Ronyó	5.54	9.07	6.93	9.26	9.63	10.18
Bufeta urinària	38.21	41.30	42.47	35.68	42.70	38.87
Encèfal i sistema nerviós	9.11	9.07	6.80	8.10	8.72	9.05
Tiroides	4.82	4.37	4.57	4.05	7.81	5.51
Limfoma Hodgkinià	2.68	3.02	2.74	2.12	3.27	2.26
Limfoma No Hodgkinià	13.57	17.80	13.07	15.43	16.90	17.52
Mieloma	5.18	5.88	4.70	5.01	6.91	6.64
Leucèmia	12.86	13.60	11.11	10.80	11.45	17.52

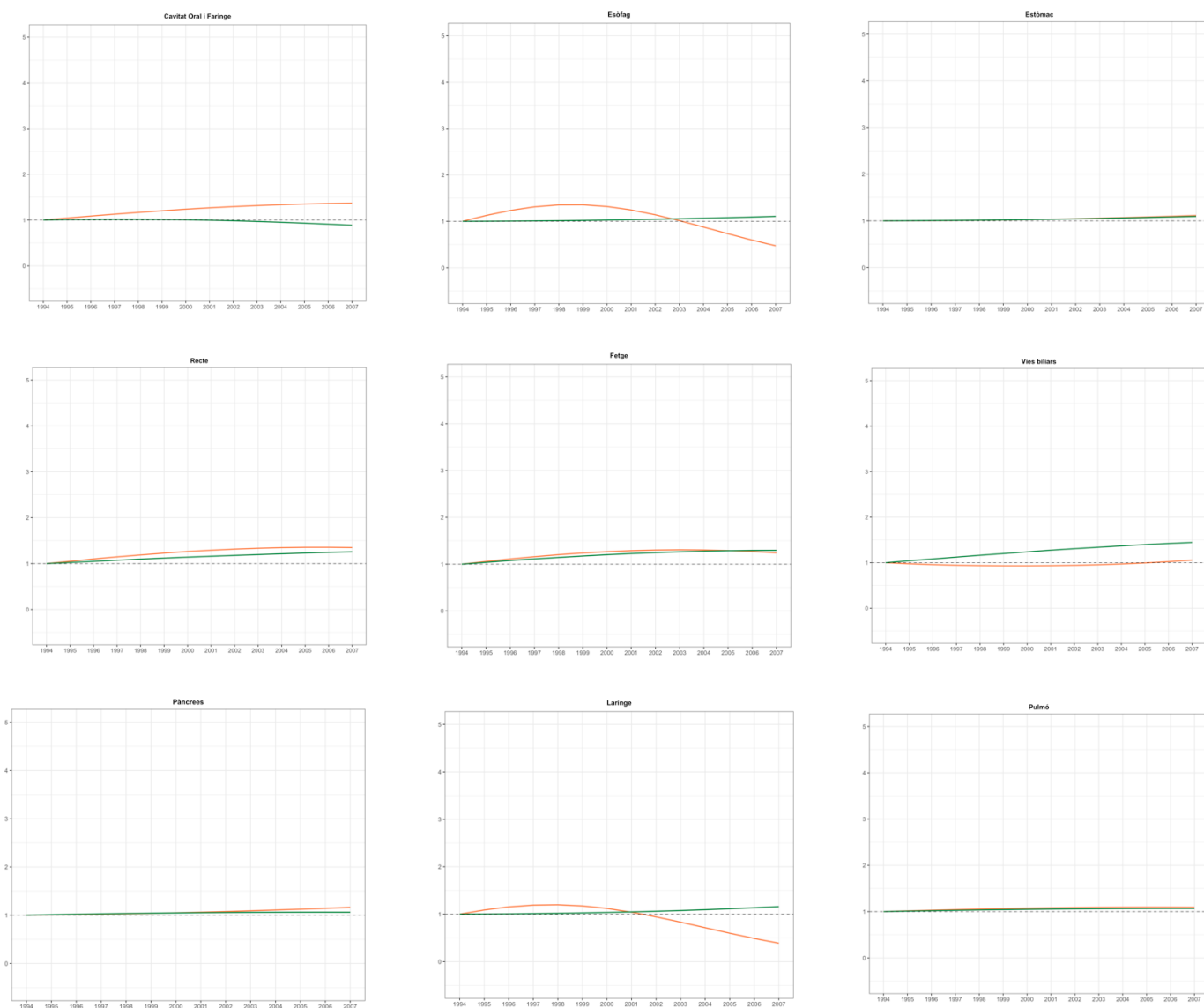
ANNEX 2. TAXES BRUTES DE MORTALITAT

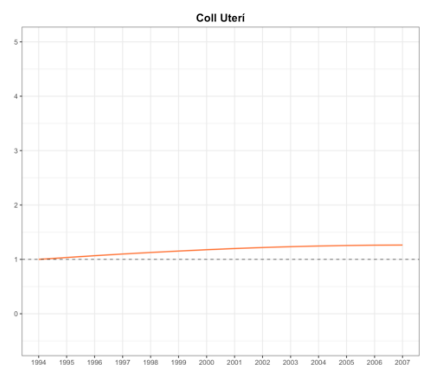
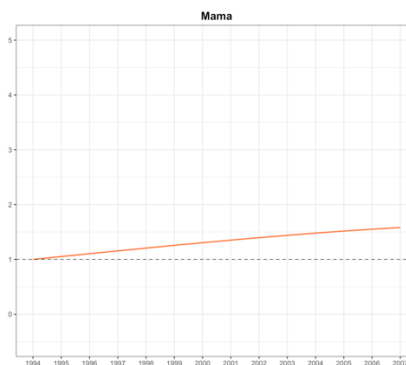
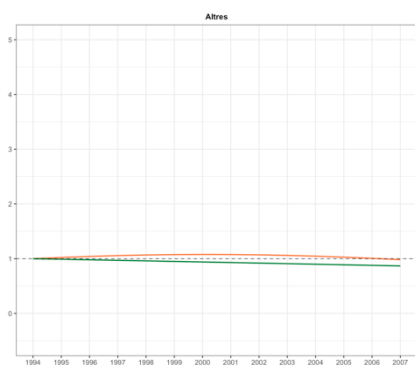
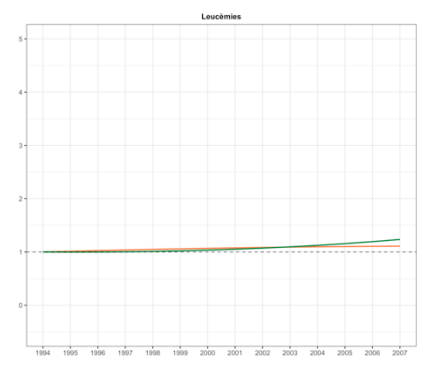
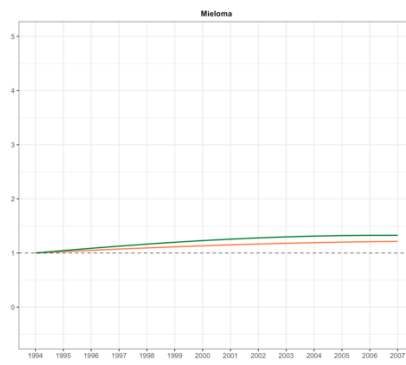
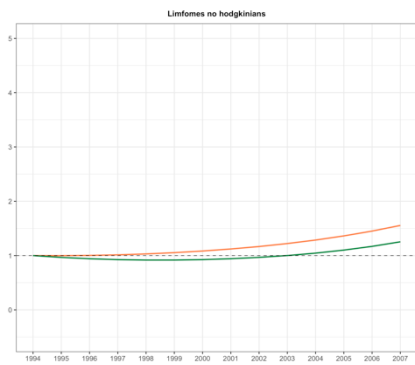
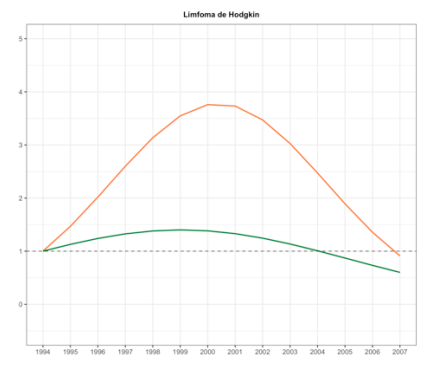
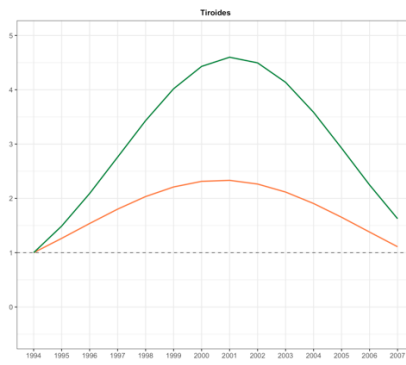
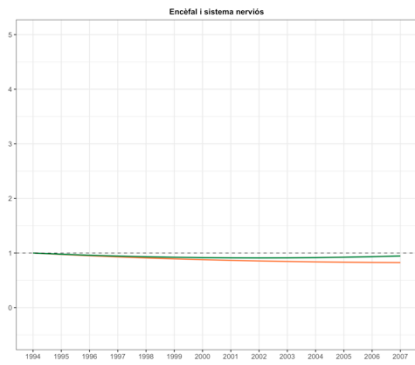
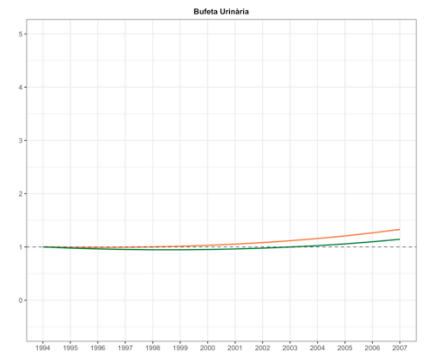
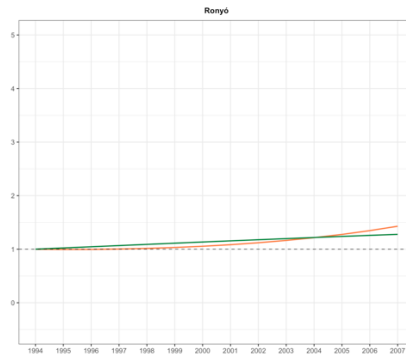
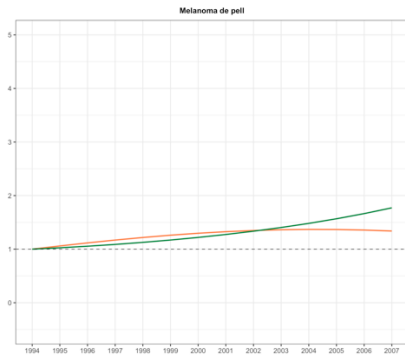
La següent taula mostra els resultats de les taxes brutes de casos incidents per 100.000 habitants a les províncies de Tarragona i Girona per aquells tumors que són comuns en homes i dones i pels anys 1994, 2000 i 2007.

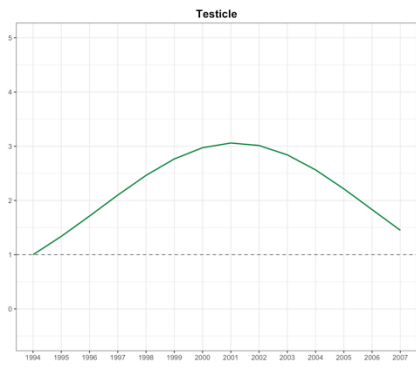
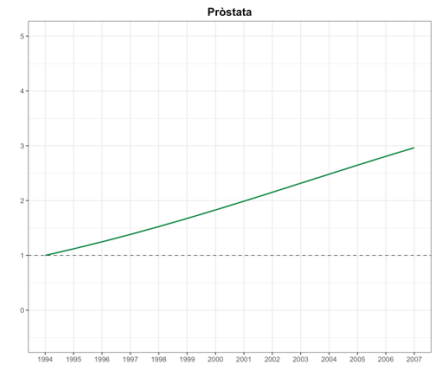
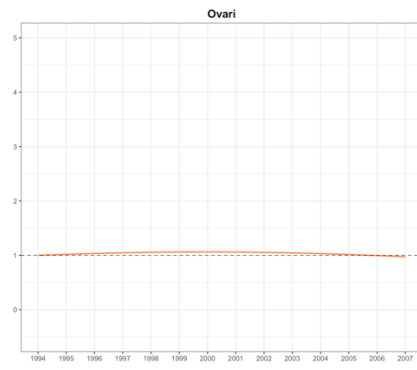
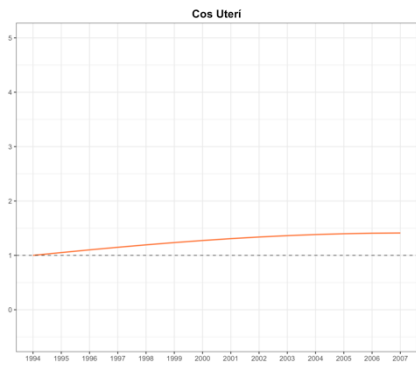
Tipus tumoral	Tarragona			Girona		
	1994	2000	2007	1994	2000	2007
Cavitat Oral i Faringe	5.54	6.04	6.01	4.24	3.82	4.38
Esòfag	3.04	4.03	2.35	5.40	3.82	3.96
Estómac	13.93	13.26	7.84	13.31	12.36	10.74
Còlon	18.75	22.67	23.92	22.95	22.71	21.20
Recte	6.96	4.70	5.36	6.75	7.09	4.81
Fetge	5.18	8.39	8.63	13.50	8.36	9.19
Vies biliars	2.68	3.36	2.35	3.09	2.91	3.96
Pàncrees	10.54	9.23	7.97	10.41	11.27	8.34
Laringe	4.29	3.86	2.74	4.82	4.00	2.83
Pulmó	35.89	40.46	38.42	42.43	44.88	41.27
Melanoma de pell	1.43	1.85	0.78	0.96	2.00	1.27
Ronyó	0.00	0.00	0.00	0.00	0.00	0.00
Bufeta urinària	0.00	0.00	0.00	0.00	0.00	0.00
Encèfal i sistema nerviós	2.14	2.52	2.22	2.70	3.09	4.66
Tiroides	10.00	9.57	7.32	7.91	8.36	9.19
Limfoma Hodgkinià	5.71	5.04	3.14	5.59	5.81	5.79
Limfoma No Hodgkinià	0.54	0.50	0.78	0.19	0.55	0.28
Mieloma	0.89	1.34	0.26	0.00	0.00	0.00
Leucèmia	0.00	0.00	0.00	0.00	0.00	0.00

ANNEX 3. EFECTE DE L'ANY SOBRE LA RIM

Els següents gràfics mostren per a cada tipus tumoral l'efecte de l'any sobre el RIM, on el 1994 és sempre el valor basal 1. Així, tendències positives estan relacionades amb un augment de la supervivència.

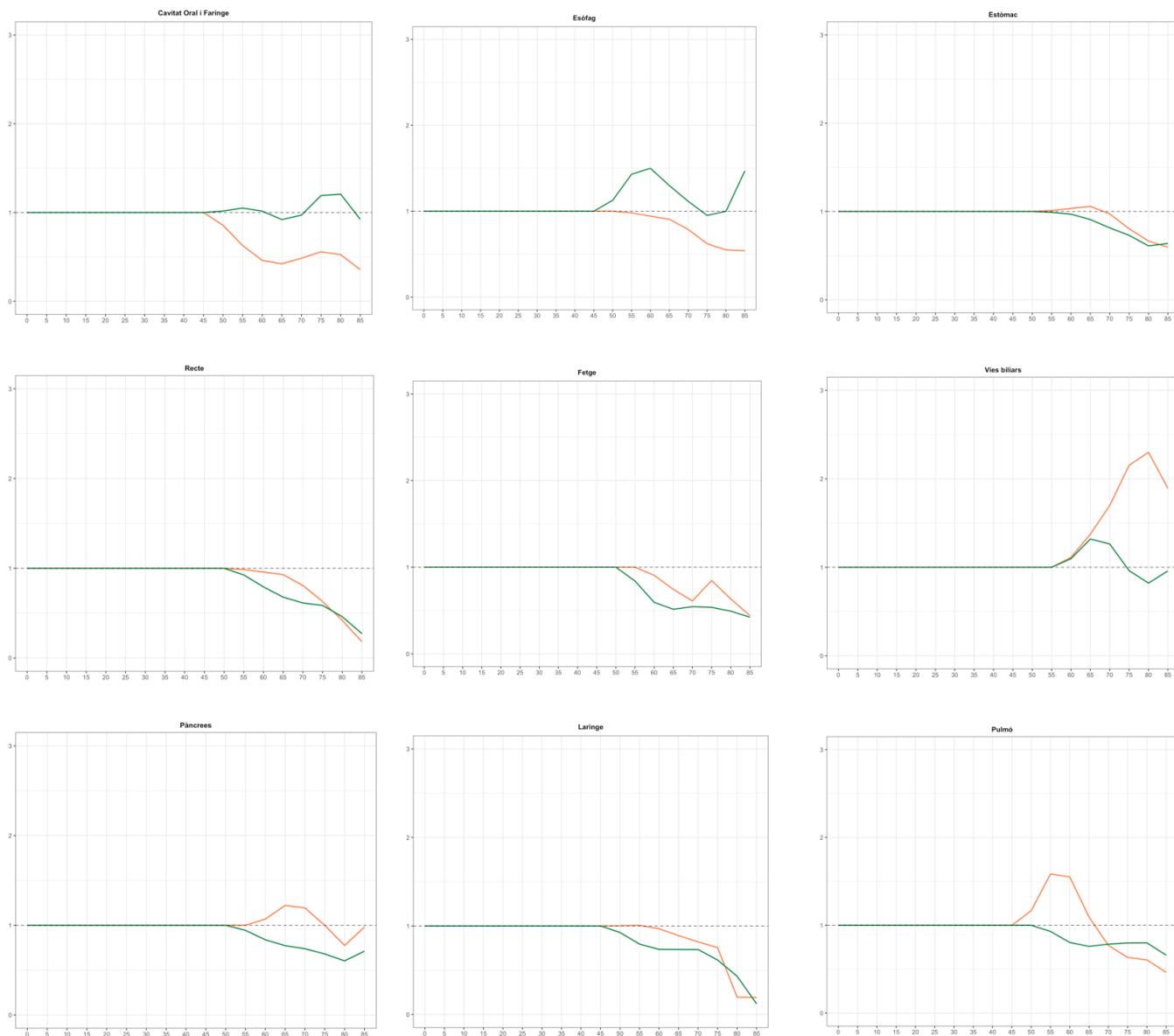


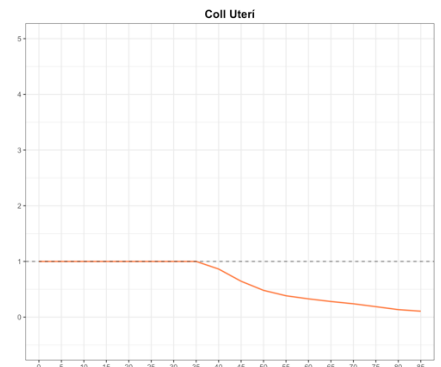
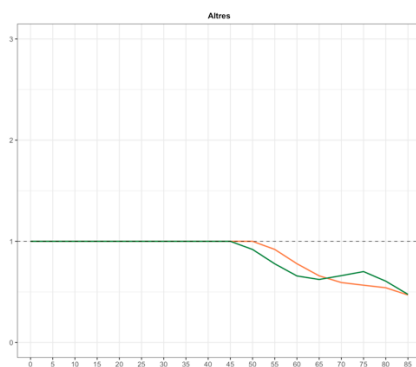
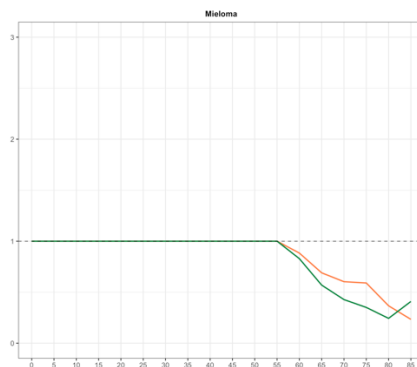
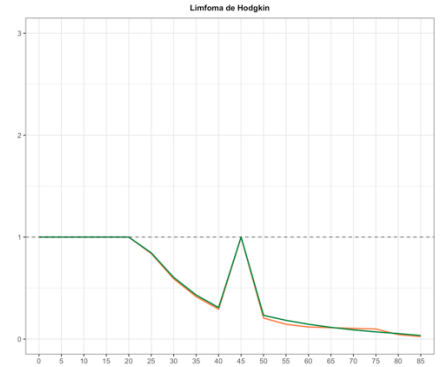
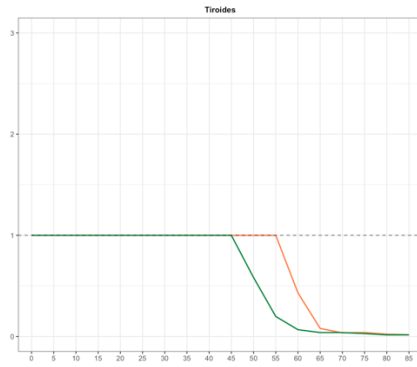
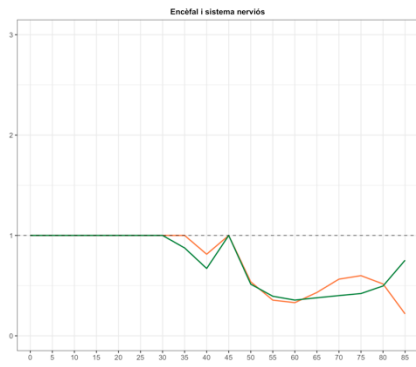
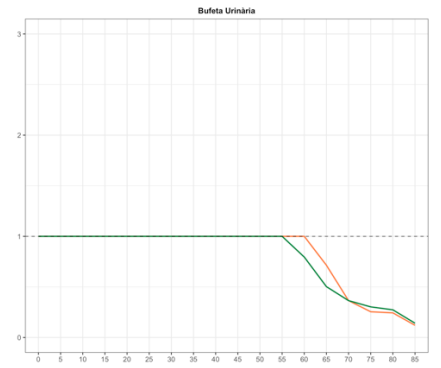
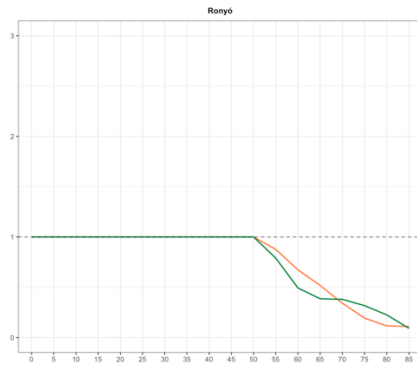
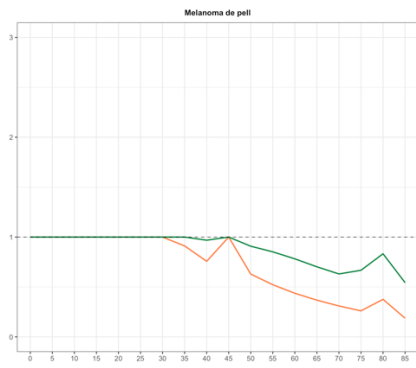


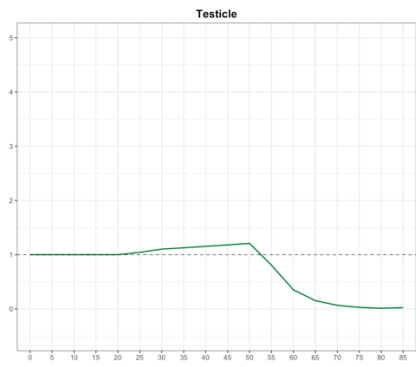
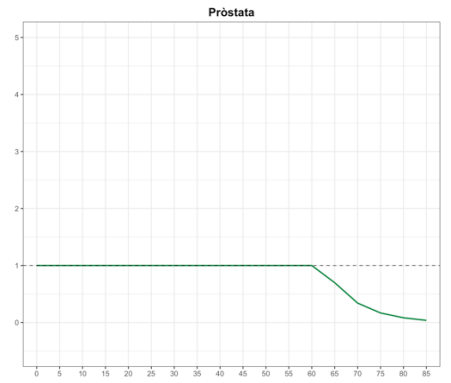
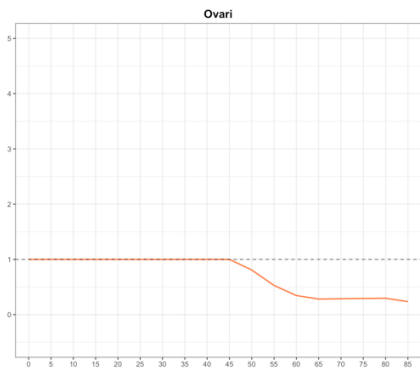
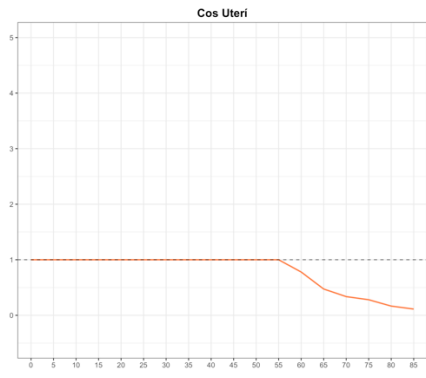


ANNEX 4. EFECTE DE L'EDAT SOBRE LA RIM

Els següents gràfics mostren per a tots els tipus tumorals l'efecte de l'edat sobre el valor de la RIM, de forma que les primeres edats, per tal com s'ha definit el model a la secció 5.1, és sempre 1 i representa el valor basal.



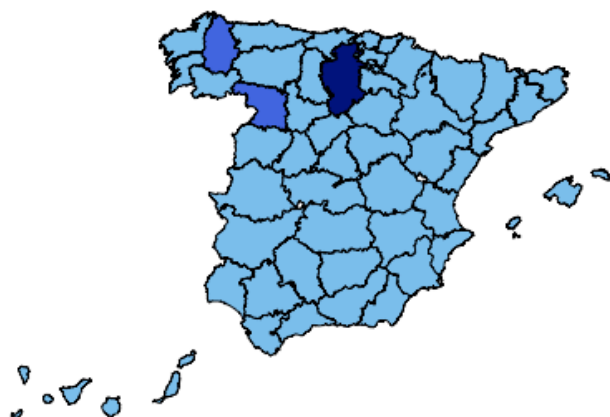




ANNEX 5. TAXES BRUTES I MAPES D'ESPANYA

En aquest apartat es presenten mapes per a alguns dels tumors més coneguts per l'any 2013, a més a més, s'afegeixen també les respectives taules en quant a les taxes brutes per aquest mateix any.

Cavitat Oral i Faringe en homes



Cavitat Oral i Faringe en dones



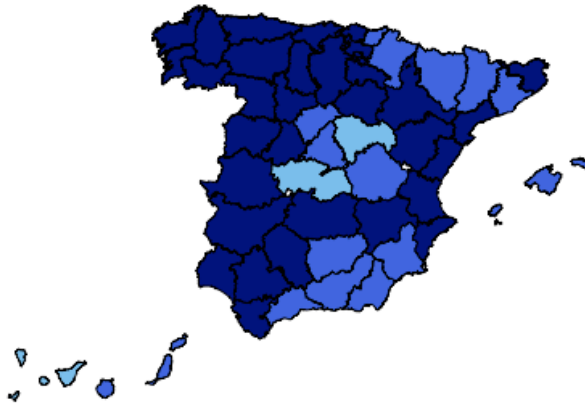
Pulmó en homes



Pulmó en dones



Bufeta Urinària en homes



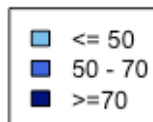
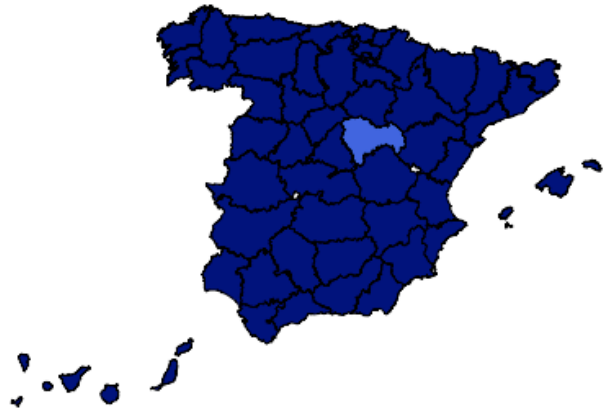
Bufeta Urinària en dones



Pròstata



Mama



Cavitat Oral i Faringe

Província	Homes			Dones		
	N	IC	TB	N	IC	TB
Álava/Araba	43	(17-86)	27.14	9	(3-18)	5.61
Albacete	37	(14-75)	18.46	29	(13-50)	14.40
Alicante/Alacant	213	(92-413)	23.09	107	(58-176)	11.48
Almería	73	(29-145)	20.87	8	(2-16)	2.33
Ávila	23	(8-49)	27.61	3	(0-8)	4.03
Badajoz	84	(35-166)	24.51	18	(7-33)	5.09
Illes Balears	118	(50-232)	21.16	26	(12-45)	4.64
Barcelona	775	(343-1496.02)	29.08	269	(152-434)	9.60
Burgos	139	(59-273)	75.29	23	(10-41)	12.74
Cáceres	70	(29-139)	34.43	9	(2-18)	4.24
Cádiz	176	(74-344)	28.28	59	(30-99)	9.49
Castellón/Castelló	84	(35-165)	28.85	15	(6-28)	5.21
Ciudad Real	54	(21-107)	20.70	22	(9-39)	8.30
Córdoba	122	(52-241)	31.20	40	(20-67)	9.81
A Coruña	241	(105-467)	44.12	63	(33-105)	10.69
Cuenca	13	(3-29)	12.41	2	(0-6)	2.21
Girona	58	(37-81)	15.49	34	(19-53)	9.12
Granada	89	(37-176)	19.51	49	(24-82)	10.49
Guadalajara	27	(9-55)	20.39	9	(2-17)	6.88
Guipúzcoa/Gipuzkoa	111	(47-219)	32.08	55	(28-91)	15.10
Huelva	78	(31-157)	30.25	11	(4-22)	4.33
Huesca	30	(10-61)	26.22	22	(9-39)	19.71
Jaén	44	(17-89)	13.46	21	(9-38)	6.41
León	110	(46-216.02)	46.06	20	(8-35)	7.84
Lleida	69	(28-136)	31.55	17	(7-31)	7.91
La Rioja	44	(16-92)	28.09	12	(4-23)	7.48
Lugo	99	(41-193)	59.20	3	(0-8)	1.67
Madrid	489	(216-948)	15.91	212	(118-345)	6.39
Málaga, Ceuta y Melilla	191	(84-372)	21.69	73	(38-122)	8.12
Murcia	149	(64-289)	20.26	59	(30-99)	8.15

Navarra	79	(32-158)	24.81	27	(12-47)	8.35
Ourense	46	(18-93)	29.60	16	(6-29)	9.44
Asturias	214	(93-413.02)	41.99	67	(35-112)	12.08
Palencia	33	(12-66)	39.66	4	(0-10)	4.74
Las Palmas	267	(116-516)	48.45	51	(26-86)	9.23
Pontevedra	124	(52-242)	26.74	58	(30-97)	11.83
Salamanca	47	(18-93)	27.50	41	(20-70)	23.23
Santa Cruz de Tenerife	189	(81-369)	37.82	38	(18-64)	7.37
Cantabria	111	(46-218)	38.50	28	(13-49)	9.20
Segovia	13	(4-29)	16.63	20	(9-36)	25.62
Sevilla	262	(115-509)	27.59	72	(38-120)	7.30
Soria	20	(6-42)	42.32	0	(0-0)	0.00
Tarragona	59	(38-81)	14.59	33	(18-51)	8.32
Teruel	17	(5-36)	23.60	10	(3-19)	13.88
Toledo	76	(31-150)	21.28	17	(7-31)	4.89
Valencia/València	318	(139-619)	25.50	95	(51-157)	7.38
Valladolid	93	(39-184)	35.96	32	(15-54)	11.68
Vizcaya/Bizkaia	234	(101-457)	42.28	81	(43-134)	13.76
Zamora	55	(22-110)	59.00	3	(0-8)	3.56
Zaragoza	85	(35-169)	17.80	19	(8-35)	3.96

Pulmó

Província	Homes			Dones		
	N	IC	TB	N	IC	TB
Álava/Araba	164	(119-216)	103.66	41	(23-65)	25.38
Albacete	210	(154-276)	105.62	35	(19-56)	17.60
Alicante/Alacant	821	(624-1056)	89.07	276	(175-408)	29.72
Almería	258	(192-337)	73.82	62	(37-96)	18.21
Ávila	80	(55-108.02)	94.23	10	(4-18)	12.00
Badajoz	413	(311-536)	120.70	67	(40-104)	19.29
Illes Balears	434	(327.98-562.02)	77.84	142	(88-213)	25.58
Barcelona	2729	(2089-3492.02)	102.40	779	(501-1139)	27.82
Burgos	174	(127-229)	94.14	55	(32-86)	30.19
Cáceres	290	(216-377)	141.96	58	(34-89)	28.35
Cádiz	589	(447-761)	94.87	140	(86-212)	22.35
Castellón/Castelló	253	(187-329)	87.18	49	(28-76)	16.82
Ciudad Real	323	(241-419)	124.84	62	(36-95)	23.70
Córdoba	373	(279-483)	95.02	66	(39-101)	16.24
A Coruña	682	(518-877)	124.81	174	(109-258.02)	29.54
Cuenca	84	(58-115)	78.99	8	(2-15)	7.72
Girona	336	(289-384)	89.74	93	(68-120)	25.02
Granada	388	(292-505)	84.91	75	(44-114)	16.17
Guadalajara	101	(72-136)	77.27	28	(14-46)	22.37
Guipúzcoa/Gipuzkoa	331	(246-431)	95.85	144	(89-216)	39.81
Huelva	229	(170-300)	88.60	34	(19-55)	12.94
Huesca	93	(66-126)	81.94	27	(13-43)	24.46
Jaén	247	(183-322)	75.38	34	(18-54)	10.28
León	311	(233-403)	130.73	75	(44-114)	30.14
Lleida	179	(131-236)	81.39	52	(29-81)	24.32
La Rioja	138	(100-184)	87.69	34	(18-54)	21.32
Lugo	204	(150-267)	122.39	47	(26-73.02)	26.65
Madrid	2459	(1883-3146)	79.98	940	(604-1379)	28.33
Málaga, Ceuta y Melilla	826	(629-1061)	93.61	214	(134-316.02)	23.79
Murcia	609	(463-786)	82.92	118	(72-177)	16.24

Navarra	334	(251-434)	105.53	118	(72-178)	36.81
Ourense	206	(152-270)	131.94	68	(40-103)	40.68
Asturias	734	(558-946)	144.12	244	(153-362)	44.07
Palencia	93	(66-126)	111.86	16	(7-28)	18.88
Las Palmas	431	(325-558.02)	78.35	191	(119-285)	34.80
Pontevedra	485	(366-626)	104.93	167	(103-247)	34.13
Salamanca	213	(158-280)	125.93	37	(20-58)	20.95
Santa Cruz de Tenerife	435	(327-561)	86.93	163	(102-243)	32.01
Cantabria	321	(240-418)	111.49	122	(74.98-183)	40.58
Segovia	71	(48-97)	87.60	9	(3-17)	11.37
Sevilla	904	(687.98-1164)	95.25	197	(123-294)	19.95
Soria	41	(26-58)	86.84	14	(6-24)	30.56
Tarragona	355	(307-405)	88.47	106	(79-136)	26.68
Teruel	81	(56-110)	113.17	14	(6-25)	20.37
Toledo	347	(260-449)	97.79	49	(28-76)	14.21
Valencia/València	1356	(1035-1737)	108.65	364	(232-535)	28.28
Valladolid	325	(243-423)	125.13	98	(59-150)	36.28
Vizcaya/Bizkaia	656	(498-846)	118.77	222	(139-329.02)	37.51
Zamora	98	(69-132)	105.12	19	(9-32)	20.13
Zaragoza	586	(444-757)	122.54	172	(107-256)	34.95

Bufeta Urinària

Província	Homes			Dones		
	N	IC	TB	N	IC	TB
Álava/Araba	123	(82-172)	77.75	26	(12-45)	16.09
Albacete	176	(120-243.02)	88.52	8	(2-16)	4.02
Alicante/Alacant	649	(460-889)	70.41	215	(129-334)	23.15
Almería	184	(127-257)	52.65	31	(15-53)	9.10
Ávila	73	(47-105)	85.98	10	(4-20)	12.00
Badajoz	307	(215-424)	89.72	56	(31-90)	16.13
Illes Balears	331	(232-455)	59.37	151	(88-236.02)	27.20
Barcelona	1690	(1206-2292.02)	63.41	320	(191-497)	11.43
Burgos	141	(96-197)	76.29	38	(19-64)	20.86
Cáceres	218	(150-302)	106.72	37	(19-62)	18.09
Cádiz	530	(374-726)	85.36	62	(34-101)	9.90
Castellón/Castelló	261	(182.98-360)	89.94	35	(17-59)	12.02
Ciudad Real	227	(158-314)	87.73	14	(5-25)	5.35
Córdoba	292	(203-400)	74.39	55	(29-90)	13.53
A Coruña	457	(322-628)	83.63	62	(34-99)	10.53
Cuenca	63	(39-92)	59.24	17	(7-30)	16.40
Girona	341	(285-399)	91.08	58	(39-79)	15.60
Granada	236	(163-326)	51.65	45	(24-74)	9.70
Guadalajara	59	(37-85)	45.14	20	(8-35)	15.98
Guipúzcoa/Gipuzkoa	235	(162-327)	68.05	92	(51-147)	25.43
Huelva	292	(204-403)	112.98	55	(29-89)	20.93
Huesca	64	(41-93)	56.39	7	(2-15.02)	6.34
Jaén	203	(139-282)	61.95	30	(15-51)	9.07
León	224	(155-310)	94.16	47	(24-78)	18.89
Lleida	117	(78-166)	53.20	29	(14-49)	13.56
La Rioja	152	(104-213)	96.58	29	(14-49)	18.19
Lugo	173	(117-243)	103.79	18	(7-33)	10.21
Madrid	2064	(1471-2811.02)	67.13	548	(333-847)	16.51
Málaga, Ceuta y Melilla	503	(355-688)	57.00	163	(96-255)	18.12
Murcia	428	(302-587)	58.28	86	(48-137)	11.83

Navarra	176	(122-244)	55.61	44	(23-72)	13.73
Ourense	216	(149-298)	138.35	35	(18-58)	20.94
Asturias	533	(378-731)	104.65	115	(67-183)	20.77
Palencia	71	(45-104)	85.40	9	(3-17)	10.62
Las Palmas	325	(226-451)	59.08	65	(36-105)	11.84
Pontevedra	398	(279-549)	86.11	36	(18-60)	7.36
Salamanca	182	(125-253)	107.61	22	(10-39)	12.46
Santa Cruz de Tenerife	243	(169-336)	48.56	46	(24-76)	9.03
Cantabria	367	(259-504)	127.47	41	(21-67)	13.64
Segovia	45	(28-67)	55.52	7	(2-15)	8.84
Sevilla	849	(601.98-1159)	89.45	153	(90-240)	15.49
Soria	37	(22-56)	78.37	6	(1-13)	13.10
Tarragona	294	(246-344)	73.27	53	(35-72)	13.34
Teruel	51	(31-75)	71.25	3	(0-8)	4.37
Toledo	173	(118-242)	48.75	37	(18-63)	10.73
Valencia/València	1107	(791-1508)	88.70	250	(148-390)	19.42
Valladolid	256	(178-355)	98.57	57	(30-94)	21.10
Vizcaya/Bizkaia	558	(395-764)	101.03	98	(56-156)	16.56
Zamora	91	(60-129)	97.61	12	(4-22)	12.71
Zaragoza	342	(239-471)	71.52	108	(62-171)	21.95

Pròstata				Mama		
Homes				Dones		
Província	N	IC	TB	N	IC	TB
Álava/Araba	138	(71-237)	87.23	294	(188-435)	181.97
Albacete	365	(195-618)	183.57	268	(172-398)	134.75
Alicante/Alacant	1679	(915-2835)	182.16	1589	(1045-2312)	171.10
Almería	434	(232.98-738)	124.18	400	(257-587)	117.48
Ávila	142	(74-242)	167.25	72	(43-111)	86.40
Badajoz	638	(346-1075)	186.46	582	(379-854)	167.61
Illes Balears	665	(359-1123)	119.27	747	(487.98-1091)	134.55
Barcelona	3979	(2184-6656)	149.30	4081	(2689.97-5927)	145.75
Burgos	334	(179.98-566)	180.71	242	(15-358)	132.84
Cáceres	286	(152.98-485)	140.01	270	(174-400)	131.99
Cádiz	1037	(564-1740)	167.02	1063	(697-1549)	169.74
Castellón/Castelló	412	(216-706)	141.97	323	(208-474)	110.89
Ciudad Real	391	(208-664)	151.12	355	(228-523)	135.73
Córdoba	497	(265-845)	126.61	637	(415.98-933)	156.71
A Coruña	1077	(585-1813)	197.10	1073	(704-1568)	182.17
Cuenca	113	(55-200)	106.26	117	(71-178)	112.90
Girona	521	(398-654)	139.16	535	(457-616)	143.94
Granada	650	(354-1095)	142.24	712	(464-1043)	153.47
Guadalajara	213	(112-364)	162.96	85	(51.98-128)	67.91
Guipúzcoa/Gipuzkoa	455	(244-770)	131.76	444	(288-649)	122.75
Huelva	503	(265-854)	194.61	517	(334.98-759)	196.72
Huesca	207	(105-357)	182.38	168	(105-248)	152.20
Jaén	481	(257-815)	146.80	393	(254-578)	118.79
León	570	(308-957)	239.60	383	(248-561)	153.90
Lleida	411	(221-692)	186.88	303	(193-446)	141.69
La Rioja	228	(117-395)	144.87	194	(122-288)	121.67
Lugo	350	(187-597)	209.98	221	(140-329)	125.30
Madrid	3628	(1979-6089)	118.00	4326	(2844.97-6294)	130.37
Málaga, Ceuta y Melilla	1323	(721-2238)	149.93	1299	(849-1894)	144.38
Murcia	827	(450-1394)	112.61	964	(630-1407)	132.63

Navarra	668	(359-1134)	211.06	332	(214-488)	103.58
Ourense	493	(264-836)	315.76	253	(162-372)	151.37
Asturias	1359	(738-2282)	266.84	957	(625-1398)	172.86
Palencia	190	(99-324)	228.54	152	(95-227)	179.36
Las Palmas	936	(507-1578)	170.15	774	(505-1132)	141.04
Pontevedra	932	(510-1567)	201.65	754	(493-1098)	154.11
Salamanca	362	(193-616)	214.03	299	(194-440)	169.32
Santa Cruz de Tenerife	496	(263-844)	99.12	865	(566-1264)	169.89
Cantabria	530	(283-900)	184.08	476	(309-697)	158.34
Segovia	154	(79-267)	190.00	108	(66-164)	136.44
Sevilla	1486	(803-2508)	156.57	1468	(964-2134)	148.67
Soria	102	(51-178)	216.05	72	(43-110)	157.18
Tarragona	428	(313-551)	106.67	621	(528-715)	156.32
Teruel	145	(74-255)	202.58	75	(45-115)	109.13
Toledo	461	(249-780)	129.91	357	(230-524)	103.52
Valencia/València	2031	(1110-3418)	162.74	2037	(1338-2967)	158.26
Valladolid	565	(305-952)	217.54	422	(273-622)	156.21
Vizcaya/Bizkaia	903	(491-1514)	163.49	994	(647-1452)	167.97
Zamora	203	(107-346)	217.76	89	(53-135)	94.28
Zaragoza	915	(502-1532)	191.33	742	(485-1083)	150.78

ANNEX 6. CODI D'R

1. Estimació de la RIM pel període 1994-2007 a la secció 5.3.

```
# Execució llibreries
library(Rcpp)
library(abind)
library(plyr)
library(boot)

library(reshape)
library(coda)
library(R2WinBUGS)

# Lectura de les bases de dades
setwd("/Volumes/ANDREA_USB/PFG")
incidencia <- read.csv("incidencia.csv", header = TRUE, sep = ";", fileEncoding = "latin1")
mortalidad <- read.csv("mortalidad.csv", header = TRUE, sep = ";", fileEncoding = "latin1")
parametros.provincia <- read.csv("parametros.provincia.csv", header = TRUE, sep = ";", fileEncoding = "latin1")
parametros.tumor <- read.csv("parametros.tumor.csv", header = TRUE, sep=";", fileEncoding = "UTF-16LE")
poblacion <- read.csv("poblacion.csv", header = TRUE, sep = ";", fileEncoding = "latin1")
# A continuació, es selecciona aquell interval d'anys que interessa
incidencia <- incidencia[incidencia$ANY>=1994 & incidencia$ANY<=2007,]
mortalidad <- mortalidad[mortalidad$ANY>=1994,]
poblacion <- poblacion[poblacion$ANY>=1994,]
# Per a la preparació d'un nou càlcul de la variable 101, cal fer alguns canvis en les bdd, de manera que:
mortalidad1 <- subset(mortalidad, mortalidad$PROVINCIA %in% incidencia$PROVINCIA)
mortalidad2 <- mortalidad[(mortalidad$PROVINCIA==7 | mortalidad$PROVINCIA==35 | mortalidad$PROVINCIA==38 | mortalidad$PROVINCIA==100),]
mortalidad <- rbind(mortalidad1, mortalidad2)
# Població
poblacion1 <- subset(poblacion, poblacion$PROVINCIA %in% incidencia$PROVINCIA)
poblacion2 <- poblacion[(poblacion$PROVINCIA==7 | poblacion$PROVINCIA==35 | poblacion$PROVINCIA==38 | poblacion$PROVINCIA==100),]
poblacion <- rbind(poblacion1, poblacion2)
# Ara que ja s'ha seleccionat tot allò que es necessita, es mira el format que tenen les bases de dades resultants i es comprova que tot estigui de forma correcta.
for(i in 5:ncol(incidencia)){
  incidencia[,i] <- as.numeric(gsub(",", ".", gsub("\\.", "", incidencia[,i])))
}

for(i in 5:ncol(mortalidad)){
  mortalidad[,i] <- as.numeric(gsub(",", ".", gsub("\\.", "", mortalidad[,i])))
}

for(j in 4:ncol(poblacion)){
  poblacion[,j] <- as.numeric(gsub(",", ".", gsub("\\.", "", poblacion[,j])))
}
colnames(incidencia)
colnames(mortalidad)
colnames(poblacion)
mortalidad.12.13 <- read.csv("mortalitat_TG_12-13.csv", header = TRUE, sep = ";", fileEncoding = "latin1")
colnames(mortalidad.12.13)[2] <- c("ANY")
m <- rbind(mortalidad,mortalidad.12.13)
mortalidad <- m
# A partir d'aquí s'adjunten les funcions necessàries per a la preparació de dades (i la creació de la província 101)
# Preparació de dades per a anàlisi RIM
preparacion.datos.tumor<-function(icdo)
{
  if (icdo < 200) {
    datos.tumor <- datos.rim(icdo,"M")
  }
}
```

```

if (icdo > 200) {
  datos.tumor <- datos.rim(icdo,"F")
}
datos.tumor <- datos.tumor[(datos.tumor$X != 0 | datos.tumor$Y != 0),]
attach(datos.tumor)
datos.tumor$Y <- correccion.rim.zero()
datos.tumor$EDAD <- correccion.rim.edad(icdo)
datos.tumor$ANY <- as.numeric(datos.tumor$ANY) - 1994
datos.tumor$ANY2 <- as.numeric(datos.tumor$ANY) * as.numeric(datos.tumor$ANY)
detach()
q1 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"P10"]
q2 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"T2"]
q3 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"T3"]
q4 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"P90"]
attach(datos.tumor)
datos.tumor$EDAD1 <- pmax(0,EDAD-q1)
datos.tumor$EDAD2 <- pmax(0,EDAD-q2)
datos.tumor$EDAD3 <- pmax(0,EDAD-q3)
datos.tumor$EDAD4 <- pmax(0,EDAD-q4)
detach()
rm(q1,q2,q3,q4)
datos.tumor
}

datos.rim <- function(icdo,sexo)
{
  corteedad <- parametros.tumor[parametros.tumor$CODIGO == icdo,"CORTEEDAD"]
  incidencia.rim <- salida.datos("I",icdo,sexo)
  mortalidad.rim <- salida.datos("M",icdo,sexo)
  if (icdo == 653 | icdo == 654 | icdo == 699)
  {
    mortalidad.655 <- salida.datos("M",655,"F")
    mortalidad.etq <- mortalidad.655[,1:4]
    mortalidad.etq$ICDO <- rep(icdo,dim(mortalidad.etq)[1])
    mortalidad.655 <- mortalidad.655[,5:22]
    mortalidad.655 <- mortalidad.655/3
    mortalidad.655 <- cbind(mortalidad.etq,mortalidad.655)
    mortalidad.rim <- rbind(mortalidad.rim,mortalidad.655)
    attach(mortalidad.rim)
    mortalidad.rim <- aggregate(mortalidad.rim[,5:22], list(PROVINCIA=PROVINCIA, ANY=ANY, SEXO=SEXO,
ICDO=ICDO), sum)
    detach()
  }
  datos.rim <- merge(incidencia.rim, mortalidad.rim,by=c("PROVINCIA","ANY","SEXO","ICDO"))
  datos.rim
  datos.rim$TOT <- rowSums(datos.rim[,5:23])
  datos.rim$p <- (datos.rim$TOT-datos.rim$SE)/datos.rim$TOT
  datos.rim[,24:41] <- datos.rim[,24:41]*datos.rim$p
  kf1 <- 5 + corteedad - 1
  ki1 <- 5 + corteedad
  kf2 <- 24 + corteedad - 1
  ki2 <- 24 + corteedad
  datos.rim$I00 <- rowSums(datos.rim[5:kf1])
  datos.rim$M00 <- rowSums(datos.rim[24:ki2])
  aux1 <- names(datos.rim[1:2])
  aux2 <- names(datos.rim[ki1:22])
  aux3 <- names(datos.rim[ki2:41])
  aux <- cbind(t(aux1),"I00",t(aux2),"M00",t(aux3))
  datos.rim <- datos.rim[,aux]
  datos.rim.t <- melt(datos.rim, id=c("PROVINCIA","ANY"))
  n <- dim(datos.rim.t)[1]
  kf <- n/2
  ki <- kf+1
  datos.rim.t1 <- datos.rim.t[1:kf,]
  datos.rim.t1$EDAD <- as.numeric(substr(as.character(datos.rim.t1$variable),2,3))
  datos.rim.t2 <- datos.rim.t[ki:n,]
  datos.rim.t2$EDAD <- as.numeric(substr(as.character(datos.rim.t2$variable),2,3))
  datos.rim <- merge(datos.rim.t1, datos.rim.t2,by=c("PROVINCIA","ANY","EDAD"))
  datos.rim$X <- datos.rim$value.x
  datos.rim$Y <- datos.rim$value.y
  aux <- c("PROVINCIA","ANY","EDAD","X","Y")
  datos.rim <- datos.rim[,aux]
  datos.rim
}

```

```

correcion.rim.edad <- function(icdo)
{
  corteedad <- parametros.tumor[parametros.tumor$CODIGO == icdo,"CORTEEDAD"]
  prov <- (EDAD*5) - 2.5
  prov[EDAD==0] <- (corteedad*5)-2.5
  prov[EDAD==18] <- 90
  prov
}

correcion.rim.zero <- function()
{
  prov <- Y
  prov[Y==0] <- exp(-3)
  prov
}

salida.datos <- function(tipo,icdo,sexo)
{
  salida <- salida.datos.registro(tipo,1, icdo,sexo)
  salida <- rbind(salida,salida.datos.registro(tipo,2, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,12, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,13, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,16, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,17, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,18, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,20, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,26, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,30, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,31, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,33, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,43, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,48, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,50, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,54, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,55, icdo,sexo))
  salida <- rbind(salida,salida.datos.registro(tipo,56, icdo,sexo))
  if (icdo != 655) {
    modelo <- parametros.tumor[parametros.tumor$CODIGO == icdo,"TIPO"]
    if(modelo == "SP") {
      attach(salida)
      if (tipo == "I") {
        salida.agr <- aggregate(salida[,5:23], list(ANY=ANY, SEXO=SEXO, ICDO=ICDO), sum)
      }
      if (tipo == "M") {
        salida.agr <- aggregate(salida[,5:22], list(ANY=ANY, SEXO=SEXO, ICDO=ICDO), sum)
      }
      detach()
      n <- dim(salida.agr)[1]
      PROVINCIA <- rep(101,n)
      salida <- cbind(PROVINCIA,salida.agr)
    }
  }
  salida
}

# Funcio de sortida de dades d'incidència, mortalitat, població.
salida.datos.registro <- function(tipo, provincia, icdo,sexo)
{
  ANYi <- 1994
  ANYf <- 2007
  if (tipo=="I") {
    salida <- incidencia[incidencia$PROVINCIA == provincia & incidencia$SEXO == sexo & incidencia$ICDO ==
icdo & incidencia$ANY >=ANYi & incidencia$ANY <=ANYf,]
  }
  if (tipo=="M") {
    if (provincia <=50) {
      salida <- mortalidad[mortalidad$PROVINCIA == provincia & mortalidad$SEXO == sexo & mortalidad$ICDO
== icdo & mortalidad$ANY >=ANYi & mortalidad$ANY <=ANYf,]
    }
    if (provincia == 54) {
      salida <- correcion.mortalidad.provincia.tumor(provincia, 7, icdo, sexo, ANYi, ANYf)
    }
    if (provincia == 55) {

```

```

    salida <- correcion.mortalidad.provincia.tumor(provincia, 35, icdo, sexo, ANYi, ANYf)
  }
  if (provincia == 56) {
    salida <- correcion.mortalidad.provincia.tumor(provincia, 38, icdo, sexo, ANYi, ANYf)
  }
}
if (tipo=="P") {
  salida <- poblacion[poblacion$PROVINCIA == provincia & poblacion$SEXO == sexo & poblacion$ANY >=ANYi
& poblacion$ANY <=ANYf,]
}
salida
}

```

Funcions correccio mortalitat per illes

```

correcion.mortalidad.provincia.tumor <- function(isla, provincia, icdo, sexo, ANYi=1994, ANYf=2007)
{
  mortalidad.provincia <- mortalidad[mortalidad$PROVINCIA == provincia & mortalidad$SEXO == sexo &
mortalidad$ICDO == icdo & mortalidad$ANY >=ANYi & mortalidad$ANY <=ANYf,5:22]
  lista.ANY <- mortalidad[mortalidad$PROVINCIA == provincia & mortalidad$SEXO == sexo & mortalidad$ICDO
== icdo & mortalidad$ANY >=ANYi & mortalidad$ANY <=ANYf,2]
  poblacion.isla <- poblacion[poblacion$PROVINCIA == isla & poblacion$SEXO == sexo & poblacion$ANY >=ANYi
& poblacion$ANY <=ANYf,]
  poblacion.provincia <- poblacion[poblacion$PROVINCIA == provincia & poblacion$SEXO == sexo &
poblacion$ANY >=ANYi & poblacion$ANY <=ANYf,]
  i <- dim(mortalidad.provincia)[1]
  PROVINCIA <- c(rep(isla, i))
  ANY <- lista.ANY
  SEXO <- c(rep(sexo,i))
  ICDO <- c(rep(icdo,i))
  for (ANY in ANYi:ANYf) {
    if (any(lista.ANY==ANY)==FALSE) {
      poblacion.isla <- poblacion.isla[poblacion.isla$ANY != ANY,]
      poblacion.provincia <- poblacion.provincia[poblacion.provincia$ANY != ANY,]
    }
  }
  poblacion.isla <- poblacion.isla[,4:21]
  poblacion.provincia <- poblacion.provincia[,4:21]
  peso <- poblacion.isla/poblacion.provincia
  mortalidad.isla <- mortalidad.provincia*peso
  mortalidad.isla <- cbind(PROVINCIA, ANY, SEXO, ICDO, mortalidad.isla)
  mortalidad.isla
}

```

Codi creacio variable 101

```

correcion.mortalidad.provincia <- function(isla, provincia, ANYi=1994 ,ANYf=2007)
{
  salida <- correcion.mortalidad.provincia.tumor(isla, provincia, 114,"M",ANYi ,ANYf)
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 115,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 116,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 118,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 119,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 122,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 123,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 125,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 132,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 134,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 143,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 161,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 162,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 164,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 167,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 173,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 171,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 181,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 182,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 190,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 191,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 199,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 100,"M",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 614,"F",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 615,"F",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 616,"F",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 618,"F",ANYi ,ANYf))
  salida <- rbind(salida,correcion.mortalidad.provincia.tumor(isla, provincia, 619,"F",ANYi ,ANYf))
}

```



```

salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 622,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 623,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 625,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 632,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 634,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 643,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 650,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 653,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 654,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 656,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 664,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 667,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 673,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 671,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 681,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 682,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 690,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 691,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 699,"F",ANYi ,ANYf))
salida <- rbind(salida,correccion.mortalidad.provincia.tumor(isla, provincia, 600,"F",ANYi ,ANYf))
salida
}

```

```

# Funció quartils

```

```

cuartiles <- function(icdo)
{
  datos <- mortalidad[mortalidad$PROVINCIA==101 & mortalidad$ICDO==icdo,]
  datos <- aggregate(datos[,4:22],list(ICDO=datos$ICDO),sum)
  datos <- datos[,3:20]
  if (icdo == 653 | icdo == 654 | icdo == 699 ) {
    datos2 <- mortalidad[mortalidad$PROVINCIA==101 & mortalidad$ICDO==655,]
    datos2 <- aggregate(datos2[,4:22],list(ICDO=datos2$ICDO),sum)
    datos2 <- datos2[,3:20]
    datos2 <- datos2/3
    datos <- datos + datos2
  }
  suma <- sum(datos)
  frecuencia <- (datos/suma)*100
  delta1 <- 100
  delta2 <- 100
  delta3 <- 100
  delta4 <- 100
  for (i in 1:18) {
    suma.sup <- sum(frecuencia[,1:i])
    if (abs(10-suma.sup) < delta1) {
      delta1 <- abs(10-suma.sup)
      cortedad <- i
      p10 <- i*5
    }
    if (abs(33.3333-suma.sup) < delta2) {
      delta2 <- abs(33.3333-suma.sup)
      t2 <- i*5
    }
    if (abs(66.6666-suma.sup) < delta3) {
      delta3 <- abs(66.6666-suma.sup)
      t3 <- i*5
    }
    if (abs(90-suma.sup) < delta4) {
      delta4 <- abs(90-suma.sup)
      p90 <- i*5
    }
  }
  aux <- c(icdo, cortedad,p10,t2,t3,p90)
  aux
}

```

```

# Codi per simular

```

```

simulacion.icdo <- function(icdo, tipo = "P2", cadenas=3, iteraciones=50000, calentamiento=10000,
intervalo=5)
{
  setwd("C:/Users/user/Desktop/PFG/Resultats")
  prep.openbugs(icdo)
  nombre.tumor <- paste0("tumor-",icdo, ".txt")
  datos <- read.csv2(nombre.tumor, header=TRUE)
}

```

```

datos.sim <- simulacion(datos, tipo, cadenas, iteraciones, calentamiento, intervalo)
result.sim <- datos.sim$BUGSoutput$sims.matrix
nombre.tumor2 <- paste0("simulacion-", icdo, ".txt")
write.csv2(result.sim, nombre.tumor2, row.names=FALSE)
print(datos.sim, dig=6)
}
prep.openbugs <- function(icdo)
{
nombre.tumor <- paste0("tumor-", icdo, ".txt")
datos.tumor <- preparacion.datos.tumor(icdo)
write.csv2(datos.tumor, file=nombre.tumor, row.names=FALSE)
}

simulacion <- function(datos, tipo = "P2", cadenas=3, iteraciones=50000, calentamiento=10000,
intervalo=5)
{
if (tipo == "SE") {
aux <- c("PROVINCIA", "ANY", "ANY2", "X", "Y")
datos <- datos[,aux]
datos <- aggregate(datos, list(PROVINCIA=datos$PROVINCIA, ANY=datos$ANY, ANY2=datos$ANY2), sum)
datos <- datos[,aux]
rm(aux)
print(datos)
}
J <- nrow(datos)
X <- datos$X
Y <- datos$Y
if (tipo == "P2" | tipo == "P1" | tipo == "SE" | tipo == "SP") {
any <- datos$ANY
}
if (tipo == "P2" | tipo == "SE" | tipo == "SP") {
any2 <- datos$ANY2
}
if (tipo != "SE") {
edad1 <- datos$EDAD1
edad2 <- datos$EDAD2
edad3 <- datos$EDAD3
edad4 <- datos$EDAD4
}
ID <- datos$PROVINCIA
ID <- factor(ID)
if (tipo == "P2") {
data <- list(J=J, X=X, Y=Y, ID=ID, any=any, any2=any2, edad1=edad1, edad2=edad2, edad3=edad3,
edad4=edad4)
parameters <- c("alpha", "a", "ANY", "ANY2", "EDAD1", "EDAD2", "EDAD3", "EDAD4", "tau", "sigma")
inits <- function() {list(alpha=rnorm(15), a=rnorm(1), ANY=rnorm(1), ANY2=rnorm(1,0,0.1),
EDAD1=rnorm(1), EDAD2=rnorm(1), EDAD3=rnorm(1), EDAD4=rnorm(1), tau=1)}
datos.sim <- jags(data, inits, parameters, "model-02.bug", n.chains=cadenas, n.iter=iteraciones,
n.burnin=calentamiento, n.thin=intervalo)
rm(datos, J, X, Y, any, any2, edad1, edad2, edad3, edad4, ID, data, parameters, inits)
}
if (tipo == "P1") {
data <- list(J=J, X=X, Y=Y, ID=ID, any=any, edad1=edad1, edad2=edad2, edad3=edad3, edad4=edad4)
parameters <- c("alpha", "a", "ANY", "EDAD1", "EDAD2", "EDAD3", "EDAD4", "tau", "sigma")
inits <- function() {list(alpha=rnorm(15), a=rnorm(1), ANY=rnorm(1), EDAD1=rnorm(1),
EDAD2=rnorm(1), EDAD3=rnorm(1), EDAD4=rnorm(1), tau=1)}
datos.sim <- jags(data, inits, parameters, "model-01.bug", n.chains=cadenas, n.iter=iteraciones,
n.burnin=calentamiento, n.thin=intervalo)
rm(datos, J, X, Y, any, edad1, edad2, edad3, edad4, ID, data, parameters, inits)
}
if (tipo == "P0") {
data <- list(J=J, X=X, Y=Y, ID=ID, edad1=edad1, edad2=edad2, edad3=edad3, edad4=edad4)
parameters <- c("alpha", "a", "EDAD1", "EDAD2", "EDAD3", "EDAD4", "tau", "sigma")
inits <- function() {list(alpha=rnorm(15), a=rnorm(1), EDAD1=rnorm(1), EDAD2=rnorm(1),
EDAD3=rnorm(1), EDAD4=rnorm(1), tau=1)}
datos.sim <- jags(data, inits, parameters, "model-00.bug", n.chains=cadenas, n.iter=iteraciones,
n.burnin=calentamiento, n.thin=intervalo)
rm(datos, J, X, Y, edad1, edad2, edad3, edad4, ID, data, parameters, inits)
}
if (tipo == "SE") {
data <- list(J=J, X=X, Y=Y, ID=ID, any=any, any2=any2)
parameters <- c("alpha", "a", "ANY", "ANY2", "tau", "sigma")
inits <- function() {list(alpha=rnorm(15), a=rnorm(1), ANY=rnorm(1), ANY2=rnorm(1,0,0.1), tau=1)}
}
}

```

```

    datos.sim <- jags(data, inits, parameters, "model-se.bug", program = "OpenBUGS", n.chains=cadenas,
n.iter=iteraciones, n.burnin=calentamiento, n.thin=intervalo)
    rm(datos,J,X,Y,any,any2,ID,data,parameters,inits)
  }
  if (tipo == "SP") {
    data <- list(J=J, X=X, Y=Y, any=any, any2=any2, edad1=edad1, edad2=edad2, edad3=edad3,
edad4=edad4)
    parameters <- c("a", "ANY", "ANY2", "EDAD1", "EDAD2", "EDAD3", "EDAD4")
    inits <- function() {list (a=rnorm(1), ANY=rnorm(1), ANY2=rnorm(1,0,0.1), EDAD1=rnorm(1),
EDAD2=rnorm(1), EDAD3=rnorm(1), EDAD4=rnorm(1))}
    datos.sim <- jags(data, inits, parameters, "model-sp.bug", program = "OpenBUGS", n.chains=cadenas,
n.iter=iteraciones, n.burnin=calentamiento, n.thin=intervalo)
    rm(datos,J,X,Y,any,any2,edad1,edad2,edad3,edad4,data,parameters,inits)
  }
  datos.sim
}

```

Creació província 101

```

incidencia.101 <- incidencia[incidencia$PROVINCIA != 100 & incidencia$ANY >=1994 & incidencia$ANY
<=2007,]
attach(incidencia.101)
incidencia.global <- aggregate(incidencia.101[,5:23], list(ANY=ANY, SEX0=SEX0, ICDO=ICDO), sum)
detach()
n <- dim(incidencia.global)[1]
PROVINCIA <- rep(101,n)
incidencia.global <- cbind(PROVINCIA,incidencia.global)
incidencia <- rbind(incidencia,incidencia.global)
rm(n, PROVINCIA, incidencia.global,incidencia.101)

mortalidad.101 <-mortalidad[mortalidad$PROVINCIA != 100 & mortalidad$ANY >=1994 & mortalidad$ANY <=2007,]
mortalidad.101 <-mortalidad[mortalidad$PROVINCIA != 7 & mortalidad$PROVINCIA != 35 & mortalidad$PROVINCIA
!= 38,]
mortalidad.101 <- rbind(mortalidad.101,correccion.mortalidad.provincia(54,7,1994,2007))
mortalidad.101 <- rbind(mortalidad.101,correccion.mortalidad.provincia(55,35,1994,2007))
mortalidad.101 <- rbind(mortalidad.101,correccion.mortalidad.provincia(56,38,1994,2007))
attach(mortalidad.101)
mortalidad.global <- aggregate(mortalidad.101[,5:22], list(ANY=ANY, SEX0=SEX0, ICDO=ICDO), sum)
detach()
n <- dim(mortalidad.global)[1]
PROVINCIA <- rep(101,n)
mortalidad.global <- cbind(PROVINCIA,mortalidad.global)
mortalidad <- rbind(mortalidad,mortalidad.global)
rm(n, PROVINCIA, mortalidad.global,mortalidad.101)

poblacion.101 <-poblacion[poblacion$PROVINCIA != 100 & poblacion$ANY >=1994 & poblacion$ANY <=2007,]
poblacion.101 <-poblacion[poblacion$PROVINCIA != 7 & poblacion$PROVINCIA != 35 & poblacion$PROVINCIA !=
38,]
attach(poblacion.101)
poblacion.global <- aggregate(poblacion.101[,4:21], list(ANY=ANY, SEX0=SEX0), sum)
detach()
n <- dim(poblacion.global)[1]
PROVINCIA <- rep(101,n)
poblacion.global <- cbind(PROVINCIA,poblacion.global)
poblacion <- rbind(poblacion,poblacion.global)
rm(n, PROVINCIA, poblacion.global,poblacion.101)

```

2. Estimació incidència Tarragona i Girona per al període 2009-2013 a la secció 6.1.

```
coe.u <- funcion(provincia,icdo,escenario,any)
{
  modelo <- parametros.tumor[parametros.tumor$CODIGO == icdo,"TIPO"]
  q1 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"P10"]
  q2 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"T2"]
  q3 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"T3"]
  q4 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"P90"]
  setwd("/Users/mac/Desktop/PFG/Resultats")
  #setwd("/Volumes/ANDREA_USB")
  nombre.tumor <- paste0("simulacion-",icdo, ".txt")
  result.sim <- read.csv2(nombre.tumor, header=TRUE)
  if(modelo=="P2" | modelo=="SE") {
    if(provincia==17) { coe.a <- result.sim$alpha.4. }
    if(provincia==43) { coe.a <- result.sim$alpha.11. }
    if(provincia==100) { coe.a <- result.sim$a }
  }
  if(modelo=="SP") {
    coe.a <- result.sim$a
  }
  ANY <- result.sim$ANY
  ANY2 <- result.sim$ANY2
  if(modelo=="P2" | modelo=="SP")
  {
    EDAD1 <- result.sim$EDAD1
    EDAD2 <- result.sim$EDAD2
    EDAD3 <- result.sim$EDAD3
    EDAD4 <- result.sim$EDAD4
  }
  if(modelo=="SE")
  {
    EDAD1 <- rep(0,24000)
    EDAD2 <- rep(0,24000)
    EDAD3 <- rep(0,24000)
    EDAD4 <- rep(0,24000)
  }
  if (escenario=="A")
  {
    coe.any <- (2007-1994)*ANY + ((2007-1994)*(2007-1994))*ANY2
  }
  if (escenario=="B")
  {
    coe.any <- (2007-1994)*ANY + ((2007-1994)*(2007-1994))*ANY2
    coe.any <- coe.any*((any-1994)/(2007-1994))
  }
  if (escenario=="C")
  {
    coe.any <- (any-1994)*ANY + ((any-1994)*(any-1994))*ANY2
  }
  coe.edad1 <- EDAD1*pmax(0,2.5-q1)
  coe.edad2 <- EDAD2*pmax(0,2.5-q2)
  coe.edad3 <- EDAD3*pmax(0,2.5-q3)
  coe.edad4 <- EDAD4*pmax(0,2.5-q4)
  coe.edad.01 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
  coe.01 <- coe.a + coe.any + coe.edad.01
  coe.edad1 <- EDAD1*pmax(0,7.5-q1)
  coe.edad2 <- EDAD2*pmax(0,7.5-q2)
  coe.edad3 <- EDAD3*pmax(0,7.5-q3)
  coe.edad4 <- EDAD4*pmax(0,7.5-q4)
  coe.edad.02 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
  coe.02 <- coe.a + coe.any + coe.edad.02
  coe.edad1 <- EDAD1*pmax(0,12.5-q1)
  coe.edad2 <- EDAD2*pmax(0,12.5-q2)
  coe.edad3 <- EDAD3*pmax(0,12.5-q3)
  coe.edad4 <- EDAD4*pmax(0,12.5-q4)
  coe.edad.03 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
  coe.03 <- coe.a + coe.any + coe.edad.03
  coe.edad1 <- EDAD1*pmax(0,17.5-q1)
  coe.edad2 <- EDAD2*pmax(0,17.5-q2)
  coe.edad3 <- EDAD3*pmax(0,17.5-q3)
  coe.edad4 <- EDAD4*pmax(0,17.5-q4)
  coe.edad.04 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
}
```



```

coe.edad2 <- EDAD2*pmax(0,82.5-q2)
coe.edad3 <- EDAD3*pmax(0,82.5-q3)
coe.edad4 <- EDAD4*pmax(0,82.5-q4)
coe.edad.17 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.17 <- coe.a + coe.any + coe.edad.17
coe.edad1 <- EDAD1*pmax(0,90-q1)
coe.edad2 <- EDAD2*pmax(0,90-q2)
coe.edad3 <- EDAD3*pmax(0,90-q3)
coe.edad4 <- EDAD4*pmax(0,90-q4)
coe.edad.18 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.18 <- coe.a + coe.any + coe.edad.18
coe.u <-
cbind(coe.01,coe.02,coe.03,coe.04,coe.05,coe.06,coe.07,coe.08,coe.09,coe.10,coe.11,coe.12,coe.13,coe.14,c
oe.15,coe.16,coe.17,coe.18)
  coe.u
}

coe.sigma <- function(icdo)
{
  setwd("/Users/mac/Desktop/PFG/Resultats")
  nombre.tumor <- paste0("simulacion-",icdo,".txt")
  result.sim <- read.csv2(nombre.tumor, header=TRUE)
  modelo <- parametros.tumor[parametros.tumor$CODIGO == icdo,"TIPO"]
  coe.sigma <- result.sim$sigma
  coe.sigma
}

proyeccion.N <- function(provincia,icdo,escenario,any,n.sim=1)
{
  set.seed(112358)
  u <- coe.u(provincia,icdo,escenario,any)
  sigma <- coe.sigma(icdo)
  muertos <- mortalidad[mortalidad$PROVINCIA==provincia & mortalidad$ICDO==icdo &
mortalidad$ANY==any,5:22]
  if(dim(muertos)[1]==0){
    nombres <- colnames(muertos)
    muertos <- as.data.frame(matrix(0, ncol = 18, nrow = 1))
    colnames(muertos) <- nombres
  }
  for (i in 1:n.sim) {
    RIM <- exp(u)
    ESP.01<-RIM[,1]*muertos[,1]
    ESP.02<-RIM[,2]*muertos[,2]
    ESP.03<-RIM[,3]*muertos[,3]
    ESP.04<-RIM[,4]*muertos[,4]
    ESP.05<-RIM[,5]*muertos[,5]
    ESP.06<-RIM[,6]*muertos[,6]
    ESP.07<-RIM[,7]*muertos[,7]
    ESP.08<-RIM[,8]*muertos[,8]
    ESP.09<-RIM[,9]*muertos[,9]
    ESP.10<-RIM[,10]*muertos[,10]
    ESP.11<-RIM[,11]*muertos[,11]
    ESP.12<-RIM[,12]*muertos[,12]
    ESP.13<-RIM[,13]*muertos[,13]
    ESP.14<-RIM[,14]*muertos[,14]
    ESP.15<-RIM[,15]*muertos[,15]
    ESP.16<-RIM[,16]*muertos[,16]
    ESP.17<-RIM[,17]*muertos[,17]
    ESP.18<-RIM[,18]*muertos[,18]
    ITER <- rep(i,24000)
    SIM <- rep(1:24000)
    PROVINCIA <- rep(provincia,24000)
    ICDO <- rep(icdo,24000)
    ANY <- rep(any,24000)
    ESCENARIO <- rep(escenario,24000)
    N.01 <- rpois(24000,ESP.01)
    N.02 <- rpois(24000,ESP.02)
    N.03 <- rpois(24000,ESP.03)
    N.04 <- rpois(24000,ESP.04)
    N.05 <- rpois(24000,ESP.05)
    N.06 <- rpois(24000,ESP.06)
    N.07 <- rpois(24000,ESP.07)
    N.08 <- rpois(24000,ESP.08)
    N.09 <- rpois(24000,ESP.09)
  }
}

```

```

N.10 <- rpois(24000,ESP.10)
N.11 <- rpois(24000,ESP.11)
N.12 <- rpois(24000,ESP.12)
N.13 <- rpois(24000,ESP.13)
N.14 <- rpois(24000,ESP.14)
N.15 <- rpois(24000,ESP.15)
N.16 <- rpois(24000,ESP.16)
N.17 <- rpois(24000,ESP.17)
N.18 <- rpois(24000,ESP.18)
if (i==1) {CASOS <-
cbind(ITER,SIM,PROVINCIA,ICDO,ANY,ESCENARIO,N.01,N.02,N.03,N.04,N.05,N.06,N.07,N.08,N.09,N.10,N.11,N.12,N
.13,N.14,N.15,N.16,N.17,N.18)}
if (i>1) {
aux <-
cbind(ITER,SIM,PROVINCIA,ICDO,ANY,ESCENARIO,N.01,N.02,N.03,N.04,N.05,N.06,N.07,N.08,N.09,N.10,N.11,N.12,N
.13,N.14,N.15,N.16,N.17,N.18)
CASOS <- rbind(CASOS,aux)
}
}
CASOS
}
proyeccion <- function(icdo,n.sim=1)
{
resultat <- proyeccion.N(17,icdo,"A",2009,n.sim)
resultat <- rbind(resultat,proyeccion.N(17,icdo,"A",2010,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"A",2011,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"A",2012,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"A",2013,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"A",2009,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"A",2010,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"A",2011,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"A",2012,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"A",2013,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"B",2009,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"B",2010,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"B",2011,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"B",2012,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"B",2013,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"B",2009,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"B",2010,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"B",2011,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"B",2012,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"B",2013,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"C",2009,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"C",2010,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"C",2011,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"C",2012,n.sim))
resultat <- rbind(resultat,proyeccion.N(17,icdo,"C",2013,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"C",2009,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"C",2010,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"C",2011,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"C",2012,n.sim))
resultat <- rbind(resultat,proyeccion.N(43,icdo,"C",2013,n.sim))
setwd("/Users/mac/Desktop/PFG")
nombre.archivo <- paste0("Resultat-tumor-",icdo,".txt")
write.csv2(resultat,file=nombre.archivo,row.names=FALSE)
}

icdo <- parametros.tumor$CODI
icdo <- icdo[-c(6,24,30,length(icdo))]
for(i in 1:length(icdo)){
codi <- icdo[i]
proyeccion(codi)
}

```


3. Procés de selecció del millor escenari a la secció 6.2

```

incidencia_TG <- read.csv2("incidencia_TG_08-13.csv")
resultats.validacio <- llista.validacio()
llista.descripcio.validacio()

llista.validacio <- function()
{
n <- nrow(parametros.tumor)
for (i in 1:n)
  {
    if (parametros.tumor$TIPO[i] != 0)
      {
        icdo <- parametros.tumor$CODIGO[i]
        if (icdo == 114 ) { salida <- validacio.escenari(icdo) }
        if (icdo != 114 ) { salida <- rbind(salida, validacio.escenari(icdo)) }
      }
  }
salida
}
descripcio.validacio <- function(icdo)
{
dades <- resultats.validacio[resultats.validacio$ICDO==icdo,]
tumor <- as.character(parametros.tumor[parametros.tumor$CODIGO==icdo, "VALOR"])
if (icdo<200) {sexe <- "Homes"}
if (icdo>200) {sexe <- "Dones"}
escenari <- summary(as.factor(dades$ESCENARIO))
p.escenari <- round(summary(as.factor(dades$ESCENARIO))/24000,4)*100
obs <- mean(dades$DIF_TOT_OBS)
esp.a <- as.numeric(quantile(dades$DIF_TOT_ESP_A,0.5))
esp.a_l <- as.numeric(quantile(dades$DIF_TOT_ESP_A,0.025))
esp.a_u <- as.numeric(quantile(dades$DIF_TOT_ESP_A,0.975))
esp.b <- as.numeric(quantile(dades$DIF_TOT_ESP_B,0.5))
esp.b_l <- as.numeric(quantile(dades$DIF_TOT_ESP_B,0.025))
esp.b_u <- as.numeric(quantile(dades$DIF_TOT_ESP_B,0.975))
esp.c <- as.numeric(quantile(dades$DIF_TOT_ESP_C,0.5))
esp.c_l <- as.numeric(quantile(dades$DIF_TOT_ESP_C,0.025))
esp.c_u <- as.numeric(quantile(dades$DIF_TOT_ESP_C,0.975))
dif.a <- as.numeric(quantile(dades$DIF_TOT_A,0.5))
dif.a_l <- as.numeric(quantile(dades$DIF_TOT_A,0.025))
dif.a_u <- as.numeric(quantile(dades$DIF_TOT_A,0.975))
dif.b <- as.numeric(quantile(dades$DIF_TOT_B,0.5))
dif.b_l <- as.numeric(quantile(dades$DIF_TOT_B,0.025))
dif.b_u <- as.numeric(quantile(dades$DIF_TOT_B,0.975))
dif.c <- as.numeric(quantile(dades$DIF_TOT_C,0.5))
dif.c_l <- as.numeric(quantile(dades$DIF_TOT_C,0.025))
dif.c_u <- as.numeric(quantile(dades$DIF_TOT_C,0.975))
abs.a <- mean(dades$ABS_A)
abs.b <- mean(dades$ABS_B)
abs.c <- mean(dades$ABS_C)
cat("\n")
cat("-----\n")
cat("\n")
cat("TIPUS TUMORAL:  ", tumor, " - ", sexe)
cat("\n")
cat("-----\n")
cat("\n")
if (which.max(escenari)==1)
{
cat("El escenari triat es el A \n")
cat("El nombre de casos observats en el peri?de 2009-2013 es:          ", round(obs,0), "\n")
cat("El nombre de casos esperats en el peri?de 2009-2013 es:          ", round(esp.a,2), " (",
", round(esp.a_l,2), " - ", round(esp.a_u,2), " )", "\n")
cat("El error en la estimaci? en el peri?de 2009-2013 a l'escenari A es: ", round(dif.a,2), "%", " (",
", round(dif.a_l,2), "% - ", round(dif.a_u,2), "% )", "\n")
}
if (which.max(escenari)==2)
{
cat("El escenari triat es el B \n")
cat("El nombre de casos observats en el peri?de 2009-2013 es:          ", round(obs,0), "\n")
cat("El nombre de casos esperats en el peri?de 2009-2013 es:          ", round(esp.b,2), " (",
", round(esp.b_l,2), " - ", round(esp.b_u,2), " )", "\n")
cat("El error en la estimaci? en el peri?de 2009-2013 a l'escenari B es: ", round(dif.b,2), "%", " (",

```



```

",round(dif.b_l,2),"% - ",round(dif.b_u,2),"% )","\n")
}
if (which.max(escenari)==3)
{
cat("El escenari triat es el C \n")
cat("El nombre de casos observats en el peri?de 2009-2013 es: ",round(obs,0),"\n")
cat("El nombre de casos esperats en el peri?de 2009-2013 es: ",round(esp.c,2)," (
",round(esp.c_l,2)," - ",round(esp.c_u,2)," )","\n")
cat("El error en la estimaci? en el peri?de 2009-2013 a l'escenari C es: ",round(dif.c,2),"%"," (
",round(dif.c_l,2),"% - ",round(dif.c_u,2),"% )","\n")
}
cat("-----\n")
cat("Nombre de simulacions favorables","\n")
cat(" A B C","\n")
cat(escenari,"\n")
cat("\n")
cat("Percentatge","\n")
cat(" A B C","\n")
cat(p.escenari,"\n")
cat("\n")
cat("-----\n")
cat("El error en la estimaci? en el peri?de 2009-2013 a l'escenari A es: ",round(dif.a,2),"%","\n")
cat("El error en la estimaci? en el peri?de 2009-2013 a l'escenari B es: ",round(dif.b,2),"%","\n")
cat("El error en la estimaci? en el peri?de 2009-2013 a l'escenari C es: ",round(dif.c,2),"%","\n")
cat("\n")
cat("-----\n")
cat("El valor mitja de desviaci? en l'escenari A es:
",round(abs.a,2),"%","\n")
cat("El valor mitja de desviaci? en l'escenari B es:
",round(abs.b,2),"%","\n")
cat("El valor mitja de desviaci? en l'escenari C es:
",round(abs.c,2),"%","\n")
cat("\n")
cat("\n")
}

llista.descripcio.validacio <- function()
{
n <- nrow(parametros.tumor)
for (i in 1:n)
{
if (parametros.tumor$TIPO[i] != 0)
{
icdo <- parametros.tumor$CODIGO[i]
descripcio.validacio(icdo)
}
}
}

validacio.escenari <- function(icdo)
{
corteedad <- parametros.tumor[parametros.tumor$CODIGO == icdo,"CORTEEDAD"]
setwd("/Volumes/ANDREA_USB/Resultats")
nombre.tumor <- paste0("Resultat-tumor-",icdo,".txt")
dades <- read.csv2(nombre.tumor, header=TRUE)

casos <- incidencia_TG[incidencia_TG$ICDO==icdo & incidencia_TG$ANY>=2009 & incidencia_TG$ANY<=2013,]
DIF_TOT_OBS <- sum(casos[,5:23])
result.N <- dades[dades$ICDO==icdo & dades$ESCENARIO=="A" & dades$ANY>=2009 & dades$ANY<=2013,]
DIF_TOT_ESP_A <- casos.esperats(result.N)
result.N <- dades[dades$ICDO==icdo & dades$ESCENARIO=="B" & dades$ANY>=2009 & dades$ANY<=2013,]
DIF_TOT_ESP_B <- casos.esperats(result.N)
result.N <- dades[dades$ICDO==icdo & dades$ESCENARIO=="C" & dades$ANY>=2009 & dades$ANY<=2013,]
DIF_TOT_ESP_C <- casos.esperats(result.N)

casos <- incidencia_TG[incidencia_TG$PROVINCIA==17 & incidencia_TG$ICDO==icdo &
incidencia_TG$ANY==2009,]
result.N <- dades[dades$PROVINCIA==17 & dades$ICDO==icdo & dades$ESCENARIO=="A" & dades$ANY==2009,]
ADJ_G2009A <- validacio.escenari.any.ajust(result.N,casos,17,icdo,"A",2009,corteedad)
ABS_G2009A <- validacio.escenari.any.abs(result.N,casos,17,icdo,"A",2009)

result.N <- dades[dades$PROVINCIA==17 & dades$ICDO==icdo & dades$ESCENARIO=="B" & dades$ANY==2009,]
ADJ_G2009B <- validacio.escenari.any.ajust(result.N,casos,17,icdo,"B",2009,corteedad)

```



```

ABS_T2009C <- validacio.escenari.any.abs(result.N,casos,43,icdo,"C",2009)

casos <- incidencia_TG[incidencia_TG$PROVINCIA==43 & incidencia_TG$ICDO==icdo &
incidencia_TG$ANY==2010,]
result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="A" & dades$ANY==2010,]
ADJ_T2010A <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"A",2010,cortedad)
ABS_T2010A <- validacio.escenari.any.abs(result.N,casos,43,icdo,"A",2010)

result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="B" & dades$ANY==2010,]
ADJ_T2010B <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"B",2010,cortedad)
ABS_T2010B <- validacio.escenari.any.abs(result.N,casos,43,icdo,"B",2010)

result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="C" & dades$ANY==2010,]
ADJ_T2010C <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"C",2010,cortedad)
ABS_T2010C <- validacio.escenari.any.abs(result.N,casos,43,icdo,"C",2010)

casos <- incidencia_TG[incidencia_TG$PROVINCIA==43 & incidencia_TG$ICDO==icdo &
incidencia_TG$ANY==2011,]
result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="A" & dades$ANY==2011,]
ADJ_T2011A <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"A",2011,cortedad)
ABS_T2011A <- validacio.escenari.any.abs(result.N,casos,43,icdo,"A",2011)

result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="B" & dades$ANY==2011,]
ADJ_T2011B <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"B",2011,cortedad)
ABS_T2011B <- validacio.escenari.any.abs(result.N,casos,43,icdo,"B",2011)

result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="C" & dades$ANY==2011,]
ADJ_T2011C <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"C",2011,cortedad)
ABS_T2011C <- validacio.escenari.any.abs(result.N,casos,43,icdo,"C",2011)

casos <- incidencia_TG[incidencia_TG$PROVINCIA==43 & incidencia_TG$ICDO==icdo &
incidencia_TG$ANY==2012,]
result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="A" & dades$ANY==2012,]
ADJ_T2012A <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"A",2012,cortedad)
ABS_T2012A <- validacio.escenari.any.abs(result.N,casos,43,icdo,"A",2012)

result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="B" & dades$ANY==2012,]
ADJ_T2012B <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"B",2012,cortedad)
ABS_T2012B <- validacio.escenari.any.abs(result.N,casos,43,icdo,"B",2012)

result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="C" & dades$ANY==2012,]
ADJ_T2012C <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"C",2012,cortedad)
ABS_T2012C <- validacio.escenari.any.abs(result.N,casos,43,icdo,"C",2012)

casos <- incidencia_TG[incidencia_TG$PROVINCIA==43 & incidencia_TG$ICDO==icdo &
incidencia_TG$ANY==2013,]
result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="A" & dades$ANY==2013,]
ADJ_T2013A <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"A",2013,cortedad)
ABS_T2013A <- validacio.escenari.any.abs(result.N,casos,43,icdo,"A",2013)

result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="B" & dades$ANY==2013,]
ADJ_T2013B <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"B",2013,cortedad)
ABS_T2013B <- validacio.escenari.any.abs(result.N,casos,43,icdo,"B",2013)

result.N <- dades[dades$PROVINCIA==43 & dades$ICDO==icdo & dades$ESCENARIO=="C" & dades$ANY==2013,]
ADJ_T2013C <- validacio.escenari.any.ajust(result.N,casos,43,icdo,"C",2013,cortedad)
ABS_T2013C <- validacio.escenari.any.abs(result.N,casos,43,icdo,"C",2013)

ADJ_A <- ADJ_G2009A + ADJ_G2010A + ADJ_G2011A + ADJ_G2012A + ADJ_G2013A + ADJ_T2009A + ADJ_T2010A +
ADJ_T2011A + ADJ_T2012A + ADJ_T2013A
ADJ_B <- ADJ_G2009B + ADJ_G2010B + ADJ_G2011B + ADJ_G2012B + ADJ_G2013B + ADJ_T2009B + ADJ_T2010B +
ADJ_T2011B + ADJ_T2012B + ADJ_T2013B
ADJ_C <- ADJ_G2009C + ADJ_G2010C + ADJ_G2011C + ADJ_G2012C + ADJ_G2013C + ADJ_T2009C + ADJ_T2010C +
ADJ_T2011C + ADJ_T2012C + ADJ_T2013C
ABS_A <- (ABS_G2009A + ABS_G2010A + ABS_G2011A + ABS_G2012A + ABS_G2013A + ABS_T2009A + ABS_T2010A +
ABS_T2011A + ABS_T2012A + ABS_T2013A)/10
ABS_B <- (ABS_G2009B + ABS_G2010B + ABS_G2011B + ABS_G2012B + ABS_G2013B + ABS_T2009B + ABS_T2010B +
ABS_T2011B + ABS_T2012B + ABS_T2013B)/10
ABS_C <- (ABS_G2009C + ABS_G2010C + ABS_G2011C + ABS_G2012C + ABS_G2013C + ABS_T2009C + ABS_T2010C +
ABS_T2011C + ABS_T2012C + ABS_T2013C)/10
DIF_TOT_A <- ((DIF_TOT_ESP_A - DIF_TOT_OBS) / DIF_TOT_OBS)*100
DIF_TOT_B <- ((DIF_TOT_ESP_B - DIF_TOT_OBS) / DIF_TOT_OBS)*100
DIF_TOT_C <- ((DIF_TOT_ESP_C - DIF_TOT_OBS) / DIF_TOT_OBS)*100
ESCENARIO <- rep("A",24000)

```

```

for(i in 1:24000)
{
  if(ADJ_B[i]<ADJ_A[i]) { ESCENARIO[i] <- "B" }
  if(ADJ_C[i]<ADJ_B[i] & ADJ_C[i]<ADJ_A[i] ) { ESCENARIO[i] <- "C" }
}
ICDO <- rep(icdo,24000)
result <- data.frame(ICDO, DIF_TOT_OBS, DIF_TOT_ESP_A, DIF_TOT_A, DIF_TOT_ESP_B, DIF_TOT_B,
DIF_TOT_ESP_C, DIF_TOT_C, ADJ_A, ADJ_B, ADJ_C, ABS_A, ABS_B, ABS_C, ESCENARIO)
result
}

casos.esperats <- function(result.N)
{
  result.N <- aggregate(result.N[,7:24],by=list(SIM=result.N$SIM),FUN="sum")
  result.N <- result.N[,2:19]
  N <- rowSums(result.N)
  N
}

validacio.escenari.any.ajust <- function(result.N,casos,provincia,icdo,escenario,any,cortedad)
{
  kf1 <- cortedad - 1
  ki1 <- cortedad
  result.N <- result.N[,7:24]
  result.N$N.00 <- rowSums(result.N[1:kf1])
  aux <- names(result.N[ki1:18])
  aux <- cbind("N.00",t(aux))
  result.N <- result.N[,aux]
  SE <- casos[,23]
  casos <- casos[,5:22]
  N <- sum(casos)
  casos <- casos*((N+SE)/N)
  casos$I00 <- rowSums(casos[1:kf1])
  aux <- names(casos[ki1:18])
  aux <- cbind("I00",t(aux))
  casos <- casos[,aux]
  casos <- casos[rep(seq_len(nrow(casos)), each=24000),]
  dif <- ((casos-result.N)^2)/result.N
  dif[is.na(dif)] <- 0
  dif <- do.call(data.frame,lapply(dif,function(x) replace(x,is.infinite(x),0)))
  valor <- rowSums(dif)
  valor
}

validacio.escenari.any.abs <- function(result.N,casos,provincia,icdo,escenario,any)
{
  result.N <- result.N[,7:24]
  N.pr <- rowSums(result.N)
  casos <- casos[,5:23]
  N.real <- sum(casos)
  N.real <- rep(N.real,24000)
  dif <- abs(N.pr-N.real)
  p.dif <- (dif/N.real)*100
  p.dif
}

```

4. Estimació de la incidència per les províncies sense RCBP a la secció 7

```

mort9411 <- read.csv("/Users/mac/Desktop/PFG/mortalidad.csv", header = TRUE, sep = ";", fileEncoding =
"latin1")
mort1213 <- read.csv("/Users/mac/Desktop/PFG/mortalitat.2012-2013.csv", header = TRUE, sep = ";",
fileEncoding = "latin1")
colnames(mort1213)[2] <- "ANY"
mortalitat <- rbind(mort9411,mort1213)
mortalitat <- subset(mortalitat, mortalitat$ANY>=1994)
head(mortalitat)
mortalidad <- mortalitat

provincies <- subset(parametros.provincia,
(parametros.provincia$CODIGO>2 & parametros.provincia$CODIGO<16) |

```

```

        parametros.provincia$CODIGO==19 | (parametros.provincia$CODIGO>20 &
parametros.provincia$CODIGO<26) |
        (parametros.provincia$CODIGO>26 & parametros.provincia$CODIGO<30) |
        (parametros.provincia$CODIGO>31 & parametros.provincia$CODIGO<33) |
        (parametros.provincia$CODIGO>33 & parametros.provincia$CODIGO<43) |
        (parametros.provincia$CODIGO>43 & parametros.provincia$CODIGO<48) |
        (parametros.provincia$CODIGO>48 & parametros.provincia$CODIGO<54) )

parametros.tumor$ESCENARIO <- c("A", "C", "C", "B", "A", "0", "B", "A", "B", "C", "B", "A", "C", "C",
"B", "A", "B", "C", "C", "A", "C", "A", "A", "0", "C", "C", "B", "B", "A", "0", "B", "C", "C", "C", "B",
"C", "B", "C", "C", "A", "A", "A", "B", "C", "C", "A", "A", "A", "B", "0")

coe.u <- funcion(provincia,icdo,escenario,any)
{
  modelo <- parametros.tumor[parametros.tumor$CODIGO == icdo,"TIPO"]
  q1 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"P10"]
  q2 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"T2"]
  q3 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"T3"]
  q4 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"P90"]
  setwd("/Users/mac/Desktop/PFG/Resultats")
  nombre.tumor <- paste0("simulacion-",icdo,".txt")
  result.sim <- read.csv2(nombre.tumor, header=TRUE)

  coe.a <- result.sim$a
  coe.p <- rnorm(24000,mean=0, sd=coe.sigma(icdo))

  ANY <- result.sim$ANY
  ANY2 <- result.sim$ANY2

  EDAD1 <- result.sim$EDAD1
  EDAD2 <- result.sim$EDAD2
  EDAD3 <- result.sim$EDAD3
  EDAD4 <- result.sim$EDAD4

  if (escenario=="A")
  {
    coe.any <- (2007-1994)*ANY + ((2007-1994)*(2007-1994))*ANY2
  }
  if (escenario=="B")
  {
    coe.any <- (2007-1994)*ANY + ((2007-1994)*(2007-1994))*ANY2
    coe.any <- coe.any*((any-1994)/(2007-1994))
  }
  if (escenario=="C")
  {
    coe.any <- (any-1994)*ANY + ((any-1994)*(any-1994))*ANY2
  }

  coe.edad1 <- EDAD1*pmax(0,2.5-q1)
  coe.edad2 <- EDAD2*pmax(0,2.5-q2)
  coe.edad3 <- EDAD3*pmax(0,2.5-q3)
  coe.edad4 <- EDAD4*pmax(0,2.5-q4)
  coe.edad.01 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
  coe.01 <- coe.a + coe.any + coe.edad.01 + coe.p
  coe.edad1 <- EDAD1*pmax(0,7.5-q1)
  coe.edad2 <- EDAD2*pmax(0,7.5-q2)
  coe.edad3 <- EDAD3*pmax(0,7.5-q3)
  coe.edad4 <- EDAD4*pmax(0,7.5-q4)
  coe.edad.02 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
  coe.02 <- coe.a + coe.any + coe.edad.02 + coe.p
  coe.edad1 <- EDAD1*pmax(0,12.5-q1)
  coe.edad2 <- EDAD2*pmax(0,12.5-q2)
  coe.edad3 <- EDAD3*pmax(0,12.5-q3)
  coe.edad4 <- EDAD4*pmax(0,12.5-q4)
  coe.edad.03 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
  coe.03 <- coe.a + coe.any + coe.edad.03 + coe.p
  coe.edad1 <- EDAD1*pmax(0,17.5-q1)
  coe.edad2 <- EDAD2*pmax(0,17.5-q2)
  coe.edad3 <- EDAD3*pmax(0,17.5-q3)
  coe.edad4 <- EDAD4*pmax(0,17.5-q4)
  coe.edad.04 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
  coe.04 <- coe.a + coe.any + coe.edad.04 + coe.p
  coe.edad1 <- EDAD1*pmax(0,22.5-q1)
  coe.edad2 <- EDAD2*pmax(0,22.5-q2)

```

```

coe.edad3 <- EDAD3*pmax(0,22.5-q3)
coe.edad4 <- EDAD4*pmax(0,22.5-q4)
coe.edad.05 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.05 <- coe.a + coe.any + coe.edad.05 + coe.p
coe.edad1 <- EDAD1*pmax(0,27.5-q1)
coe.edad2 <- EDAD2*pmax(0,27.5-q2)
coe.edad3 <- EDAD3*pmax(0,27.5-q3)
coe.edad4 <- EDAD4*pmax(0,27.5-q4)
coe.edad.06 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.06 <- coe.a + coe.any + coe.edad.06 + coe.p
coe.edad1 <- EDAD1*pmax(0,32.5-q1)
coe.edad2 <- EDAD2*pmax(0,32.5-q2)
coe.edad3 <- EDAD3*pmax(0,32.5-q3)
coe.edad4 <- EDAD4*pmax(0,32.5-q4)
coe.edad.07 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.07 <- coe.a + coe.any + coe.edad.07 + coe.p
coe.edad1 <- EDAD1*pmax(0,37.5-q1)
coe.edad2 <- EDAD2*pmax(0,37.5-q2)
coe.edad3 <- EDAD3*pmax(0,37.5-q3)
coe.edad4 <- EDAD4*pmax(0,37.5-q4)
coe.edad.08 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.08 <- coe.a + coe.any + coe.edad.08 + coe.p
coe.edad1 <- EDAD1*pmax(0,42.5-q1)
coe.edad2 <- EDAD2*pmax(0,42.5-q2)
coe.edad3 <- EDAD3*pmax(0,42.5-q3)
coe.edad4 <- EDAD4*pmax(0,42.5-q4)
coe.edad.09 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.09 <- coe.a + coe.any + coe.edad.09 + coe.p
coe.edad1 <- EDAD1*pmax(0,47.5-q1)
coe.edad2 <- EDAD2*pmax(0,47.5-q2)
coe.edad3 <- EDAD3*pmax(0,47.5-q3)
coe.edad4 <- EDAD4*pmax(0,47.5-q4)
coe.edad.10 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.10 <- coe.a + coe.any + coe.edad.10 + coe.p
coe.edad1 <- EDAD1*pmax(0,52.5-q1)
coe.edad2 <- EDAD2*pmax(0,52.5-q2)
coe.edad3 <- EDAD3*pmax(0,52.5-q3)
coe.edad4 <- EDAD4*pmax(0,52.5-q4)
coe.edad.11 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.11 <- coe.a + coe.any + coe.edad.11 + coe.p
coe.edad1 <- EDAD1*pmax(0,57.5-q1)
coe.edad2 <- EDAD2*pmax(0,57.5-q2)
coe.edad3 <- EDAD3*pmax(0,57.5-q3)
coe.edad4 <- EDAD4*pmax(0,57.5-q4)
coe.edad.12 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.12 <- coe.a + coe.any + coe.edad.12 + coe.p
coe.edad1 <- EDAD1*pmax(0,62.5-q1)
coe.edad2 <- EDAD2*pmax(0,62.5-q2)
coe.edad3 <- EDAD3*pmax(0,62.5-q3)
coe.edad4 <- EDAD4*pmax(0,62.5-q4)
coe.edad.13 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.13 <- coe.a + coe.any + coe.edad.13 + coe.p
coe.edad1 <- EDAD1*pmax(0,67.5-q1)
coe.edad2 <- EDAD2*pmax(0,67.5-q2)
coe.edad3 <- EDAD3*pmax(0,67.5-q3)
coe.edad4 <- EDAD4*pmax(0,67.5-q4)
coe.edad.14 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.14 <- coe.a + coe.any + coe.edad.14 + coe.p
coe.edad1 <- EDAD1*pmax(0,72.5-q1)
coe.edad2 <- EDAD2*pmax(0,72.5-q2)
coe.edad3 <- EDAD3*pmax(0,72.5-q3)
coe.edad4 <- EDAD4*pmax(0,72.5-q4)
coe.edad.15 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.15 <- coe.a + coe.any + coe.edad.15 + coe.p
coe.edad1 <- EDAD1*pmax(0,77.5-q1)
coe.edad2 <- EDAD2*pmax(0,77.5-q2)
coe.edad3 <- EDAD3*pmax(0,77.5-q3)
coe.edad4 <- EDAD4*pmax(0,77.5-q4)
coe.edad.16 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.16 <- coe.a + coe.any + coe.edad.16 + coe.p
coe.edad1 <- EDAD1*pmax(0,82.5-q1)
coe.edad2 <- EDAD2*pmax(0,82.5-q2)
coe.edad3 <- EDAD3*pmax(0,82.5-q3)
coe.edad4 <- EDAD4*pmax(0,82.5-q4)

```



```

coe.edad.17 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.17 <- coe.a + coe.any + coe.edad.17 + coe.p
coe.edad1 <- EDAD1*pmax(0,90-q1)
coe.edad2 <- EDAD2*pmax(0,90-q2)
coe.edad3 <- EDAD3*pmax(0,90-q3)
coe.edad4 <- EDAD4*pmax(0,90-q4)
coe.edad.18 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.18 <- coe.a + coe.any + coe.edad.18 + coe.p
coe.u <-
cbind(coe.01,coe.02,coe.03,coe.04,coe.05,coe.06,coe.07,coe.08,coe.09,coe.10,coe.11,coe.12,coe.13,coe.14,coe.15,coe.16,coe.17,coe.18)
  coe.u
}

coe.sigma <- function(icdo)
{
  setwd("/Users/mac/Desktop/PFG/Resultats")
  nombre.tumor <- paste0("simulacion-",icdo,".txt")
  result.sim <- read.csv2(nombre.tumor, header=TRUE)
  modelo <- parametros.tumor[parametros.tumor$CODIGO == icdo,"TIPO"]
  coe.sigma <- result.sim$sigma
  coe.sigma
}

projeccion.N <- function(provincia,icdo,escenario,any,n.sim=1)
{
  set.seed(112358)
  u <- coe.u(provincia,icdo,escenario,any)
  sigma <- coe.sigma(icdo)
  muertos <- mortalidad[mortalidad$PROVINCIA==provincia & mortalidad$ICDO==icdo &
mortalidad$ANY==any,5:22]
  if(dim(muertos)[1]==0){
    nombres <- colnames(muertos)
    muertos <- as.data.frame(matrix(0, ncol = 18, nrow = 1))
    colnames(muertos) <- nombres
  }

  for (i in 1:n.sim) {
    alpha.01 <- rnorm(24000,1,sigma)
    alpha.02 <- rnorm(24000,1,sigma)
    alpha.03 <- rnorm(24000,1,sigma)
    alpha.04 <- rnorm(24000,1,sigma)
    alpha.05 <- rnorm(24000,1,sigma)
    alpha.06 <- rnorm(24000,1,sigma)
    alpha.07 <- rnorm(24000,1,sigma)
    alpha.08 <- rnorm(24000,1,sigma)
    alpha.09 <- rnorm(24000,1,sigma)
    alpha.10 <- rnorm(24000,1,sigma)
    alpha.11 <- rnorm(24000,1,sigma)
    alpha.12 <- rnorm(24000,1,sigma)
    alpha.13 <- rnorm(24000,1,sigma)
    alpha.14 <- rnorm(24000,1,sigma)
    alpha.15 <- rnorm(24000,1,sigma)
    alpha.16 <- rnorm(24000,1,sigma)
    alpha.17 <- rnorm(24000,1,sigma)
    alpha.18 <- rnorm(24000,1,sigma)
    alpha <-
cbind(alpha.01,alpha.02,alpha.03,alpha.04,alpha.05,alpha.06,alpha.07,alpha.08,alpha.09,alpha.10,alpha.11,
alpha.12,alpha.13,alpha.14,alpha.15,alpha.16,alpha.17,alpha.18)
    alpha[alpha < 0] <- 0
    RIM <- exp(u) * alpha
    ESP.01<-RIM[,1]*muertos[,1]
    ESP.02<-RIM[,2]*muertos[,2]
    ESP.03<-RIM[,3]*muertos[,3]
    ESP.04<-RIM[,4]*muertos[,4]
    ESP.05<-RIM[,5]*muertos[,5]
    ESP.06<-RIM[,6]*muertos[,6]
    ESP.07<-RIM[,7]*muertos[,7]
    ESP.08<-RIM[,8]*muertos[,8]
    ESP.09<-RIM[,9]*muertos[,9]
    ESP.10<-RIM[,10]*muertos[,10]
    ESP.11<-RIM[,11]*muertos[,11]
    ESP.12<-RIM[,12]*muertos[,12]
    ESP.13<-RIM[,13]*muertos[,13]
  }
}

```

```

ESP.14<-RIM[,14]*muertos[,14]
ESP.15<-RIM[,15]*muertos[,15]
ESP.16<-RIM[,16]*muertos[,16]
ESP.17<-RIM[,17]*muertos[,17]
ESP.18<-RIM[,18]*muertos[,18]
ITER <- rep(i,24000)
SIM <- rep(1:24000)
PROVINCIA <- rep(provincia,24000)
ICDO <- rep(icdo,24000)
ANY <- rep(any,24000)
ESCENARIO <- rep(escenario,24000)
N.01 <- rpois(24000,ESP.01)
N.02 <- rpois(24000,ESP.02)
N.03 <- rpois(24000,ESP.03)
N.04 <- rpois(24000,ESP.04)
N.05 <- rpois(24000,ESP.05)
N.06 <- rpois(24000,ESP.06)
N.07 <- rpois(24000,ESP.07)
N.08 <- rpois(24000,ESP.08)
N.09 <- rpois(24000,ESP.09)
N.10 <- rpois(24000,ESP.10)
N.11 <- rpois(24000,ESP.11)
N.12 <- rpois(24000,ESP.12)
N.13 <- rpois(24000,ESP.13)
N.14 <- rpois(24000,ESP.14)
N.15 <- rpois(24000,ESP.15)
N.16 <- rpois(24000,ESP.16)
N.17 <- rpois(24000,ESP.17)
N.18 <- rpois(24000,ESP.18)
if (i==1) {CASOS <-
cbind(ITER, SIM, PROVINCIA, ICDO, ANY, ESCENARIO, N.01, N.02, N.03, N.04, N.05, N.06, N.07, N.08, N.09, N.10, N.11, N.12, N
.13, N.14, N.15, N.16, N.17, N.18)}
  if (i>1) {
    aux <-
cbind(ITER, SIM, PROVINCIA, ICDO, ANY, ESCENARIO, N.01, N.02, N.03, N.04, N.05, N.06, N.07, N.08, N.09, N.10, N.11, N.12, N
.13, N.14, N.15, N.16, N.17, N.18)
    CASOS <- rbind(CASOS, aux)
  }
}
CASOS
}

projeccion <- funcion(icdo, n.sim=1)
{
  escenario <- parametros.tumor[ parametros.tumor$CODIGO==icdo, 10]
  anys <- 2009:2013
  provincias.sr <- provincias$CODIGO
  for(i in 1:length(provincias.sr)){ # es selecciona la provincia
    prov <- provincias.sr[i]
    for(j in 1:length(anys)){ # del 2009 al 2013
      any <- anys[j]
      if(i==1 & j==1){
        resultat <- projeccion.N(prov, icdo, escenario, any, n.sim)
      }else{
        resultat <- rbind(resultat, projeccion.N(prov, icdo, escenario, any, n.sim))
      }
    }
  }
}
setwd("/Users/mac/Desktop/PFG/Resultats_0913_SENSE_RCBP")
#setwd("/Volumes/ANDREA_USB")
nombre.archivo <- paste0("Resultat-tumor-SenseRCBP-", icdo, ".txt")
write.csv2(resultat, file=nombre.archivo, row.names=FALSE)
}

```


5. Estimació de la incidència per les províncies amb RCBP a la secció 7

```

mort9411 <- read.csv("/Users/mac/Desktop/PFG/mortalidad.csv", header = TRUE, sep = ";", fileEncoding =
"latin1")
mort1213 <- read.csv("/Users/mac/Desktop/PFG/mortalitat.2012-2013.csv", header = TRUE, sep = ";",
fileEncoding = "latin1")
colnames(mort1213)[2] <- "ANY"
mortalitat <- rbind(mort9411,mort1213)
mortalitat <- subset(mortalitat, mortalitat$ANY>=1994)
head(mortalitat)
mortalidad <- mortalitat
# Selecciono les províncies amb RCBP a excepció de Girona i Tarragona
provincies <- subset(parametros.provincia,
parametros.provincia$CODIGO<=2 |
parametros.provincia$CODIGO==16 |
parametros.provincia$CODIGO==18 |
parametros.provincia$CODIGO==20 |
parametros.provincia$CODIGO==26 |
parametros.provincia$CODIGO==30 |
parametros.provincia$CODIGO==31 |
parametros.provincia$CODIGO==33 |
parametros.provincia$CODIGO==48 |
parametros.provincia$CODIGO==54 |
parametros.provincia$CODIGO==55 |
parametros.provincia$CODIGO==56 )

parametros.tumor$ESCENARIO <-c("A", "C", "C", "B", "A", "0", "B", "A", "B", "C", "B", "A", "C", "C", "B",
"A", "B", "C", "C", "A", "C", "A", "A", "0", "C", "C", "B", "B", "A", "0", "B", "C", "C", "C", "B", "C",
"B", "C", "C", "A", "A", "A", "B", "C", "C", "A", "A", "A", "B", "0")

coe.u <- funcion(provincia,icdo,escenario,any)
{
modelo <- parametros.tumor[parametros.tumor$CODIGO == icdo,"TIPO"]
q1 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"P10"]
q2 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"T2"]
q3 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"T3"]
q4 <- parametros.tumor[parametros.tumor$CODIGO == icdo,"P90"]
setwd("/Users/mac/Desktop/PFG/Resultats")
nombre.tumor <- paste0("simulacion-",icdo,".txt")
result.sim <- read.csv2(nombre.tumor, header=TRUE)

coe.a <- result.sim$a
coe.p <- rnorm(24000,mean=0, sd=coe.sigma(icdo))

ANY <- result.sim$ANY
ANY2 <- result.sim$ANY2

EDAD1 <- result.sim$EDAD1
EDAD2 <- result.sim$EDAD2
EDAD3 <- result.sim$EDAD3
EDAD4 <- result.sim$EDAD4

if (escenario=="A")
{
coe.any <- (2007-1994)*ANY + ((2007-1994)*(2007-1994))*ANY2
}
if (escenario=="B")
{
coe.any <- (2007-1994)*ANY + ((2007-1994)*(2007-1994))*ANY2
coe.any <- coe.any*((any-1994)/(2007-1994))
}
if (escenario=="C")
{
coe.any <- (any-1994)*ANY + ((any-1994)*(any-1994))*ANY2
}

coe.edad1 <- EDAD1*pmax(0,2.5-q1)
coe.edad2 <- EDAD2*pmax(0,2.5-q2)
coe.edad3 <- EDAD3*pmax(0,2.5-q3)
coe.edad4 <- EDAD4*pmax(0,2.5-q4)
coe.edad.01 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.01 <- coe.a + coe.any + coe.edad.01 + coe.p
coe.edad1 <- EDAD1*pmax(0,7.5-q1)
coe.edad2 <- EDAD2*pmax(0,7.5-q2)

```

```

coe.edad3 <- EDAD3*pmax(0,7.5-q3)
coe.edad4 <- EDAD4*pmax(0,7.5-q4)
coe.edad.02 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.02 <- coe.a + coe.any + coe.edad.02 + coe.p
coe.edad1 <- EDAD1*pmax(0,12.5-q1)
coe.edad2 <- EDAD2*pmax(0,12.5-q2)
coe.edad3 <- EDAD3*pmax(0,12.5-q3)
coe.edad4 <- EDAD4*pmax(0,12.5-q4)
coe.edad.03 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.03 <- coe.a + coe.any + coe.edad.03 + coe.p
coe.edad1 <- EDAD1*pmax(0,17.5-q1)
coe.edad2 <- EDAD2*pmax(0,17.5-q2)
coe.edad3 <- EDAD3*pmax(0,17.5-q3)
coe.edad4 <- EDAD4*pmax(0,17.5-q4)
coe.edad.04 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.04 <- coe.a + coe.any + coe.edad.04 + coe.p
coe.edad1 <- EDAD1*pmax(0,22.5-q1)
coe.edad2 <- EDAD2*pmax(0,22.5-q2)
coe.edad3 <- EDAD3*pmax(0,22.5-q3)
coe.edad4 <- EDAD4*pmax(0,22.5-q4)
coe.edad.05 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.05 <- coe.a + coe.any + coe.edad.05 + coe.p
coe.edad1 <- EDAD1*pmax(0,27.5-q1)
coe.edad2 <- EDAD2*pmax(0,27.5-q2)
coe.edad3 <- EDAD3*pmax(0,27.5-q3)
coe.edad4 <- EDAD4*pmax(0,27.5-q4)
coe.edad.06 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.06 <- coe.a + coe.any + coe.edad.06 + coe.p
coe.edad1 <- EDAD1*pmax(0,32.5-q1)
coe.edad2 <- EDAD2*pmax(0,32.5-q2)
coe.edad3 <- EDAD3*pmax(0,32.5-q3)
coe.edad4 <- EDAD4*pmax(0,32.5-q4)
coe.edad.07 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.07 <- coe.a + coe.any + coe.edad.07 + coe.p
coe.edad1 <- EDAD1*pmax(0,37.5-q1)
coe.edad2 <- EDAD2*pmax(0,37.5-q2)
coe.edad3 <- EDAD3*pmax(0,37.5-q3)
coe.edad4 <- EDAD4*pmax(0,37.5-q4)
coe.edad.08 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.08 <- coe.a + coe.any + coe.edad.08 + coe.p
coe.edad1 <- EDAD1*pmax(0,42.5-q1)
coe.edad2 <- EDAD2*pmax(0,42.5-q2)
coe.edad3 <- EDAD3*pmax(0,42.5-q3)
coe.edad4 <- EDAD4*pmax(0,42.5-q4)
coe.edad.09 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.09 <- coe.a + coe.any + coe.edad.09 + coe.p
coe.edad1 <- EDAD1*pmax(0,47.5-q1)
coe.edad2 <- EDAD2*pmax(0,47.5-q2)
coe.edad3 <- EDAD3*pmax(0,47.5-q3)
coe.edad4 <- EDAD4*pmax(0,47.5-q4)
coe.edad.10 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.10 <- coe.a + coe.any + coe.edad.10 + coe.p
coe.edad1 <- EDAD1*pmax(0,52.5-q1)
coe.edad2 <- EDAD2*pmax(0,52.5-q2)
coe.edad3 <- EDAD3*pmax(0,52.5-q3)
coe.edad4 <- EDAD4*pmax(0,52.5-q4)
coe.edad.11 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.11 <- coe.a + coe.any + coe.edad.11 + coe.p
coe.edad1 <- EDAD1*pmax(0,57.5-q1)
coe.edad2 <- EDAD2*pmax(0,57.5-q2)
coe.edad3 <- EDAD3*pmax(0,57.5-q3)
coe.edad4 <- EDAD4*pmax(0,57.5-q4)
coe.edad.12 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.12 <- coe.a + coe.any + coe.edad.12 + coe.p
coe.edad1 <- EDAD1*pmax(0,62.5-q1)
coe.edad2 <- EDAD2*pmax(0,62.5-q2)
coe.edad3 <- EDAD3*pmax(0,62.5-q3)
coe.edad4 <- EDAD4*pmax(0,62.5-q4)
coe.edad.13 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.13 <- coe.a + coe.any + coe.edad.13 + coe.p
coe.edad1 <- EDAD1*pmax(0,67.5-q1)
coe.edad2 <- EDAD2*pmax(0,67.5-q2)
coe.edad3 <- EDAD3*pmax(0,67.5-q3)
coe.edad4 <- EDAD4*pmax(0,67.5-q4)

```

```

coe.edad.14 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.14 <- coe.a + coe.any + coe.edad.14 + coe.p
coe.edad1 <- EDAD1*pmax(0,72.5-q1)
coe.edad2 <- EDAD2*pmax(0,72.5-q2)
coe.edad3 <- EDAD3*pmax(0,72.5-q3)
coe.edad4 <- EDAD4*pmax(0,72.5-q4)
coe.edad.15 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.15 <- coe.a + coe.any + coe.edad.15 + coe.p
coe.edad1 <- EDAD1*pmax(0,77.5-q1)
coe.edad2 <- EDAD2*pmax(0,77.5-q2)
coe.edad3 <- EDAD3*pmax(0,77.5-q3)
coe.edad4 <- EDAD4*pmax(0,77.5-q4)
coe.edad.16 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.16 <- coe.a + coe.any + coe.edad.16 + coe.p
coe.edad1 <- EDAD1*pmax(0,82.5-q1)
coe.edad2 <- EDAD2*pmax(0,82.5-q2)
coe.edad3 <- EDAD3*pmax(0,82.5-q3)
coe.edad4 <- EDAD4*pmax(0,82.5-q4)
coe.edad.17 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.17 <- coe.a + coe.any + coe.edad.17 + coe.p
coe.edad1 <- EDAD1*pmax(0,90-q1)
coe.edad2 <- EDAD2*pmax(0,90-q2)
coe.edad3 <- EDAD3*pmax(0,90-q3)
coe.edad4 <- EDAD4*pmax(0,90-q4)
coe.edad.18 <- coe.edad1+coe.edad2+coe.edad3+coe.edad4
coe.18 <- coe.a + coe.any + coe.edad.18 + coe.p
coe.u <-
cbind(coe.01,coe.02,coe.03,coe.04,coe.05,coe.06,coe.07,coe.08,coe.09,coe.10,coe.11,coe.12,coe.13,coe.14,c
oe.15,coe.16,coe.17,coe.18)
  coe.u
}

coe.sigma <- function(icdo)
{
  setwd("/Users/mac/Desktop/PFG/Resultats")
  nombre.tumor <- paste0("simulacion-",icdo,".txt")
  result.sim <- read.csv2(nombre.tumor, header=TRUE)
  coe.sigma <- result.sim$sigma
  coe.sigma
}

projeccion.N <- function(provincia,icdo,escenario,any,n.sim=1)
{
  set.seed(112358)
  u <- coe.u(provincia,icdo,escenario,any)
  sigma <- coe.sigma(icdo)
  muertos <- mortalidad[mortalidad$PROVINCIA==provincia & mortalidad$ICD0==icdo &
mortalidad$ANY==any,5:22]

  if(icdo<200){
    sexo <- "M"
  }else if(icdo>200){
    sexo <- "F"
  }

  if (provincia == 54){
    muertos <- correccion.mortalidad.provincia.tumor(provincia, 7, icdo, sexo, any)[5:22]
  }
  if (provincia == 55){
    muertos <- correccion.mortalidad.provincia.tumor(provincia, 35, icdo, sexo, any)[5:22]
  }
  if (provincia == 56){
    muertos <- correccion.mortalidad.provincia.tumor(provincia, 38, icdo, sexo, any)[5:22]
  }

  if(dim(muertos)[1]==0){
    nombres <- colnames(muertos)
    muertos <- as.data.frame(matrix(0, ncol = 18, nrow = 1))
    colnames(muertos) <- nombres
  }

  for (i in 1:n.sim) {
    RIM <- exp(u)
    ESP.01<-RIM[,1]*muertos[,1]
  }
}

```

```

ESP.02<-RIM[,2]*muertos[,2]
ESP.03<-RIM[,3]*muertos[,3]
ESP.04<-RIM[,4]*muertos[,4]
ESP.05<-RIM[,5]*muertos[,5]
ESP.06<-RIM[,6]*muertos[,6]
ESP.07<-RIM[,7]*muertos[,7]
ESP.08<-RIM[,8]*muertos[,8]
ESP.09<-RIM[,9]*muertos[,9]
ESP.10<-RIM[,10]*muertos[,10]
ESP.11<-RIM[,11]*muertos[,11]
ESP.12<-RIM[,12]*muertos[,12]
ESP.13<-RIM[,13]*muertos[,13]
ESP.14<-RIM[,14]*muertos[,14]
ESP.15<-RIM[,15]*muertos[,15]
ESP.16<-RIM[,16]*muertos[,16]
ESP.17<-RIM[,17]*muertos[,17]
ESP.18<-RIM[,18]*muertos[,18]
ITER <- rep(i,24000)
SIM <- rep(1:24000)
PROVINCIA <- rep(provincia,24000)
ICDO <- rep(icdo,24000)
ANY <- rep(any,24000)
ESCENARIO <- rep(escenario,24000)
N.01 <- rpois(24000,ESP.01)
N.02 <- rpois(24000,ESP.02)
N.03 <- rpois(24000,ESP.03)
N.04 <- rpois(24000,ESP.04)
N.05 <- rpois(24000,ESP.05)
N.06 <- rpois(24000,ESP.06)
N.07 <- rpois(24000,ESP.07)
N.08 <- rpois(24000,ESP.08)
N.09 <- rpois(24000,ESP.09)
N.10 <- rpois(24000,ESP.10)
N.11 <- rpois(24000,ESP.11)
N.12 <- rpois(24000,ESP.12)
N.13 <- rpois(24000,ESP.13)
N.14 <- rpois(24000,ESP.14)
N.15 <- rpois(24000,ESP.15)
N.16 <- rpois(24000,ESP.16)
N.17 <- rpois(24000,ESP.17)
N.18 <- rpois(24000,ESP.18)
if (i==1) {
  CASOS <-
cbind(ITER,SIM,PROVINCIA,ICDO,ANY,ESCENARIO,N.01,N.02,N.03,N.04,N.05,N.06,N.07,N.08,N.09,N.10,N.11,N.12,N
.13,N.14,N.15,N.16,N.17,N.18)
}
if (i>1) {
  aux <-
cbind(ITER,SIM,PROVINCIA,ICDO,ANY,ESCENARIO,N.01,N.02,N.03,N.04,N.05,N.06,N.07,N.08,N.09,N.10,N.11,N.12,N
.13,N.14,N.15,N.16,N.17,N.18)
  CASOS <- rbind(CASOS,aux)
}
}
CASOS
}

projeccion <- funcion(icdo,n.sim=1)
{
  escenario <- parametros.tumor[parametros.tumor$CODIGO==icdo, 10]
  anys <- 2009:2013
  provincias <- provincias$CODIGO
  for(i in 1:length(provincias)){ # es selecciona la provincia
    prov <- provincias[i]
    for(j in 1:length(anys)){ # del 2009 al 2013
      any <- anys[j]
      if(i==1 & j==1){
        resultat <- projeccion.N(prov,icdo,escenario,any,n.sim)
      }else{
        resultat <- rbind(resultat, projeccion.N(prov,icdo,escenario,any,n.sim))
      }
    }
  }
}

setwd("/Users/mac/Desktop/PFG/Resultats_0913_AMB_RCBP")

```

```

#setwd("/Volumes/ANDREA_USB")
nombre.archivo <- paste0("Resultat-tumor-AmbRCBP-",icdo,".txt")
write.csv2(resultat,file=nombre.archivo,row.names=FALSE)
}

icdo <- parametros.tumor$CODI
icdo <- icdo[-c(6,24,30,length(icdo))]
icdo <- icdo[-c(1,2,3,4,5,6,7,8,9,10,11,12,13,17,18)]
icdo <- icdo[-c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16)]
for(i in 1:length(icdo)){
  codi <- icdo[i]
  proyeccion(codi)
}

correccion.mortalidad.provincia.tumor <- function(isla, provincia, icdo, sexo, ANYi=1994, ANYf=2007)
{
  mortalidad.provincia <- mortalidad[mortalidad$PROVINCIA == provincia & mortalidad$SEXO == sexo &
mortalidad$ICDO == icdo & mortalidad$ANY == any, 5:22]
  lista.ANY <- mortalidad[mortalidad$PROVINCIA == provincia & mortalidad$SEXO == sexo & mortalidad$ICDO
== icdo & mortalidad$ANY == any, 2]
  poblacion.isla <- poblacion[poblacion$PROVINCIA == isla & poblacion$SEXO == sexo & poblacion$ANY ==
any,]
  poblacion.provincia <- poblacion[poblacion$PROVINCIA == provincia & poblacion$SEXO == sexo &
poblacion$ANY == any,]
  i <- dim(mortalidad.provincia)[1]
  PROVINCIA <- c(rep(isla, i))
  ANY <- lista.ANY
  SEXO <- c(rep(sexo,i))
  ICDO <- c(rep(icdo,i))

  if (any(lista.ANY==ANY)==FALSE) {
    poblacion.isla <- poblacion.isla[poblacion.isla$ANY != ANY,]
    poblacion.provincia <- poblacion.provincia[poblacion.provincia$ANY != ANY,]
  }

  poblacion.isla <- poblacion.isla[,4:21]
  poblacion.provincia <- poblacion.provincia[,4:21]
  peso <- poblacion.isla/poblacion.provincia
  mortalidad.isla <- mortalidad.provincia*peso
  mortalidad.isla <- cbind(PROVINCIA, ANY, SEXO, ICDO, mortalidad.isla)
  mortalidad.isla
}

```

6. Creació mapes i taula corresponent a la secció 7

Aquestes dues funcions permeten obtenir els mapes i la taula per a tots els tipus de tumors introduint el seu codi identificador, havent llegit primer les dades.

```
library(plyr)
# Lectura de dades
pob <- read.csv("/Users/mac/Desktop/PFG/poblacion.csv", header = TRUE, sep = ";", fileEncoding =
"latin1")
pob <- pob[pob$ANY>=2009,]
for(j in 4:ncol(pob)){
  pob[,j] <- as.numeric(gsub(",", ".", gsub("\\.", "", pob[,j])))
}
pob1213 <- read.csv(paste0("/Users/mac/Desktop/PFG/Poblacion 2012-2013.csv"), header = TRUE, sep =
";",fileEncoding = "latin1")
colnames(pob1213)[c(2,3)] <- c("ANY","SEX0")
pob1213 <- pob1213[,1:21]
poblacio <- rbind.fill(pob,pob1213)

creacio.taula <- function(icdo){
  if (icdo < 200) {
    sex <- "M"
  }
  if (icdo > 200) {
    sex <- "F"
  }
  # Provincies amb RCBP
  p1 <- read.csv(paste0("/Users/mac/Desktop/PFG/Resultats_0913_AMB_RCBP/Resultat-tumor-AmbRCBP-", icdo,
".txt"), header = TRUE, sep = ";")
  # Provincies sense RCBP
  p2 <- read.csv(paste0("/Users/mac/Desktop/PFG/Resultats_0913_SENSE_RCBP/Resultat-tumor-SenseRCBP-",
icdo, ".txt"), header = TRUE, sep = ";")
  # Provincies Tgn i Grn amb RCBP
  p3 <- read.csv(paste0("/Users/mac/Desktop/PFG/Resultats/Resultat-tumor-", icdo, ".txt"), header = TRUE,
sep = ";")
  escenari <- p1$ESCENARIO[1]
  if( escenari == "A"){
    esc <- "A"
  }else if(escenari=="B"){
    esc <- "B"
  }else{
    esc <- "C"
  }
  p3 <- subset(p3, p3$ESCENARIO==esc)
  data <- rbind(p1,p2,p3)
  data$PROVINCIA <- as.factor(data$PROVINCIA)
  provincies <- levels(data$PROVINCIA)
  anys <- 2009:2013
  df <- as.data.frame(matrix(0, nrow=length(provincies), ncol = 21))
  for(i in 1:length(provincies)){ # totes les provincies
    v <- c()
    prov <- provincies[i]
    v <- c(v, prov)
    for(j in 1:length(anys)){
      any <- anys[j]
      prov.any <- subset(data, data$PROVINCIA==prov & data$ANY==any)
      prov.any$SUMA <- rowSums(prov.any[,7:24])
      prov.any.poblacio <- subset(poblacio, poblacio$PROVINCIA==prov & poblacio$ANY==any &
poblacio$SEX0==sex)
      prov.any.poblacio$SUMA <- rowSums(prov.any.poblacio[,4:21])
      p.prov.any <- sum(prov.any.poblacio$SUMA)
      prov.any$SUMA <- rowSums(prov.any[,7:24])
      n.prov.any <- round(mean(prov.any$SUMA),2)
      tb <- round(n.prov.any/p.prov.any*100000,2)
      ic.prov.any <- round(quantile(prov.any$SUMA, probs = c(0.025,0.975)),2)
      ic.prov.any.l <- as.vector(ic.prov.any[1])
      ic.prov.any.u <- as.vector(ic.prov.any[2])
      v <- c(v, n.prov.any, ic.prov.any.l, ic.prov.any.u, tb)
    }
    v <- as.numeric(v)
    df[i,] <- v
  }
}
```

```

colnames(df) <- c("PROVINCIA", "N.2009", "ICL.2009", "ICU.2009", "TB.2009", "N.2010", "ICL.2010", "ICU.2010",
"TB.2010", "N.2011", "ICL.2011", "ICU.2011", "TB.2011", "N.2012", "ICL.2012", "ICU.2012",
"TB.2012", "N.2013", "ICL.2013", "ICU.2013", "TB.2013")
print(df)
}

# Mapa
install.packages('maptools', repos = "http://cran.fhcrc.org/")

creacio.mapa <- function(icdo){
  suppressMessages(library(maptools))
  Map <- readShapePoly("/Volumes/ANDREA_USB/provincies_SRCBP/PROV.shp")

  df <- creacio.taula(icdo)

  if (icdo < 200) {
    sex <- "M"
  }
  if (icdo > 200) {
    sex <- "F"
  }
  # Provincies amb RCBP
  p1 <- read.csv(paste0("/Users/mac/Desktop/PFG/Resultats_0913_AMB_RCBP/Resultat-tumor-AmbRCBP-", icdo,
".txt"), header = TRUE, sep = ";")
  # Provincies sense RCBP
  p2 <- read.csv(paste0("/Users/mac/Desktop/PFG/Resultats_0913_SENSE_RCBP/Resultat-tumor-SenseRCBP-",
icdo, ".txt"), header = TRUE, sep = ";")
  # Provincies Tgn i Grn amb RCBP
  p3 <- read.csv(paste0("/Users/mac/Desktop/PFG/Resultats/Resultat-tumor-", icdo, ".txt"), header = TRUE,
sep = ";")
  escenari <- p1$ESCENARIO[1]
  if( escenari == "A"){
    esc <- "A"
  }else if(escenari=="B"){
    esc <- "B"
  }else{
    esc <- "C"
  }
  p3 <- subset(p3, p3$ESCENARIO==esc)
  data <- rbind(p1,p2,p3)
  data$PROVINCIA <- as.factor(data$PROVINCIA)
  provincies <- levels(data$PROVINCIA)
  provincies[1:9] <- c("01", "02", "03", "04", "05", "06", "07", "08", "09")
  anys <- 2009:2013
  for(j in 1:length(anys)){
    any <- anys[j]
    df4color <- as.data.frame(matrix(0, nrow=58, ncol = 2))
    df4color[,1] <- Map$PROV
    colnames(df4color) <- c("PROVINCIA", "COLOR")
    cols <- colors()
    for(i in 1:length(provincies)){
      prov <- provincies[i]
      provtb <- provincies[i]
      if(prov=="01"){
        provtb <- 1
      }else if(prov=="02"){
        provtb <- 2
      }else if(prov=="03"){
        provtb <- 3
      }else if(prov=="04"){
        provtb <- 4
      }else if(prov=="05"){
        provtb <- 5
      }else if(prov=="06"){
        provtb <- 6
      }else if(prov=="07"){
        provtb <- 7
      }else if(prov=="08"){
        provtb <- 8
      }else if(prov=="09"){
        provtb <- 9
      }
    }

    if(any==2013){

```

```

    col <- ncol(df)
  }else if(any==2012){
    col <- ncol(df)-4
  }else if(any==2011){
    col <- ncol(df)-8
  }else if(any==2010){
    col <- ncol(df)-12
  }else if(any==2009){
    col <- ncol(df)-16
  }

  tb <- df[df$PROVINCIA==provtb, col]

  if(tb <= 50){
    col <- cols[591]
  }else if(tb>50 && tb<70){
    col <- cols[562]
  }else if(tb>=70){
    col <- cols[490]
  }
  df4color[which(df4color$PROVINCIA==prov),2] <- col
}

colors <- df4color$COLOR
plot(Map , col=colors)
v <- colors()
cols <- c(v[591],v[562],v[490])
legend("bottomright",legend=c("<= 50", "50 - 70", ">=70"), fill = cols, cex=0.8)
title(paste0(any, " ", icdo, " ", sex))
}
}

```