

Màster en Estadística i Investigació Operativa

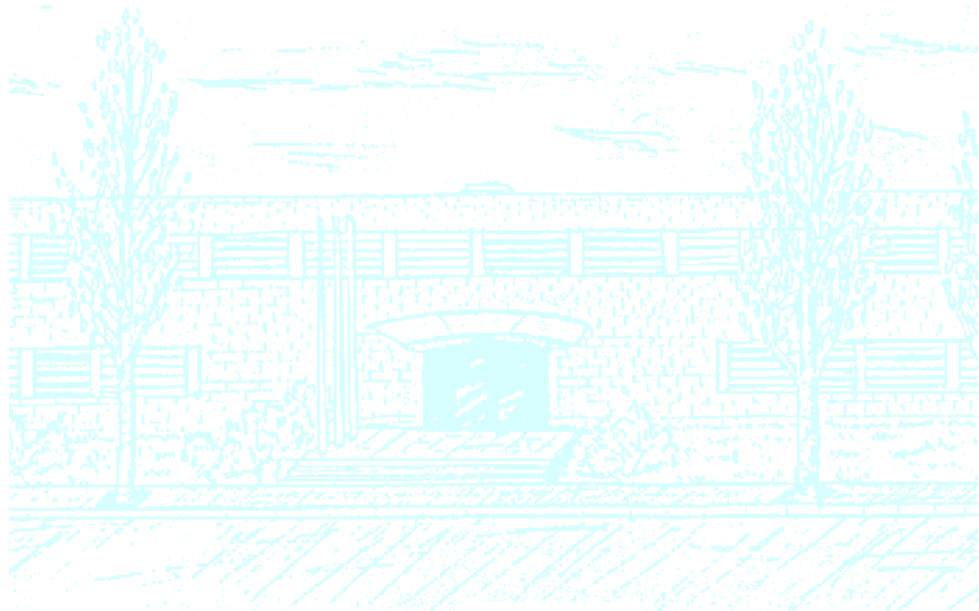
Títol: Eines estadístiques enfocades al control epidemiològic. Aplicació amb la mortalitat per càncer a Catalunya (1995-2004).

Autor: Laura Esteban Etchamendi


Directors: Ramón Clèries Soler i Josepa Ribes Puig

Departament: Estadística i Investigació Operativa

Convocatòria: Juny 2010



Facultat de Matemàtiques
i Estadística



EINES ESTADÍSTIQUES ENFOCADES AL CONTROL EPIDEMIOLÒGIC

Aplicació amb la mortalitat per càncer a Catalunya (1995-2004)

Estudiant: Laura Esteban Etchamendi

Directors: Ramón Clèries Soler i Josepa Ribes Puig

Ponent: Klaus Langohr

Treball Final de Màster

Màster en Estadística i Investigació Operativa

Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya

AGRAÏMENTS

Quan va sorgir la idea d'aquest treball, tot i que estàvem convençuts de la seva finalitat i utilitat, jo pensava que era impossible veure'l acabat. I bé, mai es pot donar per acabat un projecte, doncs sempre cal fer diferents aportacions per tal de millorar el que ja hi ha o bé es plantegen nous reptes a partir d'aquest, però, si que és cert que ja tenim uns resultats acords al que preteníem i amb els quals ens podem recolzar i esperem que siguin de gran utilitat, no només dins del nostre equip, sinó que també per a altres possibles usuaris.

Amb tot i això, aquest projecte no hagués estat possible sense l'ajuda de molta gent i és per això que vull agrair-los-hi l'esforç dedicat, ja sigui acadèmicament, laboralment o emocionalment.

En primer lloc, m'agradaria agrair a en Ramón Clèries i na Josepa Ribes del Registre de Càncer de Catalunya (RCC), Pla Director d'Oncologia (PDO), per ser els meus directors i haver-me guiat el treball durant aquest temps, per tots els consells que m'han donat i la col·laboració que han demostrat, des de un punt de vista laboral (i no tant laboral). Així mateix, també vull donar-li les gràcies a en Klaus Langohr, el meu tutor/ponent d'aquest projecte pel seu interès i optimisme i per la seva ajuda i col·laboració en el projecte.

D'altra banda, m'agradaria donar les gràcies a l'equip del Registre de Mortalitat de Catalunya per la cessió de les dades que han fet possible l'exemple d'aquesta aplicació.

Agrair a l'equip de treball del qual formo part, al PDO, amb especial, als membres del RCC per tots els recolzaments i idees que m'han aportat. Molt particularment, al Jordi Gálvez, per l'ajuda de la implementació de les funcions via web, a la Mery Nomen, pel mini curs personal que em va fer sobre el funcionament del Reference Manager ("Ojo Mery, encara no havia nascut en Joel!") i a la Laura Pareja, per veure des de altres punts de vista les meues idees tant laborals com emocionals, per l'ajuda que m'ha mostrat en aquest treball i per "burxar-me" tant per a que em posés les piles! També haig de fer un agraïment als meus companys de l'antic SERC de l'ICO, amb molts dels quals encara mantenim el contacte i en algun moment o altre em mantingut una "xarla" sobre el projecte.

Així mateix, donar les gràcies a l' Esther Gràcia, per moltes converses que em mantingut i per estar batallant "mano a mano" les dues amb el nostre projecte (Esther, haurem de programar el viatge que tenim pendent des de fa molt de temps sense cap destí concret!), a la Intza Aoiz, companya de pis amb la quina sempre he pogut expressar-me molt obertament sobre diferents temes (Sempre em diu: "Trabajo de tortuga, Laura, pasito a pasito siempre se llega al fin!"), a les ex companyes de pis, en particular a la Maria Lourdes Cos, pel seu recolzament i a la Judith Maldonado per fer-me veure la vida des de un punt de vista més optimista. També a diversos amics: Les Núries, els Jordis, la Montse, els Alberts, el Joan, la Diana, el Guillem, el Quique, la Encar, etc. per aguantar sempre els meus comentaris tipus "No puc que haig d'estudiar, no puc que haig de fer projecte" i per haver-me recolzat en tot moment.

I per finalitzar vull agrair a la meua família, i en particular als meus pares, el meu germà i els meus iaïos, el recolzament moral i la insistència en què fes el projecte, i sobretot, en el fet de que sempre han estat al meu costat.

LLISTA D'ABREVIATURES

AIC	Akaike Information Criterion (Criteri d'Informació d'Akaike)
ASR:	Age Standardized Rate (Taxa estandarditzada per edat)
CR:	Crude Rate (Taxa Crua)
CumR:	Cumulative Rate (Taxa Acumulada)
EAPC:	Estimated Annual Percent Change (Estimació del Percentatge de Canvi Anual)
GOF:	Goodness of fit (Bondat d'Ajust)
IDESCAT:	Institut d'Estadística de Catalunya
INE:	Institut Nacional d'Estadística
MLG:	Models Lineals Generalitzats
RCC:	Registre de Càncer de Catalunya
RMC:	Registre de Mortalitat de Catalunya
SART:	Statistical Analysis of Rates and Trends
SEER:	Surveillance Epidemiology and End Results
SIMR:	Standardized Incidence or Mortality Ratio (Raó d'Incidència o Mortalitat Estandarditzada)
SIR:	Standardized Incidence Ratio (Raó d'Incidència Estandarditzada)
SMR:	Standardized Mortality Ratio (Raó de Mortalitat Estandarditzada)
SR:	Specific Rate
TR:	Truncate Rate (Taxa Truncada)
WHO:	World Health Organization (Organització Mundial de la Salut)

ÍNDEX

1. INTRODUCCIÓ	1
1.1. OBJECTIUS DE LA VIGILÀNCIA	2
1.1.1. Descriptiva epidemiològica dels problemes de salut	2
1.1.2. Enllaços de la vigilància als serveis	3
1.1.3. Enllaços de la vigilància a la recerca	3
1.1.4. Avaluació de les intervencions	4
1.1.5. Planificació i projeccions	4
1.1.6. Educació i política	4
1.2. EXEMPLE DE MESURA DE VIGILÀNCIA	4
1.2.1. Registre de càncer	5
1.2.2. Altres registres i necessitat d'automatització d'eines estadístiques	6
1.3. EPIDEMIOLOGIA DEL CÀNCER	7
1.3.1. Antecedents històrics	7
1.3.2. Incidència i mortalitat per càncer	8
1.4. ESTRUCTURA DEL PROJECTE	10
2. JUSTIFICACIÓ	12
3. OBJECTIUS	14
3.1. APLICACIÓ DESCRIPTIVA	14
3.2. APLICACIÓ ESTIMACIÓ DEL PERCENTATGE DEL CANVI ANUAL	14
3.3. APLICACIÓ ESPERATS	14
3.4. APLICACIÓ DE RAÓ DE INCIDÈNCIA O MORTALITAT ESTANDARDITZADA	14
4. MATERIALS I MÈTODES ESTADÍSTICS	15
4.1. FONTS D'INFORMACIÓ I PROGRAMARI	15
4.1.1. Fonts d'informació	15
4.1.2. Programari	17
4.2. DESCRIPCIÓ DELS MÈTODES ESTADÍSTICS PER A CADA APLICACIÓ	18
4.3. EL CÀLCUL DE TAXES	19
4.3.1. Taxes brutes (Taxes crues)	19
4.3.2. Estandardització de taxes (Taxes ajustades)	21
4.3.3. Taxes truncades	24
4.3.4. Taxes acumulades	25

4.4.	MODELS LINEALS GENERALITZATS.....	26
4.4.1.	Distribucions de la família exponencial i funció de versemblança	26
4.4.2.	Funció de log-versemblança i les seves propietats	27
4.4.3.	Exemples de distribucions de la família exponencial	28
4.4.4.	Funcions link pels Models Lineals Generalitzats	31
4.4.5.	Bondat d'ajust.....	32
4.4.6.	Criteri d'Informació d'Akaike.....	34
4.4.7.	Models log-lineals.....	34
4.4.8.	Distribució binomial negativa.....	36
4.4.9.	Selecció de models log-lineals	36
4.4.10.	Estimació del percentatge de canvi anual i els seus intervals de confiança.....	37
4.5.	CÀLCUL DE LA RAO D'INCIDÈNCIA O MORTALITAT ESTANDARDITZADA	39
5.	APLICACIÓ PEL CàLCUL AUTOMATITZAT VIA WEB.....	41
5.1.	APLICACIÓ DESCRIPTIVA.....	42
5.2.	APLICACIÓ DE L'ESTIMACIÓ DEL PERCENTATGE DE CANVI ANUAL.....	47
5.3.	APLICACIÓ ESPERATS.....	49
5.4.	APLICACIÓ DE LA RAÓ D'INCIDÈNCIA O MORTALITAT ESTANDARDITZADA	51
6.	RESULTATS.....	54
6.1.	APLICACIÓ DESCRIPTIVA.....	54
6.1.1.	Exemple	54
6.1.2.	Altres utilitats de l'aplicació.....	59
6.2.	APLICACIÓ DE L'ESTIMACIÓ DEL PERCENTATGE DE CANVI ANUAL.....	60
6.2.1.	Exemple	60
6.2.2.	Altres utilitats de l'aplicació.....	64
6.3.	APLICACIÓ ESPERATS.....	64
6.3.1.	Exemple	65
6.3.2.	Altres utilitats de l'aplicació.....	70
6.4.	APLICACIÓ DE LA RAÓ D'INCIDÈNCIA O MORTALITAT ESTANDARDITZADA	70
6.4.1.	Exemple 1: Evolució temporal del risc de morir per càncer a Catalunya.....	70
6.4.2.	Exemple 2: Comparació de la mortalitat per càncer a Catalunya amb la d'Espanya en el mateix període.....	74
6.4.3.	Altres utilitats de l'aplicació.....	78
7.	DISCUSSIÓ.....	79
8.	CONCLUSIONS.....	84
9.	BIBLIOGRAFIA	85
10.	ANNEXES.....	90
	ANNEX 1.....	90
	ANNEX 2.....	109

1. INTRODUCCIÓ

Els professionals que dirigeixen programes per a la prevenció i control de malalties en una àrea geogràfica determinada necessiten informació sobre indicadors de salut poblacionals. El procediment que s'utilitza per a recol·lectar, administrar, analitzar, interpretar i informar sobre una malaltia o un esdeveniment relacionat amb la salut s'anomena *vigilància* ó *vigilància en la salut pública* ja que aquests sistemes funcionen majoritàriament a través d'agències de salut públiques [1;2]. Aquests poden proporcionar les bases per reduir la morbiditat i la mortalitat i millorar la salut de la població [2].

Les dades difoses per un sistema de vigilància de la salut pública es poden utilitzar immediatament per tal de prendre mesures de salut pública, planificació, avaluació de programes i per formular hipòtesis de recerca. Més específicament, aquests sistemes s'han desenvolupat per fer front a una àmplia gamma de necessitats de salut pública [2], per exemple:

- a) Com a guia d'acció immediata per a casos de salut pública important.
- b) Com a mesura de càrrega d'una malaltia.
- c) Com a monitorització de les tendències de la càrrega d'una malaltia.
- d) Com a guia de planificació, implementació i avaluació de programes per prevenir i controlar malalties, lesions o exposicions adverses.
- e) Per avaluar les polítiques en aquest àmbit.
- f) Per a detectar canvis en la pràctica de la salut i els efectes d'aquests canvis.
- g) Per a prioritzar l'assignació de recursos per a la salut.
- h) Per a descriure el curs clínic d'una malaltia.
- i) I per a proporcionar una base per a la investigació epidemiològica.

Quan apareix un problema nou de salut pública, la implementació ràpida de la vigilància és necessària per tal de que hi hagi una reacció efectiva i primerenca. Amb el temps, els sistemes de vigilància s'utilitzen per a identificar canvis en la naturalesa o per abastir els problemes de salut i valorar l'eficàcia de les intervencions en salut pública [1].

Les poblacions sota vigilància es defineixen segons les necessitats d'informació de programes de prevenció o control. Per exemple, com una part d'un programa hospitalari de vigilància i prevenció d'infeccions hospitalàries, la població objectiu seria els pacients que reben atenció en aquest hospital. Per als organismes de salut pública, la població

sota vigilància, generalment, representa els residents dins de la seva jurisdicció política, que pot ser una ciutat, regió o nació [1].

En comparació amb les dades dels estudis d'investigació dirigits, l'avantatge de les dades de vigilància és sovint la seva puntualitat i la seva amplitud en el temps, cobertura geogràfica, o el nombre de persones representades. Per ser eficaç, la vigilància ha de ser el més simplificat possible. Les dades de vigilància poden ser menys detallades o precises que les obtingudes amb els estudis d'investigació. Hi ha diversos mètodes que s'utilitzen per l'obtenció d'informació de vigilància, des de la tradicional notificació de casos fins a l'adaptació de les dades recollides fonamentalment per a altres finalitats, com per exemple els registres d'atenció mèdica informatitzats [1].

1.1. OBJECTIUS DE LA VIGILÀNCIA

El principal objectiu de la vigilància epidemiològica consisteix en controlar la incidència o la prevalença de problemes de salut específica, per tal de documentar els seus efectes en poblacions definides i per a caracteritzar les persones afectades i les persones amb major risc. A nivell de la comunitat, la vigilància pot guiar els departaments de salut en la prestació de serveis a les persones, en general, les dades de vigilància es poden utilitzar per informar i avaluar els programes de salut pública. Les tendències detectades a través de la vigilància es poden utilitzar per a preveure les futures tendències, i així poder ajudar en la planificació de la salut. A més de proporcionar informació bàsica sobre els problemes de salut, la vigilància pot conduir a una hipòtesi o a la identificació de participants per realitzar investigacions epidemiològiques més detallades. Per ser eficaços, les dades de vigilància hauran d'estar degudament comunicades a tots els constituents de l'àmbit que les poden utilitzar [1].

1.1.1. DESCRIPTIVA EPIDEMIOLÒGICA DELS PROBLEMES DE SALUT

Un dels objectius principals en la majoria de sistemes de vigilància epidemiològica és la monitorització de les tendències de les taxes d'ocurrència d'una malaltia. Amb les tendències es pot detectar un augment d'esdeveniments de salut adversos que poden alertar als organismes de salut en quant a la necessitat de realitzar més recerca. La vigilància pot ajudar a l'avaluació de la importància dels canvis d'incidència documentats o percebuts quan es sospita d'un brot o clúster d'una malaltia. Alternativament, la identificació d'un decrement de les tendències poden proporcionar una indicació de l'èxit de les intervencions realitzades, tot i que podrà ser necessari realitzar estudis més detallats per avaluar formalment els programes utilitzats [1].

La vigilància també pot orientar les activitats de prevenció abans que es defineixi l'etiologia d'una malaltia. La informació sobre les característiques comuns de les persones amb problemes de salut permet que s'identifiquin els grups de major risc de la malaltia, mentre que la informació sobre exposicions o comportaments dóna una idea de etiologies o maneres de propagació [1].

Atès que les dades es poden recollir segons les característiques de la pròpia malaltia (per exemple: durada, gravetat, mètode de diagnòstic, tractament o evolució, entre d'altres), això permet mesurar l'efecte de la malaltia i la identificació de grups en els que aquesta pot ser més greu [1].

Descrivint on hi ha la majoria de casos d'una malaltia o on es troben les majors taxes, ens permet orientar a les intervencions en salut pública. Per exemple, la representació de les dades de vigilància mitjançant mapes pot il·lustrar clústers geogràfics destacant diferències regionals en prevalença o incidència i generant o recolzant hipòtesis pel que fa l'etiologia [1].

1.1.2. ENLLAÇOS DE LA VIGILÀNCIA ALS SERVEIS

La vigilància és una part integral de la prestació de serveis preventius i terapèutics dels departaments de salut a nivell de la comunitat. A nivell local, no només proporciona dades agregades per als planificadors de la salut, sinó que també serveix per iniciar les accions preventives i terapèutiques [1].

1.1.3. ENLLAÇOS DE LA VIGILÀNCIA A LA RECERCA

Tot i que les dades de vigilància poden ser útils en la caracterització dels problemes de salut epidemiològica bàsica, rarament proporcionen detall suficient com per a provar hipòtesis epidemiològiques més a fons. Aquestes dades ens permeten realitzar comparacions entre diferents grups, que poden ser definits, entre altres, per edat, sexe, data d'incidència i localització geogràfica. No proporcionen, però, un grup de comparació de les persones sense el problema de salut que s'estudia. No obstant, la vigilància facilita un pont important als investigadors subministrant claus per tal de realitzar una investigació més a fons [1].

1.1.4. AVALUACIÓ DE LES INTERVENCIIONS

Els encarregats de planificar la salut necessiten saber si les intervencions són efectives i la realització d'una avaluació de les intervencions a escala completa no sempre és viable. La vigilància, mitjançant la realització d'un gràfic de tendències pel nombre de casos o taxes d'esdeveniments, proporciona una avaluació relativament ràpida i suficient de l'efecte dels esforços de la intervenció. En alguns casos, els canvis de tendències de la malaltia i les intervencions que estan tan lligades fan que la vigilància, per si sola, pugui facilitar documentació simple i convincent dels efectes que ha tingut la intervenció [1].

1.1.5. PLANIFICACIÓ I PROJECCIONS

La planificació és necessària per anticipar-se a les futures demandes dels serveis de salut. Les tendències observades en la incidència d'una malaltia, juntament amb alguna altra informació sobre la població a risc o la història natural de la malaltia es pot utilitzar per tal d'anticipar l'efecte d'una malaltia o l'atenció necessària per a aquesta. Les tècniques per a predir les tendències d'una malaltia utilitzant les dades de vigilància, poden anar des de l'aplicació de models epidemiològics complexos fins a estratègies relativament simples, tals com l'aplicació de les taxes actuals de la malaltia a futures estimacions de la població per tal d'estimar un futur impacte d'aquesta malaltia en la població [1].

1.1.6. EDUCACIÓ I POLÍTICA

El valor educatiu de la vigilància s'estén des de l'alerta sobre problemes de salut de la comunitat als metges, fins a informar als polítics sobre la necessitat de prevenció o de recursos per l'atenció. La vigilància i altres evidències epidemiològiques o científiques proporcionen una perspectiva essencial en la configuració de la política de la salut pública i ha de ser integrat de manera efectiva amb altres perspectives que sovint porten a la presa de decisions polítiques [1].

1.2. EXEMPLE DE MESURA DE VIGILÀNCIA

La principal mesura de vigilància és el registre de malalties. Aquest és un sistema per a recopilar i mantenir, en un expedient estructurat, informació sobre persones que tenen una malaltia comú o una condició adversa a la salut [3].

A continuació es presenta en què consisteix un registre de càncer com a exemple de mesura de vigilància.

1.2.1. REGISTRE DE CÀNCER

El registre de càncer té un paper crucial en el control del càncer. La seva principal funció és el manteniment d'un arxiu o registre de tots els casos de càncer que es donen en una població definida i en un període de temps determinat. En el registre es documenten de forma contínua i sistemàtica les dades personals dels pacients amb càncer, característiques clíniques i patològiques sobre els càncers, tot extret de diferents fonts d'informació [4].

El registre analitza i interpreta les dades periòdicament i proporciona informació de la incidència i les característiques específiques dels càncers en diversos segments de la població resident i diferents variacions temporals de la incidència. Aquesta informació és la primera font, no tan sols per a la investigació epidemiològica per a determinants del càncer, sinó que també per a la planificació i avaluació dels serveis de salut per a la prevenció, diagnòstic i tractament del càncer [4].

Els registres de càncer també es poden utilitzar per a monitoritzar grups professionals i cohorts d'individus exposats a diferents agents carcinògens i com a font de subjectes per a estudis clínics i epidemiològics. Els registres basats en hospitals tenen un paper important en la cura dels pacients amb càncer, ajudant als metges en el seguiment dels seus casos i en el subministrament de les dades estadístiques sobre els resultats de la teràpia. La investigació epidemiològica, basada en el registre global de càncer, segueix sent la més vàlida i eficaç per a planificar i avaluar tots els aspectes de la lluita contra el càncer [4].

La majoria dels registres de càncer poblacionals (que abasteixen tota una població) que actualment estan en funcionament, i dels quals les seves dades es publiquen en la "International Agency for Research on Cancer" (IARC) pertanyen a Europa i Amèrica del Nord. Existeix una necessitat urgent de creació de més registres de càncer en els països en vies de desenvolupament d'Àsia, Àfrica com d'Amèrica del Sud, on el càncer està reconegut com un problema de salut important [4]. En el darrer volum (Vol IX) del Cancer Incidence in 5 Continents [5], tenim que hi participen 11 registres espanyols: Albacete, Astúries, País Basc, les Illes Canàries, Conca, Girona, Granada, Múrcia, Navarra, Tarragona i Saragossa.

1.2.2. ALTRES REGISTRES I NECESSITAT D'AUTOMATITZACIÓ D'EINES ESTADÍSTIQUES

Si s'observen els sistemes de vigilància en salut pública, pràcticament tots tenen objectius molt similars, on moltes vegades, tant sols varia la definició del cas del registre. A tall d'exemple podem tenir:

- El registre de mortalitat: registra les morts que es donen en una àrea concreta, la causa de la mort (que inclou, a més de malalties, accidents) i l'àrea sanitària de residència del difunt entre altres variables.
- El registre del Síndrome de la Immunodeficiència Adquirida (SIDA): registra casos de la SIDA diagnosticats en un àrea concreta, segons l'edat, el sexe, l'any de diagnòstic, el mecanisme de transmissió i altres variables.
- El registre de malalts renals: recull informació de tots els malalts que reben tractament substitutiu renal en un àrea concreta.

Tot i que variï la definició de cas, moltes vegades es vol arribar a uns resultats mitjançant la mateixa eina estadística. És a dir, per tal de realitzar una bona descriptiva epidemiològica es voldrà obtenir uns indicadors descriptius com són el nombre de casos, el percentatge, la taxa crua, la taxa ajustada, i en ocasions també pot interessar la taxa trucada o l'acumulada. Aquests indicadors permeten realitzar comparacions entre diferents períodes o entre diferents àrees i la detecció de l'existència de canvis de tendència.

Així mateix, moltes vegades també interessa realitzar estimacions en un àrea concreta (per exemple estimar la incidència del càncer a Catalunya mitjançant els dos registres poblacionals dels que es disposa) o bé prediccions futures per a intentar prevenir o valorar la situació de salut pública que es pot donar.

Dins del Pla Director d'Oncologia (PDO) de Catalunya del Departament de Salut de la Generalitat de Catalunya s'ubica el Registre de Càncer de Catalunya (RCC), que es troba en fase inicial [6], i on es planteja l'automatització d'algunes eines estadístiques per tal d'agilitzar els resultats d'indicadors descriptius, tendències, estimacions i projeccions. Aquest projecte es planteja per la necessitat d'actualització constant de les dades ja que el càncer és una patologia d'interès cada vegada més creixent donada la seva incidència en augment. Així mateix, a tall d'exemple, en aquest treball s'apliquen aquestes eines sobre dades de mortalitat per càncer a Catalunya en el període 1995-2004.

1.3.EPIDEMIOLOGIA DEL CÀNCER

1.3.1. ANTECEDENTS HISTÒRICS

En termes generals, el càncer es defineix com la reproducció i creixement descontrolat d'una cèl·lula i la tendència a la invasió de teixits propers formant metàstasis (disseminació) en altres regions de l'organisme.

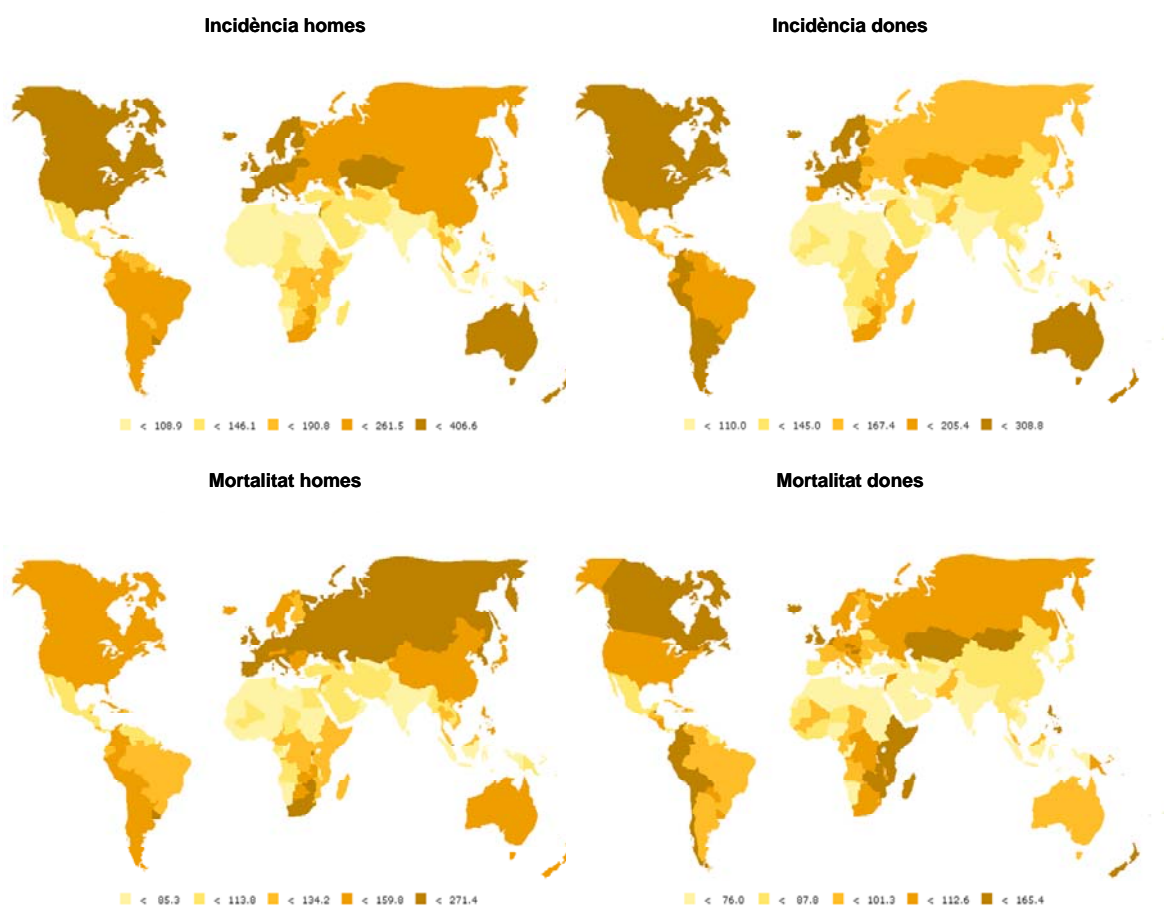
La descripció més antiga del càncer en els humans es va trobar en un paper egipci escrit entre 3000-1500 aC. referint-se als tumors de mama. Aproximadament el 400 aC. a Hipòcrates (pare de la medicina) se li atribueix ser el primer en reconèixer la diferència entre tumors malignes (càncer) i benignes. En el segle XVIII a Reims (França) es funda el primer hospital de càncer però amb la creença errònia de que el càncer es una malaltia contagiosa. En l'any 1839, el ginecòleg Recamier va descriure la invasió del torrent sanguini per cèl·lules del càncer, el que es va anomenar metàstasis. Va ser en 1895 quan Roentgen va descobrir els rajos X els quals encara s'utilitzen actualment tant pel diagnòstic del càncer com pel tractament (radioteràpia). Un important avenç va ésser quan en l'any 1953, Francis Crick i James Watson van desxifrar l'estructura de l'ADN (Àcid desoxiribonucleic), que va permetre iniciar estudis i comprendre les causes del càncer a nivell molecular i l'elaboració de nous tractaments basats en aquest coneixement. En els últims 50 anys s'estan produint molts avenços en quant a la comprensió d'aquesta malaltia [7].

Sota el nom de càncer englobem tot un conjunt de més de cent malalties diferents, la qual cosa dóna una idea de la complexitat d'aquest problema. El càncer ha estat present a la humanitat des de l'antiguitat, i és difícil imaginar que pugui desaparèixer. Tot i això, hi ha nombrosos arguments que demostren que és possible lluitar contra el càncer de forma eficaç. De fet, les importants variacions de la incidència i el nombre de casos nous per any d'aquesta malaltia que s'observen entre els països són indicatius de la importància que poden tenir els factors de risc relacionats amb la manera de viure de cada individu, amb el seu medi ambient i amb l'herència com a causes de càncer. L'impacte del càncer a la població es pot mesurar aplicant diferents indicadors. Les principals mesures són la incidència, la mortalitat i la supervivència [7].

1.3.2. INCIDÈNCIA I MORTALITAT PER CÀNCER

El càncer és una patologia d'interès cada vegada més creixent, degut a la seva freqüència i una de les causes amb major mortalitat. En la Figura 1.1, podem apreciar la incidència i la mortalitat per càncer a nivell geogràfic en l'any 2002, destacant la gran freqüència d'aquesta patologia en els països desenvolupats.

Figura 1.1. Taxa estandarditzada per 100000 persones-any per a totes les localitzacions del càncer excepte pell no melanoma en el mon. [8]



Els països amb major taxa estandarditzada d'incidència del càncer en el 2002, diferenciant per sexe, van ser: Estats Units, Hongria i Nova Zelanda en els homes i Estats Units, Israel i Nova Zelanda en les dones. Mentre que els corresponents amb la menor incidència varen ser: Congo Brazzaville, Gàmbia i Níger en els homes i Oman, Tunísia i Gàmbia en les dones [8].

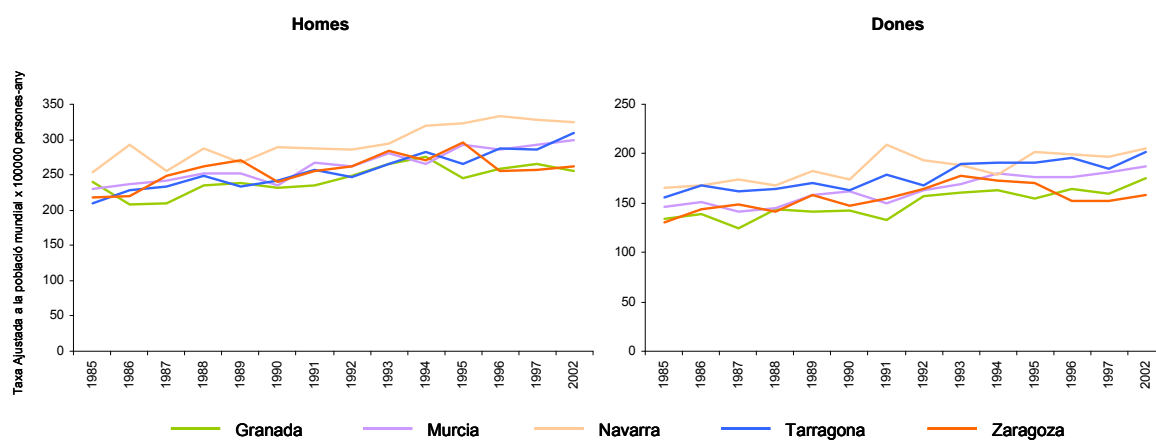
Els països amb major taxa estandarditzada de mortalitat en el mateix any varen ser: Hongria, Eslovàquia i Kazakhstan en els homes i Zimbabwe, Dinamarca y Ruanda en les

dones. Les taxes més baixes es van correspondre amb els països: Congo Brazzaville, Vanuatu i Fiji en els homes i Oman, Tadjikistan i Turquia en les Dones [8].

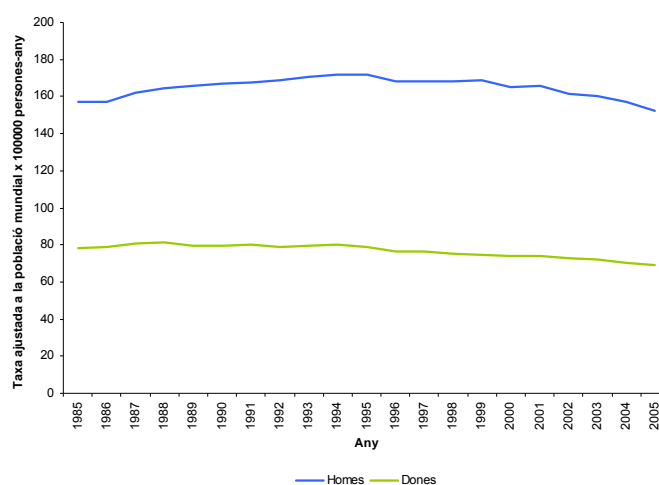
El càncer és la principal causa de mortalitat a nivell mundial, se li atribueixen 7.9 milions de defuncions en el 2007 (aproximadament un 13% del total). La major part de la mortalitat anual per càncer obeeix al càncer de pulmó, estómac, fetge, colon i mama. Les freqüències dels diferents tipus de càncer varien segons el sexe, així doncs, en els homes tenim que els tipus de càncer més freqüents són pulmó, estómac, fetge, colon i recte, esòfag i pròstata mentre que en les dones són mama, pulmó, estómac, colon i recte i coll uterí [9].

En el 2007, a Catalunya, el 27.5% de les morts van estar causades per tumors, essent aquesta, la segona causa de mortalitat després de les malalties del sistema circulatori [10]. Segons les dades dels registres de càncer espanyols la incidència del càncer en el període 1985-2002 va incrementar tant en homes com en dones (Figura 1.2), tenint en compte, però, la diferència de la magnitud [5;11].

Figura 1.2. Tendència de la Incidència per càncer excepte pell no melanoma a Espanya per a registres poblacionals. (1985-2002) [5;11]



En quant a la mortalitat per càncer en el període 1985-2005 a Espanya, en els homes s'observa un creixement fins l'any 1995, a partir del qual comença a decreixer i en les dones al principi es manté constant i és a partir del 1995 que també comença a decreixer (Figura 1.3). També cal remarcar la diferència de la magnitud de la mortalitat entre homes i dones.

Figura 1.3. Tendència de la Mortalitat per càncer a Espanya. (1985-2005) [12]

1.4. ESTRUCTURA DEL PROJECTE

Aquest treball presenta una aplicació web que permet a l'usuari carregar una base de dades a partir de la qual s'hi podran calcular una sèrie d'indicadors epidemiològics poblacionals. Aquesta aplicació rep el nom de SART (*Statistical Analysis of Rates and Trends*) i es troba disponible a <http://rht.iconcologia.net/stats/sart.html>.

Cal esmentar que aquest projecte s'ha desenvolupat en el marc del Grup d'Investigació Emmergent. Registre de Càncer de Catalunya. SGR 2009-2013 de la AGAUR (Agència de Gestió d'Ajuts Universitaris i de Recerca).

El document que presentem està estructurat en 10 capítols, el contingut dels quals es descriu a continuació.

El primer capítol es correspon amb la **introducció** on s'ha especificat quins són els objectius de la vigilància epidemiològica així com diferents exemples de mesura de vigilància. En aquest apartat també s'ha introduït breument en què consisteix l'epidemiologia del càncer i finalment s'especifica el contingut del treball.

El segon capítol, **justifica** la necessitat de crear una eina estadística via web que ens permeti extreure indicadors epidemiològics poblacionals. A més a més, descriu altres eines existents i es justifica la importància de la creació d'aquesta nova aplicació.

El tercer capítol conté quin és l'**objectiu** principal d'aquest projecte i descriu les quatre aplicacions via web que es desenvoluparan: descriptiva, estimació del percentatge de

canvi anual de la incidència i/o mortalitat, el càlcul de casos esperats i, la raó d'incidència o mortalitat estandarditzada d'una malaltia determinada (en el nostre cas, càncer).

En el quart capítol, **material i mètodes estadístics**, s'especifica quines són les dades utilitzades per tal d'exemplificar els resultats que pot retornar l'aplicació via web, així com el programari emprat en el present projecte. D'altra banda, es descriuen els mètodes estadístics utilitzats per a dur a terme les quatre aplicacions descrites: càlcul de taxes, els models lineals generalitzats i el càlcul de la raó d'incidència o mortalitat estandarditzada.

El cinqué capítol es correspon a l'**aplicació pel càlcul automatitzat via web**, en el qual es descriu cadascuna de les diferents aplicacions i s'especifiquen com han de ser els fitxers d'entrada i quins paràmetres es necessiten per tal de que es puguin executar correctament cadascuna d'elles.

En el capítol sisé, es mostren els **resultats** a partir de dades de la mortalitat per càncer a Catalunya. En aquest apartat, es descriu detalladament els fitxers d'entrada, la petició de dades, els resultats i la interpretació per a cadascuna de les aplicacions via web.

La **discussió**, corresponent al seté capítol, compara l'aplicació SART amb altres eines estadístiques disponibles actualment, ressaltant les avantatges i les limitacions d'aquesta aplicació estadística per al càlcul via web presentada en aquest projecte.

En el capítol vuitè, s'enuncien les **conclusions** més rellevants que es desprenen d'aquest projecte.

El capítol novè conté la **Bibliografia** utilitzada i el capítol desè els **Annexes**, on es troba el codi per a cadascuna de les funcions programades en R i altres anàlisis estadístics no inclosos en el projecte.

2.JUSTIFICACIÓ

El càlcul d'indicadors poblacionals de salut i els models estadístics que s'empren s'actualitzen a mesura que va passant el temps. Per tant, donada la rapidesa de nous mètodes estadístics que sorgeixen i, de vegades, la dificultat en aplicar-los als sistemes de vigilància epidemiològica, fa que no sempre hi hagi una homogeneïtzació en el seu ús ni en la seva implementació. Així doncs, es fa necessària l'automatització d'aquests processos, facilitant la seva implementació i la seva projecció a altres àmbits diferents al càncer.

L'automatització dels processos de càlcul permet, a més d'estalviar temps computacional als professionals, assegurar una homogeneïtzació en l'ús d'eines estadístiques i per tant, facilitar la comparació d'indicadors de salut.

D'eines estadístiques per a calcular mesures d'associació n'hi ha diverses, entre les quals es troben EPIZOO [13], EpiInfo [14;15] i EpiData [14;16]. En canvi, n'hi ha poques que calculin indicadors poblacionals que s'adeqüin als nostres objectius. Es pot trobar algun programari que permeti calcular taxes d'una malaltia determinada, però no permeten analitzar-les segons les variables d'interès en un període de temps i nivell de confiança determinats.

Entre aquests programaris es troba Orius [17], desenvolupat per "Health Canada" [18] i per la "North American Association of Central Cancer Registries" (NAACCR) [19]. A nivell predefinit sobre una àrea concreta, com per exemple el càncer, si que es troben eines disponibles, aquest és el cas de *Surveillance Epidemiology and End Results* (SEER) que distribueixen lliurement el software SEER*Stat 6.5.2 [20]. D'altra banda es requereix de la instal·lació del programa per tal de poder realitzar els anàlisis, a més d'estar limitat en l'àmbit del càncer.

Per tal que siguin eines estadístiques accessibles per qualsevol usuari, convé que l'aplicació sigui fàcil d'utilitzar i estigui disponible al públic. La necessitat d'anàlisi d'indicadors poblacionals genera la demanda d'un programari web que permeti dur a terme aquests càlculs de forma remota, sense la necessitat d'instal·lar el programa en el nostre ordinador. En aquest sentit l'equip del RCC va desenvolupar l'any 2005 l'aplicació de càlcul via web de la supervivència relativa anomenada WAERS (*Web Assisted Estimation of Relative Survival*) [21-23] per tal de facilitar als investigadors la possibilitat

de dur a terme aquest càlcul sense la necessitat d'instal·lar en el seu ordinador software ja existents tal com CANSURV [24] , SURVSoft [25] o SURV3 [26].

Per a mostrar la utilitat de les eines desenvolupades en el present projecte s'analitzen les dades de mortalitat per càncer a Catalunya.

3. OBJECTIUS

L'objectiu principal d'aquest projecte és desenvolupar una aplicació web de fàcil accés que permeti dur a terme el càlcul d'indicadors poblacionals d'una malaltia o esdeveniment determinat a qualsevol professional de la salut o de l'àmbit social.

L'aplicació constarà de quatre mòduls:

3.1. APLICACIÓ DESCRIPTIVA

Calcular el nombre de casos, la taxa crua, la taxa ajustada pel mètode directe, la taxa truncada i la taxa acumulada de la incidència i mortalitat per a una o més variables determinades (localització tumoral, zona geogràfica, ètnia, sexe, grups d'edat, entre d'altres).

3.2. APLICACIÓ ESTIMACIÓ DEL PERCENTATGE DEL CANVI ANUAL

Calcular el percentatge de canvi anual (i el corresponent interval de confiança) de les taxes d'incidència / mortalitat d'una malaltia determinada en un període predeterminat.

3.3. APLICACIÓ ESPERATS

Per a un període i una àrea geogràfica determinada, realitzar una predicció del nombre de casos esperats d'una malaltia per a cada variable d'interés, tenint en compte la tendència temporal d'aquest conjunt en un període precedent. Per a fer la predicció es tria un model estadístic entre 4 de proposats.

3.4. APLICACIÓ DE RAÓ DE INCIDÈNCIA O MORTALITAT ESTANDARDITZADA

Calcular la raó de incidència o mortalitat estandarditzada i el seu interval de confiança (amb el nivell de confiança que es desitgi) per a cadascuna de les variables sol·licitades.

4. MATERIALS I MÈTODES ESTADÍSTICS

4.1. FONTS D'INFORMACIÓ I PROGRAMARI

4.1.1. FONTS D'INFORMACIÓ

Les dades utilitzades en aquesta anàlisi s'extreuen de diferents fonts d'informació: el Registre de Mortalitat de Catalunya (RMC) [27], l'Institut Nacional d'Estadística (INE) [28], l'Institut d'Estadística de Catalunya (IDESCAT) [29] i la *World Health Organization* (WHO) [12].

Per tal de dur a terme les anàlisis, les dades de mortalitat individualitzades provinents del RMC s'agreguen a nivell d'unitat territorial, en aquest cas Catalunya, i per grups d'edat quinquennals (en total 18 grups d'edat, veure en la Taula 4.1). A més, només es seleccionen les causes de mort corresponents a les malalties neoplàsiques del període 1995-2004. Atès que la mortalitat a Catalunya es va codificar fins l'any 1998 amb la Classificació Internacional de Malalties (CIM) en la versió 9 [30] i a partir d'aquest any, amb la CIM versió 10 [30;31], les localitzacions tumorals s'han codificat diferent segons el període de temps, essent necessari realitzar una conversió prèvia de les dades anteriors a 1998 a la CIM versió 10. La Taula 4.2 mostra com s'ha agrupat les diferents localitzacions tumorals i a quins codis de la CIM corresponen.

Taula 4.1. Grups d'edat utilitzats en l'anàlisi.

Grup d'edat	Anys corresponents
1	0-4
2	5-9
3	10-14
4	15-19
5	20-24
6	25-29
7	30-34
8	35-39
9	40-44
10	45-49
11	50-54
12	55-59
13	60-64
14	65-69
15	70-74
16	75-79
17	80-84
18	>85

Taula 4.2. Conversió de la versió 9 a la 10 de la Classificació Internacional de Malalties (CIM)

Localització Tumoral	CIM10	CIM9
Cavitat Oral i Faringe	C00-C14	140-149
Esòfag	C15	150
Estómac	C16	151
Colorrectal	C18-C21	153-154
Còlon	C18	153
Recte	C19-C21	154
Fetge	C22	155
Pàncreas	C25	157
Laringe	C32	161
Pulmó	C33-C34	162
Melanoma de pell	C43	172
Mama (dones)	C50	174
Coll uterí	C53	180
Endometri	C54	182
Úter	C53-C55	179, 180, 182
Ovari	C56-C57	183-184
Pròstata	C61	185
Testicle	C62	186
Ronyó	C64-C66, C68	189
Bufeta de l'orina	C67	188
Sistema Nerviós Central	C70-C72	191-192
Tiroides	C73	193
Limfoma de Hodgkin	C81	201
Limfoma no Hodgkin	C82-C85, C96	200, 202
Malaltia immunoproliferativa o Mieloma múltiple	C88, C90	203, 238.6, 273.3
Leucèmies	C91-C95	204-208
Totes les localitzacions	C00-C97	140-208
Totes les localitzacions excepte pell no Melanoma	C00-C97 menys C44	140-208 menys 173

Les dades demogràfiques de Catalunya partir de l'any 1995 fins al 2004 s'extreuen de l'IDESCAT [29] i pertanyen a les estimacions de la població per grups d'edat a 1 de Juliol dels anys compresos en aquest període de temps. Per a aquest treball s'han emprat els 18 grups d'edat descrits a la Taula 4.1.

Per a l'anàlisi de les projeccions de la mortalitat per càncer, són necessàries les dades de la població catalana del període 2010-2015. Atès que aquesta és una estimació futura de la distribució per grups d'edat de la població catalana i està subjecta a l'evolució dels naixements, defuncions i migracions, s'ha de triar un possible escenari per a aquesta distribució. En aquest treball s'ha seleccionat l'escenari mitjà-alt que assumeix una combinació de fecunditat mitjana, una alta esperança de vida i lleugers canvis migratoris [29]. Aquest escenari pretén reflectir l'evolució demogràfica catalana d'acord amb les dades recents i és l'escenari recomanat per l'IDESCAT.

Per tal de valorar si una població presenta més risc d'una malaltia respecte a una població de referència es pot emprar la Raó d'Incidència o Mortalitat Estandarditzada

(*Standardized Incidence or Mortality Ratio*, SIR o SMR respectivament). Amb el SIR o l'SMR es comparen el nombre de casos observats en una població respecte a uns casos esperats que es calculen en base a una població de referència. En aquest treball es presenten dos exemples del càlcul de la SMR (mortalitat). En el primer exemple, es vol avaluar si existeix més (o menys) risc de morir per càncer a Catalunya en el període 2000-2004 respecte al període 1995-1999. En aquest cas, la població de referència serà la del període 1995-1999. En el segon exemple es vol determinar si hi ha diferències en quan al risc de morir per càncer entre dues àrees geogràfiques, Catalunya i Espanya, durant el període 2000-2004. En aquest darrer cas, la població de referència serà la espanyola en el període 2000-2004. Atès que per al càlcul de les SMR es necessiten les taxes de mortalitat per càncer en les poblacions de referència, la població catalana durant el període 1995-1999 i la espanyola durant el període 2000-2004 són requerides en el primer i segon exemple, respectivament. D'altra banda, en ambdós exemples s'ha de disposar del nombre de casos morts per càncer observats i la població a risc de Catalunya durant el període 2000-2004. Les taxes de mortalitat de la població espanyola s'han extret del WHO (casos observats de mortalitat) [12] i de l'INE (població) [28], mentre que les dades referents a Catalunya s'han extret del RMC (casos observats de mortalitat) [27] i de l'IDESCAT (població) [29].

Les localitzacions tumorals detallades a la Taula 4.2 són les que s'han fet servir per al primer exemple. Per al segon exemple, atès que la base de dades de la mortalitat de la WHO no disposa de informació per a totes les localitzacions tumorals, s'han exclòs de l'anàlisi els dos grups "Endometri" i "Totes les localitzacions tumorals excepte pell no Melanoma". La codificació del càncer d'ovari i de ronyó és diferent entre el RMC i la WHO. A la base de dades del RMC els codis del càncer d'ovari són C56 i C57 mentre que en la WHO només es considera el codi C56. Pel càncer de ronyó es consideren els codis del C64 fins al C66 i el C68 en el RMC mentre que, la WHO només considera el codi C64. Per tant, caldrà interpretar amb cautela els resultants SMRs per a aquestes localitzacions tumorals.

4.1.2. PROGRAMARI

En aquest projecte s'han implementat un conjunt d'aplicacions que permeten automatitzar els càlculs estadístics plantejats. Per a fer-ho possible, s'ha optat per l'entorn de programació que ofereix el paquet estadístic R. Es tracta de programari lliure (<http://www.r-project.org>) que ofereix molts avantatges des del punt de vista metodològic. En l'aplicació d'esperats s'empren funcions de la llibreria MASS d'aquest entorn. Les

aplicacions es programen en PHP i HTML. Per a treure els resultats en HTML s'utilitza una llibreria de l'R anomenada R2HTML.

A més s'utilitza el programa Excel 2003 (Microsoft® Office Excel 2003) per a extreure resultats per a l'exemple aplicat.

4.2 DESCRIPCIÓ DELS MÈTODES ESTADÍSTICS PER A CADA APLICACIÓ

Els resultats s'obtenen per a cada localització tumoral i sexe especificats en el fitxer d'entrada, de manera que, si un fitxer conté 4 grups per a homes i total (la suma d'homes i dones) és realitzarà els anàlisis per a cadascun dels grups sol·licitats.

L'aplicació Descriptiva retorna per a cada grup, el nombre de casos, el percentatge (opcional), la taxa crua, la taxa ajustada a la població que es desitgi (pel mètode d'estandardització directe), la taxa truncada (en els grups d'edat que es vulgui) i la acumulada (també en els grups desitjats) de la incidència/mortalitat.

L'aplicació de l'Estimació del Percentatge de Canvi Anual (*Estimated Annual Percent Change*, EAPC) retorna, per a cada sexe i grups sol·licitats, una taula amb el seu EAPC i l'interval de confiança corresponents. L'EAPC es calcula a partir d'un Model Lineal Generalitzat amb distribució Normal per a les taxes ajustades.

L'aplicació Esperats retorna una base de dades amb el nombre de casos projectats per a cada sexe, grup d'edat, any o període, grup (localització tumoral en el nostre cas) i població a risc prèviament seleccionats. A més també retorna una taula especificant els models emprats, el p-valor del test de bondat d'ajust, la significació estadística i el model estadístic seleccionat. Les projeccions (o casos esperats) s'obtenen a partir d'un model lineal generalitzat considerant una distribució Binomial Negativa o Poisson per al nombre de casos, incloent en el model l'efecte edat i una única pendent, model "age-drift" [32], i un model amb l'efecte edat i una pendent per edat, anomenat "age-specific-slope" [33]. En total es fan servir 4 models diferents.

L'aplicació Raó d'Incidència o Mortalitat Estandaritzada (*Standardized Incidence or Mortality Ratio*, SIMR) retorna per a cadascun dels sexes i grup la SIR o la SMR, dependent de si les dades són de incidència o mortalitat, així com el seu respectiu interval de confiança.

A continuació es donaran els detalls tècnics que s'utilitzen per tal d'obtenir els resultats que s'han especificat en aquest apartat.

4.3. EL CÀLCUL DE TAXES

És molt important tenir una definició clara de l'esdeveniment resultat. En epidemiologia un esdeveniment es defineix com l'ocurrència de qualsevol fenomen de salut que pugui ser caracteritzat. En aquest projecte, a tall d'exemple, l'esdeveniment es defineix com la "causa de mort per localització tumoral".

La incidència, mortalitat i recaiguda d'una malaltia es mesuren en termes de taxes, per exemple, com el nombre de casos nous diagnosticats per 100000 persones-any durant el temps d'observació. Les taxes de malalties varien àmpliament d'acord amb el sexe, l'edat, el temps de calendari i altres variables demogràfiques.

Les taxes, amb diferència de les freqüències, impliquen un element del temps. La taxa d'ocurrència d'un element en una població és el nombre d'esdeveniments que es donen durant un període de temps específic, dividit per la quantitat total del temps d'observació acumulat durant aquest interval [34]. Per a una taxa d'incidència, els esdeveniments són els casos diagnosticats de nou d'una malaltia en una població lliure d'aquesta malaltia en un període de temps determinat. El denominador de les taxes es pot calcular a partir del temps que aporta cadascun dels individus a risc, habitualment anys. Per a una taxa de mortalitat, els esdeveniments són els casos morts a causa d'una malaltia determinada en una població i en un període de temps determinats.

La taxa d'incidència/mortalitat anual per a un any de calendari particular és el nombre de casos nous diagnosticats durant l'any (o casos morts), dividit entre una aproximació de les persones-any d'observació, tal com la població a mitjans d'any.

4.3.1. TAXES BRUTES (TAXES CRUES)

La taxa crua (de l'anglès *Crude Rate*, CR) es calcula simplement dividint el nombre de casos nous durant un període de temps donat (o casos morts), entre el nombre corresponent d'individus a risc en la població a estudiar.

Suposem que per a un període o any concret es té el següent:

- X_T : Nombre de casos en el que es dona l'esdeveniment d'estudi per a tots els grups d'edat.
- N_T : Població a risc.

Aleshores, la CR es defineix com:

$$CR = \frac{X_T}{N_T} \cdot 100000$$

Normalment, tal i com ja s'ha comentat anteriorment, en molts estudis s'expressen les taxes en 100000 persones-any i aquest és el cas de la majoria d'estudis sobre càncer. En l'aplicació descriptiva, les taxes doncs, venen multiplicades per 100000 persones-any i cal que l'usuari ho tingui present per tal de realitzar una bona interpretació dels resultats.

D'altra banda, de forma similar, es calcula la Taxa Específica (*Specific Rate*, SR) per grup d'edat, suposant el següent:

- Grups d'edat: $i=1, 2, \dots, k$ on k es correspon al major grup d'edat. Com a exemple prenem $i=1, \dots, 18$ com a 18 grups d'edat quinquennals (0-4, 5-9, ..., 80-84, +85).
- X_i : Nombre de casos en el que es dona l'esdeveniment d'estudi per al grup d'edat i -èssim. Òbviament es satisfà $\sum_{i=1}^k X_i = X_T$.
- N_i : Població a risc per al grup d'edat i -èssim. De la mateixa manera, es satisfà:

$$\sum_{i=1}^k N_i = N_T.$$

Aleshores, la SR per grup d'edat és:

$$SR_i = \frac{X_i}{N_i} \cdot 100000, \text{ on } i \text{ per exemple és } i=1, \dots, 18.$$

Una observació important que cal remarcar és que $\sum_{i=1}^k SR_i \neq CR$, atès que

$$CR = \frac{X_T}{N_T} = \frac{1}{N_T} \sum_{i=1}^k X_i = \frac{1}{N_T} \sum_{i=1}^k SR_i \cdot N_i = \sum_{i=1}^k SR_i \frac{N_i}{N_T} = \sum_{i=1}^k SR_i P_i,$$

on P_i és el pes del grup d'edat i -èssim en la població en estudi.

Per tant, la CR d'una població és una mesura ponderada de les taxes específiques per grups d'edat on els pesos estan formats per les proporcions que representen els efectius

de cada grup d'edat sobre la població total. Cal remarcar que la CR no es pot utilitzar per aspectes comparatius entre dues poblacions diferents atès que l'estructura de la població pot ser diferent en les poblacions que es vol comparar o inclús diferents segons el període estudiat, per això cal realitzar una estandardització de taxes.

4.3.2. ESTANDARDITZACIÓ DE TAXES (TAXES AJUSTADES)

Les comparacions entre taxes de diferents poblacions es troben influenciades per la distribució de les variables que les componen com poden ser l'edat, el sexe, l'ocupació i la classe social entre d'altres. Així doncs, si per exemple es considera la CR de la mortalitat, s'observa que aquesta està influenciada per l'estructura de les edats de la població a la que es refereix. La comparació de la mortalitat entre dues o més poblacions no s'ha de realitzar amb la CR, donat que aquesta queda influenciada per la proporció d'individus que té cadascuna de les poblacions en els diferents grups d'edat.

La estandardització o ajust de taxes es basa en el càlcul d'una mesura de resum de les taxes específiques, és a dir, de les taxes de cadascun dels estrats de la variable. El sexe s'ha pres com a variable d'estratificació degut a les grans diferències que hi ha entre aquesta variable i les malalties i és per aquest motiu, que ja no es té en compte a l'hora d'estandarditzar.

El procés d'estandardització consisteix en ponderar les taxes específiques per cada estrat per una sèrie estàndard i així es controla la diferent composició dels grups d'edat en les dues poblacions. En el cas d'aquest projecte s'estandarditza per grups d'edat i per tant es controla la diferent composició dels grups d'edat en les dues poblacions:

$$R_{sd} = \sum_{i=1}^k SR_i w_i, \text{ on } w_i \text{ es correspondria a una sèrie estàndard de pesos.}$$

En aquest treball s'han emprat els mètodes d'estandardització d'ús més freqüent en epidemiologia: són el mètode directe i el mètode indirecte. Tot i que existeixen mètodes alternatius d'estandardització descrits per Yerushalmy [35] i per Yule [36], aquests no s'han aplicat en aquest treball.

MÈTODE DIRECTE

El mètode directe consisteix en construir la Taxa estandarditzada per edat (*Age Standardized Rate, ASR*) a partir d'una mitjana ponderada de les taxes específiques per grup d'edat, on cada grup d'edat té un pes segons la distribució d'edat de la població de referència. Aquesta serà la incidència (mortalitat) que s'observaria si la població tingués l'estructura d'edat de la població estàndard. Els pesos més comuns que s'utilitzen mundialment són els de la població estàndard de Segi (1960) [37] i que posteriorment van ser modificats per Doll et al.(1966) [38], basats en l'experiència de l'estructura de l'edat de la població de 46 països. La distribució d'aquests pesos es pot veure en la Taula 4.3. D'altra banda, tot i que en l'anàlisi d'aquest projecte s'han utilitzat aquests pesos, els pesos són un paràmetre de les funcions i per tant no delimita que l'usuari pugui introduir aquells que cregui més convenients pel seu estudi. La ASR es pot calcular com:

$$ASR = \sum_{i=1}^k \frac{X_i w_i^*}{N_i} \cdot 100000 = \sum_{i=1}^k SR_i w_i^* \quad (1)$$

On:

- i representa cada grup d'edat, essent $i=k$ el major dels grups d'edat.
- X_i : El nombre de casos de la població d'estudi en el grup d'edat i -èssim.
- N_i : La mida de la població d'estudi en el grup d'edat i -èssim.
- w_i^* : Pes del grup d'edat i -èssim en la població mundial, tal que $\sum_{i=1}^k w_i^* = 1$.
- SR_i : Taxa específica de la població d'estudi pel grup d'edat i -èssim.

Taula 4.3. Pesos de la població estàndard de Segi.

Grups d'edat *	Pes
0-4	0,120
5-9	0,100
10-14	0,090
15-19	0,090
20-24	0,080
25-29	0,080
30-34	0,060
35-39	0,060
40-44	0,060
45-49	0,060
50-54	0,050
55-59	0,040
60-64	0,040
65-69	0,030
70-74	0,020
75-79	0,010
80-84	0,005
>84	0,005

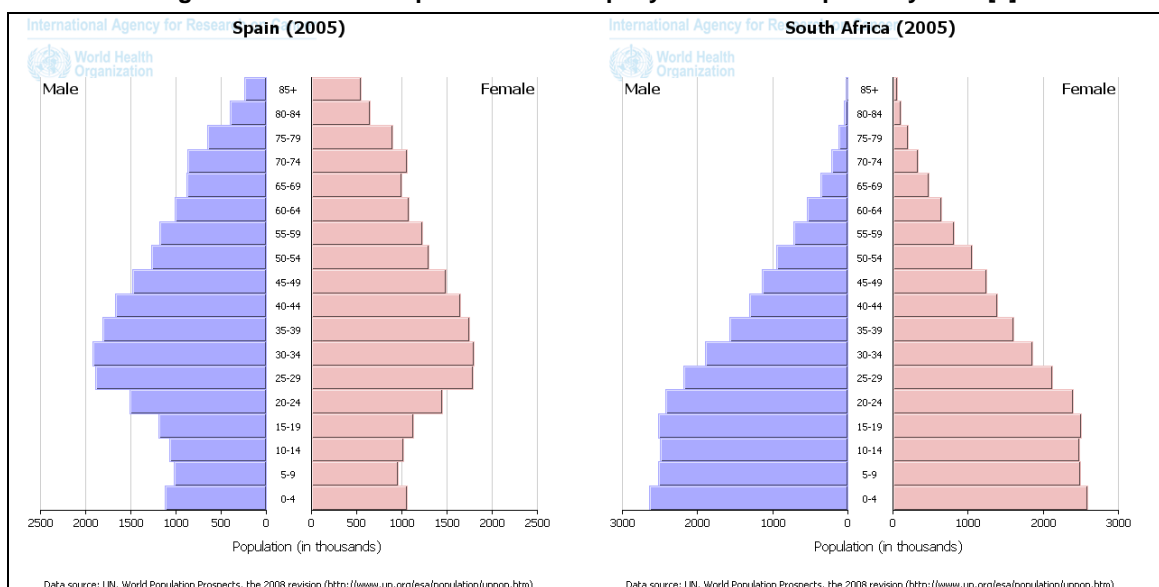
Per tal d'exemplificar la importància de l'estandardització de taxes quan es comparen diferents països, en la Taula 4.4, podem observar la CR i la ASR a la població mundial (sèrie estàndard de pesos) de dos països amb distribució poblacional diferent (Figura 4.1).

Taula 4.4. Descriptiva de la mortalitat per càncer a Espanya i a Sud Àfrica en l'any 2005 [9].

País	Sexe	Morts per càncer	CR	ASR*
Espanya	Homes	60942	287.6	154.4
	Dones	35928	164.3	70.1
Sud Àfrica	Homes	16829	71.1	132.5
	Dones	16351	67.0	85.8

* Taxa ajustada a la població mundial.

Figura 4.1. Distribució poblacional d'Espanya i Sud Àfrica per l'any 2005 [9].



Si es comparés directament la CR d'Espanya amb la de Sud Àfrica (Taula 4.4), observariem que en els homes es més elevada en Espanya que en Sud Àfrica, mentre que, si comparem la ASR, aquesta diferència no és tant gran, doncs en l'any 2005 l'ASR és de 154.4 per 100000 homes-any en Espanya i 132.5 per 1000000 homes-any. Així doncs, no es pot comparar directament la CR, doncs aquests dos països tenen piràmides poblacionals molt diferents (Figura 4.1) i és necessari estandarditzar per comparar-los. En les dones, aquesta diferència encara és més rellevant, ja que la CR és molt més elevada en Espanya que en Sud Àfrica, mentre que quan es realitza l'ajust per grups d'edat, s'obté tot el contrari.

MÈTODE INDIRECTE

Aquest mètode utilitza les taxes d'incidència o mortalitat d'una població estàndard i les aplica sobre la població d'estudi, obtenint la incidència o mortalitat esperada per edats en la població d'estudi [39]. Generalment, aquesta mesura no s'utilitza directament com un indicador epidemiològic, atès que solsament ens permet calcular els casos esperats (E) que es donarien en la població en estudi si aquesta tingués les taxes de la població de referència o estàndard. Aquests E si que tenen una utilitat comparativa quan es comparen amb els casos observats en la població en estudi mitjançant, per exemple, la SIR o la SMR.

Per tal de calcular E, ho podem fer de la següent forma:

$$E = \sum_{i=1}^k SR_i^* N_i \quad (2),$$

on:

- i representa cada grup d'edat, essent i=k el major dels grups d'edat.
- N_i : Població a risc en el grup d'edat i-èssim.
- SR_i^* : Taxa específica de la població de referència pel grup d'edat i-èssim. Aquesta taxa no ha d'estar en 100000 persones-any.
- E: Valors esperats.

4.3.3. TAXES TRUNCADES

A la pràctica, en els estudis epidemiològics sol ser interessant definir grups d'edat per construir indicadors com són les taxes truncades (*Truncate Rates*, TR). Es a dir, si estem interessats en estudiar uns grups d'edat específics ja que la malaltia en qüestió es dona amb més incidència en aquests grups, aleshores podem truncar i obtenir l'indicador específic per a aquest grup. Un exemple, serien els indicadors pel càncer pediàtric, on la TR està basada en el grup d'edat de 0 a 14 anys.

Tot i que en aquest projecte es presenten les TR de 35 a 64 anys, en l'aplicació descriptiva és possible seleccionar aquells grups d'edat d'interès, fixats prèviament per l'usuari, pel càlcul de la TR.

El càlcul de les TR consisteix en seleccionar els grups d'edat pertinents i calcular de nou els seus pesos mitjançant la població de referència, i forçant que sumin la unitat [40].

Suposem ara que G és un subconjunt del conjunt de grups d'edat ($G \subset \{1, 2, \dots, k\}$, on k és el grup d'edat més gran). Es defineixen aleshores els pesos

modificats com: $w_i^* = \frac{w_i}{\sum_{i \in G} w_i}$, $i \in G$. Es satisfà que $\sum_{i \in G} w_i^* = 1$. Aleshores es defineix la

taxa truncada (Truncate Rate, TR) com:

$$TR = \sum_i SR_i \cdot w_i^*, \text{ on } i \in G \quad (3).$$

4.3.4. TAXES ACUMULADES

Les taxes acumulades (*Cumulative Rate*, CumR) es poden interpretar com un cas específic de les taxes truncades, amb la particularitat que el subconjunt de grups d'edats seleccionats correspon a una acumulació dels grups d'edat. Tot i que els nostres resultats es basen en una CumR de 0 a 74 anys, l'aplicació permet seleccionar els grups d'edat que es volen acumular.

La CumR és una aproximació de la probabilitat de desenvolupar una malaltia (ó la mort) durant un període determinat. Aquesta taxa té l'avantatge de resumir les taxes específiques per edat independentment de l'estructura d'edat de la població, i mostra la probabilitat que té un individu de desenvolupar (o morir) una malaltia. Aquest càlcul és teòric i assumeix que les taxes específiques per edat són estables per a un individu.

La CumR d'incidència o mortalitat sobreestima la probabilitat de desenvolupar un event durant la vida quan les taxes són elevades. D'altra banda, quan les taxes d'incidència o mortalitat són baixes, la CumR no està afectada significativament per aquest biaix i per tant, es pot interpretar la CumR com el risc que té un individu de morir o incidir per una malaltia abans de certa edat.

La CumR es calcula de la següent forma [11]:

$$CumR = g \cdot \sum_{i=1}^a SR_i \cdot 100 \cdot \frac{1}{100000} \quad (4)$$

Essent:

- i : Grup d'edat i -èssim.
- a : Es correspon al grup d'edat fins on es vol acumular. Per a agilitzar l'aplicació s'han considerat grups d'edat homogenis. Per tant s'acumula des de $i=1$ fins $i=a$.
- g : Amplada de cada grup d'edat. En principi són homogenis, per tant, la dimensió serà la mateixa per a cada grup d'edat, per exemple, 5 en el cas de grups d'edat quinquennals.
- SR_i : Taxa específica en el grup d'edat i -èssim.

La CumR s'expressa en forma de percentatge i és per això que (4) es mostra multiplicat per 100.

4.4. MODELS LINEALS GENERALITZATS

Els models lineals generalitzats (MLG) són una extensió dels models lineals que engloben tant els models amb variable resposta contínua com categòrica. Això dona pas a considerar altres distribucions per tal de poden tenir en compte variables resposta contínues o categòriques on a més de la distribució Normal es poden considerar la Binomial i la Poisson, entre d'altres. El lligam entre la resposta i els predictors pot ser log lineal, logit, probit, logístic i de regressió lineal [41].

Per tal de poder utilitzar MLG, s'ha de donar el següent:

Y_i ha de ser una variable aleatòria amb distribució pertanyent a la **família exponencial** [41], tal que:

$$E(Y_i) = \mu_i \text{ i } \eta_i = g(\mu_i) = \sum_j^p x_{ij} \beta_j,$$

essent x_1, x_2, \dots, x_p les covariables que produeixen el predictor lineal η i anomenant a $g(\cdot)$ funció link, essent aquesta alguna funció diferenciable i monòtona [41].

4.4.1. DISTRIBUCIONS DE LA FAMÍLIA EXPONENCIAL I FUNCÍO DE VERSEMBLANÇA

Sigui Y una variable aleatòria tal que la seva funció de probabilitat (si és discreta) o la seva funció de densitat de probabilitat (si és contínua), en qualsevol cas, depèn d'un únic paràmetre θ . La distribució pertany a la família exponencial si es pot escriure de la següent forma [41]:

$$f_Y(y, \theta, \phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right)$$

On $a(\cdot)$, $b(\cdot)$ i $c(\cdot)$ són funcions específiques.

- Si ϕ és conegut aleshores el model pertany a la família exponencial i al únic paràmetre θ se'l denomina paràmetre canònic.
- Si ϕ no és conegut aleshores el model pot pertànyer o no a la família exponencial (amb 2 paràmetres).
- La funció $a(\phi)$ sol ser de la forma: $a(\phi) = \frac{\phi}{w}$, $w \in \mathfrak{R}$, on ϕ es denomina paràmetre de dispersió i és constant en les observacions i $w \in \mathfrak{R}$ que és un pes a priori conegut que varia d'observació a observació.
- La funció $b(\cdot)$ es denomina funció acumulant. La funció acumulant es el logaritme de la funció generatriu dels moments avaluada en l'origen. La derivada primera, segona i tercera d'aquesta funció faciliten els moments centrats d'ordre 1, 2 i 3 de la variable respecte al paràmetre θ (mitjana o variància, entre d'altres).

4.4.2. FUNCIÓ DE LOG-VERSEMBLANÇA I LES SEVES PROPIETATS

Sigui la funció de log-versemblança d'una observació donada y pertanyent a la família exponencial i considerant la funció dependent de θ i ϕ tenim que s'escriu [41]:

$$\ell(\theta, \phi, y) = \log f_Y(y, \theta, \phi) = \frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)$$

La mitjana i la variància de Y es pot trobar fàcilment a través de les relacions conegudes següents:

$$E\left(\frac{\partial \ell}{\partial \theta}\right) = 0$$

i

$$E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right) + E\left(\frac{\partial \ell}{\partial \theta}\right)^2 = 0$$

Per tant, si es fan les respectives derivades de la funció de log-versemblança s'obté que:

$$\frac{\partial l}{\partial \theta} = \{y - b'(\theta)\}/a(\phi)$$

i

$$\frac{\partial^2 l}{\partial \theta^2} = -b''(\theta)/a(\phi)$$

I d'aquestes equacions s'obté el següent:

$$0 = E\left(\frac{\partial l}{\partial \theta}\right) = \{\mu - b'(\theta)\}/a(\phi) \quad \text{i per tant} \quad E(Y) = \mu = b'(\theta).$$

I de forma semblant s'obté que:

$$0 = -\frac{b''(\theta)}{a(\phi)} + \frac{\text{var}(Y)}{a^2(\phi)} \quad \text{amb el que} \quad \text{Var}(Y) = b''(\theta) a(\phi)$$

A la funció $b''(\theta)$ se l'anomena funció de variància i és una funció que depèn de μ i normalment s'expressa com $V(\mu)$.

4.4.3. EXEMPLES DE DISTRIBUCIONS DE LA FAMÍLIA EXPONENCIAL

Existeixen moltes distribucions pertanyents a la família exponencial, com per exemple la distribució Normal, la Poisson, la Binomial, la Gamma i la Gaussiana inversa. En aquest projecte es treballa amb la distribució Normal, la distribució Poisson i la Binomial Negativa. Veiem com s'escriuen aquestes distribucions en forma de família exponencial i quina és la seva funció de log versemblança, així com l'esperança i variància de cadascuna d'elles.

DISTRIBUCIÓ NORMAL

La seva funció de densitat de probabilitat és:

$$f_Y(y, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y - \mu)^2}{2\sigma^2}\right)$$

Expressant-ho en forma de la família exponencial es té:

$$\begin{aligned} f_Y(y, \theta, \phi) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) = \exp\left(\ln(1) - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2 - 2y\mu + \mu^2}{2\sigma^2}\right) = \\ &= \exp\left(-\frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2} + \frac{2y\mu}{2\sigma^2} - \frac{\mu^2}{2\sigma^2}\right) = \exp\left(\frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}\right) = \\ &= \exp\left(\frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{1}{2}\left(\ln(2\pi\sigma^2) - \frac{y^2}{\sigma^2}\right)\right) \end{aligned}$$

Aleshores :

$$\phi = \sigma^2; \theta = \mu; a(\phi) = \phi; b(\theta) = \frac{\theta^2}{2} \text{ i } c(y, \phi) = -\frac{1}{2}\left(\ln(2\pi\phi) - \frac{y^2}{\phi}\right)$$

Per tant, s'obté el següent:

$$E(Y) = b'(\theta) = \theta = \mu$$

$$Var(Y) = b''(\theta) a(\phi) = 1 \phi = \phi \sigma^2$$

DISTRIBUCIÓ POISSON

La seva funció de probabilitat és:

$$f_Y(y, \lambda) = \frac{\lambda^y}{y!} \exp(-\lambda)$$

Expressant-ho en forma de la família exponencial tenim el següent:

$$f_Y(y, \theta, \phi) = \frac{\lambda^y}{y!} \exp(-\lambda) = \exp(y \ln(\lambda) - \ln(y!) - \lambda) = \exp\left(\frac{y \ln(\lambda) - \lambda}{1} - \ln(y!)\right)$$

Aleshores tenim que:

$$\phi = 1; \theta = \ln(\lambda); a(\phi) = 1; b(\theta) = \exp(\theta) \text{ i } c(y, \phi) = -\ln(y!)$$

Per tant, s'obté el següent:

$$E(Y) = b'(\theta) = \exp(\theta) = \exp(\ln(\lambda)) = \lambda$$

$$Var(Y) = b''(\theta) a(\phi) = \exp(\theta) 1 = \exp(\theta) = \lambda$$

DISTRIBUCIÓ BINOMIAL NEGATIVA

La seva funció de probabilitat és:

$$f_Y(y, p) = \binom{y+r-1}{r-1} p^r (1-p)^y = \frac{(y+r-1)!}{(r-1)!(y+r)!} p^r (1-p)^y$$

Expressant-ho en forma de la família exponencial tenim el següent:

$$\begin{aligned} f_Y(y, \theta, \phi) &= \frac{(y+r-1)!}{(r-1)!(y+r)!} p^r (1-p)^y = \exp\left(\ln\left(\frac{(y+r-1)!}{(r-1)!(y+r)!}\right) + r \ln(p) + y \ln(1-p)\right) = \\ &= \exp\left(\frac{y \ln(1-p) - (-r \ln(p))}{1} + \ln\left(\frac{(y+r-1)!}{(r-1)!(y+r)!}\right)\right) \end{aligned}$$

Aleshores tenim que:

$$\phi = 1; \theta = \ln(1-p); a(\phi) = 1; b(\theta) = -r \ln(1 - \exp(\theta)) \text{ i } c(y, \phi) = \ln\left(\frac{(y+r-1)!}{(r-1)!(y+r)!}\right)$$

Per tant, s'obté el següent:

$$E(Y) = b'(\theta) = -r \frac{-\exp(\theta)}{1 - \exp(\theta)} = r \frac{\exp(\theta)}{1 - \exp(\theta)} = r \frac{(1-p)}{1 - (1-p)} = r \frac{(1-p)}{p}$$

$$Var(Y) = b''(\theta) a(\phi) = b''(\theta) 1 = r \left[\frac{\exp(\theta)(1 - \exp(\theta)) - \exp(\theta)(-\exp(\theta))}{(1 - \exp(\theta))^2} \right] =$$

$$= r \left[\frac{(1-p)(1-(1-p)) - (1-p)(-(1-p))}{(1-(1-p))^2} \right] = \frac{r(1-p)}{p^2}$$

4.4.4. FUNCIONS LINK PELS MODELS LINEALS GENERALITZATS

Sigui Y un vector d'observacions, tals que el vector $E[Y] = \mu$ és el valor esperat de les dades, està funcionalment relacionat amb el predictor lineal η a través de la denominada

funció link que es denota com $g(\cdot)$, de manera que $\eta = g(\mu) = \sum_j^p x_j \beta_j$.

Quan es tracta amb comptatges i la distribució és la Poisson, es té que $\mu > 0$. En aquest cas el link identitat no és pot fer servir sense assumir certes restriccions sobre els predictors atès que es poden obtenir η negatives mentre que les μ no ho són. Els models utilitzats per a comptatges contenen efectes multiplicatius dels regressors i per tal d'expressar aquests com efectes additius es realitza la transformació logarítmica (*link log*), $\eta = g(\mu) = \log(\mu)$ de tal forma que la inversa és $\mu = g^{-1}(\eta) = \exp(\eta)$ [41]. Aquesta transformació és la que s'utilitza en aquest treball ja que es treballa amb comptatges, per tant, el link utilitzat es correspon amb el link log.

A tall d'exemple i tot i que en aquest estudi no s'ha fet servir, la mitjana ha de prendre valors entre 0 i 1 per a la distribució binomial, i atès que el predictor lineal pot prendre qualsevol valor real, no sembla coherent prendre la identitat com a funció link. En aquest cas, es solen considerar principalment tres funcions [41]:

- La funció de *link logit*: $\eta = \log(\mu/(1-\mu))$. Aquest link s'utilitza amb els models de regressió logística.
- La funció de *link probit*: $\eta = \Phi^{-1}(\mu)$ on $\Phi(\cdot)$ és la funció de distribució Gaussiana. Els models utilitzats són els que s'anomenen models probit.
- La funció de *link log-log complementària*: $\eta = \log(-\log(1-\mu))$

Les funcions de link de la família de potències són importants per a observacions amb mitjana positiva i es defineixen per a cada observació com: $\eta = (\mu^\lambda - 1) / \lambda$ tals que en el cas de $\lambda \rightarrow 0$ llavors $\eta = \log(\mu)$ [41].

4.4.5. BONDAT D'AJUST

El procés d'ajustar un model a les dades es pot considerar una forma de substituir un conjunt de valors de dades y per un conjunt de valors predits $\hat{\mu}$ que es deriven d'un model que, generalment, conté un nombre relativament petit de paràmetres. Per tal de comparar la bondat d'ajust del model es pot avaluar la diferència entre els valors predits (μ 's) i els observats (y 's). Si la discrepància entre aquests valors no és rellevant es pot considerar que el model ajusta a les dades, metre que si és rellevant, s'hauria de considerar un model alternatiu al plantejat o comprovar si alguna de les observacions es un "outlier". Les mesures de discrepància o de bondat d'ajust es poden definir de diferents maneres. En primer lloc, farem referència a aquelles que estan formades pel logaritme de raons de versemblança, anomenades deviança [41].

Donades n observacions es poden ajustar models que continguin fins a n paràmetres. El model nul és aquell més senzill que només conté un paràmetre que representa la mitjana comú per totes les observacions i que per tant assigna cada possible variació en les dades a la component aleatòria. En l'extrem oposat es troba el model complet o maximal amb n paràmetres, un per observació i per tant, en aquest model, els valors ajustats coincideixen exactament amb els observats sense deixar pas a cap component aleatòria (tota la variació l'assigna a la component sistemàtica).

A la pràctica, el model nul és massa simple i el model complet no és informatiu per què aquest no resumeix les dades, però és interessant ja que ens dona la base per a mesurar la discrepància per a un model intermedi amb p paràmetres.

És convenient expressar la log-versemblança en termes del paràmetre μ (valor de la mitjana) enlloc del paràmetre canònic θ . Sent $l(\hat{\mu}, \phi; y)$ la log versemblança maximitzada sobre les β per a valors fixats del paràmetre de dispersió ϕ . La màxima versemblança que es pot assolir en un model complet amb n paràmetres és $l(y, \phi; y)$ que normalment és finita. La discrepància d'un ajust es proporcional a dues vegades la diferència entre la màxima log-versemblança assolida i la que ha estat assolida pel model sota investigació. Si es denota per $\hat{\theta} = \theta(\hat{\mu})$ i $\tilde{\theta} = \theta(y)$ les estimacions dels paràmetres canònics sota els dos models, la discrepància, assumint $a_i(\phi) = \phi / w_i$, es pot escriure com:

$$\sum 2 w_i \left(y_i \left(\tilde{\theta}_i - \hat{\theta}_i \right) - b \left(\tilde{\theta}_i \right) + b \left(\hat{\theta}_i \right) \right) / \phi = D(y; \hat{\mu}) / \phi,$$

on $D(y; \hat{\mu})$ es coneix com la deviança pel model actual i només està en funció de les dades. A més,

$$D^*(y; \hat{\mu}) = D(y; \hat{\mu}) / \phi$$

de manera que la deviança escalada $D^*(y; \hat{\mu})$ és la deviança expressada com un múltiple del paràmetre de dispersió.

En aquest projecte, s'utilitza la distribució de Poisson i la Binomial Negativa per a extreure els valors esperats, per tant, la forma de la deviança per a aquestes distribucions és:

$$2 \sum_{i=1}^n \left(y \log(y / \hat{\mu}) - (y - \hat{\mu}) \right)$$

Una altra mesura important de discrepància és l'estadístic X^2 de Pearson [41] que té la forma següent:

$$X^2 = \sum (y - \hat{\mu})^2 / V(\hat{\mu}),$$

On $V(\hat{\mu})$ és la funció de la variància estimada per la distribució que ens interessa.

Tant la deviança com l'estadístic X^2 de Pearson segueixen distribucions exactes χ^2 per la teoria Normal dels models lineals (suposant que el model és cert), i per a les altres distribucions es disposa dels resultats asimptòtics. No obstant, els resultats asimptòtics no s'haurien d'utilitzar quan s'analitzen dades amb poca mida mostral i en aquest cas es pot demostrar que tant D com X^2 tenen propietats de distribució millors. La deviança té l'avantatge general, com a mesura de discrepància, de que és additiva per a conjunts de models niats si s'utilitzen les estimacions màxim-versemblants, mentre que el X^2 en general no. D'altra banda, moltes vegades es prefereix l'estadístic X^2 per què té una interpretació més directa. En aquest treball avaluarem la bondat d'ajust dels models (*Goodness of fit* -GOF-) a partir de l'estadístic X^2 [41].

4.4.6. CRITERI D'INFORMACIÓ D'AKAIKE

El Criteri d'Informació d'Akaike (Akaike Information Criterion -AIC-) és una eina per a la selecció de models. Donat un conjunt de dades, el millor model dels ajustats és aquell que té menor valor del AIC.

El AIC està basat en el concepte d'entropia, doncs ofereix una mesura relativa de la informació perduda quan s'utilitza un model donat per a descriure la realitat. Aquest criteri descriu la compensació entre biaix i variància. S'expressa de la següent forma:

$$AIC = 2k - 2\ln(L)$$

On:

- k és el nombre de paràmetres del model estadístic.
- L és el valor màxim de la funció de versemblança pel model estimat.

4.4.7. MODELS LOG-LINEALS

Els models log-lineals s'acostumen a utilitzar quan s'estudien comptatges. En aquest tipus de models la dependència de $\mu_i = E(Y_i)$ sobre el vector de covariables x_i s'assumeix que és multiplicativa i usualment s'escriu en forma logarítmica:

$$\log \mu_i = \eta_i x_i ; \quad i = 1, \dots, n.$$

Així doncs, el model log-lineal és un cas concret dels MLG on el lligam entre la resposta i el predictor és el logaritme i on les dades, Y_i , es corresponen a una variable aleatòria de la família exponencial (usualment, la distribució de Poisson o la Binomial Negativa, veure apartat 4.4.3).

Cal remarcar que en l'exemple d'aplicació d'aquest projecte es defineix la variable resposta Y com el nombre de morts per càncer i s'observa independentment per als valors específics de les variables predictorres: any i grup d'edat. La variable any pren els valors 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003 i 2004 i es pren com a variable contínua. El grup d'edat pot prendre els valors de 0-4 anys, 5-9 anys, 10-14, 15-

19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84 i >85 anys i es considera com a variable categòrica.

A més a més, aquests models tenen en compte un paràmetre conegut que s'anomena *offset* i que en els models que considerem és funció de la mida de la població en l'estrat, és a dir, en un any i un grup d'edat concret.

En aquest projecte, es consideren dues distribucions diferents pel nombre de morts per càncer. Així doncs, es considera $Y_i \sim \text{Poisson}(\lambda_i)$ o bé $Y_i \sim \text{Binomial Negativa}(p_i)$. A més a més es proposen dos tipus de models diferents [33]:

$$\ln(E(Y_{it})) = \ln(P_{it}) + \alpha_i + \beta t \quad (5),$$

i

$$\ln(E(Y_{it})) = \ln(P_{it}) + \alpha_i + \beta_i t \quad (6),$$

on el model (5) és conegut com edat-tendència (age-drift en anglès) [32] i el model (6) és conegut com el de pendent específica per edat (age-specific slope en anglès) [33]:

- Y_{it} és el nombre de morts per càncer en el grup d'edat i -èssim i l'any t .
- $E(Y_{it})$ és el nombre de morts per càncer esperats en el grup d'edat i -èssim i l'any t .
- P_{it} és la població a risc en el grup d'edat i -èssim i l'any t .
- α_i són els coeficients dels grups d'edat.
- β és el coeficient de la variable temps que representa una única pendent per a tots els grups d'edat.
- β_i són els coeficients de la variable temps que es corresponen a les pendents per a cada grup d'edat considerat.
- El terme $\ln(P_{it})$ és el que anomenem *offset*. Quan s'introdueix aquest terme estem modelant, indirectament, el valor esperat de la taxa de mortalitat: $y_{it} = E[Y_{it}]/P_{it}$.
- S'introdueix l'any re-escalat. Per exemple, si estem treballant amb els anys 1995, 1996, 1997..., aquests es transformen extraient l'any mínim, per tant, es treballa amb 0, 1, 2...

També cal considerar, que tot i que en l'exemple que farem servir es treballa amb anys, també es pot treballar amb períodes i per tant agafaríem com a referència un any central. Així doncs, si treballem amb els anys centrals 1997, 2002, 2007 quan ho re-escalem, ens quedaríem amb 0, 5 i 10. D'altra banda, en l'exemple es treballen amb 18 grups d'edat, però a l'hora de realitzar altres anàlisis, pot ser que els grups d'edat estiguin definits d'una altra forma. A més a més, cal considerar que en l'exemple treballem amb casos morts per càncer, però en general es pot parlar de casos, ja sigui incidents o morts o inclús de taxes d'incidència i/o mortalitat.

4.4.8. DISTRIBUCIO BINOMIAL NEGATIVA

La distribució Binomial Negativa és una distribució amb forma més generalitzada utilitzada per variables resposta de comptatges que permet major dispersió en la variància dels comptatges. A la pràctica és força comú trobar-nos amb que la variància de la variable resposta de comptatge és major que la seva mitjana i això s'anomena sobre dispersió. Si una variable de comptatge és sobre dispersa, la regressió de Poisson subestima els errors estàndard de les variables predictores. Quan la sobre dispersió és evident, una solució és especificar que els errors tenen una distribució binomial negativa. La regressió de la Binomial Negativa dona els mateixos coeficients que la regressió de Poisson però major errors estàndard. La interpretació dels resultats és la mateixa que per la regressió de Poisson [41],[42].

4.4.9. SELECCIÓ DE MODELS LOG-LINEALS

En l'apartat de Models log-lineals es presenten els diferents models que s'utilitzen com a candidats pel càlcul dels casos esperats i que tenen en compte diferents pendents per a cadascun dels grups d'edat. En total s'ajusten 4 models per a cadascun dels grups i sexe. A la Taula 4.5 es pot veure la relació dels diferents models.

Taula 4.5. Especificació dels models ajustats en l'estudi.

Distribució de Y_{ij}	Especificació del model	Nom del model en l'aplicació web
$Y_{it} \sim BinNeg(\mu_{it}, \nu)$	$\ln(E(Y_{it})) = \ln(P_{it}) + \alpha_i + \beta t$	Model_nb_drift
	$\ln(E(Y_{it})) = \ln(P_{it}) + \alpha_i + \beta_i t$	Model_nb_I
$Y_{it} \sim Poiss(\lambda_{it})$	$\ln(E(Y_{it})) = \ln(P_{it}) + \alpha_i + \beta t$	Model_p_drift
	$\ln(E(Y_{it})) = \ln(P_{it}) + \alpha_i + \beta_i t$	Model_p_I

Una vegada ajustats els 4 models, el procediment per a seleccionar els models en els que ens basarem per a fer les estimacions per a cada combinació de grup i sexe és el següent:

- Pas 1) S'ajusten tots els models per a cadascuna de les malalties i sexe.
- Pas 2) Es pre-selecciona el model amb mínim AIC.
- Pas 3) Es poden donar diferents situacions, que consistirien en:
 - 3.1) El model pre-seleccionat ajusta (el p-valor de la GOF > 0.05). Aquest és el model final. El procés acaba i el mètode de selecció ens indica que s'ha escollit el model amb el menor AIC ("Minimum AIC").
 - 3.2) El model pre-seleccionat no ajusta bé, per tant, s'avaluen altres models.
 - 3.2.1) Si algun o alguns models ajusten, es selecciona el que té major p-valor. El procés acaba i el mètode de selecció que s'ha escollit el model amb màxima GOF ("Maximum GOF").
 - 3.2.2) Si no ajusta cap dels models, es selecciona el que té menor AIC. El procés acaba i el mètode de selecció ens indica que el model no ajusta ("The model does not fit").

Quan es triï un model amb distribució Binomial Negativa en lloc d'un model amb distribució Poisson, és probable que sigui a causa d'un efecte de sobredispersió. Quan s'avaluen els models, és possible que el menor AIC mostri un model amb distribució Poisson, i al no ajustar degut a la sobredispersió, s'esculli un entre els models que ajusten i que tingui menor AIC.

Una vegada seleccionat el millor model, les estimacions no sempre són les que es prenen com a definitives degut a que s'ha considerat un criteri mitjançant el qual es posen zeros en els valors projectats tals que tinguin un z-valor d'algun dels coeficients en valor absolut menor a 0.01. Els coeficients amb un z-valor menor a 0.01 es corresponen a aquells coeficients amb una estimació de l'error elevada el qual perjudica a l'hora de fer les estimacions. D'aquesta forma controlem que els valors de les estimacions no es disparin.

4.4.10. ESTIMACIÓ DEL PERCENTATGE DE CANVI ANUAL I ELS SEUS INTERVALS DE CONFIANÇA

L'anàlisi de tendències avalua l'evolució de les taxes al llarg del temps. La pràctica d'aquestes anàlisis ha mostrat que hi pot haver un ampli ventall d'escenaris. El més

senzill és la situació en la que la incidència o mortalitat d'una malaltia està estabilitzada, mostrant una evolució plana. Alternativament, es pot donar també la situació d'una tendència creixent o decreixent de la incidència o mortalitat, essent aquests canvis lineals o exponencials. En aquests casos, la pràctica també ha mostrat que les taxes solen créixer o decreixer fins un cert moment, on s'observa una asímptota que indica l'estabilització de la malaltia. Finalment, també es poden donar casos de canvis de tendència ó desestabilitzacions de les taxes. En qualsevol cas, resulta fonamental realitzar una bona descriptiva de l'evolució de les taxes per poder entendre millor l'ajust dels models pertinents. En els casos de canvi de tendència, els models s'han d'ajustar per trams, doncs aquests poden estar provocats per canvis en els sistemes de codificació, millores en els tractaments de la malaltia, implantació de noves tècniques de diagnòstic o de programes de cribatge, entre d'altres. Per tant, l'avaluació de tendències és un pas que cal fer prèviament a l'anàlisi de projeccions i que cal tenir en compte per tal d'obtenir una bona predicció. Per a avaluar aquesta tendència, entre d'altres eines, s'utilitza el que s'anomena EAPC.

El EAPC es calcula utilitzant un MLG considerant la distribució Gaussiana per al logaritme de la taxa estandarditzada per edat (ASR). Es calcula el EAPC sota l'assumpció de linealitat en escala logarítmica, el qual, és equivalent a una assumpció de canvi constant [40;43].

Es considera:

ASR_t = Taxa estandarditzada per edat per a un conjunt d'anys on $t=0, 1, \dots, n$.

El MLG assumeix que el valor esperat de la resposta està relacionat amb el temps mitjançant una expressió logarítmica:

$$\ln(E(ASR_t)) = \lambda_t = \gamma + \beta t \quad (7),$$

essent $\hat{E}[ASR_t] = e^{\gamma + \hat{\beta}t}$ el valor estimat per a la ASR mitjançant el model.

Llavors el EAPC es calcula com:

$$EAPC = \left(\frac{\hat{ASR}_{t+1} - \hat{ASR}_t}{\hat{ASR}_t} \right) \cdot 100 = \left(\frac{\hat{ASR}_{t+1}}{\hat{ASR}_t} - 1 \right) \cdot 100 = \left(\frac{e^{\gamma + \hat{\beta}(t+1)}}{e^{\gamma + \hat{\beta}t}} - 1 \right) \cdot 100 = (e^{\hat{\beta}} - 1) \cdot 100 \quad (8)$$

Sigui S_{β}^{\wedge} el error estàndard de β^{\wedge} , llavors els $(1 - \alpha)$ % límits de l'interval de confiança

(IC) per a l'EAPC es calculen com $e^{\beta^{\wedge} \pm \left(Z_{\frac{1-\alpha}{2}} \right) \cdot S_{\beta}^{\wedge}}$.

4.5. CÀLCUL DE LA RAO D'INCIDÈNCIA O MORTALITAT ESTANDARDITZADA

Aquest estadístic es va introduir l'any 1786 per J.N.Tetens (nascut en 1736 a Alemanya va ésser filòsof, matemàtic, científic i estadista) en la segona part de la seva obra anomenada "*Einleitung zur Berechnung der Leibrenten und Anwartschaften*" publicada en Leipzig [44;45]. En 1855 el mateix estadístic es va tornar a utilitzar per W.H. Farr en la memòria anual del Registre General de Gran Bretanya per a comparar la mort en diferents grups ocupacionals [39].

Atès que en aquest treball es calcula de la mateixa forma la SIR i la SMR, tan sols depèn de si les dades són d'incidència o de mortalitat, d'ara en endavant anomenarem indistintament SIMR tant a la SIR com a la SMR.

La SIMR és una mesura resum de l'estandardització de taxes (d'incidència ó mortalitat) pel mètode indirecte. Donat un període de temps en estudi, la SIMR es calcula com la raó entre el nombre de casos observats (O_l) i el nombre de casos esperats (E_l) per la malaltia l - èsima en estudi,

$$SIMR_l = \frac{O_l}{E_l} \quad (9),$$

tenint en compte que els E_l es calculen segons (2).

Una SIMR de 1 significa que no hi ha diferències entre el nombre de casos observats en la població d'estudi i el nombre de casos esperats en la població de referència. Una SIMR major de 1 ens indica quantes vegades es moren o incideixen més els casos observats respecte als esperats en la població de referència (increment del risc de morir per pertànyer a la població d'estudi). Una SIMR menor de 1, indica quantes vegades menys es moren o incideixen els casos observats respecte els esperats en la població de referència (els individus de l'estudi, es moren menys que els de la població de referència).

L'interval de confiança per a la SIMR es calcula mitjançant l'assumpció de la distribució de Poisson per al nombre de casos observats. A partir del límit inferior i superior de l'interval de confiança per a una observació provinent d'una distribució de Poisson, es divideixen els límits de l'interval per el nombre de casos esperats [39;46-48].

Primer, es calcula l'interval de confiança del 100 (1- α)% pel nombre de casos observats, de la següent forma:

- Si $X > 0$, aleshores, el límit inferior (lower limit -LL-) i el límit superior (upper limit -UL-) es corresponen a:

$$LL = \frac{\chi^2\left(\frac{\alpha}{2}, 2 \cdot X\right)}{2} \quad ; \quad UL = \frac{\chi^2\left((1-\alpha) + \frac{\alpha}{2}, 2 \cdot (X+1)\right)}{2}$$

- Si $X = 0$, aleshores, LL i UL són:

$$LL = 0 \quad ; \quad UL = -\log(\alpha)$$

I a continuació es calcula el límit inferior per a la SIMR ($LL(SIMR)$) i el límit superior per a la SIMR ($UL(SIMR)$) com segueix:

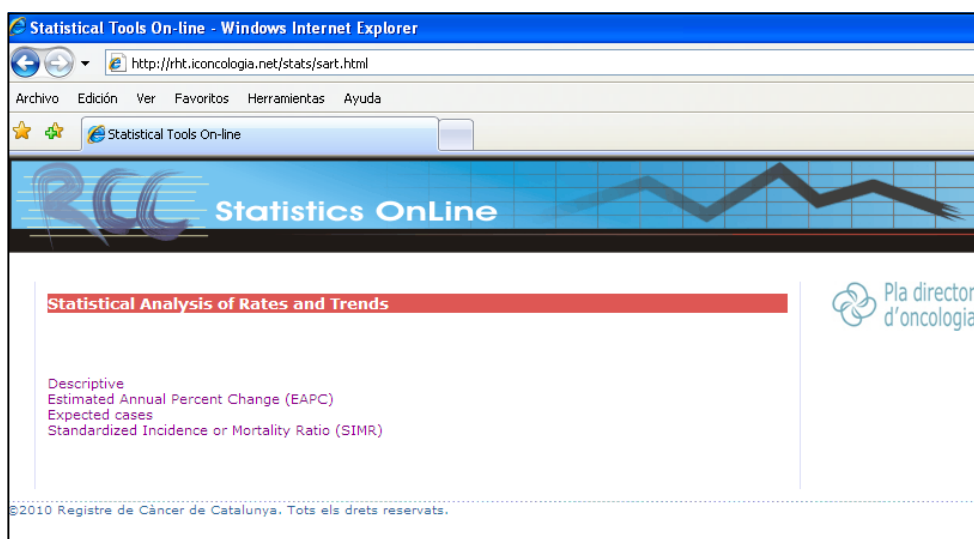
$$LL(SIMR) = \frac{LL}{E} \quad ; \quad UL(SIMR) = \frac{UL}{E}$$

5. APLICACIÓ PEL CÀLCUL AUTOMATITZAT VIA WEB

Tal i com s'ha comentat anteriorment, en aquest projecte es programen quatre funcions amb R (les podem trobar en l'Annex 1) de les quals posteriorment es fan les aplicacions via web mitjançant PHP i HTML . L'aplicació s'anomena "Statistical Analysis of Rates and Trends (SART)", i es pot trobar en la pàgina: <http://rht.iconcologia.net/stats/sart.html>.

El text està en anglès per tal de facilitar l'accés internacional. A continuació es mostra quin aspecte té cada pantalla, així com quins paràmetres d'entrada es necessiten en cada cas.

Figura 5.1. Pàgina principal on es troben les aplicacions.



En la primera pàgina (Figura 5.1) es mostren les quatre opcions a triar, la descriptiva de les dades ("Descriptive"), el Percentatge de Canvi Anual Estimad ("Estimated Annual Percent Change"), la Raó d'Incidència o Mortalitat Estandarditzada ("Standardized Incidence or Mortality Ratio") i el càlcul de casos Esperats ("Expected cases"). En cadascuna de les opcions s'ha de carregar un conjunt de dades, així com fixar els paràmetres d'ajust.

Els resultats s'obtenen per a cadascun dels grups (en el nostre cas, localitzacions tumorals i sexe) especificats en el fitxer d'entrada.

5.1. APLICACIÓ DESCRIPTIVA

L'aplicació Descriptiva retorna per a cada grup prèviament especificat, el nombre de casos, el percentatge (opcional), la CR, la ASR ajustada a la població seleccionada (mètode d'estandardització directe), la TR (en els grups desitjats) i la CumR (també en els grups desitjats) .

La Figura 5.2 i 5.3. ens mostren el pas 1 i el pas 2 respectivament de l'aplicació descriptiva.

Figura 5.2. Pàgina de l'aplicació descriptiva.

Descriptive (beta version)

This application allows user automatic calculation of a table describing the percentage (%), number of cases (N), crude rate (CR), age standardized rate (ASR), as user-defined reference population, the truncated ASR rate (TR) and the cumulative rate (CumR). Both rates may be truncated or accumulated for the age groups that are more interesting for the study. The period of analysis can be selected by the user.

Methodology (PDF).

Files must be ASCII type, ";" separated and unquoted values (i.e: ".txt" or ".csv" files). Variables included in these files must be in the following order:

Data file (click here to view an example):
a) sex: 1=Male; 2=Female; 0=Both sexes. It is not necessary to include all the categories.
b) age.group: ID number of the age group. Age group must be matched with that from the age group file (defined below).
c) year: Year or reference year (in case of period).
d) group: Name of the groups to study. For example, when studying cancer sites it could be "Bladder", "Pancreas", "All sites",... Or when studying a pathology by geographical areas, it could be "Italy", "France", "Spain", etc.
e) cases: Number of cases in each group.
f) population: Population at risk corresponding to each group.

Age group file. Age groups must be consecutive and homogeneous (click here to view an example):
a) ged: ID number of the age group.
b) ged.min: Minimum age group.
c) ged.max: Maximum value of the age group. **If it has no maximum value, e.g. 85+, should be entered as "NA".**

Weights file (click here to view an example):
a) age.group: ID number of the age group. Age group must be matched with that from the age group file (defined below).
b) W: Weights of the reference population.
Note: You can rename the W column to your preference. Column values must sum 1.

NOTE: This process could take a few minutes!

Title (ex: Spain Cancer incidence):

Name of data file:

Name of age groups file:

Name of weights file:

En el pas 1 (Figura 5.2) es dona un títol (*Title*) per a l'anàlisi descriptiva, es carrega el fitxer de dades (*Name of data file*), el fitxer de grups d'edat (*Name of age groups file*) i el de pesos (*Name of weights file*).

Un cop introduïdes les dades en el pas 1 (Figura 5.2.), passem a configurar el pas 2 (Figura 5.3).

Figura 5.3. Pàgina de l'aplicació descriptiva (pas 2).

The screenshot shows a web browser window titled "Registre Hospitalari de Tumors ICO/CSUB - Windows Internet Explorer". The address bar shows the URL "http://rht.iconcologia.net/stats/configure.php". The page content is titled "Descriptive Configuration" and includes the following fields and options:

- Starting year // Ending year:** Two dropdown menus, both set to "1985".
- Groups to truncate:** Two dropdown menus, both set to "1 (0-4)".
- Groups to accumulate:** Two dropdown menus, both set to "1 (0-4)".
- Denominator group, used to calculate percentages (ex: All sites but non melanoma of skin):** A dropdown menu set to "No Percentage".
- Groups excluded from percentage calculation (ex: "All sites", "Colon", use ctrl+click for multiple selection):** A list box containing "Bufeta de l'orina", "Cavitat Oral i Faringe", "Coll Uteri", and "Còlon".
- Select graphs to obtain:** A list of checkboxes:
 - Number of cases ranking
 - Crude rate ranking
 - Age standardized rate ranking
 - Truncate rate ranking
 - Cumulative rate ranking
- The results will be sent to the following e-mail:** An empty text input field.
- Buttons:** "Get Results" and "Reset".

At the bottom of the page, there is a copyright notice: "© 2008-2009 Institut Català d'Oncologia. Tots els drets reservats."

En el pas 2 (Figura 5.3) es selecciona l'any d'inici (*Starting year*) i l'any de finalització (*Ending year*) del període d'estudi. Així mateix, es selecciona els grups d'edat pels quals es vol realitzar la TR (*Groups to truncate*) especificant el grup d'edat mínim i el màxim per on es vol truncar i els grups d'edat per calcular la CumR (*Groups to accumulate*). També hi ha l'opció d'escollir el grup pel qual es vol calcular els percentatges (*Denominator group, used to calculate percentages*) i d'excloure aquells grups que no es vol calcular percentage (*Groups excluded form percentage calculation*). També es poden seleccionar els gràfics (tipus rànquing) que a l'usuari estigui interessat. Finalment, només resta especificar l'adreça del correu electrònic (e-mail) on es vol rebre les dades i obtenir els resultats (*Get Results*).

Per tant, els paràmetres d'entrada per a aquesta aplicació són els següents:

1. **Títol (Title):** text que es desitgi introduir sobre l'anàlisi que es realitza.
2. **Un fitxer de dades (Data file)** que conté les següents columnes:
 - a. *sex*: 1 = Homes; 2 = Dones; 0 = Ambdós sexes. No es necessari introduir totes les categories, únicament les que l'usuari desitgi analitzar.
 - b. *age.group*: Numero identificador del grup d'edat. Els grups d'edat han de coincidir amb els del fitxer "age group file" (definit a sota).
 - c. *year*: Any o any de referència (en el cas de període).
 - d. *group*: Noms dels grups en estudi. Per exemple, quan estudiem localitzacions per càncer serà: "Pulmó", "Pàncrees", "Totes les localitzacions", entre d'altres. O quan estudiem una patologia concreta on el grup siguin les àrees podem tenir: "Itàlia", "França", "Espanya", entre d'altres. Així mateix, també podríem estudiar una patologia per raça, amb el qual tindríem: "Blancs", "Negres", "Hispànics", "Asiàtics", entre d'altres. Per tant, no només ens basem en l'exemple d'estudi, en el que s'analitza morts per càncer i per tant els grups són el conjunt de localitzacions, sinó que també es poden analitzar altre tipus de grups.
 - e. *cases*: Nombre de casos en cadascun dels grups.
 - f. *population*: Població a risc corresponent a cada grup.

Aquest fitxer té l'aspecte de la Taula 5.1 i pot ser en format ".txt" o ".csv" :

Taula 5.1. Exemple del fitxer de dades.

sex	age.group	year	group	cases	population
2	1	1995	Bladder	0	133142
2	2	1995	Bladder	0	140745
2	3	1995	Bladder	0	174170
...
1	16	1999	All sites excluding skin non-melanoma	1565	88320
1	17	1999	All sites excluding skin non-melanoma	1057	46132
1	18	1999	All sites excluding skin non-melanoma	955	32705

3. **Un fitxer amb els grups d'edat (Age group file)** que conté les següents columnes:
 - a. *ged*: Numero identificador del grup d'edat.
 - b. *ged.min*: Mínim valor del grup d'edat.
 - c. *ged.max*: Màxim valor del grup d'edat. Si no està definit excitament el valor màxim, per exemple, els majors de 85 anys, aleshores s'ha d'introduir "NA".

És important que els grups d'edat siguin consecutius i homogenis, sinó es complicaria el càlcul de la CumR.

Aquest fitxer, al igual que l'anterior, també pot tenir format “.txt” o “.csv” i té l'aspecte de la Taula 5.2.

Taula 5.2. Exemple del fitxer grups d'edat.

ged	ged.min	ged.max
1	0	4
2	5	9
3	10	14
4	15	19
5	20	24
6	25	29
7	30	34
8	35	39
9	40	44
10	45	49
11	50	54
12	55	59
13	60	64
14	65	69
15	70	74
16	75	79
17	80	84
18	85	NA

4. Fitxer de pesos (*Weights file*) que conté les columnes:

- a. *age.group*: Número identificador del grup d'edat. El grup d'edat ha de coincidir amb els que s'han definit en el fitxer “*Age group file*” (definit anteriorment).
- b. *W*: Pesos corresponents a la població de referència.

Nota: Donat que els pesos es poden introduir segons la petició de l'usuari també es dona l'opció de donar-los-hi el nom desitjat. Aquesta columna ha de sumar 1.

Igual que els altres dos fitxers, aquest també pot ser “.txt” o “.csv” i un exemple el podem veure en la Taula 5.3.

Taula 5.3. Exemple del fitxer de pesos.

age.group	W
1	0.12
2	0.1
3	0.09
4	0.09
5	0.08
6	0.08
7	0.06
8	0.06
9	0.06
10	0.06
11	0.05
12	0.04
13	0.04
14	0.03
15	0.02
16	0.01
17	0.005
18	0.005

A continuació es defineixen una sèrie de paràmetres que s'han considerat per que l'usuari realitzi les anàlisis de la forma més adient segons els objectius marcats. Així mateix, seguint aquestes recomanacions l'usuari no haurà d'entrar cada vegada fitxers i "retrocedir" per a canviar els paràmetres.

- 5. Any d'inici i any final (*Starting year and Ending year*):** Es pot analitzar les dades per a més d'un període, especificant l'any d'inici i l'any final dels períodes que es vol analitzar. Si només es vol realitzar l'anàlisi per a un període concret, només caldrà indicar l'any d'inici i l'any final del mateix.
- 6. Grups d'edat pels quals es vol truncar (*Groups to truncate*):** Es demana per a quins grups d'edat es vol truncar. Per exemple, en el cas d'estudiar el càncer, pot ser que estiguem estudiant el càncer infantil i per tant seran uns grups d'edat corresponents als menors de 15 anys.
- 7. Grups d'edat pels que es vol acumular (*Groups to accumulate*):** Es demana per a quins grups d'edat es vol calcular la CumR. Aquesta taxa ens permetrà calcular el risc acumulat. De vegades és més interessant acumular per uns grups d'edat que per uns altres, per exemple, davant d'una situació on la piràmide poblacional sigui molt envellida és més interessant realitzar una CumR fins a 79 anys en lloc de fins als 74.

8. **Grup a partir del qual es vol calcular el percentatge pels altres grups (*Denominator group, used to calculate percentages*):** Aquest paràmetre ens permet triar per a quin grup volem calcular el percentatge. Si no s'especifica res, no es calcularà el percentatge.
9. **Grups que es volen excloure del càlcul del percentatge (*Groups excluded from percentage calculation*):** En el cas que es calculin percentatges, hi ha grups pels quals no interessa calcular-los i per tant, aquest paràmetre permet extreure aquells grups que no volem conèixer el percentatge. En el resultat, s'especificarà com "NA".
10. **Rànquing dels indicadors desitjats:** L'usuari pot seleccionar el gràfic del rànquing segons qualsevol dels següents paràmetres: nombre de casos, CR, ASR, TR i/o CumR.
11. **Correu electrònic (*E-mail*):** l'usuari ha d'introduir el correu electrònic per tal que l'aplicació li envii els resultats sol·licitats.

5.2. APLICACIÓ DE L'ESTIMACIÓ DEL PERCENTATGE DE CANVI ANUAL

L'aplicació EAPC (Figura 5.4) retorna una taula especificant per les variables prèviament seleccionades per l'usuari l'EAPC de les taxes estandarditzades durant el període d'estudi amb els respectius intervals de confiança.

Per a fer funcionar aquesta aplicació és necessari introduir un fitxer de dades ("*Name of data file*") similar al de l'aplicació Descriptiva, un fitxer de pesos per grup d'edat ("*Name of weights file*"), l'any inici i any final del període que es vol analitzar ("*Starting//Ending year*") i el valor de confiança amb el que es vol realitzar l'estudi. Finalment només queda introduir l'adreça electrònica per rebre els resultats.

Figura 5.4. Pàgina de l'aplicació EAPC.

Estimated Annual Percent Change (beta version)

This web application allows user to obtain the Estimated Annual Percent Change (EAPC) of the disease rates and its % confidence interval for the sex and groups defined by the user. Under the assumption of linearity on the log scale, which is equivalent to a constant change assumption, the EAPC is calculated using a generalized linear model considering a Gaussian distribution for the age-standardized rate. Let ASR_{t+1} be the age-adjusted rate at period T for which we assume Gaussian distribution. Under the assumption of linearity on the log-scale the model $\ln(E(ASR_{t+1})) = \alpha + \beta T$ is fitted, being $(e^{\beta} - 1) \cdot 100$ the EAPC. The user should select the confidence level of the EAPC confidence interval.

Methodology (PDF).

Files must be ASCII type ",", separated and unquoted values (i.e: ".txt" or ".csv" files).

Variables included in these files must be in the following order:

Data file (click here to view an example):

- sex: 1=Male ; 2=Female ; 0=Both sexes. It is not necessary to include all the categories.
- age.group: ID of the age group.
- year: Year or reference year (in case of period). You need more than one year or period to make the calculations.
- group: Name of the groups to study. For example, when studying cancer sites it could be "Bladder", "Pancreas", "All sites"... Or when studying a pathology by geographical areas it could be "Italy", "France", "Spain"...
- cases: Number of cases in each group.
- population: Population at risk corresponding to each group.

Weights file (click here to view an example):

- age.group: Age grouping must be matched with that from the age.group of data file (defined above).
- W: Weights of the reference population.

Note: You can rename the "W" column to your preference. Column values must sum 1.

Starting // Ending year: The starting and the ending year or reference year for the analysis.

Confidence value (1- α): The confidence value to calculate the confidence interval. It has to be between 0 and 1 (usually 0.95).

NOTE: This process could take a few minutes!

Name of data file:

Name of weights file:

Starting // Ending year (ex: 1995 // 2004):

Confidence value:

The results will be sent to the following e-mail:

Els paràmetres d'entrada de l'aplicació EAPC són:

1. **Un fitxer de dades (*Data file*):** Fitxer idèntic al definit en l'aplicació Descriptiva.
2. **Un fitxer amb els pesos (*Weights file*):** Fitxer idèntic al definit en l'aplicació Descriptiva
3. **Any d'inici i any final (*Starting year and Ending year*):** Es demana l'any d'inici i final del període on es vol realitzar l'anàlisi. S'ha de tenir en compte de prendre més de dos anys, sinó no es pot efectuar l'anàlisi per manca de graus de llibertat.

4. **Valor de confiança (*Confidence value*):** El valor de confiança ($1 - \alpha$) que es vol prendre per a obtenir els intervals de confiança pel EAPC. S'ha d'introduir en tant per 1. Normalment es pren 0.95.
5. **Correu electrònic (*E-mail*):** Idèntic al de l'aplicació descriptiva.

L'aplicació EAPC retorna per a cada grup especificat l'EAPC i l'interval de confiança respectiu (segons el nivell de confiança que s'introdueixi). També retorna els gràfics pertinents on es pot veure el creixement o el decreixement de les taxes i si aquest canvi és o no significatiu.

5.3. APLICACIÓ ESPERATS

L'aplicació Esperats (Figura 5.5) retorna una base de dades amb el nombre de casos projectats per a cada sexe, grup d'edat, any o període, grup (localització tumoral) i població a risc corresponents. També retorna una taula amb cadascun dels models que s'ha utilitzat per a cada grup, el valor d'AIC, el p-valor del test de bondat d'ajust, la significació estadística i el model elegit. Les projeccions (o casos esperats) s'obtenen a partir d'un MLG considerant una distribució Binomial Negativa o de Poisson segons el nombre de casos (incidents o morts) i considerant el model "*age-drift*" o "*age-specific-slope*". En total s'han provat 4 models diferents.

En l'aplicació cal introduir un fitxer de dades ("*Name of data file*") i un fitxer de població ("*Name of population file*") i el correu electrònic de l'usuari.

Figura 5.5. Pàgina de l'aplicació esperats.

Expected cases (beta version)

The purpose of this application is to produce the expected number of cases (EC) for a certain population based on the time trends of the data file (DF) population. This certain population could be the same of the DF but for another time period. In this situation, the EC are the projected ones for our population in a future time period. This application can be also used to predict the number of cases in another area taking into account the time trend of the DF population. For example, if our DF contains incidence data of a certain region of Catalonia with cancer registry and using these data we want to predict the incidence for another region in a certain period which does not have cancer registry, we need an expected population file which must have the combination sex, age-group, time period and person-years at risk during the period and geographic zone of this area.

Methodology (PDF).

Files must be ascii type, ";" separated and unquoted values (i.e: ".txt" or ".csv" files).

Variables included in those files are:

Data file (click here to view an example):
a) sex: 1=Male ; 2=Female ; 0=Both sexes. It is not necessary to include all the categories.
b) age.group: ID number of the age group. Age groups must be consecutive and homogeneous in order to use the descriptive function afterwards.
c) year: Year or reference year (in case of period).
d) group: Name of the groups to study. For example, when studying cancer sites it could be "Bladder", "Pancreas", "All sites"... Or when studying a pathology by geographical areas it could be "Italy", "France", "Spain"....
e) cases: Number of cases in each group.
f) population: Population at risk corresponding to each group.

Expected population file (click here to view an example):
a) sex: 1=Male ; 2=Female ; 0=Both sexes. The same categories as Data file.
b) age.group: It has to match the age values from the age group of data file (defined above).
c) year: Year or reference year (in case of period).
d) population: Population at risk to study projected values. It would be a future period or/and other zone.

NOTE: This process could take a few minutes because it tests 4 models for each group and sex.

Name of data file:

Name of population file:

The results will be sent to the following e-mail:

Els paràmetres per a l'aplicació esperats són els següents:

1. **Un fitxer de dades (*Data file*):** Fitxer idèntic al de l'aplicació Descriptiva.
2. **Un fitxer de la població esperada (*Population file*):** Aquest fitxer conté les següents columnes:
 - a. *sex*: 1 = Homes; 2 = Dones; 0 = Ambdós sexes.
 - b. *age.group*: Numero identificador del grup d'edat. Els grups d'edat han de coincidir amb els del fitxer "data file".
 - c. *year*: Any o any de referència (en el cas de període).

- d. *population*: Població a risc del lloc i període en el que es vol projectar. Si és un període futur, haurà de ser la població estimada d'aquell període.

Aquest fitxer pot tenir format “.txt” o “.csv”. En la Taula 4.8 es pot veure un exemple.

Taula 5.4. Exemple del fitxer de població esperada.

sex	age.group	year	population
1	1	2010	238540
1	2	2010	211672
1	3	2010	178636
...
2	16	2015	138975
2	17	2015	135362
2	18	2015	152490

3. **Correu electrònic (*E-mail*):** Idèntic al de l'aplicació descriptiva.

5.4. APLICACIÓ DE LA RAÓ D'INCIDÈNCIA O MORTALITAT ESTANDARDITZADA

Aquesta aplicació (Figura 5.6) permet calcular la raó d'incidència o mortalitat estandarditzada per al conjunt de causes en l'estudi tenint en compte una població de referència i el seu interval de confiança.

L'aplicació demana un fitxer de dades (“*Name of data file*”), un fitxer amb la població de referència (“*Name of reference population file*”), el valor de confiança amb el qual es vol treballar (“*Confidence value*”) i el correu electrònic de l'usuari.

Figura 5.6. Pàgina de l'aplicació SIMR

SIMR: Standardized Incidence or Mortality Ratio (beta version)

This tool allows automated calculation of the standardized incidence or mortality ratio (SIMR) and its confidence interval. The SIMR is the ratio of cases observed versus the expected.

[Methodology \(PDF\)](#).

Files must be ASCII type, " separated and unquoted values (i.e: ".txt" or ".csv" files).

Variables included in these files must be in the following order:

Data file (click here to view an example):

- sex: 1 = Male; 2 = Female; 0 = Both sexes. It is not necessary to include all the categories.
- age.group: ID number of the age group.
- year: Year or reference year (in case of period).
- group: Name of the groups to study. For example, when studying cancer sites it could be "Bladder", "Pancreas", "All sites"... Or when studying a pathology by geographical areas it could be "Italy", "France", "Spain"...
- cases: Number of cases in each group.
- population: Population at risk corresponding to each group.

Reference population file (click here to view an example):

- sex: 1 = Male; 2 = Female; 0 = Both sexes. The same categories as Data file.
- age.group: It has to match the age values from the age group of data file (defined above).
- year: Year or reference year (in case of period).
- group: Name of the groups to study. For example, when studying cancer sites it could be "Bladder", "Pancreas", "All sites"... Or when studying a pathology by geographical areas it could be "Italy", "France", "Spain"...
- cases: Number of cases in each group.
- population: Population at risk corresponding to each group.

Confidence value(1-0): The confidence value to calculate the confidence interval. It has to be between 0 and 1 (usually 0.95).

NOTE: This process could take a few minutes!

Name of data file:

Name of reference population file:

Confidence value:

The results will be sent to the following e-mail:

A aquesta aplicació, se li passen tres paràmetres:

- 1. Un fitxer de dades (*Data file*):** Fitxer idèntic al de l'aplicació Descriptiva.
- 2. Un fitxer amb la població de referència (*Reference population file*):** Aquest fitxer ha de contenir les següents columnes:
 - sex: 1 = Homes; 2 = Dones; 0 = Ambdós sexes. No és necessari incloure cada categoria de sexe, però sí que hi ha d'haver les mateixes categories que en el fitxer "data file" (definit a dalt).
 - age.group: Numero identificador del grup d'edat. Els grups d'edat han de coincidir amb els del fitxer "data file" (definit a dalt).
 - year. Any o any de referència (en el cas de període).

- d. *group*: Noms dels grups en estudi. S'ha de correspondre amb el "group" del fitxer "data file"
- e. *cases*: casos de la població de referència.
- f. *Population*: Població a risc de la població de referència.

Aquest fitxer pot tenir format ".txt" o ".csv" i té l'aspecte que es mostra en la Taula 4.9:

Taula 5.5. Exemple del fitxer de la població de referència.

Sex	age.group	year	group	cases	population
1	1	2003	Cavitat Oral i Faringe	0	1057260
1	1	2003	Esòfag	0	1057260
1	1	2003	Estómac	0	1057260
...
2	18	2004	Malaltia immunoproliferativa o Mieloma múltiple	159	540992
2	18	2004	Leucèmies	270	540992
2	18	2004	Totes les localitzacions	7209	540992

3. Correu electrònic (*E-mail*): Idèntic al de l'aplicació descriptiva.

L'aplicació SIMR retorna per a cada sexe i grups prèviament especificats la SIMR i el seu respectiu interval de confiança (segons la confiança que s'introdueixi). També retorna els gràfics pertinents on es pot veure l'excés o disminució del risc i si aquest és o no estadísticament significatiu.

6. RESULTATS

Totes les aplicacions, com ja s'ha comentat anteriorment, es troben a la web de "SART" (<http://rht.iconcologia.net/stats/sart.html>) i retornen els resultats mitjançant un correu electrònic a l'adreça que prèviament l'usuari a introduït com a paràmetre.

En aquest apartat, per a cadascuna de les aplicacions del "SART" es presenten exemples que contenen l'objectiu, la preparació del fitxer d'entrada, la petició de les dades, els resultats i la interpretació d'aquests.

6.1. APLICACIÓ DESCRIPTIVA

L'aplicació descriptiva retorna un document adjunt amb format ".pdf" i un altre amb "html" que contenen les taules dels resultats sol·licitats segons les peticions especificades per l'usuari. Així mateix, també es retornen, amb format ".jpg" els gràfics que l'usuari a sol·licitat en l'aplicació.

6.1.1. EXEMPLE

OBJECTIU: Es vol descriure la mortalitat per càncer a Catalunya segons sexe i localització tumoral en el període 1995-2004.

FITXER D'ENTRADA: Tenint en compte l'objectiu descrit, entrem a la pàgina on s'ubica la l'aplicació descriptiva (<http://rht.iconcologia.net/stats/csd.html>) i inclouem els fitxers d'entrada preparats prèviament seguint els criteris definits a l'apartat 5.1, (Aplicació descriptiva).

PETICIÓ DE LES DADES: Una vegada inclosos els fitxers d'entrada l'usuari configura la sortida dels resultats. S'ha seleccionat el període d'estudi (1995-2004), els grups d'edat per a obtenir la taxa de mortalitat per càncer truncada desitjada (35-64 anys) i els grups d'edat per a la CumR (0-74 anys). Així mateix, l'aplicació permet seleccionar el denominador pel càlcul del percentatge (Totes les localitzacions excepte la pell no Melanoma) i les localitzacions tumorals que volem excloure del percentatge (Totes les localitzacions). Es selecciona el rànquing per al nombre de casos i la ASR i, per acabar, especifiquem el nostre correu electrònic. (Veure l'apartat 5.1, Aplicació descriptiva).

RESULTATS: Aquesta aplicació retorna un fitxer “html” i un “.pdf” on apareix cadascuna de les taules. En aquest exemple s’obtenen les taules 6.1 i 6.2 (una per cadascun dels sexes introduït al fitxer d’entrada). Els fitxers “.jpg” contenen les figures dels rànquings de les localitzacions tumorals (Figures 6.1, 6.2, 6.3 i 6.4).

Taula 6.1. Descriptive Results. Catalonia cancer mortality descriptive. Period: 1995-2004 . Age from 0 to + 85 years. Women

	%	N	CR	ASR(W)	TR(W)	CumR
Bufeta de l'orina	2.13	1184	3.69	1.09	0.59	0.09
Cavitat Oral i Faringe	1.09	608	1.89	0.85	1.38	0.09
Coll Uteri	1.53	851	2.65	1.51	3.15	0.16
Colon	11.96	6654	20.73	8.00	9.87	0.82
Colorrectal	15.36	8546	26.63	10.27	12.72	1.05
Endometri	2.36	1312	4.09	1.66	1.91	0.20
Esofag	0.71	397	1.24	0.49	0.60	0.05
Estomac	6.18	3438	10.71	4.03	4.68	0.40
Fetge	4.69	2610	8.13	2.97	2.37	0.34
Laringe	0.17	93	0.29	0.15	0.31	0.02
Leucemies	3.97	2208	6.88	3.14	3.23	0.30
Limfoma de Hodgkin	0.34	188	0.59	0.31	0.31	0.03
Limfoma no Hodgkin	3.41	1899	5.92	2.52	3.01	0.27
Malaltia immunoproliferativa o Mieloma multiple	2.60	1444	4.50	1.73	1.89	0.19
Mama	17.86	9935	30.95	15.58	29.64	1.72
Melanoma de pell	0.93	517	1.61	0.81	1.44	0.08
Ovari	5.08	2826	8.80	4.26	7.23	0.50
Pancrees	5.47	3043	9.48	3.64	4.00	0.40
Pulmo	5.75	3198	9.96	4.76	8.41	0.53
Recte	3.40	1892	5.89	2.27	2.85	0.23
Ronyo	1.66	926	2.88	1.20	1.46	0.12
Sistema Nervios Central	2.82	1568	4.89	2.74	4.32	0.31
Tiroides	0.46	255	0.79	0.30	0.32	0.03
Uter	5.03	2797	8.71	3.97	6.22	0.44
Totes les localitzacions	NA	55812	173.88	74.66	104.66	7.90
Totes les localitzacions excepte pell no Melanoma	100.00	55634	173.33	74.51	104.60	7.90

%: Percentage, N: Number of cases , CR: Crude Rate x 100000 person-years, ASR: Age Standardized Rate x 100000 person-years, TR: Truncated Rate x 100000 person-years by the 35-64 age group, CumR: Cumulative Rate (percent) by the 0-74 age group, (.): Reference Population.

Taula 6.2. Descriptive Results. Catalonia cancer mortality descriptive. Period: 1995-2004 . Age from 0 to + 85 years. Men

	%	N	CR	ASR(W)	TR(W)	CumR
Bufeta de l'orina	5.54	5075	16.47	8.25	7.29	0.84
Cavitat Oral i Faringe	3.09	2828	9.18	6.11	13.56	0.72
Colon	8.70	7970	25.87	13.39	14.25	1.39
Colorrectal	11.71	10731	34.83	18.19	19.78	1.93
Esofag	2.78	2545	8.26	5.27	10.60	0.63
Estomac	5.90	5401	17.53	9.50	11.92	1.05
Fetge	5.39	4936	16.02	8.66	9.90	1.07
Laringe	2.60	2382	7.73	4.73	8.68	0.57
Leucemies	2.95	2704	8.78	5.13	4.77	0.49
Limfoma de Hodgkin	0.29	268	0.87	0.61	0.93	0.06
Limfoma no Hodgkin	2.24	2056	6.67	3.90	5.51	0.42
Malaltia immunoproliferativa o Mieloma multiple	1.53	1397	4.53	2.32	2.32	0.25
Melanoma de pell	0.67	615	2.00	1.22	2.09	0.13
Pancrees	3.81	3490	11.33	6.40	9.10	0.74
Prostata	9.16	8394	27.25	12.10	4.12	0.87
Pulmo	27.73	25400	82.45	47.41	71.28	5.77
Recte	3.01	2761	8.96	4.80	5.53	0.54
Ronyo	1.88	1725	5.60	3.05	3.92	0.34
Sistema Nervios Central	2.18	1995	6.48	4.26	6.65	0.48
Testicle	0.04	40	0.13	0.11	0.16	0.01
Tiroides	0.13	122	0.40	0.22	0.22	0.03
Totes les localitzacions	NA	91868	298.21	165.36	216.16	18.27
Totes les localitzacions excepte pell no Melanoma	100.00	91606	297.36	164.95	215.85	18.24

%: Percentage, N: Number of cases , CR: Crude Rate x 100000 person-years, ASR: Age Standardized Rate x 100000 person-years, TR: Truncated Rate x 100000 person-years by the 35-64 age group, CumR: Cumulative Rate (percent) by the 0-74 age group, (.): Reference Population.

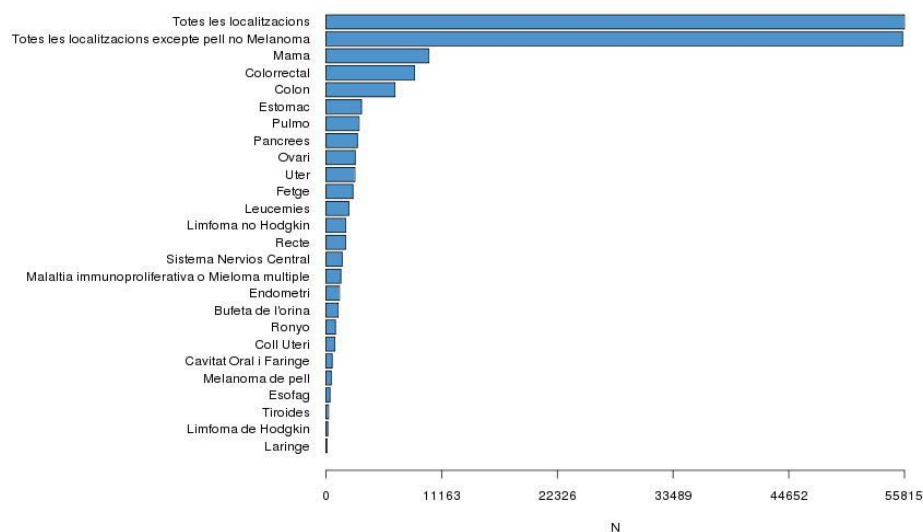
Figura 6.1. Number of cases. Women, period 1995-2004.

Figura 6.2. Age Standardized Rate per 100000 persons-year. Women, period 1995-2004.

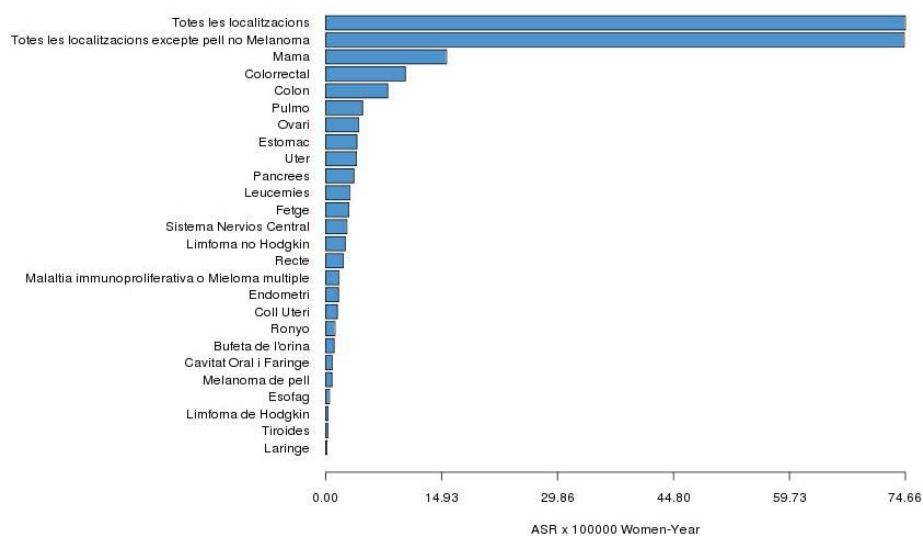


Figura 6.3. Number of cases. Men, period 1995-2004.

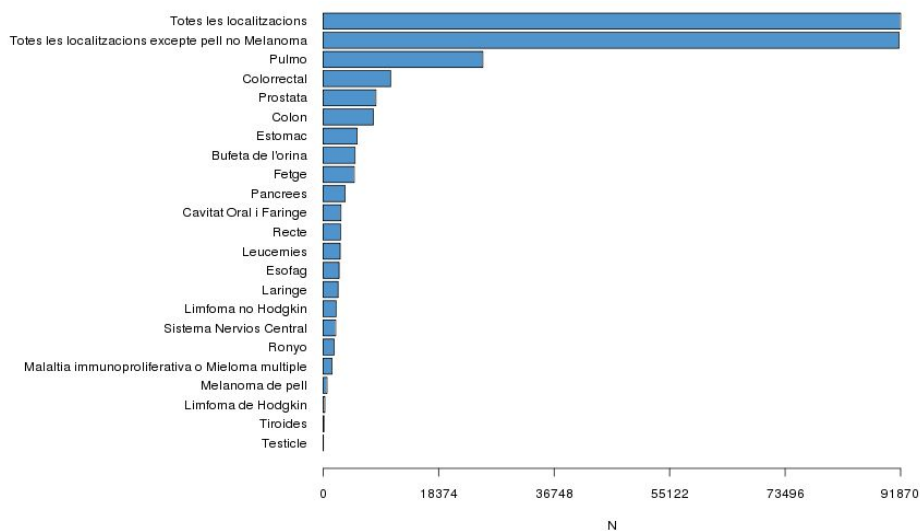
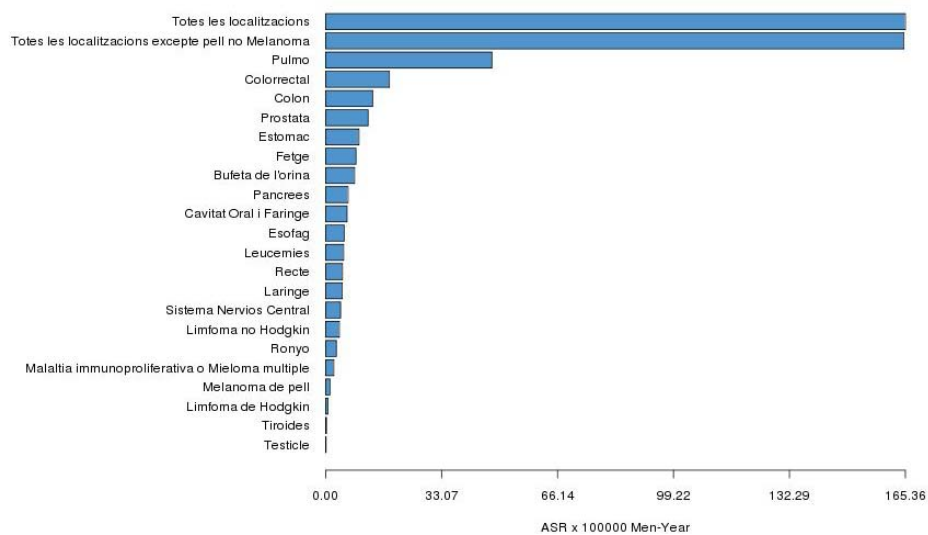


Figura 6.4. Age Standardized Rate per 100000 persons-year. Men, period 1995-2004.



INTERPRETACIÓ:

El percentatge (%) expressa la freqüència relativa que té un succés respecte a un altre, per tant, d'aquí es pot extreure quines són les localitzacions tumorals més freqüents. Les xifres absolutes (N) donen una idea sobre la magnitud o volum real del succés. La CR és un quocient que representa la probabilitat matemàtica d'ocurrència d'un succés en una població i temps definit. La ASR, s'utilitza per a comparar xifres observades en àrees ó períodes. La ASR reflecteix quina taxa s'observaria en la població estàndard (o de referència) de certa estructura d'edats si en aquesta existís una força de mortalitat com la de la població en estudi. La TR s'interpreta com l'ajustada però tenint en compte certs grups d'edat. La CumR és una estimació del risc acumulat (CRisk). Quan els valors són petits (<10%) es correspon amb el risc acumulat, i de fet, es pot convertir fàcilment mitjançant la formula: $CRisk=1-exp(CumR)$. El risc acumulat s'interpreta com el risc de que un individu desenvolupi o mori per certa causa durant un període de temps.

Segons l'exemple, a Catalunya en el període 1995-2004 van morir 55634 dones per càncer, essent el càncer de mama la causa de mortalitat per càncer més freqüent (9935; 17.86%). En segona posició hi ha el càncer colorrectal (8546; 15.36%), seguit del d'estómac (3438; 6.18%) i pulmó (3198; 5.75%). La CR de mortalitat per càncer a Catalunya va ésser de 173.33 per 100000 dones-any, la ASR ajustada a la població mundial de 74.51 per 100000 dones-any i la TR entre 35 i 64 anys de 104.60. Les taxes ajustades a la població mundial de les principals localitzacions tumorals varen ser les dels càncers de mama (ASR=15.58), colorrectal (ASR=10.27), pulmó (ASR= 4.76), ovari (ASR=4.26) i el d'estómac (ASR= 4.03). El risc de que una dona mori per càncer al llarg de la seva vida a Catalunya és del 7.9%, sent la probabilitat de morir per càncer de mama del 1.72% (Taula 6.1 i Figures 6.1 i 6.2).

En canvi, en el període 1995-2004 van morir 91606 homes per càncer, essent el càncer de pulmó (25400; 27.73%) la principal causa tumoral de mortalitat, seguida pel càncer colorrectal (10731; 11.71%), el de pròstata (8394; 9.16%) i el d'estómac (5401; 5.90%). La CR de mortalitat de càncer va ser de 297.36 per 100000 homes-any, mentre que la ASR va ser de 164.95 per 100000 homes-any i la truncada entre 35 i 64 anys, de 215.85. Segons localització tumoral i ajustant a la població mundial, les principals causes de mort varen ser els càncers de pulmó (47.41 per 100000 homes-any), colorrectal (ASR=18.19), pròstata (ASR= 12.10), estómac (ASR=9.50) i càncer de fetge (ASR= 8.66. El risc de que una home mori per càncer en el període 1995-2004 és del 18.24%, sent la probabilitat de morir per càncer de pulmó del 5.77% (Taula 6.2 i Figures 6.3 i 6.4.

6.1.2. ALTRES UTILITATS DE L'APLICACIÓ

En l'exemple exposat es pren com a grup d'anàlisi les diferents localitzacions tumorals. En l'Annex 2.1 es mostra una anàlisi estadística realitzada sobre una base de dades procedent de l'IDESCAT on analitzem la procedència dels immigrants residents a Catalunya en el període 2000-2004.

Tal i com s'ha descrit prèviament en l'apartat 5.1 (Aplicació descriptiva de material i mètodes), l'usuari pot analitzar les dades modificant la selecció de paràmetres. Així, per exemple, es pot realitzar l'anàlisi per diferents períodes o anys concrets. Pot ser interessant comparar el període 1995-1999 i el 2000-2004 per tal de veure possibles canvis entre aquests dos períodes. Això es realitza introduint els mateixos fitxers de dades i canviant els paràmetres d'any inici i any final (Els resultats es mostren en l'Annex 2.2 i 2.3).

En l'exemple actual es prenen com a mostra l'edat compresa entre els 35 i 64 anys per al càlcul de la TR, però podríem seleccionar un altre grup de mortalitat com per exemple el grup corresponent a la mortalitat infantil (0 i 14 anys). Els resultats es mostren en l'Annex 2.2 i 2.3. Si la població d'estudi presenta una piràmide poblacional envellida, podríem seleccionar per calcular la TR el grup d'edat corresponent als 0 a 79 anys. Un exemple es mostra en l'Annex 2.4.

Si a l'usuari no li interessa que l'aplicació calculi el percentatge, existeix la opció de no seleccionar-ho. D'altra banda, si a l'usuari li interessa calcular el percentatge però no per tots els grups, aquest podrà triar aquells grups pels quals **no** vol que es calculi.

En l'Annex 2.5, es poden trobar altres possibles resultats sobre les dades de mortalitat per càncer a Catalunya que es desprenen de l'aplicació descriptiva.

6.2. APLICACIÓ DE L'ESTIMACIÓ DEL PERCENTATGE DE CANVI ANUAL

L'aplicació del EAPC retorna un document adjunt amb format “.pdf” i un altre en “html” amb els resultats sol·licitats per l'usuari. També retorna els gràfics corresponents a l'EAPC amb els seus corresponents intervals de confiança.

6.2.1. EXEMPLE

OBJECTIU: Es vol analitzar l'EAPC en les taxes de mortalitat per càncer a Catalunya en el període 1995-2004 segons sexe i localització tumoral.

FITXER D'ENTRADA: Tenint en compte l'objectiu descrit, entrem a la pàgina on s'ubica l'aplicació EAPC (<http://rht.iconcologia.net/stats/eapc.html>) on s'inclouran els fitxers d'entrada preparats prèviament seguint els criteris definits a l'apartat 5.2 (Aplicació de l'estimació del percentatge de canvi anual).

PETICIÓ DE LES DADES: Una vegada inclosos els fitxers d'entrada l'usuari configura la sortida dels resultats. Es selecciona la confiança amb la qual es vol treballar (en aquest exemple el 95%, valor 0.95) per al càlcul dels intervals de confiança, i incloem l'adreça del correu electrònic (Veure l'apartat 5.2, Aplicació de l'estimació del percentatge de canvi anual).

RESULTATS: Aquesta aplicació retorna per e-mail un fitxer “html” i un “.pdf” on apareixen les taules (una taula per a cada sexe, veure Taules 6.3 i 6.4) i diferents arxius “.jpg” amb les figures dels EAPC i els respectius intervals de confiança, una per a cada sexe (Figures 6.5 i 6.6).

Taula 6.3. EAPC Results. Estimated Annual Percent Change for the period 1995 - 2004 . 95 % Confidence Interval. Women

	EAPC	LCI	UCI
Bufeta de l'orina	-0.92	-3.15	1.37
Cavitat Oral i Faringe	0.93	-1.65	3.58
Coll Uteri	-5.31	-7.75	-2.81
Colon	-0.70	-1.76	0.37
Colorrectal	-1.01	-1.69	-0.32
Endometri	-1.67	-3.44	0.13
Esofag	0.18	-3.14	3.62
Estomac	-3.21	-3.92	-2.49
Fetge	-3.08	-4.48	-1.67
Laringe	8.99	0.48	18.22
Leucemies	0.29	-1.16	1.77
Limfoma de Hodgkin	-5.19	-10.78	0.74
Limfoma no Hodgkin	0.98	-1.19	3.20
Malaltia immunoproliferativa o Mieloma multiple	-4.18	-5.63	-2.70
Mama	-2.69	-3.84	-1.52
Melanoma de pell	0.82	-2.15	3.88
Ovari	-1.12	-2.53	0.30
Pancrees	3.11	1.16	5.09
Pulmo	4.95	2.88	7.06
Recte	-2.12	-3.76	-0.44
Ronyo	-2.53	-5.51	0.55
Sistema Nervios Central	1.45	-0.20	3.12
Tiroides	-1.21	-5.54	3.32
Totes les localitzacions	-1.16	-1.75	-0.56
Totes les localitzacions excepte pell no Melanoma	-1.16	-1.75	-0.57
Uter	-3.91	-4.48	-3.35

EAPC: Estimated Annual Percent Change, LCI: Lower Confidence Interval, UCI: Upper Confidence Interval

Taula 6.4. EAPC Results. Estimated Annual Percent Change for the period 1995 - 2004 . 95 % Confidence Interval. Men

	EAPC	LCI	UCI
Bufeta de l'orina	2.23	0.45	4.04
Cavitat Oral i Faringe	-1.41	-3.54	0.76
Colon	0.63	-0.46	1.72
Colorrectal	0.69	-0.29	1.68
Esofag	-1.89	-3.36	-0.41
Estomac	-2.50	-3.58	-1.41
Fetge	-3.16	-4.24	-2.06
Laringe	-3.22	-4.46	-1.96
Leucemies	1.30	0.00	2.61
Limfoma de Hodgkin	-5.25	-9.82	-0.46
Limfoma no Hodgkin	1.16	-1.12	3.49
Malaltia immunoproliferativa o Mieloma multiple	-2.72	-4.68	-0.73
Melanoma de pell	0.33	-4.28	5.17
Pancrees	0.61	-1.16	2.42
Pulmo	-0.39	-1.22	0.45
Recte	0.87	-0.14	1.90
Ronyo	0.97	-0.67	2.64
Sistema Nervios Central	0.90	-1.18	3.01
Tiroides	2.22	-6.42	11.66
Totes les localitzacions	-0.67	-1.20	-0.14
Totes les localitzacions excepte pell no Melanoma	-0.68	-1.21	-0.15
Prostata	-2.57	-3.59	-1.55
Testicle	-5.32	-15.33	5.88

EAPC: Estimated Annual Percent Change, LCI: Lower Confidence Interval, UCI: Upper Confidence Interval

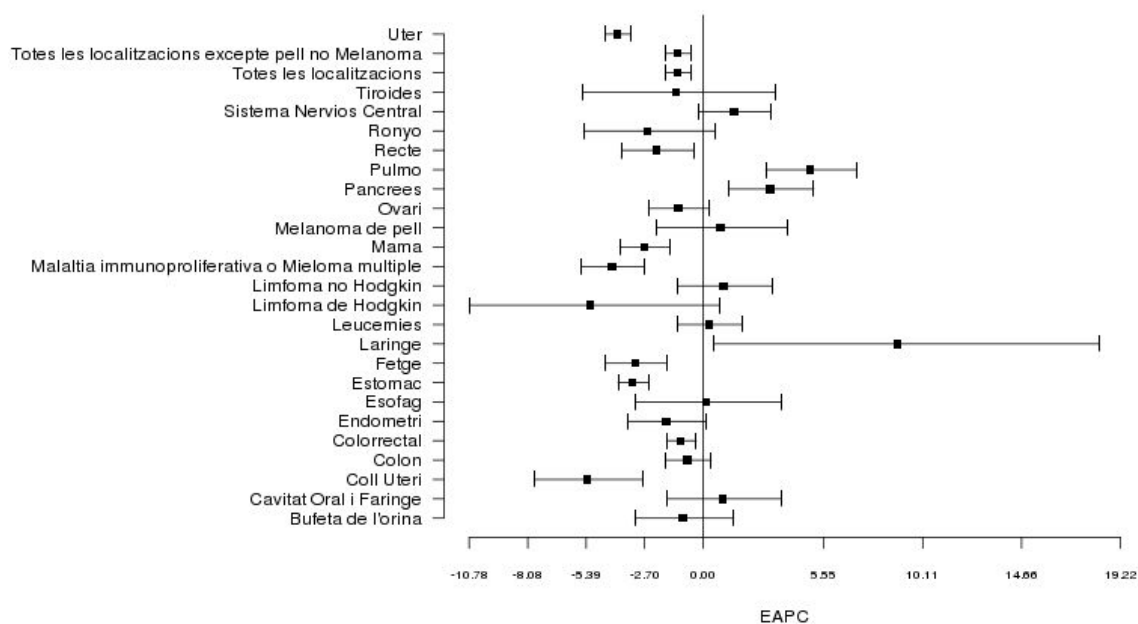
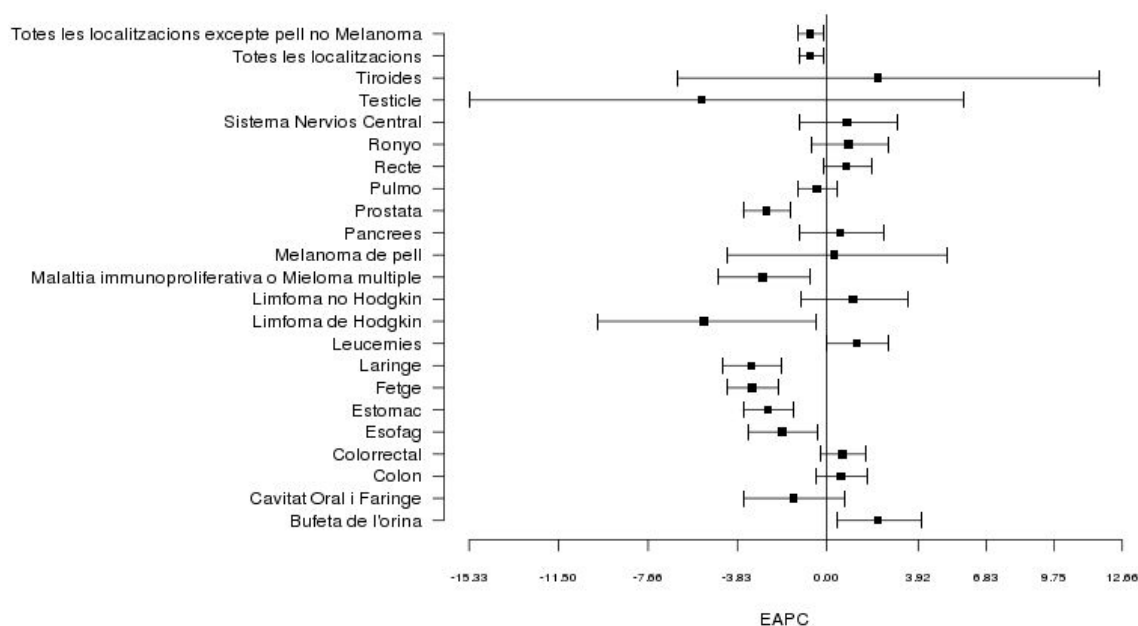
Figura 6.5. EAPC for Women. Period 1995-2004.

Figura 6.6. EAPC for Men. Period 1995-2004.



INTERPRETACIÓ:

El percentatge de canvi anual (EAPC) s'interpreta com la tendència de les taxes de mortalitat en el període d'estudi. Si el valor del EAPC és negatiu significa que les taxes de mortalitat en el període estudiat decreixen, si l'EAPC és positiu, les taxes de mortalitat augmenten. L'interval de confiança determina si la tendència és estadísticament significativa, és a dir, quan l'interval no inclou el valor 0.

Segons les dades de l'exemple, en el període 1995-2004 la mortalitat per càncer a Catalunya disminueix significativament un 1.16% anual en les dones (IC95%:1.75, -0.57). La mortalitat per càncer disminueix significativament durant el període d'estudi en les següents localitzacions tumorals: coll uterí (EAPC = -5.31%; IC95%: -7.75, -2.81), malaltia immunoproliferativa o mieloma múltiple (EAPC= -4.18%; IC95%: -5.63, 2.70), úter (EAPC= -3.91%; IC95%: -4.48, -3.35), estómac (EAPC= -3.21%; IC95%: -3.92, -2.49), fetge (EAPC= -3.08%; IC95%:-4.48, -1.67), mama [EAPC= -2.69%; IC95%: -3.84, -1.52) i colorrectal [EAPC= -1.01%; IC95%: -1.69, -0.32). D'altra banda, la mortalitat dels càncers de laringe (EAPC= 8.99%; IC95%: 0.48, 18.22), pulmó (EAPC= 4.95%; IC95%: 2.88, 7.06) i el de pàncrees (EAPC= 3.11%; IC95%: 1.16, 5.09) augmenten de forma significativa (Taula 6.3). La Figura 6.5 mostra la mateixa informació de forma més senzilla, de manera que els intervals de confiança situats a la dreta o esquerra del zero

representen respectivament augments o disminucions de la mortalitat estadísticament significatius.

Pel que fa als homes, en el període 1995-2004 la mortalitat per càncer a Catalunya decreix significativament (EAPC= -0.68%; IC95%: -1.21, -0.15). Els càncers que han mostrat una tendència de la mortalitat decreixent estadísticament significativa són el limfoma de Hodgkin (EAPC= -5.25%; IC95%: -9.82, -0.46), laringe (EAPC= -3.22%; IC95%: -0.46, -1.96), fetge (EAPC= -3.16%; IC95%: -4.24, -2.06), malaltia immunoproliferativa o mieloma múltiple (EAPC= -2.72%; IC95%: -4.68, -0.73), pròstata (EAPC= -2.57%; IC95%: -3.59, -1.55), estómac (EAPC= -2.50%; IC95%: -3.58, -1.41) i el d'esòfag (EAPC= -1.89%; IC95%: -3.36, -0.41). En canvi, pel mateix període, la mortalitat per càncer de bufeta urinària augmenta significativament (EAPC=2.23%; IC95%: 0.45, 4.04) (Taula 6.4 i Figura 6.6).

6.2.2. ALTRES UTILITATS DE L'APLICACIÓ

Tal i com ja s'ha comentat per a l'aplicació descriptiva, en l'exemple exposat es pren com a grup d'anàlisi les diferents localitzacions tumorals, però es poden analitzar qualsevol tipus de dades segons el perfil professional de l'usuari.

Com ja s'ha descrit prèviament, l'usuari pot analitzar les dades modificant els paràmetres d'entrada. Per exemple, pot ser interessant fer l'anàlisi per a diferents períodes o anys concrets, ja sigui per a comparar els dos períodes a nivell global o bé per què ens trobem en la situació on hi ha un punt de canvi en el que es marquen dues tendències diferents que es volen analitzar per separat. Un exemple que mostri això el podem consultar a l'Annex 2.6 i 2.7.

L'usuari pot analitzar les dades modificant el grau de confiança, per exemple, pot ser que per a un estudi es prengui el 95% de confiança però per a un altre, el 90%. Es pot consultar un exemple d'aquest cas particular a l'Annex 2.8. L'Annex 2.9, mostra altres resultats obtinguts a partir de les dades de mortalitat per càncer a Catalunya.

6.3. APLICACIÓ ESPERATS

L'aplicació esperats retorna dos documents adjunts, un amb format “.pdf” que conté informació sobre els models ajustats per a cadascuna de les localitzacions tumorals i

sexe i l'altre, amb format “.txt”, que conté un fitxer amb els valors esperats. Ambdós fitxers mostren els resultats segons les peticions especificades per l'usuari.

6.3.1. EXEMPLE

OBJECTIU: Es vol estimar la mortalitat per càncer a Catalunya segons sexe i localització tumoral en els anys 2010 i 2015 a partir de les dades de mortalitat del període 1995-2004.

FITXER D'ENTRADA: Tenint en compte l'objectiu descrit, entrem a la pàgina on s'ubica l'aplicació esperats (<http://rht.iconcologia.net/stats/cse.html>) on s'inclouran els fitxers d'entrada preparats prèviament seguint els criteris definits a l'apartat 5.3 (Aplicació esperats).

PETICIÓ DE LES DADES: Una vegada inclosos els fitxers d'entrada s'adjunta l'adreça del correu electrònic on s'enviaran els resultats (Veure l'apartat 5.3, Aplicació esperats).

RESULTATS: El fitxer “.pdf” especifica quin és el model seleccionat pel càlcul d'esperats i el mètode utilitzat per la selecció (Taula 6.5). El fitxer “.txt”, mostra els valors esperats (Taula 6.6).

Taula 6.5. Model selection by group and sex. Models fitted.

group	sex	model.chosen	pv.test	significant	method.AIC.GOF
1	2	Poisson age-drift	1.00	No	Minimum AIC
2	1	Poisson age-drift	0.99	No	Minimum AIC
3	2	Poisson age-drift	1.00	No	Minimum AIC
4	1	Poisson age-drift	0.96	No	Minimum AIC
5	2	Poisson age-drift	0.99	No	Minimum AIC
6	2	Poisson age-drift	0.99	No	Minimum AIC
7	1	Poisson age-drift	1.00	No	Minimum AIC
8	2	Poisson age-drift	0.99	No	Minimum AIC
9	1	Poisson age-drift	0.99	No	Minimum AIC
10	2	Poisson age-drift	1.00	No	Minimum AIC
11	2	Poisson age-drift	1.00	No	Minimum AIC
12	1	Poisson age-drift	0.99	No	Minimum AIC
13	2	Poisson age-drift	1.00	No	Minimum AIC
14	1	Poisson age-drift	0.93	No	Minimum AIC
15	2	Poisson age-drift	1.00	No	Minimum AIC
16	1	Poisson age-drift	0.36	No	Minimum AIC
17	2	Poisson age-drift	1.00	No	Minimum AIC
18	1	Negative Binomial age-drift	1.00	No	Minimum AIC
19	2	Poisson age-drift	0.35	No	Minimum AIC
20	1	Poisson age-drift	0.73	No	Minimum AIC
21	2	Poisson age-drift	1.00	No	Minimum AIC
22	1	Poisson age-drift	1.00	No	Minimum AIC
23	2	Poisson age-drift	0.99	No	Minimum AIC
24	1	Poisson age-drift	0.82	No	Minimum AIC
25	2	Poisson age-drift	1.00	No	Minimum AIC
26	1	Poisson age-drift	1.00	No	Minimum AIC
27	2	Poisson age-specific slope	0.85	No	Minimum AIC
28	2	Poisson age-drift	0.99	No	Minimum AIC
29	1	Poisson age-drift	1.00	No	Minimum AIC
30	2	Poisson age-drift	0.88	No	Minimum AIC
31	2	Poisson age-drift	0.99	No	Minimum AIC
32	1	Poisson age-drift	0.97	No	Minimum AIC
33	2	Poisson age-specific slope	0.97	No	Minimum AIC
34	1	Poisson age-specific slope	0.74	No	Minimum AIC
35	2	Poisson age-drift	1.00	No	Minimum AIC
36	1	Poisson age-drift	0.99	No	Minimum AIC
37	2	Poisson age-drift	0.65	No	Minimum AIC
38	1	Poisson age-drift	0.99	No	Minimum AIC
39	2	Poisson age-drift	1.00	No	Minimum AIC
40	1	Poisson age-drift	0.51	No	Minimum AIC
41	2	Poisson age-drift	1.00	No	Minimum AIC
42	1	Poisson age-drift	1.00	No	Minimum AIC
43	2	Poisson age-drift	0.11	No	Minimum AIC
44	1	Poisson age-specific slope	0.19	No	Minimum AIC
45	2	Poisson age-drift	0.12	No	Minimum AIC
46	1	Poisson age-specific slope	0.20	No	Minimum AIC
47	2	Poisson age-drift	1.00	No	Minimum AIC
48	1	Poisson age-drift	1.00	No	Minimum AIC
49	1	Poisson age-drift	1.00	No	Minimum AIC

group: Group of interest; sex: 0= Both sexes, 1= Men, 2= Women; model.chosen: The fitted model chosen to predict; pv.test: The p-value of the model goodness of fit test; significant: Is the preceding p.value significant?; method.AIC.GOF: The method used for model selection (AIC: Akaike Information Criterion; GOF: chi-square goodness of fit test)

Taula 6.6. Mostra del fitxer “.txt” dels valors esperats que ens retorna l'aplicació.

```

sex;age.group;year;group;cases;population
2;1;2010;Bufeta de l'orina;1;226087
2;2;2010;Bufeta de l'orina;0;201760
2;3;2010;Bufeta de l'orina;0;169395
...
2;1;2010;Colorrectal;0;226087
2;2;2010;Colorrectal;0;201760
2;3;2010;Colorrectal;0;169395
2;4;2010;Colorrectal;0;159562
2;5;2010;Colorrectal;0;173323
2;6;2010;Colorrectal;1;225698
2;7;2010;Colorrectal;2;305909
2;8;2010;Colorrectal;10;319467
2;9;2010;Colorrectal;13;288487
2;10;2010;Colorrectal;22;267192
2;11;2010;Colorrectal;36;244636
2;12;2010;Colorrectal;43;217174
2;13;2010;Colorrectal;61;206258
2;14;2010;Colorrectal;82;176133
2;15;2010;Colorrectal;99;149029
2;16;2010;Colorrectal;159;155621
2;17;2010;Colorrectal;202;126798
2;18;2010;Colorrectal;294;123646
...
1;1;2015;Sistema Nervios Central;2;222534
1;2;2015;Sistema Nervios Central;4;247956
1;3;2015;Sistema Nervios Central;3;219790
1;4;2015;Sistema Nervios Central;2;186327
1;5;2015;Sistema Nervios Central;2;179695
1;6;2015;Sistema Nervios Central;3;201022
1;7;2015;Sistema Nervios Central;5;255296
1;8;2015;Sistema Nervios Central;9;340532
1;9;2015;Sistema Nervios Central;11;352549
1;10;2015;Sistema Nervios Central;22;308910
1;11;2015;Sistema Nervios Central;26;272200
1;12;2015;Sistema Nervios Central;31;236435
1;13;2015;Sistema Nervios Central;35;201504
1;14;2015;Sistema Nervios Central;42;181839
1;15;2015;Sistema Nervios Central;43;146223
1;16;2015;Sistema Nervios Central;37;109756
1;17;2015;Sistema Nervios Central;28;91548
1;18;2015;Sistema Nervios Central;19;775931
...
1;16;2015;Testicle;0;109756
1;17;2015;Testicle;1;91548
1;18;2015;Testicle;1;77593

```

INTERPRETACIÓ:

El p-valor del test de Bondat d'Ajust informa de si hi ha prou evidència per rebutjar la hipòtesi nul·la. Tenint en compte que la hipòtesi nul·la correspon a les dades que ajusten correctament, ens interessa avaluar si el model és acceptable per a dur a terme les projeccions de la mortalitat. Pel que fa al mètode utilitzat per a la selecció del millor model (es poden veure tots els possibles models i com es seleccionen en el punt 4.4.9, Selecció dels models log-lineals), ens trobem davant de diferents situacions. La situació més favorable és aquella amb el menor valor de l'AIC i que a més a més, ajusti correctament. Com a segona opció, tindríem el model que millor ajusta i per a finalitzar, i en la pitjor de les situacions, s'obtidria un model que no ajusta i que per tant, tot i tenint el mínim AIC no permet efectuar projeccions adequades i cal explorar nous models. En quant a les estimacions es poden interpretar com el nombre de casos morts per càncer segons sexe, grup d'edat, any i localització tumoral en una població i període o any determinats. La primera columna del fitxer d'esperats es correspon al sexe, la segona, al grup d'edat, la tercera a l'any, la quarta a la localització tumoral, la cinquena al nombre de morts per càncer i la sisena a la població corresponent a l'àrea i l'any (o període) on es fan les estimacions. Per a cadascuna de les localitzacions tumorals-sexes es realitzen les estimacions de morts per càncer segons el model seleccionat.

Mitjançant les dades de l'exemple, es pot interpretar que ajustant les dades de mortalitat per càncer en el període 1995-2004 es selecciona, per a cadascuna de les localitzacions tumorals i sexe, un model amb el menor valor de l'AIC que ajusta correctament (Taula 6.5). En la majoria de casos, aquest model es correspon amb el model edat-tendència considerant la distribució de Poisson. Tan sols en 6 de les 49 situacions s'escullen altres models: el model edat-tendència considerant la distribució Binomial Negativa és el seleccionat per les estimacions de la mortalitat per càncer de laringe en els homes; el model de pendent específica per edat considerant la distribució de Poisson es selecciona per a fer les estimacions per al càncer de mama i pulmó en dones, i el de pulmó i totes les localitzacions en els homes. Als Annexes 10 i 11 es pot veure l'AIC i el p-valor del GOF per cadascuna de les localitzacions tumorals i sexe.

Per tal de poder extreure conclusions sobre els casos esperats, s'hauria d'analitzar el fitxer ".txt" que retorna l'aplicació (Taula 6.6). Tenint en compte que es disposa de l'aplicació descriptiva i que el fitxer ".txt" té el mateix format que es demana en aquesta aplicació (veure l'apartat 5.1, Aplicació descriptiva), es realitzen les anàlisis descriptives corresponents (Veure Annex 2.12). Observant l'Annex 2.12, s'aprecia que en el 2010 hi

haurà 17861 defuncions per càncer a Catalunya on 6460 es correspondran a dones (Taula A1) i 11401 a homes (Taula A2). Així mateix, en el 2015 hi haurà 19321 morts per càncer de les quals 6748 seran dones (Taula A3) i 12573 seran homes (Taula A4).

En les dones, per l'any 2010 (Annex 2.12, Taula A1), la causa principal de mortalitat per càncer serà la de mama (1061; 16.42%), seguit del càncer colorrectal (1024;15.85%), pulmó (618; 9.57%), pàncrees (513;7.94% i ovari (363;5.62%). La CR de mortalitat per càncer global serà de 172.90 per 100000 dones-any, la ASR serà de 67.79 per 100000 dones-any i la TR (dones de 35-64 anys) de 95.04. Les principals causes de mortalitat per càncer són: mama (12.06 per 100000 dones-any), colorrectal (ASR=9.47), pulmó (ASR=8.71), pàncrees (ASR=4.95) i d'ovari (ASR= 4.48). El risc de que una dona mori per càncer en l'any 2010 serà del 7.17%, essent la probabilitat de morir per càncer de mama del 1.32%.

En els homes, per l'any 2010 (Annex 2.12, Taula A2), la principal causa de mortalitat per càncer serà la de pulmó (3166; 27.77%), seguit del colorrectal (1530;13.42%), pròstata (908;7.96%), bufeta de l'orina (866;7.60%), i estómac (545;4.78%). La CR de mortalitat per càncer es de 310.94 per 100000 homes-any, la ASR de 154.25 per 100000 homes-any i la TR (de 35 a 64 anys) de 194.87 per 100000 homes-any. Segons les ASR ajustades a la població mundial les principals causes de mort per càncer són: pulmó (45.96 per 100000 homes-any), colorrectal (ASR=19.52), bufeta d'orina (ASR= 10.48), pròstata (ASR=8.96) i estómac (ASR=7.32). El risc de que una home mori per càncer a l'any 2010 és del 16.76%, essent la probabilitat de morir per càncer de pulmó del 5.46%.

En les dones, per l'any 2015 (Annex 2.12, Taula A3), la causa de mortalitat per càncer més freqüent en el 2015 serà el càncer de mama (1107; 16.40%), seguit del càncer colorrectal (1081;16.02%), pulmó (907; 13.44%), pàncrees (637;9.44%) i ovari (398;5.90%). Per tant, el rànquing corresponent a la distribució segons localització tumoral en les dones no varia entre l'any 2010 i el 2015. La CR de mortalitat per càncer serà de 173.38 per 100000 dones-any, la ASR de 64.63 per 100000 dones-any i la truncada (35 i 64 anys) de 60.67 per 100000 dones-any. Segons les ASR ajustades a la població mundial, les principals causes de mort seran: pulmó (12.86 per 100000 dones-any), mama (ASR= 10.97), colorrectal (ASR=9.08), pàncrees (ASR=5.70) i ovari (ASR=4.58). El risc de que una dona mori per càncer en l'any 2015 serà del 6.84%, essent la probabilitat de morir per càncer de mama del 1.36%.

En els homes, per l'any 2015 (Annex 2.12, Taula A4), les causes de mortalitat per càncer més freqüents serà la de pulmó (3501; 27.85%), seguit del colorrectal (1774;14.11%), bufeta d'orina (1094;8.70%), pròstata (911;7.25%) i pàncrees (581;4.62%). La CR de mortalitat per càncer serà de 328.13 per 100000 homes-any, la ASR de 149.83 per 100000 homes-any i la TR (entre 35 i 64 anys) de 186.18 per 100000 homes-any. Segons les ASR ajustades a la població mundial, les principals localitzacions tumorals seran: pulmó (45.60 per 100000 homes-any), colorrectal (ASR=20.11), bufeta d'orina (ASR=11.72), pròstata (ASR=7.75) i pàncrees (ASR=7.54). El risc de que un home mori per càncer en l'any 2015 és del 16.12%, essent la probabilitat de morir per càncer de pulmó del 5.34%.

6.3.2. ALTRES UTILITATS DE L'APLICACIÓ

Tal i com ja s'ha comentat per a la aplicació descriptiva i l'aplicació EAPC, qualsevol fitxer d'entrada degudament preparat pot ser enviat a l'aplicació.

Les estimacions demanades a l'aplicació poden ser per períodes (per exemple, estimar la mortalitat d'una població en el període 2008-2010 a partir de les dades del període 2011-2013), o bé, d'una àrea geogràfica concreta (per exemple, estimar el nombre de morts de la comarca de l'Alt Penedès a l'any 2007 a partir de les dades de Catalunya en el període 1995-2004 (veure annex 2.13).

6.4. APLICACIÓ DE LA RAÓ D'INCIDÈNCIA O MORTALITAT ESTANDARDITZADA

L'aplicació SIMR retorna dos documents, un amb format “.pdf” i l'altre amb format “html”, així com uns gràfics amb les SIMR i els corresponents intervals de confiança. Per tal d'il·lustrar l'execució de l'aplicació es presenten dos exemples.

6.4.1. EXEMPLE 1: Evolució temporal del risc de morir per càncer a Catalunya

OBJECTIU: Es vol obtenir la raó estandarditzada de la mortalitat per càncer a Catalunya segons localització tumoral i sexe en el període 2000-2004 tenint en compte la població de referència del període 1995-1999.

FITXER D'ENTRADA: Segons l'objectiu descrit, entrem a la pàgina on s'ubica l'aplicació SIMR (<http://rht.iconcologia.net/stats/simr.html>) i s'inclouran els fitxers d'entrada preparats prèviament seguint els criteris definits a l'apartat 5.4 (Aplicació SIMR).

PETICIÓ DE LES DADES: Una vegada inclosos els fitxers d'entrada configurem la sortida dels resultats. Seleccionarem el valor de confiança per treballar (0.95) i especificarem el correu electrònic (Veure l'apartat 5.4, Aplicació SIMR).

RESULTATS: Els fitxers "html" i ".pdf" mostren les taules segons el sexe. En el següent exemple, com que s'ha sol·licitats les estadístiques pels dos sexes per separat, els resultats mostraran dos taules (Taules 6.7 i 6.8). L'aplicació retorna un gràfic de la SIMR en format ".jpg" per cada sexe (Figura 6.7 i 6.8).

Taula 6.7. SIMR Results. Standardized Incidence or Mortality Ratio by sex and group. Women.

Group	CASES	SIMR	LCI	UCI
Bufeta de l'orina	609	0.93	0.86	1.01
Cavitat Oral i Faringe	319	1.00	0.90	1.12
Coll Uteri	395	0.79	0.72	0.87
Colon	3400	0.94	0.91	0.97
Colorrectal	4368	0.94	0.92	0.97
Endometri	665	0.94	0.87	1.01
Esofag	203	0.94	0.81	1.08
Estomac	1652	0.83	0.79	0.87
Fetge	1275	0.86	0.82	0.91
Laringe	53	1.22	0.91	1.59
Leucemies	1166	1.02	0.96	1.08
Limfoma de Hodgkin	90	0.84	0.68	1.04
Limfoma no Hodgkin	1051	1.13	1.06	1.20
Malaltia immunoproliferativa o Mieloma multiple	696	0.85	0.79	0.91
Mama	5003	0.93	0.90	0.96
Melanoma de pell	283	1.11	0.98	1.25
Ovari	1502	1.05	0.99	1.10
Pancrees	1717	1.17	1.11	1.22
Pulmo	1778	1.14	1.09	1.19
Recte	968	0.94	0.89	1.00
Ronyo	471	0.94	0.86	1.03
Sistema Nervios Central	850	1.11	1.03	1.18
Tiroides	123	0.84	0.70	1.01
Totes les localitzacions	28618	0.95	0.94	0.97
Totes les localitzacions excepte pell no Melanoma	28517	0.95	0.94	0.97
Uter	1336	0.83	0.79	0.88

SIMR:Standardized Incidence or Mortality Ratio, LCI:Lower Confidence Interval, UCI: Upper Confidence Interval.

Taula 6.8. SIMR Results. Standardized Incidence or Mortality Ratio by sex and group. Men.

Group	CASES	SIMR	LCI	UCI
Bufeta de l'orina	2782	1.09	1.05	1.13
Cavitat Oral i Faringe	1403	0.91	0.86	0.96
Colon	4252	1.03	1.00	1.06
Colorrectal	5744	1.04	1.01	1.07
Esofag	1274	0.93	0.88	0.98
Estomac	2662	0.88	0.85	0.92
Fetge	2395	0.86	0.83	0.90
Laringe	1163	0.88	0.83	0.93
Leucemies	1468	1.08	1.02	1.14
Limfoma de Hodgkin	120	0.74	0.61	0.89
Limfoma no Hodgkin	1141	1.14	1.07	1.21
Malaltia immunoproliferativa o Mieloma multiple	679	0.85	0.79	0.92
Melanoma de pell	336	1.10	0.99	1.22
Pancrees	1849	1.03	0.98	1.08
Prostata	4094	0.84	0.81	0.86
Pulmo	13090	0.98	0.96	0.99
Recte	1492	1.06	1.01	1.12
Ronyo	931	1.06	1.00	1.13
Sistema Nervios Central	1067	1.06	1.00	1.13
Testicle	20	0.93	0.57	1.44
Tiroides	66	1.08	0.83	1.37
Totes les localitzacions	47314	0.97	0.96	0.98
Totes les localitzacions excepte pell no Melanoma	47171	0.97	0.96	0.97

SIMR:Standardized Incidence or Mortality Ratio, LCI:Lower Confidence Interval, UCI: Upper Confidence Interval.

Figura 6.7. SIMR for Women. Period 2000-2004.

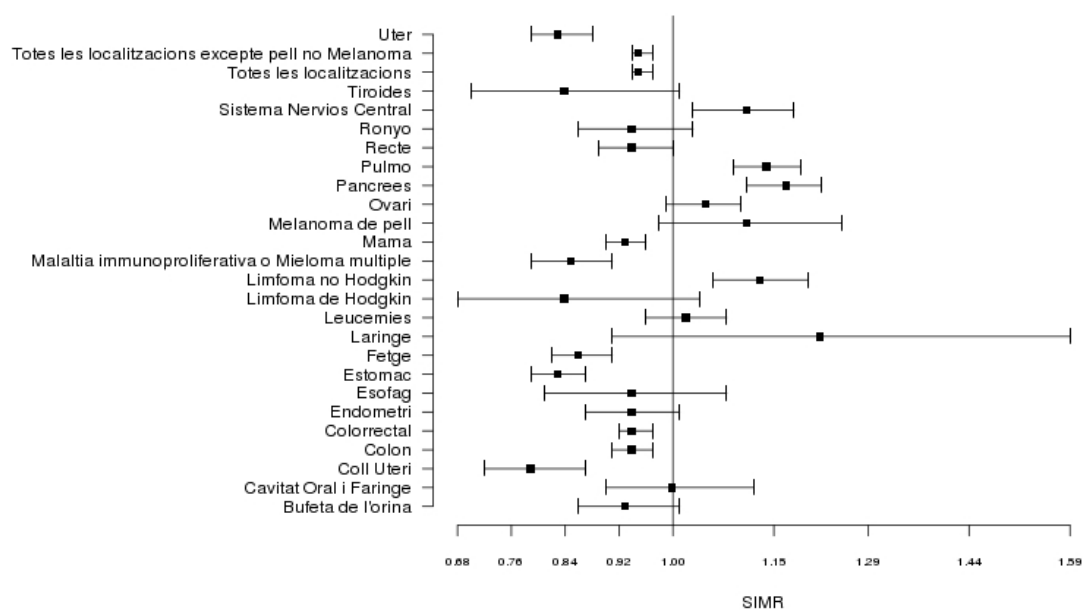
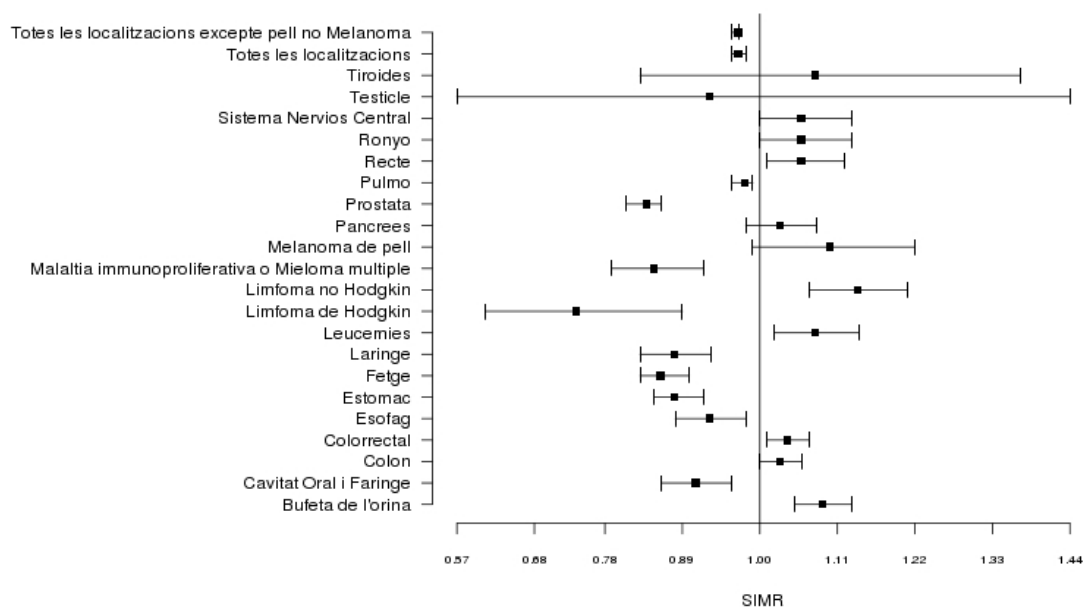


Figura 6.8. SIMR for Men. Period 2000-2004.



INTERPRETACIÓ:

La SMR mostra l'excés o disminució del risc de morir per càncer en una àrea o en un període determinats respecte a una població de referència. En aquest primer exemple, es mostra si hi ha un major o menor risc de morir per càncer a Catalunya en el període 2000-2004 respecte al període 1995-1999.

Segons els resultats de l'exemple 1, podem interpretar que en el període 2000-2004 va haver un menor risc de morir per càncer respecte a la del període 1995-1999 en ambdós sexes (SIMR dones i homes : 0.83 i 0.97, veure Taules 6.7 i 6.8 respectivament).

En quant a les dones (Taula 6.7 i Figura 6.7), el risc de morir per càncer va ser inferior en el període 2000-2004 respecte el 1995-1999 per als tumors de coll uterí, colorrectal, còlon, estómac, fetge, malaltia immunoproliferativa o mieloma múltiple, mama i úter. En canvi, detectem major risc de morir en el període 2000-2004 pel limfoma no Hodgkin, pulmó, pàncrees i sistema nerviós central.

Pel que fa als homes (Taula 6.8 i Figura 6.8), hi ha menys risc de morir en el període 2000-2004 respecte el 1995-1999 pels tumors de cavitat oral i faringe, estómac, esòfag, fetge, laringe, limfoma de Hodgkin, malaltia immunoproliferativa, mieloma múltiple, pròstata i pulmó. En canvi, es detecta un major risc de mortalitat per càncer que és estadísticament significatiu pels tumors de bufeta de l'orina, colorrectal, colon, leucèmies,

limfoma no Hodgkin, recte, ronyó i sistema nerviós central en el període 2000-2004 respecte al període 1995-1999.

6.4.2. EXEMPLE 2: Comparació de la mortalitat per càncer a Catalunya amb la d'Espanya en el mateix període.

OBJECTIU: Es vol obtenir la SMR per càncer a Catalunya segons localització tumoral i sexe en el període 2000-2004 prenent com a població de referència la mortalitat per càncer de l'Estat Espanyol en el mateix període.

FITXER D'ENTRADA: Tenint en compte l'objectiu descrit, entrem a la pàgina on s'ubica l'aplicació de la SIMR (<http://rht.iconcologia.net/stats/simr.html>) i s'inclouen els fitxers d'entrada preparats prèviament seguint els criteris definits a l'apartat 5.4 (Aplicació SIMR).

PETICIÓ DE LES DADES: Una vegada inclosos els fitxers d'entrada es configura la sortida dels resultats. S'especifica el valor de confiança es vol treballar (0.95) i s'inclou l'adreça del nostre correu electrònic (Veure l'apartat 5.4).

RESULTATS: S'obté un fitxer "html" i un ".pdf". En l'exemple, hi haurà dues taules (una per homes i l'altra per dones, Taules 6.9 i 6.10 i un gràfic de la SIMR en format ".jpg" per cada sexe (Figura 6.9 i 6.10).

Taula 6.9. SIMR Results. Standardized Incidence or Mortality Ratio by sex and group. Women.

Group	CASES	SIMR	LCI	UCI
Bufeta de l'orina	609	1.03	0.95	1.12
Cavitat Oral i Faringe	319	1.03	0.92	1.14
Coll uteri	395	0.90	0.81	0.99
Colon	3400	1.03	0.99	1.06
Colorrectal	4368	1.02	0.99	1.05
Esofag	203	1.06	0.92	1.22
Estomac	1652	0.90	0.86	0.94
Fetge	1275	1.05	1.00	1.11
Laringe	53	0.99	0.74	1.30
Leucemies	1166	1.12	1.06	1.19
Limfoma de Hodgkin	90	1.09	0.87	1.34
Limfoma no Hodgkin	1051	1.14	1.08	1.21
Malaltia immunoproliferativa o Mieloma multiple	696	1.06	0.98	1.14
Mama	5003	0.78	0.75	0.80
Melanoma de pell	283	1.05	0.93	1.18
Ovari	1502	0.83	0.79	0.87
Pancrees	1717	0.71	0.68	0.75
Pulmo	1778	1.04	0.99	1.09
Recte	968	1.00	0.94	1.07
Ronyo	471	1.10	1.00	1.21
Sistema Nervios Central	850	1.00	0.94	1.07
Tiroides	123	0.58	0.48	0.69
Totes les localitzacions	28618	1.01	1.00	1.02
Uter	1336	0.93	0.88	0.98

SIMR:Standardized Incidence or Mortality Ratio, LCI:Lower Confidence Interval, UCI: Upper Confidence Interval

Taula 6.10. SIMR Results. Standardized Incidence or Mortality Ratio by sex and group. Men.

Group	CASES	SIMR	LCI	UCI
Bufeta de l'orina	2782	1.02	0.98	1.06
Cavitat Oral i Faringe	1403	0.97	0.92	1.02
Colon	4252	1.05	1.02	1.08
Colorrectal	5744	1.05	1.02	1.07
Esofag	1274	1.03	0.97	1.09
Estomac	2662	0.92	0.89	0.96
Fetge	2395	1.05	1.01	1.09
Laringe	1163	0.88	0.83	0.93
Leucemies	1468	1.12	1.06	1.18
Limfoma de Hodgkin	120	1.12	0.93	1.34
Limfoma no Hodgkin	1141	1.14	1.07	1.21
Malaltia immunoproliferativa o Mieloma multiple	679	1.05	0.98	1.14
Melanoma de pell	336	1.05	0.94	1.16
Pancrees	1849	1.27	1.21	1.33
Prostata	4094	1.43	1.39	1.48
Pulmo	13090	1.02	1.00	1.04
Recte	1492	1.04	0.99	1.10
Ronyo	931	1.14	1.07	1.22
Sistema Nervios Central	1067	1.00	0.94	1.06
Testicle	20	0.59	0.36	0.91
Tiroides	66	1.10	0.85	1.40
Totes les localitzacions	47314	1.00	0.99	1.01

SIMR:Standardized Incidence or Mortality Ratio, LCI:Lower Confidence Interval, UCI: Upper Confidence Interval.

Figura 6.9. SIMR for Women. Period 2000-2004.

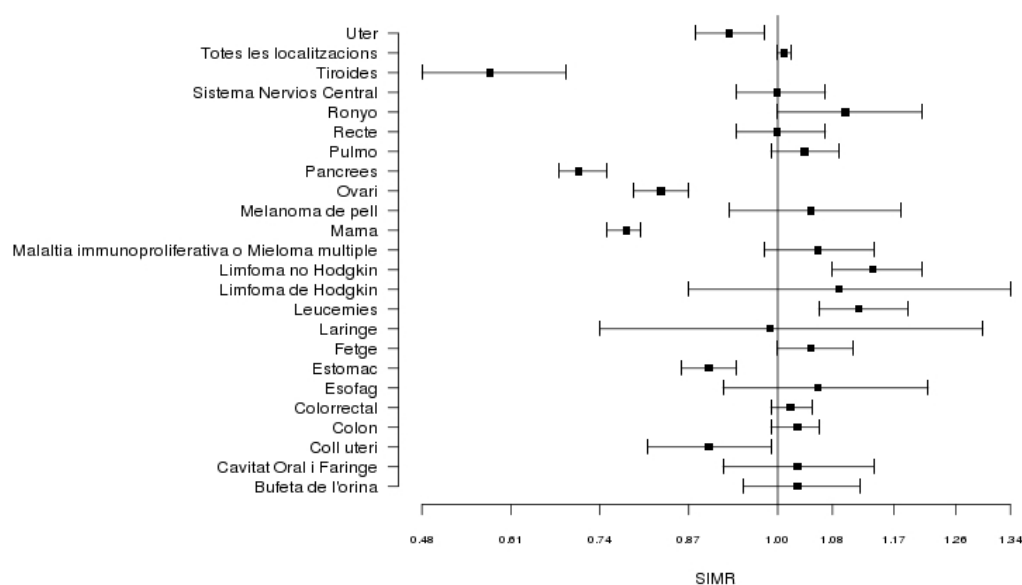
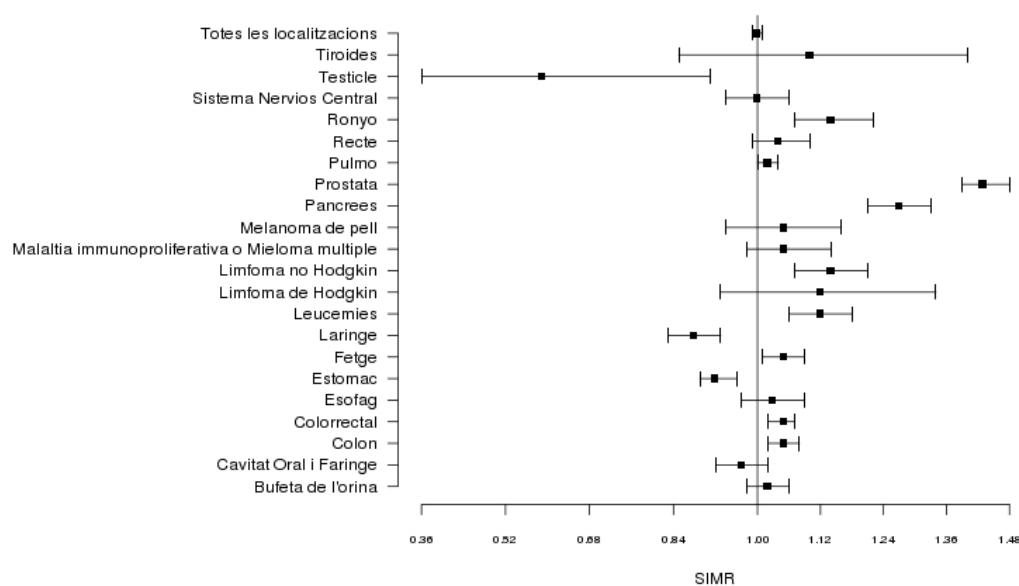


Figura 6.10. SIMR for Men. Period 2000-2004.



INTERPRETACIÓ:

En aquest exemple, es mostra l'excés o disminució del risc de la mortalitat per càncer a Catalunya en el període 2000-2004 respecte a la detectada en el mateix període a l'Estat Espanyol.

Segons els resultats de l'exemple 2 (Taula 6.9 i Figura 6.9), podem concloure que en les dones, els tumors que van presentar un menor risc de mort a Catalunya respecte a Espanya van ser coll uterí, estómac, mama ovari, pàncrees, tiroides i úter. D'altra banda a Catalunya es constata un major risc de morir per càncer de fetge, leucèmies i limfoma no Hodgkin.

Pels homes (Taula 6.10 i Figura 6.10), pel mateix període, s'observa menys risc de morir per càncer d'estómac, laringe i testicle a Catalunya respecte a Espanya i major risc de morir per càncer colorrectal (bàsicament degut al tumors de còlon), fetge, leucèmies, limfoma no Hodgkin, pròstata, pulmó, pàncrees i ronyó.

Tot i havent mostrat aquests resultats, cal tenir en compte que les dades de Catalunya són un subconjunt de les d'Espanya i això podria ocasionar que la SIMR contingüés un biaix important [49].

6.4.3. ALTRES UTILITATS DE L'APLICACIÓ

Tal i com ja s'ha comentat en les aplicacions diferents, es poden enviar dades corresponents a qualsevol perfil professional sempre i quan el fitxer d'entrada correspongui a les regles

En l'Annex 2.14 mostrem un exemple, on l'usuari ha preferit treballar amb un nivell de confiança del 90%.

7. DISCUSSIÓ

Un dels objectius principals d'aquest projecte, el qual és una de les línies del RCC del Pla Director d'Oncologia, ha estat oferir l'aplicació SART com a una eina estadística via web d'ús senzill per qualsevol usuari, independentment del seu perfil professional, que automatitzi el càlcul d'indicadors poblacionals.

L'equip del RCC té experiència en interfícies web per a càlculs estadístics automatitzats en l'àmbit de la salut pública. Exemples d'aquestes interfícies són: 1) el càlcul de supervivència observada i relativa amb l'aplicació WAERS que va permetre poder fer el càlcul de la supervivència relativa a l'Estat espanyol [21;23] atès que tenia implementades les taules de mortalitat de les províncies i comunitats autònomes, i que en una fase posterior, es va ampliar implementant les taules de mortalitat de diferents països de la Unió Europea, així com Canadà, Estats Units, Japó, Austràlia i Nova Zelanda [22] i 2) l'aplicació RiskDiff que quantifica el paper que juguen els canvis demogràfics (mida i envelliment) d'una població i els canvis en el risc de morir o desenvolupar una malaltia determinada [50] en les tendències i en les projeccions de la incidència o la mortalitat d'una malaltia.

L'RCC, amb la col·laboració del Registre de Càncer de Girona, el Registre de Càncer de Tarragona, el Registre de Càncer de Mortalitat i el Pla Director d'Oncologia va publicar l'any 2008 una monografia sobre l'impacte del càncer a Catalunya [51;52] per tal de donar a conèixer els indicadors poblacionals de càncer. Des de llavors, aquest equip és referència en la metodologia emprada en el càlcul d'alguns indicadors poblacionals i per això va sorgir la necessitat d'automatitzar aquests processos estadístics. En el present projecte, mitjançant dades procedents del Registre de Mortalitat de Catalunya, presentem el procés d'automatització del càlcul d'indicadors poblacionals via web.

Sota el nostre coneixement, les eines estadístiques que actualment es poden trobar amb lliure distribució són els següents programaris: "MIAMOD and PIAMOD" d'Itàlia [53], el "NORDCAN" dels països d'Europa del Nord [54] i el "SEER*Stat 6.5.2" d'Estats Units d'Amèrica [20]. El "MIAMOD and PIAMOD" i el "NORDCAN" només permeten fer càlculs de projeccions, no realitzen cap tipus de descriptiva, estimació del percentatge de canvi anual ni de la raó de incidència o mortalitat estandarditzada, mentre que el "SEER*Stat" permet realitzar tots aquests càlculs a excepció de les projeccions. Aquestes eines s'han

de descarregar des d'Internet i posteriorment instal·lar-les a l'ordinador. Aquests programaris presenten com a inconvenient respecte a l'eina presentada en aquest projecte que s'han d'instal·lar a nivell local, requerint per part de l'usuari certs coneixements de programació i estadística.

Amb el present projecte s'ofereix una eina estadística, SART, també de lliure disposició que introduint els fitxers d'entrada segons unes definicions preestablertes, qualsevol usuari pot triar l'aplicació pel càlcul d'indicadors poblacionals que requereixi sense la necessitat d'instal·lar els programaris al seu ordinador. A més a més, el projecte presentat mostra nombrosos exemples que serveixen com a guia d'execució i interpretació de les anàlisis estadístiques a l'usuari, facilitant la interpretació d'alguns conceptes estadístics i epidemiològics. És a dir, no es necessari tenir coneixements avançats d'estadística per realitzar i interpretar els càlculs d'indicadors.

Un altre avantatge respecte als programaris esmentats anteriorment, és que al estar disponible SART a Internet, el procés de les dades i l'entrega de resultats són ràpides. Un altre aspecte a tenir en compte és que l'ús generalitzat de l'aplicació SART permetrà una presentació estàndard de resultats en l'àmbit de la salut pública. Aquest punt és molt important, atès que impulsa l'homogeneïtzació de resultats en el camp de la salut, així com en d'altres àmbits de les ciències, ja siguin socials, demogràfiques i econòmiques, que en moltes ocasions estan relacionats amb la medicina. Altres avantatges de l'aplicació SART són, d'una banda, que estarà inclosa en una pàgina oficial de la Generalitat de Catalunya, tot i que l'adreça encara està per definir, i d'altra banda, que està escrita en anglès per tal de facilitar la seva difusió a qualsevol usuari independentment del país de procedència. Això en garanteix la seva continuïtat però a la vegada, un gran esforç per part del RCC per assegurar-ne la seva actualització.

Un dels problemes que pot generar l'automatització de processos en el càlcul d'indicadors poblacionals, com en el cas de l'aplicació SART, és que l'usuari desconfii dels resultats al desconèixer la metodologia aplicada, podent dubtar de si aquesta és la correcta pel perfil de les seves dades. Per aquesta raó, l'aplicació SART ha elaborat un breu document on s'especifica la metodologia estadística emprada, per a que l'usuari la consulti quan ho cregui oportú. D'altra banda, cal remarcar que la metodologia estadística aplicada a SART pel que fa al càlcul de les CR, ASR, TR, CumR, EAPC i la SIMR estan descrites, provades i consensuades a nivell internacional des de fa molt de temps i es poden consultar i ampliar coneixements sobre la mateixa en qualsevol publicació clàssica [40] i en pàgines web de nivell internacional [20].

On probablement pot haver-hi més discussió sobre la metodologia emprada és en el càlcul de les projeccions. Les projeccions sempre s'han d'interpretar amb certa prudència ja que s'assumeix la continuïtat de les tendències de la incidència/mortalitat observades en anys anteriors i a més, depenen de les projeccions de població futura realitzades per organismes oficials. Per a dur a terme projeccions cal triar un model estadístic adequat segons criteris de bondat d'ajust. L'aplicació SART té implementat un algorisme que permet l'elecció automàtica del millor model entre quatre possibles models, els quals duen a terme una extrapolació lineal de la tendència temporal. La metodologia NORDCAN per realitzar projeccions està basada en seleccionar el millor model entre més de 10 models possibles assumint diferents graus de polinomi en la tendència temporal [54]. La metodologia PIAMOD/MIAMOD es basa en modelitzar la raó entre incidència i mortalitat mitjançant models que a més tenen en compte les cohorts de naixement dels individus [53]. Cal reflectir doncs, que actualment no hi ha un consens internacional en quina és la millor metodologia per fer projeccions [55] i de fet, en l'últim congrés del Grup de Registres de Càncer amb Llengua Llatina que va tenir lloc a Toledo el Maig del 2010 es va concloure que, fa falta iniciar un estudi internacional per avaluar les diferents metodologies desenvolupades. L'RCC, amb l'àmplia experiència que disposa en realitzar estudis de projeccions [56-63], participarà en l'esmentat estudi internacional (conjuntament amb França, Itàlia, Bèlgica) introduint una metodologia bayesiana de projeccions segons models edat-període-cohort [61] que ha estat aplicada prèviament a la incidència i mortalitat per càncer a Catalunya [58]. Aquesta metodologia serà la quina es compararà amb les que han utilitzat altres països.

D'altra banda, el mètode de projeccions implementat a SART està basat en el mètode descrit per Dyba i Hakulinen [33] del Registre de càncer de Finlàndia. La nostra aportació ha estat el tenir en compte una possible sobredispersió a les dades i també s'han emprat models que poden corregir-la. Això comporta ampliar el ventall de possibilitats, des de el punt de vista de la metodologia de projeccions, amb la garantia de que s'ha establert que el mètode descrit per Dyba i Hakulinen pot comportar un error entre el 10 i el 20% entre els casos predits i els actuals [55]. Per tant, podem considerar que la metodologia aplicada pel càlcul de les nostres projeccions, atès que és una ampliació del mètode de Dyba i Hakulinen, pot presentar un marge d'error acceptable.

D'altra banda, i seguint amb l'aplicació de les projeccions, volem destacar que hem implementat en aquest projecte un algorisme que permet seleccionar el model d'ús adequat en cada cas d'entre els quatre models proposats, mitjançant la combinació de dos criteris, el AIC i el GOF, el millor model (Apartat 4.4.9. Selecció dels Models log-

lineals). La selecció del model a partir d'aquest algorisme permet que el model sigui el que tingui el menor valor de l'AIC. Això garanteix un model on hi participa el principi de parsimònia i per tant, assegurant que el model tingui un bon ajust i penalitzant aquells models amb major nombre de paràmetres. Tot i així, després d'aquestes precaucions, si hi ha algun cas en el que no ajusta cap dels quatre models, l'aplicació retorna a l'usuari la informació sobre aquest problema. En aquesta situació, l'usuari (fora de l'aplicació) ha de buscar diferents metodologies de modelització com per exemple Models Edad-Període-Cohort o Models amb tendència polinòmica [55].

En el treball, s'ha exposat un exemple on s'utilitza la SIMR per a calcular quins tumors presenten un major o menor risc de mortalitat a Catalunya respecte la població de referència de l'Estat Espanyol. Aquests resultats s'haurien d'interpretar també amb certa cautela atès que en l'avaluació d'un possible excés (o disminució) de mortalitat per càncer a Catalunya respecte a Espanya, hem de tenir en conte que la població catalana està inclosa en l'espanyola, i això podria originar un biaix en el sentit que podrien atenuar-se els valors SIMR estimats [49].

Una de les limitacions d'aquest projecte es que manca implementar el càlcul dels intervals de confiança per algunes estimacions, com les taxes d'incidència i mortalitat i les projeccions. Tot i que s'està treballant aquest apartat, els intervals de confiança de les projeccions requereixen una implementació que tingui en compte l'interval de predicció basat en una modificació de la matriu de variàncies-covariàncies dels paràmetres del model [33] i per tant, es preveu poder-ho tindre, però no a curt termini. Una altra possible limitació és que en l'aplicació del càlcul de valors esperats seria útil especificar per a quins anys del fitxer d'entrada es volen ajustar les dades ja que existeixen situacions on hi ha canvis de tendència (per exemple, degudes a la codificació de la malaltia) que fan que els models no ajustin correctament. És probable que, després d'analitzar les dades de forma descriptiva, es detecti on hi hagué un canvi de tendència. D'aquesta manera l'usuari pot seleccionar el període de temps posterior en el qual el model ajusti correctament.

Així mateix, també seria interessant que l'aplicació SART pogués escollir amb quin nivell de confiança es vol treballar a l'hora de realitzar el test de la Bondat d'Ajust ja que en moltes ocasions és favorable ser més estricte del que s'ha marcat (0.95) per tal de reduir la probabilitat d'un possible error tipus II. D'altra banda, el fitxer ".pdf" que retorna aquesta aplicació conté únicament una taula per a totes les localitzacions i sexe i potser seria més

interessant la presentació d'una taula per cadascun dels sexes. De tota manera, aquestes observacions són susceptibles de millora en un període relativament curt.

En resum, donat que l'aplicació SART cobreix el càlcul automatitzat d'indicadors poblacionals bàsics i que pot ser útil per a qualsevol disciplina que estudiï comportaments socials, s'espera que l'ús sigui important. Per tant, a més de les propostes de millora que ja s'han plantejat a la discussió també està en projecte la el·laboració d'una pàgina web que integri totes les publicacions estadístiques que s'han anat elaborant (WAERS, RISKdiff i SART).

8. CONCLUSIONS

En quant a les conclusions, creiem que es tracta d'una eina molt útil ja que:

1. Per primera vegada existeix una aplicació estadística, anomenada SART, via web a l'Estat Espanyol i Europa de lliure distribució que inclogui el càlcul automatitzat de taxes d'incidència/mortalitat crues, ajustades, truncades i acumulades, n'estimi la tendència de les taxes mitjançant l'estadístic EAPC, i permeti el càlcul dels casos esperats i les raons d'incidència/mortalitat estandarditzades seguint la metodologia reconeguda internacionalment.
2. S'ha proposat per primera vegada una aplicació que permet realitzar projeccions via web del nombre de casos de càncer o de qualsevol altra patologia. Aquest punt és molt important degut a la complicació metodològica que implica realitzar projeccions per la qual cosa poques vegades quedaven incloses en el camp de indicadors poblacionals
3. L'aplicació SART permet obtenir els estimadors esmentats de forma senzilla, amb rapidesa sense la necessitat d'instal·lar l'aplicació a l'ordinador de l'usuari. Això implica que l'usuari no ha de disposar de coneixements previs de programació ni de coneixements avançats d'estadística.
4. A nivell metodològic s'ha descrit un nou algorisme que selecciona un model d'entre diversos tenint en compte el Criteri d'Informació d'Akaike i la bondat d'ajust i per tant s'assegura un model que ajusti correctament i que tingui en compte el principi de parsimònia.
5. L'aplicació SART junt amb l'aplicació del càlcul de la supervivència observada i relativa (WAERS) i l'aplicació de l'avaluació dels canvis de la tendència de la incidència/mortalitat d'una malaltia que poden ser deguts a canvis demogràfics i en el risc (RiskDiff), completa un conjunt d'aplicacions per a qualsevol professional de la salut, de l'economia o de la demografia.
6. SART implica una homogeneïtzació en la metodologia pel càlcul d'indicadors poblacionals, però també en la presentació de resultats. Essent aquest punt molt important per assegurar la comparabilitat de les dades en diferents períodes de temps i d'àrees geogràfiques.

10. BIBLIOGRAFIA

1. Rothman K, Greenland S, Lash T. Modern Epidemiology. Third ed. Philadelphia: Lippincott Williams & Wilkins; 2008.
2. German R, et al. Updated Guidelines for Evaluating Public Health Surveillance Systems. Web [2001 Available from:
URL:<http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm>
3. Agency for Toxic Substances & Disease Registry. Glossary of Terms. Web [2009 Available from: URL:<http://www.atsdr.cdc.gov/glossary.html>
4. International Agency for Research on Cancer. Cancer registration: principles and methods. Lyon, France: 1991.
5. Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M et al. Cancer Incidence in Five Continents, Vol. IX. IARC Scientific Publications, Lyon, IARC ed. 2007.
6. Pla director d'oncologia. Planificació i avaluació. Generalitat de Catalunya Departament de Salut [2006 Available from:
URL:http://www.gencat.cat/salut/depsalut/pdf/pd_onco271006.pdf
7. Cancer Research UK. Learn about cancer. Web [4 A.D. Available from:
URL:<http://info.cancerresearchuk.org/cancerandresearch/learnaboutcancer/?a=5441>
8. Ferlay J, Bray F, Pisani P, Parkin D.M. GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide. IARC CancerBase No. 5. version 2.0, IARCPress ed. Lyon: 2004.
9. World Health Organization. Cancer. WHO. Web [2009 Available from:
URL:<http://www.who.int/mediacentre/factsheets/fs297/en/index.html>
10. Servei d'Informació i Estudis. Anàlisi de la mortalitat a Catalunya, 2007. Departament de Salut, Generalitat de Catalunya ed. Barcelona: 2009.
11. Parkin D, Whelan SL, Ferlay J, Storm H. Cancer Incidence in Five Continents, Vol.I to VIII. IARC CancerBase No. 7 ed. Lyon: 2005.
12. World Health Organization. Mortality database. Web [2008 [cited 2008 Sept. 11]; Available from: URL:<http://www.who.int/whosis/whosis>
13. Kouba V. EPIZOO: software for veterinary epidemiology training and problem-solving. Bull World Health Organ 1995; 73(1):77-83.
14. Kafatos G, Verlander NQ, Gelb D, Arnold E, Charlett A. A comparison of public domain software for outbreak investigations: Epi Info and EpiData. Euro Surveill 2007; 12(1):E070111.

15. United States Department of Health and Human Services. Centers for Disease Control and Prevention. What Is Epi Info™? Web [2009 Available from: URL:<http://www.cdc.gov/epiinfo/>
16. What is EpiData Software ? Web [2009 Available from: URL:<http://www.epidata.dk/>
17. On L, Semenciw RM, Mao Y. Orius software: calculation of rates and epidemiologic indicators, and preparation of graphical output. *Chronic Dis Can* 2000; 21(3):134-136.
18. Health Canada. Cancer Statistics Online. Web [2008 Available from: URL:<http://www.hc-sc.gc.ca/hpb/lcdc/webmap/>
19. The North American Association of Central Cancer Registries. Web [2007 Available from: URL:<http://www.naacr.org/>
20. Surveillance Epidemiology and End Results. National Cancer Institute. SEER*Stat 6.5.2. Web [2009 Available from: URL:<http://www.seer.cancer.gov/seerstat/>
21. Cleries R, Ribes J, Moreno V, Esteban L, Pareja L, Galvez J et al. [Relative survival computation. Comparison of methods for estimating expected survival]. *Gac Sanit* 2006; 20(4):325-331.
22. Cleries R, Valls J, Esteban L, Galvez J, Pareja L, Sanz X et al. WAERS: an application for Web-assisted estimation of relative survival. *Med Inform Internet Med* 2007; 32(3):169-175.
23. Cleries R, Ribes J, Galvez J, Melia A, Moreno V, Bosch FX. [Automatic calculation of relative survival through the web. The WAERS project of the Catalan Institute of Oncology]. *Gac Sanit* 2005; 19(1):71-75.
24. Yu B, Tiwari RC, Cronin KA, McDonald C, Feuer EJ. CANSURV: A Windows program for population-based cancer survival analysis. *Comput Methods Programs Biomed* 2005; 80(3):195-203.
25. Geiss K, Meyer M, Radespiel-Troger M, Gefeller O. SURVSOFT-Software for nonparametric survival analysis. *Comput Methods Programs Biomed* 2009; 96(1):63-71.
26. Paul Dickman. SURV3: Windows Software for Relative Survival Analysis. Finnish Cancer Registry, Helsinki, Finland [2003 Available from: URL:<http://www.cancerregistry.fi/surv3/>
27. Registre de Mortalitat de Catalunya, Servei d'Informació i Estudis, Direcció General de Recursos Sanitaris, Departament de Salut, Elaboració pròpia. Defuncions dels anys 1975-2004, per totes les causes, dels residents a Catalunya. 2006.
Ref Type: Data File
28. Instituto Nacional de Estadística. Web del Instituto Nacional de Estadística [2009 Available from: URL:<http://www.ine.es/>

29. Institut d'Estadística de Catalunya. Web de estadística oficial de Catalunya [2009 Available from: URL:<http://www.idescat.cat/>
30. International Classification of Diseases, 9th Revision, Clinical Modification. Washington, DC: Departament of Health and Human Services (DHHS No. (PHS) 80-1260).; 1979.
31. World Health Organization. International Classification of Diseases, Tenth Revision (ICD-10). WHO ed. Geneva: 1992.
32. Dyba T, Hakulinen T, Paivarinta L. A simple non-linear model in incidence prediction. *Stat Med* 1997; 16(20):2297-2309.
33. Dyba T, Hakulinen T. Comparison of different approaches to incidence prediction based on simple interpolation techniques. *Stat Med* 2000; 19(13):1741-1752.
34. Szklo M, Nieto J. *Epidemiología Intermedia. Conceptos y Aplicaciones*. Madrid, España: Ediciones Diaz de Santos, S.A.; 2003.
35. YERUSHALMY J. A mortality index for use in place of the age-adjusted death rate. *Am J Public Health Nations Health* 1951; 41(8 Pt 1):907-922.
36. Yule GU. On some points relating to vital statistics, more especially statistics on occupational mortality. *Journal of Royal Statistical Society* 1934;(97):1-84.
37. Segi M. *Cancer mortality for selected sites in 24 countries (1950-57)*. Department of Public Health. Tohoku University of Medicine. Sendai, Japan: 1960.
38. Doll R, Payne P, Waterhouse J. *Cancer Incidence in Five Continents: Technical Report*. 1966. Berlin, Springer-Verlag (for UICC).
Ref Type: Report
39. Breslow NE, Day NE. *Statistical methods in cancer research. Volume II--The design and analysis of cohort studies*. IARC Sci Publ 1987;(82):1-406.
40. Esteve J, Benhamou E, Raymond L. *Statistical methods in cancer research. Volume IV. Descriptive epidemiology*. IARC Sci Publ 1994;(128):1-302.
41. McCullagh P, Nelder J. *Generalized Linear Models*. Second ed. Chapman & Hall/CRC; 1989.
42. Chongsuvivatwong V. *Analysis of epidemiological data using R and EpiCalc*. The Epidemiology Unit - Prince of Songkla University - Thailand; 2008.
43. Average Annual Percent Change (AAPC). Nacional Cancer Institut US National Institutes of Health Statistical Research and Applications [2010 Available from: URL:<http://srab.cancer.gov/joinpoint/aapc.html>
44. Wikipedia. Johannes Nikolaus Tetens. Wikipedia La enciclopedia libre [2009 Available from: URL:http://es.wikipedia.org/wiki/Johannes_Nikolaus_Tetens
45. Keiding N. The method of expected number of deaths, 1786-1886-1986. *Int Stat Rev* 1987; 55(1):1-20.

46. Daly LE. Confidence limits made easy: interval estimation using a substitution method. *Am J Epidemiol* 1998; 147(8):783-790.
47. Guidelines for Using Confidence Intervals for Public Health Assessment. Health Data Guidelines Washington State Department of Health [2009 Available from: URL:<http://www.doh.wa.gov/data/guidelines/ConfIntguide.htm>
48. Confidence Intervals for the Mean of a Poisson Distribution. Department of Mathematics and Statistics McMaster University [1997 Available from: URL:<http://www.math.mcmaster.ca/peter/s743/poissonalpha.html>
49. Jones ME, Swerdlow AJ. Bias in the standardized mortality ratio when using general population rates to estimate expected number of deaths. *Am J Epidemiol* 1998; 148(10):1012-1017.
50. Valls J, Cleries R, Galvez J, Moreno V, Gispert R, Borrás JM et al. RiskDiff: a web tool for the analysis of the difference due to risk and demographic factors for incidence or mortality data. *BMC Public Health* 2009; 9:473.
51. Borrás JM, Borrás J, Izquierdo A, Gispert R, editors. The impact of cancer in Catalonia. *Med Clin (Bcn)*: 2008.
52. Gispert R, Cleries R, Puigdefabregas A, Freitas A, Esteban L, Ribes J. [Cancer mortality trends in Catalonia, 1985-2004]. *Med Clin (Barc)* 2008; 131 Suppl 1:25-31.
53. Istituto Superiore di Sanità. MIAMOD. Mortality, incidence analysis model. Istituto Superiore di Sanità Centro Nazionale di Epidemiologia Sorveglianza e Promozione della Salute Roma [2010 Available from: URL:<http://www.eurocare.it/MiamodPiamod/tabid/60/Default.aspx>
54. Gerda Engholm, Jacques Ferlay, Niels Christensen, Freddie Bray, Marianne L.Gjerstorff, Åsa Klint et al. NORDCAN: Cancer Incidence, Mortality, Prevalence and Prediction in the Nordic Countries, Version 3.5. Association of the Nordic Cancer Registries Danish Cancer Society [2009 Available from: URL:<http://www.ancr.nu>
55. Dyba T, Hakulinen T. Do cancer predictions work? *Eur J Cancer* 2008; 44(3):448-453.
56. Ribes J, Cleries R, Borrás J, Galceran J, Bosch FX. Time trends in incidence and mortality for chronic liver disease and liver cancer in the interval 1980-1997 in Catalonia, Spain. *Eur J Gastroenterol Hepatol* 2004; 16(9):865-872.
57. Ribes J, Navarro M, Cleries R, Esteban L, Pareja L, Binefa G et al. Colorectal cancer mortality in Spain: trends and projections for 1985-2019. *Eur J Gastroenterol Hepatol* 2008.
58. Ribes J, Cleries R, Buxo M, Ameijide A, Valls J, Gispert R. [Predictions of cancer incidence and mortality in Catalonia to 2015 by means of Bayesian models]. *Med Clin (Barc)* 2008; 131 Suppl 1:32-41.

59. Pitarque S, Cleries R, Martinez JM, Lopez-Abente G, Kogevinas M, Benavides FG. Mesothelioma mortality in men: trends during 1977-2001 and projections for 2002-2016 in Spain. *Occup Environ Med* 2007.
60. Marcos-Gragera R, Vilar-Coromina N, Galceran J, Borrás J, Cleries R, Ribes J et al. Rising trends in incidence of cutaneous malignant melanoma and their future projections in Catalonia, Spain: increasing impact or future epidemic? *J Eur Acad Dermatol Venereol* 2010.
61. Cleries R, Ribes J, Esteban L, Martinez JM, Borrás JM. Time trends of breast cancer mortality in Spain during the period 1977-2001 and Bayesian approach for projections during 2002-2016. *Ann Oncol* 2006; 17(12):1783-1791.
62. Cleries R, Martinez JM, Valls J, Pareja L, Esteban L, Gispert R et al. Life expectancy and age-period-cohort effects: analysis and projections of mortality in Spain between 1977 and 2016. *Public Health* 2009; 123(2):156-162.
63. Cleries R, Martinez JM, Escriba JM, Esteban L, Pareja L, Borrás JM et al. Monitoring the decreasing trend of testicular cancer mortality in Spain during 2005-2019 through a Bayesian approach. *Cancer Epidemiology* 2010; 34(3):244-256.

10. ANNEXES

ANNEX 1. Codi de les funcions

ANNEX 1.1. Funció descriptiva

```

SART_descriptive<-function(titol, dades, pesos, ged, periode,
trun=c(8:13), cum=c(1:15), CauPer, CauNoPer)
{
  # CONTROL DE POSSIBLES ERRORS

  if(prod(unique(dades$age.group)==pesos$age.group)==0
|prod(unique(dades$age.group)==ged$ged)==0 |
prod(pesos$age.group==ged$ged)==0)
  {
    print("Wrong definition of the input files: Data, weights or age
group.")
  }else
  {
    if(sum(pesos[2])!=1)
    {
      print("Warning: The reference weights don't sum 1.")
    }
  }

  for (a in 1:length(dades$sex))
  {
    if((dades$sex[a]!=1)&(dades$sex[a]!=2)&(dades$sex[a]!=0))
    {
      print("Error in the variable sex from input data file.")
    }
  }

  if(length(CauPer!=0))
  {
    if((CauPer %in% unique(dades$group))=="FALSE")
    {
      print("The Denominator group, used to calculate percentages, is not
in the data file.")
    }
  }

  cnp<-vector()
  if(length(CauNoPer)!=0)
  {
    for (np in 1:length(CauNoPer))
    {
      if((CauNoPer[np] %in% unique(dades$group))=="FALSE")
      {
        cnp[np]<-0
      }else
      {
        cnp[np]<-1
      }
    }
  }

  if(prod(cnp)==0)
  {

```

```

    print("Some groups excluded from percentage calculation are not in
the data file.")
}

if(((periode[1] %in% unique(dades$year))=="FALSE")|((periode[2] %in%
unique(dades$year))=="FALSE"))
{
    print("The period selected is not in the data file. The analysis is
calculated for the years comprised in the data file.")
}

tr<-vector()
for (t in 1:length(trun))
{
    if((trun[t] %in%unique(dades$age.group))=="FALSE)
    {
        tr[t]<-0
    }else
    {
        tr[t]<-1
    }
}

if(prod(tr)==0)
{
    print("Some truncate age group are not in the data file.")
}

cu<-vector()
for (d in 1:length(cum))
{
    if((cum[d] %in%unique(dades$age.group))=="FALSE)
    {
        cu[d]<-0
    }else
    {
        cu[d]<-1
    }
}

if(prod(cu)==0)
{
    print("Some cumulative age group are not in the data file.")
}

# FUNCIOÓ:

sexe<-unique(dades$sex)
output.by.sex<-list()

for(s in 1:length(sexe))
{

    locs<-as.character(unique(dades$group[dades$sex==sexe[s]]))
    dat<-dades[dades$sex==sexe[s],]
    locsaux<-vector()
    if (length(CauPer)==0)
    {
        print("No percentage was computed.")
        locsaux<-locs
    }else

```

```

{
  if(length(CauNoPer)==0)
  {
    for (i in 1:length(locs))
    {
      if(locs[i]!=CauPer)
      {
        locsaux<-c(locsaux,locs[i])
      }
    }
    locsaux<-c(locsaux,CauPer)
  }else
  {
    for (k in 1: length(locs))
    {
      trobat_a_CauNoPer<-FALSE
      for(j in 1:length(CauNoPer))
      {
        if(locs[k]==CauNoPer[j])
        {
          trobat_a_CauNoPer<-TRUE
        }
      }
      if(locs[k]!=CauPer & trobat_a_CauNoPer==FALSE)
      {
        locsaux<-c(locsaux,locs[k])
      }
    }
    locsaux<-c(locsaux,CauNoPer,CauPer)
  }
}
locs<-locsaux

output<-data.frame()

for(l in 1:length(locs))
{
  Localitzacio<-locs[l]
  loc<-dat[dat$group==locs[l],]
  loc<-loc[loc$year>=periode[1]& loc$year<=periode[2],]
  loc2<-loc[,c(2,3,5,6)]
  P<-pesos[2]

  #Nombre de casos
  N<- sum(loc$cases)

  #Taxa bruta
  TB<-((sum(loc$cases))/(sum(loc$population)))*100000

  #S'agreguen casos per grup d'edat per calcular la TA amb la
  piràmide de referència:
  cases.agreg<-
  aggregate(c(loc2$cases),by=list(age.group=loc2$age.group),FUN=sum)
  pop.agreg<-
  aggregate(c(loc2$population),by=list(age.group=loc2$age.group),FUN=sum)
  loc.agreg<-cbind(cases.agreg,pop.agreg$x)
  names(loc.agreg)<-c("age.group","cases","population")
  loc.agreg$P<-P
  loc.agreg$TE<-loc.agreg$cases/loc.agreg$population
  loc.agreg$TE.P<-loc.agreg$TE*loc.agreg$P
  TA.P<-sum(loc.agreg$TE.P)*100000

  # Taxa truncada

```

```

loc.agreg.trunc<-loc.agreg[loc.agreg$age.group %in%trun,]
loc.agreg.trunc$P<-loc.agreg.trunc$P/sum(loc.agreg.trunc$P)
loc.agreg.trunc$TE.P<-loc.agreg.trunc$TE*loc.agreg.trunc$P
TT.P<-sum(loc.agreg.trunc$TE.P)*100000

# Taxa acumulada
grups<-ged[1,3]-ged[1,2]+1
TAC<-sum((loc.agreg$TE[loc.agreg$age.group %in% cum])* grups * 100)

# La taula descriptiva
output<-rbind(output,cbind(Localitzacio,N,TB,TA.P,TT.P,TAC))
rm(Localitzacio, N, TB, TA.P,TT.P,TAC, grups, cases.agreg,
pop.agreg, loc.agreg, loc, loc2)
}
output$N<-round(as.numeric(as.character(output$N)),0)
output$TB<-round(as.numeric(as.character(output$TB)),2)
output$TA.P<-round(as.numeric(as.character(output$TA.P)),2)
output$TT.P<-round(as.numeric(as.character(output$TT.P)),2)
output$TAC<-round(as.numeric(as.character(output$TAC)),2)

# Es calcula el percentatge si hi existeix el grup per la qual es vol
fer el percentatge.

if(length(CauPer)!=0)
{
  output$Percentatge<-
round((output$N/output$N[output$Localitzacio==CauPer])*100,2)

  # Als valors que no se'ls hi vol fer percentatge, els hi posem NA
  for (k in 1:length(CauNoPer))
  {
    output$Percentatge[output$Localitzacio==CauNoPer[k]]<-NA
  }
  output<-cbind(output[,1], output[,7], output[,2:6])
  names(output)<-cbind("
", "%", "N", "CR", paste("ASR", "(", names(pesos)[2], ")", sep=""), paste("TR(", na
mes(pesos)[2], ")", sep=""), "CumR")
}else
{
  names(output)<-cbind("
", "N", "CR", paste("ASR", "(", names(pesos)[2], ")", sep=""),
paste("TR(", names(pesos)[2], ")", sep=""), "CumR")
}

output<-as.data.frame(output)
output.by.sex[[s]]<-output
rm(locs, locsaux, dat, output)

}

noms<-vector()
for(m in 1:length(sexe))
{
  if(sexe[m]==1)
  {
    noms[m]<- "Men"
  }else
  {
    if(sexe[m]==2)
    {
      noms[m]<- "Women"
    }else

```

```

        {
          noms[m]<-"Both sexes"
        }
      }
    }

names(output.by.sex)<-noms

if(is.na(ged$ged.max[length(ged$ged.max)])){
  tit.age<-paste(ged$ged.min[1]," to + ",
ged$ged.min[length(ged$ged.min)], " years.", sep="")
}else
{
  tit.age<-paste(ged$ged.min[1]," to ",
ged$ged.max[length(ged$ged.max)], " years.", sep="")
}

# Títol i peu de taula.
titol<-paste(titol, "descriptive.", "Period: ", periode[1],"-",
periode[2], ". Age from", tit.age, sep=" ")
peu.taula<- paste("%: Percentage, N: Number of cases , CR: Crude Rate x
100000 person-years,ASR: Age Standardized Rate x 100000 person-years, TR:
Truncated Rate x 100000 person-years by the
",ged$ged.min[ged$ged==min(trun)],"-",ged$ged.max[ged$ged==max(trun)],"
age group, CumR: Cumulative Rate (percent) by the
",ged$ged.min[ged$ged==min(cum)],"-",ged$ged.max[ged$ged==max(cum)]," age
group, (.): Reference Population.",sep="")
# Es mostra la sortida
llista.output<-list(titol, output.by.sex, peu.taula)
names(llista.output)<-c("Titol.Taula", "Taula","Peu.Taula")
llista.output

# ES FAN ELS GRÀFICS
graf<-llista.output
for (j in 1:length(graf$Taula))
{
  graf$Taula[[j]]<-graf$Taula[[j]][,c(1,length(graf$Taula[[j]])-
4,length(graf$Taula[[j]])-3,length(graf$Taula[[j]])-
2,length(graf$Taula[[j]])-1,length(graf$Taula[[j]])))]
}

for(s in 1:length(graf$Taula))
{
  abrev<-c(paste("N"),paste("CR x 100000 ",names(graf$Taula)[s],"-
Year",sep=""),paste("ASR x 100000 ",names(graf$Taula)[s],"-
Year",sep=""),paste("TR x 100000 ",names(graf$Taula)[s],"-
Year",sep=""),"CumR (percent)")
  for(t in 1:5)
  {
    nom_jpg<-c("N_", "CR_", "ASR_", "TR_", "CumR_")
    titles<- c(paste("Number of cases. ",names(graf$Taula)[s], " ,
period ", periode[1], "- ",periode[2],".",sep=""),paste("Crude Rate per
100000 persons-year. ",names(graf$Taula)[s], " , period ", periode[1], "-
",periode[2],".",sep=""), paste("Age Standardized Rate per 100000
persons-year. ",names(graf$Taula)[s], " , Period ", periode[1], "-
",periode[2],".",sep=""), paste("Truncate Rate per 100000 persons-year.
",names(graf$Taula)[s], " , ages ",ged$ged.min[ged$ged==min(trun)],"-
",ged$ged.max[ged$ged==max(trun)], " years. Period ", periode[1], "-
",periode[2],".",sep=""), paste("Cumulative Rate per 100000 persons-year.
",names(graf$Taula)[s], " , ages ",ged$ged.min[ged$ged==min(cum)],"-
",ged$ged.max[ged$ged==max(cum)], ". Period ", periode[1], "-
",periode[2],".",sep=""))
  }
}

```



```

    jpeg(filename=paste(nom_jpg[t],names(graf$Taula)[s],".jpg",sep=""),
quality=80, width = 950, height = 550)
    par(mar=c(4, 25, 4, 4) + 0.1)
    ord_graf<-
data.frame(group=graf$Taula[[s]][,1],indicator=graf$Taula[[s]][,t+1])
    ord_graf<-ord_graf[order(ord_graf$indicator),]
    ord_graf$group<-as.character(ord_graf$group)

    if(t==1)
    {
        a<-(trunc(max(graf$Taula[[s]][,t+1])/5)+1)
    }else{
        a<-max(graf$Taula[[s]][,t+1])/5
    }

    barplot(ord_graf$indicator,
names.arg=ord_graf$group,hORIZ=TRUE,col="steelblue3", las=1, axes=F,
main=titles[t], xlab=abrev[t], xlim=c(0,a*5))

    if(t==1)
    {
        axis(side=1,at=c(0,a,a*2,a*3,a*4,a*5))
    }else{
axis(side=1,at=c(round(0,2),round(a,2),round(a*2,2),round(a*3,2),round(a*
4,2),round(a*5,2)))
    }

    dev.off()
    rm(a, ord_graf)
    }
}
llista.output
}

```

ANNEX 1.2. Funció EAPC

```

SART_EAPC<-function(dades,wei,period,conf)
{
  #CONTROL D'ERRORS

  if(prod(unique(dades$age.group)==wei$age.group)==0)
  {
    write(file=f_errors,"Wrong definition in the input files: Data or
weights.",append=T)
    #print("Wrong definition in the input files: Data or weights.")
  }else
  {
    if(sum(wei[2])!=1)
    {
      write(file=f_errors,"Warning: The reference weights don't sum
1.",append=T)
      #print("Warning: The reference weights don't sum 1.")
    }

    if(period[1]==period[2])
    {
      write(file=f_errors,"The Estimated Annual Percent Change only makes
sense when dealing with different periods or years.",append=T)
      #print("The Estimated Annual Percent Change only makes sense when
dealing with different periods or years.")
    }else
    {
      LOC<-as.character(unique(dades$group))
      taula.EAPC<-data.frame()
      taula.sex<-list()

      for (a in 1:length(dades$sex))
      {
        if((dades$sex[a]!=1)&(dades$sex[a]!=2)&(dades$sex[a]!=0))
        {
          write(file=f_errors,"Error in the variable sex from input data
file.",append=T)
          #print("Error in the variable sex from input data file.")
        }
      }

      for(loc in 1:length(LOC))
      {
        taula.tmp<-dades[dades$group==LOC[loc],]
        SEX<-unique(taula.tmp$sex)
        for(sex in 1:length(SEX))
        {
          taula<-taula.tmp[taula.tmp$sex==SEX[sex],]
          taula<-taula[taula$year>=period[1] & taula$year<=period[2],]
          taula$weight<-rep(wei[,2],length(unique(taula$year)))
          taula$TE.W<-((taula$cases/taula$population)*100000) *
taula$weight
          taula.ag<-
aggregate(c(taula$TE.W),by=list(year=taula$year),FUN=sum)
          taula.ag<-cbind(rep(SEX[sex], dim(taula.ag)[1]),rep(LOC[loc],
dim(taula.ag)[1]),taula.ag)
          names(taula.ag)<-c("sex","group","year","TA")
          taula.ag$year<-taula.ag$year-min(taula.ag$year)
          model.tmp<-glm(TA~year,family=gaussian(link="log"),
data=taula.ag)

```

```

# Extraiem el logaritme de la mitjana.
b<-summary(model.tmp)$coeff[2]
# Extraiem l'interval de confiança del 95% (escala logarítmica)
inf.l<-b+qnorm(0.5*(1+conf), mean = 0, sd = 1, lower.tail =
TRUE, log.p = FALSE)*summary(model.tmp)$coeff[4]
sup.l<-b+qnorm(0.5*(1+conf), mean = 0, sd = 1, lower.tail =
TRUE, log.p = FALSE)*summary(model.tmp)$coeff[4]
# Escala anti-logaritme i càlcul del EAPC
mitjana<-round((exp(b)-1)*100,2)
lim.i<-round((exp(inf.l)-1)*100,2)
lim.s<-round((exp(sup.l)-1)*100,2)
taula.EAPC<-
rbind(taula.EAPC,cbind(LOC[loc],SEX[sex],as.numeric(as.character(mitjana)
),as.numeric(as.character(lim.i)),as.numeric(as.character(lim.s))))
rm(b, inf.l, lim.i, lim.s, mitjana, model.tmp,sup.l, taula,
taula.ag)
}
}

names(taula.EAPC)<-c(" ", "Sex", "EAPC", "LCI", "UCI")
taula.EAPC$Sex<-as.numeric(as.character(taula.EAPC$Sex))

for(i in 1:length(taula.EAPC$Sex))
{
  if(taula.EAPC$Sex[i]==1)
  {
    taula.EAPC$Sex[i]<-"Men"
  }else
  {
    if(taula.EAPC$Sex[i]==2)
    {
      taula.EAPC$Sex[i]<-"Women"
    }else
    {
      if(taula.EAPC$Sex[i]==0)
      {
        taula.EAPC$Sex[i]<-"Both sexes"
      }
    }
  }
}

Sexes<-unique(taula.EAPC$Sex)
EAPC<-list()

for(j in 1:length(Sexes))
{
  EAPC[[j]]<-taula.EAPC[taula.EAPC$Sex==Sexes[j],c("
", "EAPC", "LCI", "UCI")]
  EAPC[[j]]$EAPC<-round(as.numeric(as.character(EAPC[[j]]$EAPC)),2)
  EAPC[[j]]$LCI<-round(as.numeric(as.character(EAPC[[j]]$LCI)),2)
  EAPC[[j]]$UCI<-round(as.numeric(as.character(EAPC[[j]]$UCI)),2)
}

names(EAPC)<-Sexes
}

titol<-paste("Estimated Annual Percent Change for the period",
period[1], "-", period[2], ".", conf*100, "% Confidence Interval.",
sep=" ")
peu<-paste("EAPC: Estimated Annual Percent Change, LCI: Lower
Confidence Interval, UCI: Upper Confidence Interval" )
EAPC<-list(titol,EAPC,peu)

```

```

    names(EAPC)<-c("Title", "EAPC","Table_footer")
  }

  for (g in 1:length(EAPC$EAPC))
  {
    datos<-EAPC$EAPC[[g]]
    name<-names(EAPC$EAPC)[g]

    names(datos)<-c("Group","EAPC","LCI","UCI")
    datos$Group<-as.factor(as.character(datos$Group))
    a<-min(datos$LCI)
    b<-max(datos$UCI)
    jpeg(filename=paste("EAPC",name,".jpg",sep=""), quality=80, width =
800, height = 480)
    par(mar=c(5, 20, 4, 4) + 0.1)

    plot(datos$EAPC,datos$Group,pch=15,col="black",xlim=c(a,b),axes=F,xlab="E
APC",ylab="", main=paste("EAPC for", name, ". Period ",period[1],"-
",period[2],".", sep=" "))
    axis(2,labels=levels(datos[,c(1)]),at=1:length(datos$Group),las=1)
    abline(v=0)
    points(datos$LCI,datos$Group, col="black", pch="|")
    points(datos$UCI,datos$Group, col="black", pch="|")

    for(i in 1:length(datos$Group))
    {
      lines(c(datos$LCI[i],datos$UCI[i]),
c(datos$Group[i],datos$Group[i]))
    }

    q1<-b/4
    q2<-(-a)/4

    axis(side=1,at=c(round(a,2),round(a+q2,2),round(a+2*q2,2),round(a+3*q2,2)
,round(a+4*q2,2),round(1+q1,2),round(1+2*q1,2),round(1+3*q1,2),round(1+4*
q1,2)),tick=T,cex.axis=0.7)
    dev.off()
    rm(datos,name,a,b)
  }

  EAPC
}

```

ANNEX 1.3. Funció Esperats

```

SART_expected<-function(dades,pob.esperada)
{
  Ok<-0
  library(MASS)
  names(dades) <-
c('sex','age.group','year','group','cases','population')
  names(pob.esperada) <- c('sex','age.group','year','population')

  if(prod(unique(dades$age.group)==unique(pob.esperada$age.group))==0){
    write(file=f_errors,"Wrong definition on the input files: data file
or population file.",append=T)
    print("Wrong definition on the input files: data file or population
file.")
  }

  for (n in 1:length(unique(dades$sex))){
    if(unique(dades$sex)[n]!=0 & unique(dades$sex)[n]!=1 &
unique(dades$sex)[n]!=2){
      write(file=f_errors,"The sex in the data file is not properly
introduced.",append=T)
      print("The sex in the data file is not properly introduced.")
    }
  }

  for (m in 1:length(unique(pob.esperada$sex))){
    if(unique(pob.esperada$sex)[m]!=0 & unique(pob.esperada$sex)[m]!=1 &
unique(pob.esperada$sex)[m]!=2){
      write(file=f_errors,"The sex in the population file is not properly
introduced.",append=T)
      print("The sex in the population file is not properly introduced.")
    }
  }

  LOC <- as.character(unique(dades$group))

  resultats<-list()
  comptador<-1

  for (loc in 1:length(LOC))
  {
    taula.treball <- dades[dades$group==LOC[loc],]
    SEX <- unique(taula.treball$sex)
    for (sex in 1:length(SEX)){
      dades.tmp <- taula.treball[taula.treball$sex==SEX[sex],]
      new.tmp <- pob.esperada[pob.esperada$sex==SEX[sex],]
      titol<-unique(taula.treball$group)
      dades.tmp$year1<-dades.tmp$year-min(dades.tmp$year) # Centrem
l'any.

      # S'ajusten els 4 models proposats:

      model.nbl <-
glm.nb(cases~offset(log(population))+factor(age.group)+year1-
1,data=dades.tmp,control = glm.control(epsilon = 1e-8, maxit = 25, trace
= FALSE))
      model.nbl <-
glm.nb(cases~offset(log(population))+factor(age.group)+factor(age.group):
year1-1,data=dades.tmp,control = glm.control(epsilon = 1e-8, maxit = 25,
trace = FALSE))

```

```

model.pl <-
glm(cases~offset(log(population))+factor(age.group)+year1-
1,data=dades.tmp, family=poisson(),control = glm.control(epsilon = 1e-8,
maxit = 25, trace = FALSE))
model.pI <-
glm(cases~offset(log(population))+factor(age.group)+factor(age.group):yea
r1-1,data=dades.tmp, family=poisson(),control = glm.control(epsilon = 1e-
8, maxit = 25, trace = FALSE))

# Calculem els aic manualment
lik.pois<-function(y.tmp,mu)
{
  vec<-(y.tmp*log(mu)-mu-lgamma(y.tmp+1))
  sum(vec)
}

aic.pois<-function(model.tmp,y.tmp)
{
  mu.exp<-model.tmp$fitted.values
  -2*lik.pois(y.tmp,mu.exp)+2*(model.tmp$rank)
}

y.obs<-dades.tmp$cases
model.pl$aic<-aic.pois(model.pl,y.obs)
model.pI$aic<-aic.pois(model.pI,y.obs)
print(c(aic.pois(model.pl,y.obs),aic.pois(model.pI,y.obs)))
# Es posen els models en una llista:
Models<-list(model.nbl,model.nbI,model.pl, model.pI)
Models.dat<-
list(summary(model.nbl),summary(model.nbI),summary(model.pl),summary(mode
l.pI))
# Es calcula el p-valor del test "Goodness of fit".
GOF<-
c(pchisq(sum((resid(Models[[1]],"pearson"))^2),Models[[1]]$df.residual,lo
wer.tail=FALSE),pchisq(sum((resid(Models[[2]],"pearson"))^2),Models[[2]]$
df.residual,lower.tail=FALSE),
pchisq(sum((resid(Models[[3]],"pearson"))^2),Models[[3]]$df.residual,lowe
r.tail=FALSE),
pchisq(sum((resid(Models[[4]],"pearson"))^2),Models[[4]]$df.residual,lowe
r.tail=FALSE))

EstimMod<-
list(summary(model.nbl),summary(model.nbI),summary(model.pl),
summary(model.pI))
NameMod<-c("Negative Binomial age-drift","Negative Binomial age-
specific slope","Poisson age-drift","Poisson age-specific slope")

AicMod<-c(model.nbl$aic,model.nbI$aic, model.pl$aic,model.pI$aic)
print(AicMod)

AicModels<-data.frame(NameMod,AicMod,GOF)
names(AicModels)<-c("Model","AIC","pvalueGOF")

if(AicModels$pvalueGOF[which.min(AicModels$AIC)]>0.05)
{
  NomModPred<-NameMod[which.min(AicMod)]
  NomModPred<-as.character(NomModPred)
  ModSeleccionat<-Models[[which.min(AicMod)]]
  Metode<- "Minimum AIC"
}

if(AicModels$pvalueGOF[which.min(AicModels$AIC)]<=0.05)
{

```

```

if (max(AicModels$pvalueGOF)>0.05)
{
  NomModPred<-NameMod[which.max(GOF)]
  NomModPred<-as.character(NomModPred)
  ModSeleccionat<-Models[[which.max(GOF)]]
  Metode<-"Maximum GOF"
}else
{
  NomModPred<-NameMod[which.min(AicMod)]
  NomModPred<-as.character(NomModPred)
  ModSeleccionat<-Models[[which.min(AicMod)]]
  Metode<-"The model does not fit"
}
}

new.tmp$year1<-new.tmp$year-min(dades.tmp$year)

pred.interna<- exp(predict(ModSeleccionat))
pred.externa<- exp(predict(ModSeleccionat,newdata=new.tmp))

esp.interns<-
cbind(dades.tmp[1:3],rep(LOC[loc],length(pred.interna)),round(pred.interna,0),dades.tmp[6])
esp.externs<-
cbind(new.tmp[1:3],rep(LOC[loc],length(pred.externa)),round(pred.externa,0),new.tmp[4])

names(esp.interns)<-
c("sex","age.group","year","group","cases","population")
names(esp.externs)<-
c("sex","age.group","year","group","cases","population")

Mod_Pred<-list(Group=LOC[loc], Sex=SEX[sex], Dades=dades.tmp,
Est.mods=EstimMod, Mod.Aic=AicModels, Metode=Metode,
Nom.mod.sel=NomModPred, Mod.sel=ModSeleccionat, Esp.int=esp.interns,
Esp.ext=esp.externs)

zv<- summary(Mod_Pred$Mod.sel)$coefficients[,3]
pos<-vector()

for(c in 1:length(zv))
{
  if (abs(zv[c])<0.01)
  {
    pos[c]<-0
  }else
  {
    pos[c]<-1
  }
}

pos2<-vector()

if (length(pos)<(length(unique(dades.tmp$age.group))+4))
{
  pos2<-pos[1:length(unique(dades.tmp$age.group))]
}else
{
  pos2<-
pos[1:length(unique(dades.tmp$age.group))]*pos[(length(unique(dades.tmp$age.group))+1):(length(unique(dades.tmp$age.group))*2)]
}

```

```

year<-unique(Mod_Pred$Esp.ext$year)

for(y in 1:length(unique(Mod_Pred$Esp.ext$year)))
{
  Mod_Pred$Esp.ext$cases[Mod_Pred$Esp.ext$year==year[y]]<-
Mod_Pred$Esp.ext$cases[Mod_Pred$Esp.ext$year==year[y]]*pos2
}

resultats[[comptador]]<-Mod_Pred
comptador<-comptador+1

  rm(dades.tmp, new.tmp, titol, model.nbl, model.nbI, model.pl,
model.pI, EstimMod, NameMod, AicMod, AicModels, NomModPred,
ModSeleccionat, pred.interna, pred.externa , esp.interns, esp.externs,
GOF, Metode, Mod_Pred, zv, pos, pos2)
}
}

taula<-data.frame()
p.valor<-vector()

for(i in 1:length(resultats))
{
  taula[i,1]<-resultats[[i]]$Group
  taula[i,2]<-resultats[[i]]$Sex
  taula[i,3]<-resultats[[i]]$Nom.mod.sel
  p.valor[i]<-
pchisq(sum((resid(resultats[[i]]$Mod.sel,"pearson"))^2),dim(resultats[[i]]
)$Esp.int)[1]-length(resultats[[i]]$Mod.sel$coefficients),lower.tail=F)
  taula[i,4]<-round(p.valor[i],4)

  if (taula[i,4]<0.05)
  {
    taula[i,5]<-"Yes"
  }else
  {
    taula[i,5]<-"No"
  }
  taula[i,6]<-resultats[[i]]$Metode
}

names(taula)<-
c("group","sex","model.chosen","pv.test","significant","method.AIC.GOF")
peu.taula<-c("group: Group of interest; sex: 0= Both sexes, 1= Men, 2=
Women; model.chosen: The fitted model chosen to predict; pv.test: The p-
value of the model goodness of fit test; significant: Is the preceding
p.value significant?; method.AIC.GOF: The method used for model selection
(AIC: Akaike Information Criterion; GOF: chi-square goodness of fit test)")

Taula<-list(taula=taula,peu.taula=peu.taula)

esperats.int<-data.frame()
esperats.ext<-data.frame()

for(j in 1:length(resultats))
{
  esperats.int<-rbind(esperats.int,resultats[[j]]$Esp.int)
  esperats.ext<-rbind(esperats.ext,resultats[[j]]$Esp.ext)
}

ll.models<-list()
for(k in 1:length(resultats))
{

```



```
  ll.models[[k]]<-list(Group=resultats[[k]]$Group,
Sex=resultats[[k]]$Sex, Selected.Model=resultats[[k]]$Mod.sel)
}

  Avis<-Taula$taula[Taula$taula$method.select=="The model does not
fit",1:2]
  Ok<-1

list(Warning=Avis,Mod.pvalor=Taula,Mod.select=ll.models,Expected.cases=es
perats.ext,Mod.dat=Models.dat)
}
```

ANNEX 1.4. Funció SIMR

```

SART_SIMR<-function(dades, ref, conf)
{
  # CONTROL ERRORS

  if(dim(dades)[1]!=dim(ref)[1])
  {
    #write(file=f_errors,"The files do not have the same
length.",append=T)
    print("The files do not have the same length.")
  }else
  {
    for (a in 1:length(dades$sex))
    {
      if((dades$sex[a]!=1)&(dades$sex[a]!=2)&(dades$sex[a]!=0))
      {
        #write(file=f_errors,"Error in the variable sex from input data
file.",append=T)
        print("Error in the variable 'sex' from input data file.")
      }
      if((ref$sex[a]!=1)&(ref$sex[a]!=2)&(ref$sex[a]!=0))
      {
        #write(file=f_errors,"Error in the variable sex from input
reference population file.",append=T)
        print("Error in the variable 'sex' from input reference
population file.")
      }
    }
  }

  if(prod(sort(unique(dades$age.group))==sort(unique(ref$age.group)))==0)
  {
    #write(file=f_errors,"Wrong definition of the age group on the
input data file.",append=T)
    print("Wrong definition of the age group on the input data files.")
  }

  if(prod(sort(unique(dades$sex))==sort(unique(ref$sex)))==0)
  {
    #write(file=f_errors,"Wrong definition of the sex on the input data
files.",append=T)
    print("Wrong definition of the sex on the input data files.")
  }

  if(length(unique(dades$group))!=length(unique(ref$group)))
  {
    #write(file=f_errors,"Different level sets of groups.",append=T)
    print("Different level sets of groups.")
  }

  if(prod(sort(unique(dades$group))==sort(unique(ref$group)))==0)
  {
    #write(file=f_errors,"Wrong definititon of the groups on the input
files.",append=T)
    print("Wrong definititon of the groups on the input files.")
  }
}

val.conf<-conf

### function for the exact confidence interval of poisson distribution

```

```

ic.pois<-function(x.tmp,conf=val.conf)
{
  # Ref: Daly, L. Simple SAS macros for the calculation of exact
  binomial and Poisson confidence limits. Comput Biol Med 22:351-361, 1992.
  # http://www.doh.wa.gov/data/guidelines/ConfIntguide.htm
  # http://www.doh.wa.gov/data/guidelines/CI_images/CIPOISS_SAS.txt
  # http://www.math.mcmaster.ca/peter/s743/poissonalpha.html
  alpha.tmp<-1-conf
  if(x.tmp>0)
  {
    out.tmp<-c(qchisq((1-conf)/2, 2*x.tmp)/2, qchisq(conf+(1-conf)/2,
2*(x.tmp+1))/2)
  }
  if(x.tmp==0)
  {
    out.tmp<-c(0,-log(1 - (conf)))
  }
  round(out.tmp,2)
}
####

dades<-dades[order(dades$group,dades$sex,dades$year,dades$age.group),]
ref<-ref[order(ref$group,ref$sex,ref$year,ref$age.group),]
ref$cr<-ref$cases/ref$population
tot<-cbind(dades,ref$cr)
names(tot)<-c(names(dades),"cr")

sex<-unique(dades$sex)
sexe<-c()
for(m in 1:length(sex))
{
  if(sex[m]==0){ sexe[m]<-"Both sexes"}
  if(sex[m]==1){ sexe[m]<-"Men"}
  if(sex[m]==2){ sexe[m]<-"Women"}
}

SIMR<-vector("list", length(sex))
names(SIMR)<-sexe

for (k in 1:length(sex))
{
  tots_s<-tot[tot$sex==sex[k],]
  group<-unique(tots_s$group)
  SIMR[[k]]<-vector("list",length(group))
  names(SIMR[[k]])<-group

  for(l in 1:length(group))
  {
    tots_sg<-tots_s[tots_s$group==group[l],]
    tots_sg$expected<-tots_sg$population*tots_sg$cr
    SIMR_global<-sum(tots_sg$cases)/sum(tots_sg$expected)*1
    CASES_global<-sum(tots_sg$cases)
    L_SIMR_global<-
ic.pois(x.tmp=sum(tots_sg$cases))[1]/sum(tots_sg$expected)
    U_SIMR_global<-
ic.pois(x.tmp=sum(tots_sg$cases))[2]/sum(tots_sg$expected)
    SIMR_global<-
data.frame(CASES=CASES_global,SIMR=round(SIMR_global,2),LCI=round(L_SIMR_
global,2),UCI=round(U_SIMR_global,2))
    SIMR_by_age<-list()
    edats<-unique(tots_sg$age.group)
    SIMR_by_ages<-data.frame()
    LCI_SIMR_by_ages<-c()

```

```

UCI_SIMR_by_ages<-c()

for(i in 1:length(edats))
{
  seleccio<-tots_sg[tots_sg$age.group==edats[i],]
  SIMR_age<-sum(seleccio$cases)/sum(seleccio$expected) *1
  CASES_age<-sum(seleccio$cases)
  LCI_age<-
ic.pois(x.tmp=sum(seleccio$cases))[1]/sum(seleccio$expected)
  UCI_age<-
ic.pois(x.tmp=sum(seleccio$cases))[2]/sum(seleccio$expected)
  SIMR_by_ages[i,1]<-i
  SIMR_by_ages[i,2]<-CASES_age
  SIMR_by_ages[i,3]<-SIMR_age
  SIMR_by_ages[i,4]<-LCI_age
  SIMR_by_ages[i,5]<-UCI_age
  rm(seleccio, SIMR_age,LCI_age,UCI_age)
}

names(SIMR_by_ages)<-c("Age group", "CASES", "SIMR", "LCI", "UCI")
SIMR_by_ages

SIMR_by_year<-list()
anys<-unique(tots_sg$year)

SIMR_by_year<-data.frame()

for(j in 1:length(anys))
{
  seleccio<-tots_sg[tots_sg$year==anys[j],]
  SIMR_year<-sum(seleccio$cases)/sum(seleccio$expected)*100
  CASES_year<-sum(seleccio$cases)
  LCI_year<-
ic.pois(x.tmp=sum(seleccio$cases))[1]/sum(seleccio$expected)
  UCI_year<-
ic.pois(x.tmp=sum(seleccio$cases))[2]/sum(seleccio$expected)
  SIMR_by_year[j,1]<-anys[j]
  SIMR_by_year[j,2]<-CASES_year
  SIMR_by_year[j,3]<-SIMR_year
  SIMR_by_year[j,4]<-LCI_year
  SIMR_by_year[j,5]<-UCI_year
  rm(seleccio, SIMR_year,LCI_year, UCI_year)
}

names(SIMR_by_year)<-c("Year", "CASES", "SIMR", "LCI", "UCI")
SIMR_by_year

SIMR_by_age_year<-tots_sg[,1:4]
SIMR_by_age_year$CASES<-tots_sg$cases
SIMR_by_age_year$SIMR<-round(tots_sg$cases/tots_sg$expected,2)*1
SIMR_by_age_year$LCI_year<-tots_sg$expected
SIMR_by_age_year$UCI_year<-tots_sg$expected

for (j in 1:length(tots_sg$expected))
{
  SIMR_by_age_year$LCI_year[j]<-
ic.pois(x.tmp=tots_sg$cases[j])[1]/sum(tots_sg$expected[j])
  SIMR_by_age_year$UCI_year[j]<-
ic.pois(x.tmp=tots_sg$cases[j])[2]/sum(tots_sg$expected[j])
}

```

```

    SIMR[[k]][[1]]<-
list(Global.SIMR=SIMR_global,SIMR.by.ages=SIMR_by_ages,
SIMR.by.year=SIMR_by_year, SIMR.by.age.and.year=SIMR_by_age_year)
  rm(SIMR_global, SIMR_by_ages, SIMR_by_year, SIMR_by_age_year)
}
}

SIMR_global<-vector("list",length(SIMR))
names(SIMR_global)<-names(SIMR)
SIMR_age<-vector("list",length(SIMR))
names(SIMR_age)<-names(SIMR)
SIMR_year<-vector("list",length(SIMR))
names(SIMR_year)<-names(SIMR)
SIMR_age_year<-vector("list",length(SIMR))
names(SIMR_age_year)<-names(SIMR)

for(s in 1:length(SIMR))
{
  locs<-names(SIMR[[s]])
  smr_g<-data.frame()
  locs_a<-rep(locs, each = length(unique(dades$age.group)))
  smr_a<-data.frame()
  locs_y<-rep(locs, each = length(unique(dades$year)))
  smr_y<-data.frame()
  smr_a_y<-data.frame()

  for(c in 1:length(SIMR[[s]]))
  {
    smr_g<-rbind(smr_g,SIMR[[s]][[c]][[1]])
    smr_a<-rbind(smr_a,SIMR[[s]][[c]][[2]])
    smr_y<-rbind(smr_y,SIMR[[s]][[c]][[3]])
    smr_a_y<-rbind(smr_a_y,SIMR[[s]][[c]][[4]])
  }

  SIMR_global[[s]]<-cbind(Group=locs,smr_g)
  SIMR_global[[s]]<-as.data.frame(SIMR_global[[s]])

  SIMR_age[[s]]<-cbind(Group=locs_a,smr_a)
  SIMR_age[[s]]<-as.data.frame(SIMR_age[[s]])

  SIMR_year[[s]]<-cbind(Group=locs_y,smr_y)
  SIMR_year[[s]]<-as.data.frame(SIMR_year[[s]])

  SIMR_age_year[[s]]<-as.data.frame(smr_a_y[,2:8])
  names( SIMR_age_year[[s]])<-c("Age
group","Year","Group","CASES","SIMR","LCI","UCI")

  for(q in 2:4)
  {
    SIMR_global[[s]][,q]<-
round(as.numeric(as.character(SIMR_global[[s]][,q])),2)
  }

  for(q in 3:5)
  {
    SIMR_age[[s]][,q]<-
round(as.numeric(as.character(SIMR_age[[s]][,q])),2)
    SIMR_year[[s]][,q]<-
round(as.numeric(as.character(SIMR_year[[s]][,q])),2)
  }

  for(q in 4:6)
  {

```

```

    SIMR_age_year[[s]][,q]<-
round(as.numeric(as.character(SIMR_age_year[[s]][,q])),2)
}
}

Title1<-"Standardized Incidence or Mortality Ratio by sex and group"
Title2<-"Standardized Incidence or Mortality Ratio by sex, group and
age group"
Title3<-"Standardized Incidence or Mortality Ratio by sex, group and
year (or period)"
Title4<-"Standardized Incidence or Mortality Ratio by sex, group,
age.group and year (or period)"
Table_footer<-"SIMR:Standardized Incidence or Mortality Ratio,
LCI:Lower Confidence Interval, UCI: Upper Confidence Interval"

SIMR_global<-
list(Title=Title1,Table=SIMR_global,Table_footer=Table_footer)
SIMR_age<-list(Title=Title2,Table=SIMR_age,Table_footer=Table_footer)
SIMR_year<-list(Title=Title3,Table=SIMR_year,Table_footer=Table_footer)
SIMR_age_year<-
list(Title=Title4,Table=SIMR_age_year,Table_footer=Table_footer)

for (g in 1:length(SIMR_global$Table))
{
  datos<-SIMR_global$Table[[g]]
  name<-names(SIMR_global$Table)[g]
  names(datos)
  a<-min(datos$LCI)
  b<-max(datos$UCI)
  jpeg(filename=paste("SIMR",name,".jpg",sep=""), quality=100, width =
800, height = 480)
  par(mar=c(5, 20, 4, 4) + 0.1)

plot(datos$SIMR,datos$Group,pch=15,col="black",xlim=c(a,b),axes=F,xlab="S
IMR",ylab="", main=paste("SIMR for", name, ". Period
",min(unique(dades$year)),"-",max(unique(dades$year)),".", sep=" "))
  axis(2,labels=levels(datos[,c(1)]),at=1:length(datos$Group),las=1)
  abline(v=1)
  points(datos$LCI,datos$Group, col="black", pch="|")
  points(datos$UCI,datos$Group, col="black", pch="|")

  for(i in 1:length(datos$Group))
  {
    lines(c(datos$LCI[i],datos$UCI[i]),
c(datos$Group[i],datos$Group[i]))
  }

  q1<-(b-1)/4
  q2<-(1-a)/4

axis(side=1,at=c(round(a,2),round(a+q2,2),round(a+2*q2,2),round(a+3*q2,2)
,round(a+4*q2,2),round(1+q1,2),round(1+2*q1,2),round(1+3*q1,2),round(1+4*
q1,2)),tick=T,cex.axis=0.7)
  dev.off()
  rm(datos,name,a,b)
}
list(Global_SIMR=SIMR_global, SIMR_by_AGE=SIMR_age,
SIMR_by_YEAR=SIMR_year, SIMR_AGE_YEAR=SIMR_age_year,SIMR_T=SIMR)
}

```

ANNEX 2. Altres anàlisis**ANNEX 2.1. Descriptiva de la procedència dels immigrants a Catalunya (2000-2004).**

Per tal de veure un altre tipus de dades, es realitza la descriptiva sobre dades referents a la procedència del immigrants a Catalunya en l'any 2004 utilitzant 14 grups d'edat i els pesos mundials. La taxa truncada es fa per les edats compreses entre 35 i 64 anys i l'acumulada de totes les edats. El percentatge es fa sobre el total de les immigracions. A més a més, en el fitxer s'ha inclòs homes, dones i ambdós sexes, per tant, en la sortida s'obtenen 3 taules.

CS Descriptive Results

Immigrants in Catalonia descriptive. Period: 2004 - 2004 . Age from 0 to + 65 years.

Men

	%	N	CR	ASR(W)	TR(W)	CumR
Europa	22.10	78098	2355.14	2270.16	2375.90	159.56
Àsia	10.08	35608	1073.80	1046.75	912.91	71.06
Àfrica	34.68	122557	3695.86	3918.75	2782.97	254.14
Amèrica	33.08	116888	3524.90	3844.54	2625.20	248.69
Total	100.00	353395	10657.06	11086.87	8705.18	733.93

Women

	%	N	CR	ASR(W)	TR(W)	CumR
Europa	23.72	68658	2013.24	2131.17	1885.35	146.28
Àsia	6.45	18660	547.16	639.19	485.74	41.30
Àfrica	21.40	61939	1816.22	2503.17	1068.74	148.43
Amèrica	48.37	140008	4105.41	4630.72	3366.65	306.15
Total	100.00	289451	8487.48	9909.56	6812.13	642.55

Both sexes

	%	N	CR	ASR(W)	TR(W)	CumR
Europa	22.83	146756	2181.79	2199.91	2132.75	152.88
Àsia	8.44	54268	806.79	848.58	701.75	56.57
Àfrica	28.70	184496	2742.87	3229.58	1936.14	202.61
Amèrica	39.96	256896	3819.22	4229.44	2995.31	276.91
Total	100.00	642846	9557.07	10513.51	7772.90	689.41

%: Percentage, N: Number of cases , CR: Crude Rate x 100000 person-years, ASR: Age Standardized Rate x 100000 person-years, TR: Truncated Rate x 100000 person-years by the 35-64 age group, CumR: Cumulative Rate (percent) by the 0-NA age group, (·): Reference Population.

ANNEX 2.2. Descriptiva de la mortalitat per càncer en les homes. Períodes 1995-1999, 2000-2004 i 1995-2004.

Observar que es realitza l'anàlisi en diferents períodes i per a taxes truncades de 0 a 14 anys i 35-64 anys.

	1995-1999							2000-2004							1995-2004						
	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR
Bufeta de l'orina	5.16	2293	15.33	7.88	0.06	6.95	0.83	5.90	2782	17.55	8.60	0.00	7.64	0.86	5.54	5075	16.47	8.25	0.03	7.29	0.84
Cavitat Oral i Faringe	3.21	1425	9.53	6.46	0.00	14.70	0.75	2.97	1403	8.85	5.81	0.00	12.55	0.69	3.09	2828	9.18	6.11	0.00	13.56	0.72
Còlon	8.37	3718	24.86	13.17	0.00	14.04	1.37	9.01	4252	26.83	13.59	0.00	14.43	1.41	8.70	7970	25.87	13.39	0.00	14.25	1.39
Colorrectal	11.22	4987	33.34	17.79	0.00	19.48	1.87	12.18	5744	36.24	18.56	0.00	20.04	1.98	11.71	10731	34.83	18.19	0.00	19.78	1.93
Esòfag	2.86	1271	8.50	5.51	0.00	11.43	0.65	2.70	1274	8.04	5.04	0.00	9.79	0.62	2.78	2545	8.26	5.27	0.00	10.60	0.63
Estómac	6.16	2739	18.31	10.06	0.00	12.29	1.12	5.64	2662	16.80	8.97	0.00	11.56	0.98	5.90	5401	17.53	9.50	0.00	11.92	1.05
Fetge	5.72	2541	16.99	9.38	0.06	10.85	1.17	5.08	2395	15.11	7.97	0.09	8.92	0.97	5.39	4936	16.02	8.66	0.07	9.90	1.07
Laringe	2.74	1219	8.15	5.13	0.00	9.93	0.62	2.47	1163	7.34	4.37	0.00	7.53	0.52	2.60	2382	7.73	4.73	0.00	8.68	0.57
Leucèmies	2.78	1236	8.26	4.91	1.21	4.29	0.47	3.11	1468	9.26	5.33	1.32	5.25	0.52	2.95	2704	8.78	5.13	1.27	4.77	0.49
Limfoma de Hodgkin	0.33	148	0.99	0.73	0.21	1.24	0.07	0.25	120	0.76	0.50	0.00	0.65	0.04	0.29	268	0.87	0.61	0.11	0.93	0.06
Limfoma no Hodgkin	2.06	915	6.12	3.74	0.39	5.31	0.41	2.42	1141	7.20	4.03	0.17	5.69	0.44	2.24	2056	6.67	3.90	0.27	5.51	0.42
Malaltia immunoproliferativa o Mieloma múltiple	1.62	718	4.80	2.51	0.00	2.38	0.27	1.44	679	4.28	2.16	0.00	2.27	0.23	1.53	1397	4.53	2.32	0.00	2.32	0.25
Melanoma de pell	0.63	279	1.87	1.19	0.00	2.12	0.13	0.71	336	2.12	1.26	0.00	2.08	0.14	0.67	615	2.00	1.22	0.00	2.09	0.13
Pàncreas	3.69	1641	10.97	6.33	0.06	9.17	0.73	3.92	1849	11.67	6.48	0.00	9.06	0.76	3.81	3490	11.33	6.40	0.03	9.10	0.74
Pròstata	9.68	4300	28.75	13.13	0.00	4.27	0.94	8.68	4094	25.83	11.19	0.00	3.98	0.80	9.16	8394	27.25	12.10	0.00	4.12	0.87
Pulmó	27.70	12310	82.30	48.17	0.00	72.65	5.89	27.75	13090	82.59	46.65	0.05	69.92	5.64	27.73	25400	82.45	47.41	0.03	71.28	5.77
Recte	2.86	1269	8.48	4.62	0.00	5.44	0.50	3.16	1492	9.41	4.97	0.00	5.62	0.57	3.01	2761	8.96	4.80	0.00	5.53	0.54
Ronyó	1.79	794	5.31	2.98	0.03	3.93	0.33	1.97	931	5.87	3.12	0.00	3.90	0.35	1.88	1725	5.60	3.05	0.02	3.92	0.34
Sistema Nerviós Central	2.09	928	6.20	4.16	0.78	6.57	0.47	2.26	1067	6.73	4.35	0.98	6.74	0.48	2.18	1995	6.48	4.26	0.89	6.65	0.48
Testicle	0.05	20	0.13	0.11	0.06	0.12	0.01	0.04	20	0.13	0.10	0.05	0.19	0.01	0.04	40	0.13	0.11	0.05	0.16	0.01
Tiroides	0.13	56	0.37	0.21	0.00	0.24	0.03	0.14	66	0.42	0.23	0.00	0.21	0.03	0.13	122	0.40	0.22	0.00	0.22	0.03
Totes les localitzacions	-	44554	297.87	168.91	3.96	223.01	18.71	-	47314	298.53	162.05	3.57	209.70	17.85	-	91868	298.21	165.36	3.75	216.16	18.27
Totes les localitzacions excepte pell no Melanoma	100.00	44435	297.07	168.51	3.96	222.67	18.68	100.00	47171	297.62	161.62	3.57	209.43	17.82	100.00	91606	297.36	164.95	3.75	215.85	18.24

%; Percentatge, N: Nombre de casos, CR: Taxa crua x 100000 persones-any, ASR: Taxa estandaritzada x 100000 persones-anys, TR: Taxa truncada x 100000 persones-anys de 0 a 14 anys i de 35 a 64, CumR: Taxa acumulada (percentatge) de 0 a 74 anys, (-): Població de referència.

ANNEX 2.3. Descriptiva de la mortalitat per càncer en les dones. Períodes 1995-1999, 2000-2004 i 1995-2004.

Observar que es realitza l'anàlisi en diferents períodes i per a taxes truncades de 0 a 14 anys i 35-64 anys.

	1995-1999							2000-2004							1995-2004						
	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR
Bufeta de l'orina	2.12	575	3.67	1.14	0.06	0.66	0.09	2.14	609	3.71	1.04	0.00	0.53	0.09	2.13	1184	3.69	1.09	0.03	0.59	0.09
Cavitat Oral i Faringe	1.07	289	1.84	0.85	0.00	1.34	0.09	1.12	319	1.94	0.85	0.00	1.41	0.09	1.09	608	1.89	0.85	0.00	1.38	0.09
Coll Uterí	1.68	456	2.91	1.73	0.00	3.70	0.19	1.39	395	2.40	1.32	0.00	2.65	0.14	1.53	851	2.65	1.51	0.00	3.15	0.16
Còlon	12.00	3254	20.77	8.26	0.00	10.08	0.86	11.92	3400	20.69	7.76	0.00	9.65	0.79	11.96	6654	20.73	8.00	0.00	9.87	0.82
Colorrectal	15.41	4178	26.67	10.66	0.00	13.16	1.12	15.32	4368	26.59	9.89	0.00	12.27	0.99	15.36	8546	26.63	10.27	0.00	12.72	1.05
Endometri	2.39	647	4.13	1.71	0.00	1.83	0.21	2.33	665	4.05	1.60	0.00	2.00	0.18	2.36	1312	4.09	1.66	0.00	1.91	0.20
Esòfag	0.72	194	1.24	0.50	0.00	0.60	0.05	0.71	203	1.24	0.47	0.00	0.59	0.05	0.71	397	1.24	0.49	0.00	0.60	0.05
Estómac	6.59	1786	11.40	4.38	0.00	5.06	0.43	5.79	1652	10.06	3.70	0.00	4.28	0.37	6.18	3438	10.71	4.03	0.00	4.68	0.40
Fetge	4.92	1335	8.52	3.24	0.06	2.74	0.37	4.47	1275	7.76	2.71	0.00	1.98	0.31	4.69	2610	8.13	2.97	0.03	2.37	0.34
Laringe	0.15	40	0.26	0.13	0.00	0.23	0.01	0.19	53	0.32	0.18	0.00	0.38	0.02	0.17	93	0.29	0.15	0.00	0.31	0.02
Leucèmies	3.84	1042	6.65	3.12	0.65	3.42	0.30	4.09	1166	7.10	3.17	0.89	3.05	0.29	3.97	2208	6.88	3.14	0.77	3.23	0.30
Linfoma de Hodgkin	0.36	98	0.63	0.34	0.04	0.31	0.03	0.32	90	0.55	0.27	0.00	0.30	0.02	0.34	188	0.59	0.31	0.02	0.31	0.03
Linfoma no Hodgkin	3.13	848	5.41	2.44	0.16	3.02	0.26	3.69	1051	6.40	2.58	0.14	3.00	0.28	3.41	1899	5.92	2.52	0.15	3.01	0.27
Malaltia immunoproliferativa o Mieloma múltiple	2.76	748	4.77	1.91	0.00	2.16	0.21	2.44	696	4.24	1.55	0.00	1.64	0.17	2.60	1444	4.50	1.73	0.00	1.89	0.19
Mama	18.19	4932	31.48	16.54	0.09	32.47	1.83	17.54	5003	30.45	14.68	0.00	26.99	1.62	17.86	9935	30.95	15.58	0.05	29.64	1.72
Melanoma de pell	0.86	234	1.49	0.78	0.00	1.35	0.08	0.99	283	1.72	0.84	0.00	1.52	0.08	0.93	517	1.61	0.81	0.00	1.44	0.08
Ovari	4.88	1324	8.45	4.32	0.04	7.74	0.51	5.27	1502	9.14	4.20	0.00	6.74	0.49	5.08	2826	8.80	4.26	0.02	7.23	0.50
Pàncreas	4.89	1326	8.46	3.29	0.00	3.43	0.35	6.02	1717	10.45	3.99	0.00	4.60	0.44	5.47	3043	9.48	3.64	0.00	4.00	0.40
Pulmó	5.24	1420	9.06	4.23	0.00	6.96	0.47	6.23	1778	10.82	5.23	0.00	9.71	0.58	5.75	3198	9.96	4.76	0.00	8.41	0.53
Recte	3.41	924	5.90	2.40	0.00	3.08	0.25	3.39	968	5.89	2.13	0.00	2.62	0.21	3.40	1892	5.89	2.27	0.00	2.85	0.23
Ronyó	1.68	455	2.90	1.29	0.13	1.70	0.13	1.65	471	2.87	1.11	0.15	1.21	0.11	1.66	926	2.88	1.20	0.14	1.46	0.12
Sistema Nerviós Central	2.65	718	4.58	2.66	0.42	4.45	0.30	2.98	850	5.17	2.80	0.56	4.20	0.33	2.82	1568	4.89	2.74	0.49	4.32	0.31
Tiroides	0.49	132	0.84	0.32	0.00	0.34	0.04	0.43	123	0.75	0.28	0.00	0.30	0.03	0.46	255	0.79	0.30	0.00	0.32	0.03
Úter	5.39	1461	9.32	4.39	0.00	6.81	0.49	4.68	1336	8.13	3.60	0.00	5.71	0.40	5.03	2797	8.71	3.97	0.00	6.22	0.44
Totes les localitzacions	-	27194	173.56	77.05	2.72	109.19	8.19	-	28618	174.19	72.39	2.34	100.33	7.63	-	55812	173.88	74.66	2.51	104.66	7.90
Totes les localitzacions excepte pell no Melanoma	100.00	27117	173.07	76.92	2.72	109.17	8.19	100.00	28517	173.57	72.22	2.34	100.23	7.62	100.00	55634	173.33	74.51	2.51	104.60	7.90

%; Percentatge, N: Nombre de casos, CR: Taxa crua x 100000 persones-any, ASR: Taxa estandaritzada x 100000 persones-anys, TR: Taxa truncada x 100000 persones-anys de 0 a 14 anys i de 35 a 64, CumR: Taxa acumulada (percentatge) de 0 a 74 anys, (-): Població de referència.

ANNEX 2.4. Descriptiva de la mortalitat per càncer en l'any 2004 i prenent una taxa acumulada fins als 79 anys. A més, es vol extreure del percentatge "All sites including skin non-melanoma", "Hodgkin Lymphoma" i "Non Hodgkin Lymphoma".

Descriptive Results

Catalonia Cancer Mortality descriptive. Period: 2004 - 2004 . Age from 0 to + 85 years.

Women

	%	N	CR	ASR(W)	TR(W)	CumR
Bladder	2.40	139	4.08	1.18	0.69	0.18
Oral Cavity and Pharynx	1.23	71	2.08	0.95	1.64	0.14
Corpus Uteri	1.30	75	2.20	1.29	2.69	0.15
Colon	12.41	718	21.05	7.86	9.91	1.22
Rectum	3.46	200	5.86	2.09	2.33	0.36
Esophagus	0.81	47	1.38	0.54	0.60	0.08
Stomach	5.53	320	9.38	3.66	4.45	0.54
Liver	4.23	245	7.18	2.50	1.77	0.48
Larynx	0.17	10	0.29	0.17	0.39	0.02
Leukemia	3.85	223	6.54	2.80	3.51	0.38
Multiple myeloma and immunoproliferative neoplasms	2.37	137	4.02	1.48	1.55	0.28
Breast	16.64	963	28.24	13.71	24.60	1.94
Skin Melanoma	0.90	52	1.52	0.79	1.53	0.10
Ovary	5.10	295	8.65	4.06	7.18	0.62
Pancreas	6.43	372	10.91	4.21	5.11	0.73
Lung	7.41	429	12.58	6.05	11.82	0.88
Endometrium	2.07	120	3.52	1.40	1.71	0.24
Kidney	1.56	90	2.64	0.94	0.96	0.17
Brain Central Nerv Sys	2.94	170	4.98	2.62	3.81	0.39
Tyroid	0.41	24	0.70	0.30	0.36	0.05
Cervix	4.22	244	7.15	3.32	5.63	0.49
Hodgkin Lymphoma	NA	21	0.62	0.24	0.30	0.03
Non Hodgkin Lymphoma	NA	208	6.10	2.28	2.52	0.40
All sites including skin non-melanoma	NA	5801	170.10	70.71	101.08	10.83
All sites excluding skin non-melanoma	100.00	5786	169.66	70.56	101.00	10.81

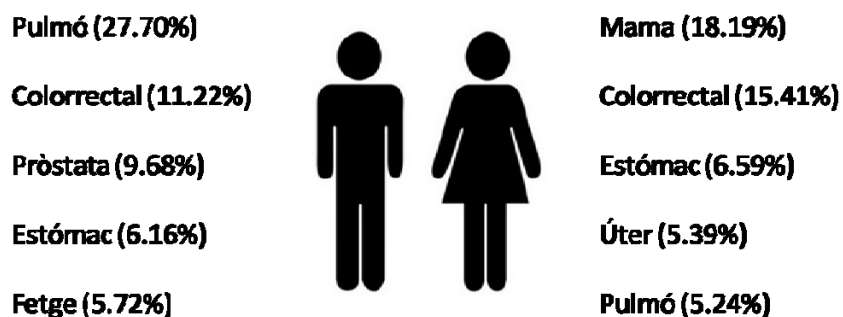
Men

	%	N	CR	ASR(W)	TR(W)	CumR
Bladder	6.48	632	19.06	9.37	8.17	1.48
Oral Cavity and Pharynx	2.84	277	8.35	5.50	11.32	0.79
Colon	9.42	919	27.71	13.96	14.61	2.31
Rectum	3.13	305	9.20	4.90	6.06	0.79
Esophagus	2.51	245	7.39	4.68	9.47	0.68
Stomach	5.71	557	16.80	8.84	10.98	1.49
Liver	4.93	481	14.51	7.98	9.67	1.40
Larynx	2.08	203	6.12	3.68	6.34	0.53
Leukemia	3.18	310	9.35	5.39	5.51	0.78
Multiple myeloma and immunoproliferative neoplasms	1.41	138	4.16	2.15	2.51	0.37
Skin Melanoma	0.63	61	1.84	1.12	2.07	0.15
Pancreas	4.30	419	12.64	7.02	9.57	1.20
Prostate	8.61	840	25.33	11.00	4.24	1.68
Lung	27.31	2664	80.34	44.89	64.97	7.85
Kidney	2.17	212	6.39	3.46	4.39	0.60
Brain Central Nerv Sys	2.23	218	6.57	4.22	6.28	0.63
Testis	0.04	4	0.12	0.08	0.14	0.01
Tyroid	0.14	14	0.42	0.23	0.28	0.03
Hodgkin Lymphoma	NA	24	0.72	0.51	0.75	0.06
Non Hodgkin Lymphoma	NA	200	6.03	3.39	4.84	0.53
All sites including skin non-melanoma	NA	9785	295.08	159.71	204.11	25.97
All sites excluding skin non-melanoma	100.00	9754	294.14	159.25	203.59	25.94

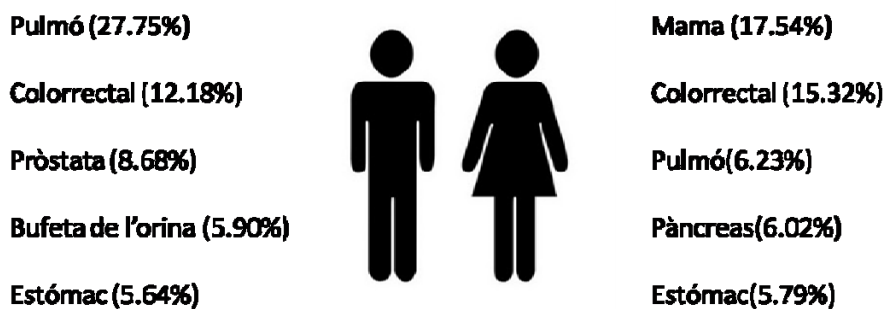
%: Percentage, N: Number of cases , CR: Crude Rate x 100000 person-years, ASR: Age Standardized Rate x 100000 person-years, TR: Truncated Rate x 100000 person-years by the 35-64 age group, CumR: Cumulative Rate (percent) by the 0-79 age group, (-): Reference Population.

ANNEX 2.5. Altres possibles resultats.

Rànquing de defuncions per localització de càncer i sexe. 1995-1999.



Rànquing de defuncions per localització de càncer i sexe. 2000-2004.



Descriptiva de la mortalitat per càncer per homes. Any 1995 i 2004.

	1995							2004						
	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR
Bufeta de l'orina	5.06	442	14.82	8.01	0.00	7.07	0.85	6.48	632	19.06	9.37	0.00	8.17	0.93
Cavitat Oral i Faringe	3.23	282	9.46	6.61	0.00	15.40	0.77	2.84	277	8.35	5.50	0.00	11.32	0.66
Còlon	8.27	723	24.24	13.42	0.00	15.12	1.41	9.42	919	27.71	13.96	0.00	14.61	1.41
Colorrectal	11.06	967	32.42	18.18	0.00	21.06	1.92	12.55	1224	36.91	18.86	0.00	20.67	1.93
Esòfag	3.10	271	9.09	6.11	0.00	12.90	0.70	2.51	245	7.39	4.68	0.00	9.47	0.56
Estómac	6.77	592	19.85	11.21	0.00	13.31	1.21	5.71	557	16.80	8.84	0.00	10.98	0.94
Fetge	5.99	524	17.57	10.15	0.00	12.27	1.26	4.93	481	14.51	7.98	0.00	9.67	0.95
Laringe	2.84	248	8.32	5.29	0.00	9.87	0.66	2.08	203	6.12	3.68	0.00	6.34	0.43
Leucèmies	2.91	254	8.52	5.15	0.75	4.93	0.49	3.18	310	9.35	5.39	1.20	5.51	0.53
Limfoma de Hodgkin	0.31	27	0.91	0.65	0.00	1.22	0.07	0.25	24	0.72	0.51	0.00	0.75	0.05
Limfoma no Hodgkin	2.04	178	5.97	3.67	0.28	4.83	0.42	2.05	200	6.03	3.39	0.19	4.84	0.35
Malaltia immunoproliferativa o Mieloma múltiple	1.70	149	5.00	2.80	0.00	2.91	0.30	1.41	138	4.16	2.15	0.00	2.51	0.24
Melanoma de pell	0.82	72	2.41	1.58	0.00	2.69	0.17	0.63	61	1.84	1.12	0.00	2.07	0.12
Pàncreas	3.76	329	11.03	6.74	0.00	10.51	0.79	4.30	419	12.64	7.02	0.00	9.57	0.83
Pròstata	9.32	815	27.33	13.20	0.00	4.48	0.99	8.61	840	25.33	11.00	0.00	4.24	0.82
Pulmó	27.23	2380	79.80	48.14	0.00	72.04	5.94	27.31	2664	80.34	44.89	0.21	64.97	5.44
Recte	2.79	244	8.18	4.75	0.00	5.94	0.52	3.13	305	9.20	4.90	0.00	6.06	0.52
Ronyó	1.76	154	5.16	2.93	0.00	3.63	0.33	2.17	212	6.39	3.46	0.00	4.39	0.38
SNC	2.24	196	6.57	4.60	0.99	7.36	0.54	2.23	218	6.57	4.22	0.41	6.28	0.48
Testicle	0.08	7	0.23	0.17	0.00	0.28	0.02	0.04	4	0.12	0.08	0.00	0.14	0.01
Tiroides	0.13	11	0.37	0.19	0.00	0.08	0.03	0.14	14	0.42	0.23	0.00	0.28	0.03
Totes les localitzacions	NA	8766	293.93	172.90	2.79	231.88	19.33	NA	9785	295.08	159.71	2.43	204.11	17.52
Totes les localitzacions excepte pell no Melanoma	100.00	8741	293.09	172.43	2.79	231.53	19.29	100.00	9754	294.14	159.25	2.43	203.59	17.49

%: Percentatge, N: Nombre de casos, CR: Taxa crua x 100000 persones-any, ASR: Taxa estandarditzada x 100000 persones-anys, TR: Taxa truncada x 100000 persones-anys de 0 a 14 anys i de 35 a 64, CumR: Taxa acumulada(percentatge) de 0 a 74 anys, (-): Població de referència.

Descriptiva de la mortalitat per càncer per dones. Any 1995 i 2004.

	1995							2004						
	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR	%	N	CR	ASR(W)	TR(W) 0-14	TR(W) 35-64	CumR
Bufeta de l'orina	2.13	116	3.72	1.21	0.00	0.83	0.08	2.40	139	4.08	1.18	0.00	0.69	0.10
Cavitat Oral i Faringe	1.01	55	1.76	0.81	0.00	1.25	0.09	1.23	71	2.08	0.95	0.00	1.64	0.10
Coll Uterí	1.96	107	3.43	2.09	0.00	4.38	0.22	1.30	75	2.20	1.29	0.00	2.69	0.14
Còlon	11.88	648	20.77	8.51	0.00	10.21	0.91	12.41	718	21.05	7.86	0.00	9.91	0.81
Colorrectal	15.01	819	26.25	10.86	0.00	13.15	1.18	15.87	918	26.92	9.95	0.00	12.24	1.03
Endometri	2.14	117	3.75	1.59	0.00	1.77	0.19	2.07	120	3.52	1.40	0.00	1.71	0.16
Esòfag	0.66	36	1.15	0.44	0.00	0.39	0.04	0.81	47	1.38	0.54	0.00	0.60	0.05
Estómac	6.54	357	11.44	4.53	0.00	4.89	0.45	5.53	320	9.38	3.66	0.00	4.45	0.38
Fetge	4.93	269	8.62	3.51	0.00	2.81	0.42	4.23	245	7.18	2.50	0.00	1.77	0.30
Laringe	0.11	6	0.19	0.11	0.00	0.32	0.01	0.17	10	0.29	0.17	0.00	0.39	0.02
Leucèmies	3.65	199	6.38	3.16	0.69	3.52	0.29	3.85	223	6.54	2.80	0.65	3.51	0.26
Limfoma de Hodgkin	0.33	18	0.58	0.38	0.00	0.17	0.03	0.36	21	0.62	0.24	0.00	0.30	0.02
Limfoma no Hodgkin	2.82	154	4.94	2.35	0.33	2.92	0.25	3.59	208	6.10	2.28	0.00	2.52	0.24
Malaltia immunoproliferativa o Mieloma múltiple	2.90	158	5.06	2.24	0.00	2.83	0.26	2.37	137	4.02	1.48	0.00	1.55	0.17
Mama	19.43	1060	33.97	18.76	0.00	39.33	2.01	16.64	963	28.24	13.71	0.00	24.60	1.53
Melanoma de pell	0.84	46	1.47	0.88	0.00	1.63	0.08	0.90	52	1.52	0.79	0.00	1.53	0.08
Ovari	5.32	290	9.29	4.97	0.17	8.87	0.60	5.10	295	8.65	4.06	0.00	7.18	0.45
Pàncreas	5.10	278	8.91	3.68	0.00	3.75	0.40	6.43	372	10.91	4.21	0.00	5.11	0.46
Pulmó	5.04	275	8.81	4.36	0.00	7.37	0.48	7.41	429	12.58	6.05	0.00	11.82	0.64
Recte	3.13	171	5.48	2.35	0.00	2.95	0.26	3.46	200	5.86	2.09	0.00	2.33	0.22
Ronyó	1.92	105	3.36	1.45	0.00	1.88	0.14	1.56	90	2.64	0.94	0.00	0.96	0.10
SNC	2.51	137	4.39	2.58	0.46	4.21	0.30	2.94	170	4.98	2.62	0.44	3.81	0.31
Tiroides	0.42	23	0.74	0.33	0.00	0.49	0.04	0.41	24	0.70	0.30	0.00	0.36	0.04
Úter	5.46	298	9.55	4.68	0.00	7.53	0.51	4.22	244	7.15	3.32	0.00	5.63	0.37
Totes les localitzacions	NA	5473	175.39	81.42	2.04	118.54	8.68	NA	5801	170.10	70.71	1.51	101.08	7.50
Totes les localitzacions excepte pell no Melanoma	100.00	5455	174.81	81.23	2.04	118.46	8.68	100.00	5786	169.66	70.56	1.51	101.00	7.49

%: Percentatge, N: Nombre de casos, CR: Taxa crua x 100000 persones-any, ASR: Taxa estandarditzada x 100000 persones-anys, TR: Taxa truncada x 100000 persones-anys de 0 a 14 anys i de 35 a 64, CumR: Taxa acumulada(percentatge) de 0 a 74 anys, (-): Població de referència.

ANNEX 2.6. Percentatge de canvi de les taxes en els homes durant els períodes 1995-1999, 2000-2004 i 1995-2004. Intervalls al 95% de confiança.

	1995-1999				2000-2004				1995-2004			
	ASR	APC	L.Lim	U.Lim	ASR	APC	L.Lim	U.Lim	ASR	APC	L.Lim	U.Lim
Bufeta de l'orina	7.88	3.15	-4.19	11.06	8.60	4.21	1.08	7.43	8.25	2.23	0.45	4.04
Cavitat Oral i Faringe	6.46	3.90	-3.34	11.68	5.81	-3.17	-4.61	-1.71	6.11	-1.41	-3.54	0.76
Còlon	13.17	2.04	-1.33	5.53	13.59	-0.65	-3.76	2.55	13.39	0.63	-0.46	1.72
Colorrectal	17.79	1.43	-1.38	4.32	18.56	-0.89	-3.80	2.10	18.19	0.69	-0.29	1.68
Esòfag	5.51	-2.56	-7.98	3.19	5.04	-2.03	-5.64	1.72	5.27	-1.89	-3.36	-0.41
Estómac	10.06	-3.36	-6.74	0.15	8.97	-2.87	-6.29	0.67	9.50	-2.50	-3.58	-1.41
Fetge	9.38	-3.32	-6.68	0.16	7.97	-2.73	-6.45	1.13	8.66	-3.16	-4.24	-2.06
Laringe	5.13	-1.26	-3.81	1.35	4.37	-7.05	-9.14	-4.92	4.73	-3.22	-4.46	-1.96
Leucèmies	4.91	-0.90	-3.93	2.22	5.33	0.84	-3.45	5.33	5.13	1.30	0.00	2.61
Limfoma de Hodgkin	0.73	-0.37	-13.08	14.20	0.50	0.16	-14.05	16.71	0.61	-5.25	-9.82	-0.46
Limfoma no Hodgkin	3.74	2.61	-2.13	7.57	4.03	-1.80	-9.77	6.87	3.90	1.16	-1.12	3.49
Malaltia immunoproliferativa o Mieloma múltiple	2.51	-3.32	-8.31	1.94	2.16	0.33	-6.78	7.99	2.32	-2.72	-4.68	-0.73
Melanoma de pell	1.19	-4.77	-22.45	16.93	1.26	-0.46	-7.55	7.17	1.22	0.33	-4.28	5.17
Pàncreas	6.33	-1.33	-4.94	2.41	6.48	3.46	-2.89	10.24	6.40	0.61	-1.16	2.42
Pròstata	13.13	-0.81	-3.58	2.05	11.19	-0.91	-2.98	1.21	12.10	-2.57	-3.59	-1.55
Pulmó	48.17	1.90	-0.55	4.41	46.65	-1.24	-2.43	-0.03	47.41	-0.39	-1.22	0.45
Recte	4.62	-0.26	-1.67	1.17	4.97	-1.53	-4.24	1.26	4.80	0.87	-0.14	1.90
Ronyó	2.98	2.00	-2.96	7.21	3.12	0.55	-4.99	6.41	3.05	0.97	-0.67	2.64
Sistema Nerviós Central	4.16	1.09	-7.54	10.52	4.35	0.49	-3.64	4.80	4.26	0.90	-1.18	3.01
Testicle	0.11	-22.44	-45.96	11.31	0.10	-3.25	-30.36	34.43	0.11	-5.32	-15.33	5.88
Tiroides	0.21	24.38	-9.92	71.74	0.23	-11.19	-22.91	2.31	0.22	2.22	-6.42	11.66
Totes les localitzacions	168.91	0.40	-1.37	2.20	162.05	-0.88	-2.04	0.29	165.36	-0.67	-1.20	-0.14
Totes les localitzacions excepte pell no Melanoma	168.51	0.38	-1.36	2.15	161.62	-0.90	-2.06	0.28	164.95	-0.68	-1.21	-0.15

ASR(W): Taxa estandarditzada x 100000 persones-any. APC: Percentatge de canvi de les taxes. L.Lim: Interval de confiança inferior. U.Lim: Interval de confiança superior.

ANNEX 2.7. Percentatge de canvi de les taxes en les dones durant els períodes 1995-1999, 2000-2004 i 1995-2004. Intervalls al 95% de confiança.

	1995-1999				2000-2004				1995-2004			
	ASR	APC	L.Lim	U.Lim	ASR	APC	L.Lim	U.Lim	ASR	APC	L.Lim	U.Lim
Bufeta de l'orina	1.14	2.38	-3.73	8.88	1.04	1.87	-4.90	9.13	1.09	-0.92	-3.15	1.37
Cavitat Oral i Faringe	0.85	-0.65	-8.66	8.07	0.85	7.30	2.67	12.14	0.85	0.93	-1.65	3.58
Coll Uterí	1.73	-3.77	-11.20	4.29	1.32	-8.72	-15.69	-1.18	1.51	-5.31	-7.75	-2.81
Còlon	8.26	1.06	-2.75	5.01	7.76	0.85	0.17	1.54	8.00	-0.70	-1.76	0.37
Colorrectal	10.66	0.17	-1.52	1.89	9.89	0.72	0.08	1.36	10.27	-1.01	-1.69	-0.32
Endometri	1.71	-1.48	-7.91	5.40	1.60	-4.60	-7.84	-1.23	1.66	-1.67	-3.44	0.13
Esòfag	0.50	2.16	-7.99	13.43	0.47	9.03	4.07	14.23	0.49	0.18	-3.14	3.62
Estómac	4.38	-3.03	-5.14	-0.87	3.70	-2.42	-4.84	0.06	4.03	-3.21	-3.92	-2.49
Fetge	3.24	-0.08	-4.58	4.62	2.71	-3.69	-5.93	-1.40	2.97	-3.08	-4.48	-1.67
Laringe	0.13	6.64	-11.87	29.03	0.18	16.32	-11.75	53.32	0.15	8.99	0.48	18.22
Leucèmies	3.12	0.94	-1.12	3.05	3.17	-0.31	-6.49	6.28	3.14	0.29	-1.16	1.77
Limfoma de Hodgkin	0.34	-1.22	-16.45	16.79	0.27	-17.41	-31.55	-0.35	0.31	-5.19	-10.78	0.74
Limfoma no Hodgkin	2.44	5.04	-1.68	12.22	2.58	-3.42	-7.24	0.55	2.52	0.98	-1.19	3.20
Malaltia immunoproliferativa o Mieloma múltiple	1.91	-4.43	-10.21	1.71	1.55	-3.59	-5.25	-1.89	1.73	-4.18	-5.63	-2.70
Mama	16.54	-3.97	-7.89	0.12	14.68	-3.46	-6.23	-0.61	15.58	-2.69	-3.84	-1.52
Melanoma de pell	0.78	3.92	-7.38	16.61	0.84	-5.25	-8.05	-2.36	0.81	0.82	-2.15	3.88
Ovari	4.32	-4.26	-7.88	-0.50	4.20	-1.42	-5.08	2.37	4.26	-1.12	-2.53	0.30
Pàncreas	3.29	-0.73	-6.84	5.79	3.99	1.08	-2.93	5.26	3.64	3.11	1.16	5.09
Pulmó	4.23	2.38	-3.77	8.94	5.23	9.49	5.43	13.71	4.76	4.95	2.88	7.06
Recte	2.40	-2.75	-9.19	4.14	2.13	0.26	-1.53	2.07	2.27	-2.12	-3.76	-0.44
Ronyó	1.29	-0.13	-7.64	8.00	1.11	-2.82	-14.12	9.96	1.20	-2.53	-5.51	0.55
Sistema Nerviós Central	2.66	5.16	0.78	9.73	2.80	0.77	-4.00	5.79	2.74	1.45	-0.20	3.12
Tiroides	0.32	6.83	-8.02	24.08	0.28	-1.43	-11.00	9.18	0.30	-1.21	-5.54	3.32
Úter	4.39	-3.35	-3.99	-2.71	3.60	-5.08	-7.39	-2.71	3.97	-3.91	-4.48	-3.35
Totes les localitzacions	77.05	-0.96	-3.33	1.46	72.39	-0.80	-2.02	0.44	74.66	-1.16	-1.75	-0.56
Totes les localitzacions excepte pell no Melanoma	76.92	-0.94	-3.29	1.46	72.22	-0.78	-1.99	0.44	74.51	-1.16	-1.75	-0.57

ASR(W): Taxa estandarditzada x 100000 persones-any. APC: Percentatge de canvi de les taxes. L.Lim: Interval de confiança inferior. U.Lim: Interval de confiança superior.

ANNEX 2.8 . Càlcul del EAPC i el seu interval de confiança al 90% en els homes pel període 2000-2004.

EAPC Results

Estimated Annual Percentatge of Change for the period 2000 - 2004 . 90 % Confidence Interval.

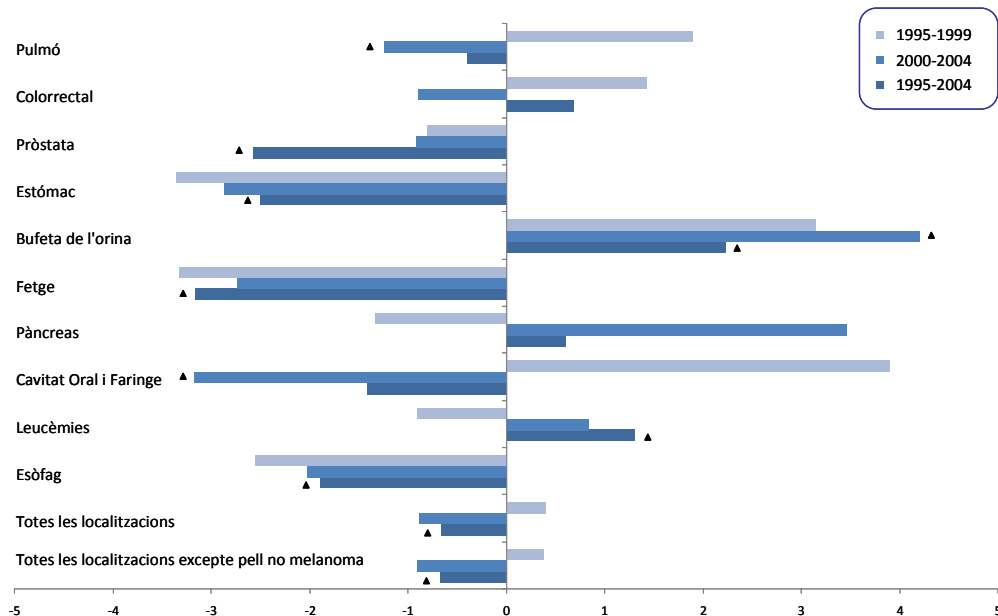
Men

	EAPC	LCI	UCI
Bladder	4.21	1.6	6.90
Oral Cavity and Pharynx	-3.17	-4.4	-1.95
Colon	-0.65	-3.3	2.03
Rectum	-1.53	-3.8	0.81
Esophagus	-2.03	-5.1	1.11
Stomach	-2.87	-5.8	0.09
Liver	-2.73	-5.9	0.50
Larynx	-7.05	-8.8	-5.26
Leukemia	0.84	-2.8	4.60
Hodgkin Lymphoma	0.16	-11.9	13.87
Non Hodgkin Lymphoma	-1.80	-8.5	5.43
Multiple myeloma and immunoproliferative neoplasms	0.33	-5.7	6.72
Skin Melanoma	-0.46	-6.4	5.90
Pancreas	3.46	-1.9	9.12
Prostate	-0.91	-2.6	0.87
Lung	-1.24	-2.2	-0.22
Kidney	0.55	-4.1	5.45
Brain Central Nerv Sys	0.49	-3.0	4.09
Testis	-3.25	-26.6	27.51
Tyroid	-11.19	-21.1	0.01
All sites including skin non-melanoma	-0.88	-1.9	0.10
All sites excluding skin non-melanoma	-0.90	-1.9	0.09

EAPC: Estimated Annual Percentage of Change, L.lim: Lower limit, U.lim: Upper limit

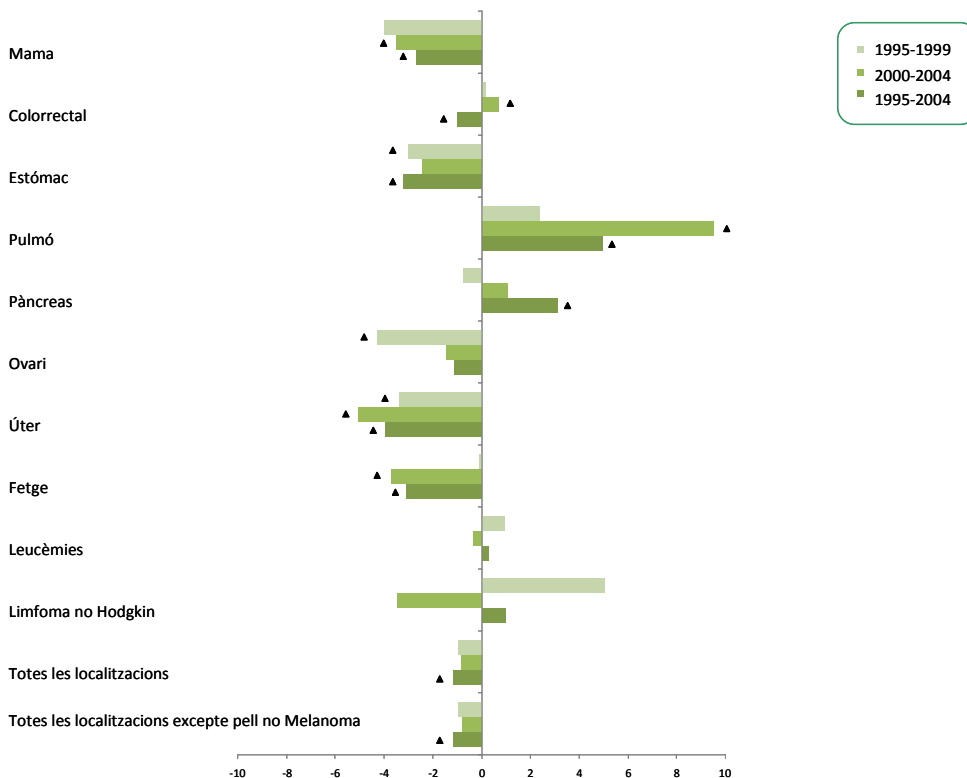
ANNEX 2.9. Altres possibles resultats

Percentatge de canvi en diferents períodes per les localitzacions més freqüents en Homes.



▲ Percentatge de canvi significatiu

Percentatge de canvi en diferents períodes per les localitzacions més freqüents en Dones.



▲ Percentatge de canvi significatiu

ANNEX 2.10. Selecció del millor model per localització i homes.

Localització	Critèri	model.nb1	model.nbl	model.p1	model.pl	Model Seleccionat
Bufeta de l'orina	AIC	779.35	796.46	779.18	794.46	Mnimum AIC
	GOF	1.00	1.00	0.99	1.00	
Cavitat Oral i Faringe	AIC	801.58	821.24	799.58	819.24	Mnimum AIC
	GOF	0.97	0.97	0.97	0.97	
Còlon	AIC	861.05	881.83	859.04	879.82	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Colorrectal	AIC	925.61	942.11	923.74	940.10	Mnimum AIC
	GOF	1.00	1.00	0.99	1.00	
Esòfag	AIC	766.81	779.52	765.47	777.52	Mnimum AIC
	GOF	1.00	1.00	0.99	1.00	
Estómac	AIC	866.83	887.73	864.82	885.73	Mnimum AIC
	GOF	0.94	0.89	0.93	0.89	
Fetge	AIC	859.21	867.14	857.69	865.14	Mnimum AIC
	GOF	0.56	0.75	0.36	0.75	
Laringe	AIC	767.11	775.17	768.91	773.70	Mnimum AIC
	GOF	1.00	1.00	0.88	0.98	
Leucèmies	AIC	904.28	922.74	902.28	920.74	Mnimum AIC
	GOF	0.73	0.71	0.73	0.71	
Limfoma de Hodgkin	AIC	544.15	563.88	542.15	561.88	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Limfoma no Hodgkin	AIC	828.78	838.76	827.21	836.77	Mnimum AIC
	GOF	0.92	0.91	0.82	0.90	
MIMP	AIC	644.78	670.09	642.78	668.09	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Melanoma de pell	AIC	624.31	653.84	622.30	651.84	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Pàncreas	AIC	804.16	828.02	802.29	826.02	Mnimum AIC
	GOF	0.99	0.96	0.97	0.96	
Pulmó	AIC	1083.71	1083.75	1084.21	1081.97	Mnimum AIC
	GOF	0.82	0.85	0.33	0.74	
Recte	AIC	757.45	779.98	755.83	777.99	Mnimum AIC
	GOF	1.00	0.99	0.99	0.99	
Ronyó	AIC	709.35	730.21	707.34	728.20	Mnimum AIC
	GOF	0.99	0.99	0.99	0.99	
Sistema Nerviós Central	AIC	874.61	894.93	872.61	892.93	Mnimum AIC
	GOF	0.52	0.40	0.51	0.40	
Tiroides	AIC	398.18	427.85	396.18	425.85	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Pròstata	AIC	732.25	753.80	730.53	751.79	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Testicle	AIC	324.42	352.39	322.41	350.38	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Totes les localitzacions	AIC	1458.45	1436.45	1459.54	1434.44	Mnimum AIC
	GOF	0.06	0.19	0.00	0.19	
Totes les localitzacions excepte pell no Melanoma	AIC	1455.47	1435.03	1455.96	1433.02	Mnimum AIC
	GOF	0.06	0.21	0.00	0.20	

■ Models que es seleccionen per cada localització i sexe; GOF: p-valor de la bondat d'ajust. AIC: Critèri d'informació d'Alkaike; model.nb1: Model edat-tendència amb distribució binomial negativa; model.nbl: Model de pendent específic per edat amb distribució binomial negativa; model.p1: Model edat-tendència amb distribució poisson; model.pl: Model de pendent específic per edat amb distribució poisson.

ANNEX 2.11. Selecció del millor model per localització i dones.

Localització	Critèri	model.nb1	model.nbl	model.p1	model.pl	Model Seleccionat
Bufeta de l'orina	AIC	555.84	578.51	553.84	576.50	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Cavitat Oral i Faringe	AIC	570.53	599.96	568.53	597.96	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Còlon	AIC	859.19	876.97	857.19	874.96	Mnimum AIC
	GOF	0.99	1.00	0.99	1.00	
Colorrectal	AIC	890.44	906.25	888.43	904.24	Mnimum AIC
	GOF	0.99	1.00	0.99	1.00	
Esòfag	AIC	507.19	536.83	505.19	534.83	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Estómac	AIC	777.40	793.39	775.40	791.39	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Fetge	AIC	696.08	724.97	694.07	722.96	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Laringe	AIC	388.31	417.33	386.31	415.33	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Leucèmies	AIC	870.27	884.47	868.26	882.47	Mnimum AIC
	GOF	0.36	0.41	0.35	0.41	
Limfoma de Hodgkin	AIC	487.89	506.10	485.89	504.09	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Limfoma no Hodgkin	AIC	753.04	771.78	751.03	769.77	Mnimum AIC
	GOF	0.99	0.99	0.99	0.99	
MIMP	AIC	627.51	647.54	625.51	645.54	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Melanoma de pell	AIC	614.96	632.10	612.96	630.10	Mnimum AIC
	GOF	0.99	1.00	0.99	1.00	
Pàncreas	AIC	751.10	770.02	749.10	768.01	Mnimum AIC
	GOF	0.99	0.99	0.99	0.99	
Pulmó	AIC	841.49	833.41	839.70	831.40	Mnimum AIC
	GOF	0.72	0.97	0.57	0.97	
Recte	AIC	693.11	716.32	691.10	714.32	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Ronyó	AIC	661.91	684.25	660.02	682.25	Mnimum AIC
	GOF	0.74	0.56	0.65	0.56	
Sistema Nervios Central	AIC	772.32	792.17	770.31	790.16	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Tiroides	AIC	459.87	487.03	457.87	485.02	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Coll Uteri	AIC	677.11	701.46	675.10	699.46	Mnimum AIC
	GOF	0.99	0.98	0.99	0.98	
Endometri	AIC	618.56	643.22	616.56	641.22	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Mama	AIC	1010.5	1009.44	1011.53	1007.44	Mnimum AIC
	GOF	0.91	0.85	0.41	0.85	
Ovari	AIC	814.23	818.71	812.40	816.71	Mnimum AIC
	GOF	0.94	0.98	0.88	0.98	
Úter	AIC	769.44	795.38	767.43	793.37	Mnimum AIC
	GOF	1.00	1.00	1.00	1.00	
Totes les localitzacions	AIC	1350.52	1354.68	1348.79	1352.66	Mnimum AIC
	GOF	0.19	0.29	0.11	0.29	
Totes les localitzacions excepte pell no Melanoma	AIC	1349.50	1352.70	1347.74	1350.68	Mnimum AIC
	GOF	0.20	0.32	0.12	0.32	

■ Models que es seleccionen per cada localització i sexe; GOF: p-valor de la bondat d'ajust. AIC: Crítèri d'informació d'Akaïke; model.nb1: Model edat-tendència amb distribució binomial negativa; model.nbl: Model de pendent específic per edat amb distribució binomial negativa; model.p1: Model edat-tendència amb distribució poisson; model.pl: Model de pendent específic per edat amb distribució poisson.

ANNEX 2.12. Descriptiva de les estimacions de la mortalitat per càncer a Catalunya en els anys 2010 i 2015 per localització i sexe.

Taula A1. SART Descriptive Results. Catalonia Cancer Mortality descriptive. Period: 2010 - 2010 . Age from 0 to + 85 years. Women

	%	N	CR	ASR(W)	TR(W)	CumR
Bufeta de l'orina	2.37	153	4.10	1.08	0.64	0.09
Cavitat Oral i Faringe	1.30	84	2.25	0.93	1.54	0.10
Coll Uteri	1.02	66	1.77	0.95	2.00	0.10
Còlon	12.48	806	21.57	7.47	9.27	0.77
Colorrectal	15.85	1024	27.41	9.47	11.82	0.97
Endometri	2.15	139	3.72	1.43	1.70	0.17
Esòfag	0.82	53	1.42	0.54	0.70	0.06
Estómac	4.86	314	8.40	2.83	3.40	0.28
Fetge	3.93	254	6.80	2.30	1.90	0.26
Laringe	0.39	25	0.67	0.43	0.66	0.04
Leucèmies	4.67	302	8.08	3.48	3.57	0.33
Limfoma de Hodgkin	0.33	21	0.56	0.29	0.32	0.03
Limfoma no Hodgkin	4.77	308	8.24	3.41	3.97	0.35
Malaltia immunoproliferativa o Mieloma múltiple	2.07	134	3.59	1.27	1.46	0.14
Mama	16.42	1061	28.40	12.06	19.59	1.32
Melanoma de pell	1.22	79	2.11	1.04	1.68	0.10
Ovari	5.62	363	9.72	4.48	7.60	0.52
Pàncreas	7.94	513	13.73	4.95	5.40	0.53
Pulmó	9.57	618	16.54	8.71	19.30	0.95
Recte	3.42	221	5.92	2.05	2.62	0.21
Ronyó	1.58	102	2.73	1.07	1.35	0.11
Sistema Nerviós Central	3.75	242	6.48	3.60	5.61	0.41
Tiroides	0.45	29	0.78	0.29	0.38	0.03
Úter	3.72	240	6.42	2.74	4.35	0.30
Totes les localitzacions	NA	6486	173.60	67.95	95.16	7.18
Totes les localitzacions excepte pell no Melanoma	100.00	6460	172.90	67.79	95.04	7.17

%: Percentage, N: Number of cases , CR: Crude Rate x 100000 person-years, ASR: Age Standardized Rate x 100000 person-years, TR: Truncated Rate x 100000 person-years by the 35-64 age group, CumR: Cumulative Rate (percent) by the 0-74 age group, (-): Reference Population.

Taula A2. SART Descriptive Results. Catalonia Cancer Mortality descriptive. Period: 2010 - 2010 . Age from 0 to + 85 years. Men.

	%	N	CR	ASR(W)	TR(W)	CumR
Bufeta de l'orina	7.60	866	23.62	10.48	9.31	1.07
Cavitat Oral i Faringe	2.75	314	8.56	5.40	12.03	0.64
Còlon	9.96	1136	30.98	14.27	15.16	1.48
Colorrectal	13.42	1530	41.73	19.52	21.24	2.06
Esòfag	2.40	274	7.47	4.54	9.13	0.54
Estómac	4.78	545	14.86	7.32	9.21	0.81
Fetge	4.01	457	12.46	6.36	7.32	0.78
Laringe	1.97	225	6.14	3.54	6.49	0.42
Leucèmies	3.62	413	11.26	5.99	5.60	0.57
Limfoma de Hodgkin	0.23	26	0.71	0.49	0.70	0.04
Limfoma no Hodgkin	2.86	326	8.89	4.90	6.78	0.52
Malaltia immunoproliferativa o Mieloma múltiple	1.26	144	3.93	1.79	1.83	0.19
Melanoma de pell	0.80	91	2.48	1.44	2.44	0.15
Pàncreas	4.36	497	13.55	7.12	10.09	0.83
Pulmó	27.77	3166	86.35	45.96	68.87	5.46
Recte	3.48	397	10.83	5.29	6.15	0.59
Ronyó	2.26	258	7.04	3.60	4.57	0.39
Sistema Nervios Central	2.49	284	7.75	4.99	7.62	0.55
Tiroides	0.18	21	0.57	0.34	0.40	0.04
Pròstata	7.96	908	24.76	8.96	3.16	0.64
Testicle	0.07	8	0.22	0.16	0.25	0.01
Totes les localitzacions	NA	11452	312.33	154.79	195.35	16.80
Totes les localitzacions excepte pell no Melanoma	100.00	11401	310.94	154.25	194.87	16.76

%: Percentage, N: Number of cases , CR: Crude Rate x 100000 person-years, ASR: Age Standardized Rate x 100000 person-years, TR: Truncated Rate x 100000 person-years by the 35-64 age group, CumR: Cumulative Rate (percent) by the 0-74 age group, (-): Reference Population.

Taula A.3. SART Descriptive Results. Catalonia Cancer Mortality descriptive. Period: 2015 - 2015 . Age from 0 to + 85 years. Women.

	%	N	CR	ASR(W)	TR(W)	CumR
Bufeta de l'orina	2.45	165	4.24	1.06	0.65	0.08
Cavitat Oral i Faringe	1.39	94	2.42	0.97	1.49	0.10
Coll Uteri	0.80	54	1.39	0.72	1.51	0.08
Còlon	12.70	857	22.02	7.19	8.90	0.74
Colorrectal	16.02	1081	27.77	9.08	11.27	0.93
Endometri	2.06	139	3.57	1.33	1.58	0.16
Esòfag	0.79	53	1.36	0.47	0.65	0.05
Estómac	4.27	288	7.40	2.35	2.76	0.23
Fetge	3.60	243	6.24	2.03	1.71	0.23
Laringe	0.47	32	0.82	0.50	0.78	0.04
Leucèmies	5.01	338	8.68	3.60	3.62	0.34
Limfoma de Hodgkin	0.31	21	0.54	0.29	0.29	0.03
Limfoma no Hodgkin	5.47	369	9.48	3.76	4.47	0.39
Malaltia immunoproliferativa o Mieloma múltiple	1.79	121	3.11	1.04	1.18	0.11
Mama	16.40	1107	28.44	10.97	16.20	1.19
Melanoma de pell	1.36	92	2.36	1.22	1.87	0.11
Ovari	5.90	398	10.23	4.58	7.63	0.53
Pàncreas	9.44	637	16.37	5.70	6.13	0.61
Pulmó	13.44	907	23.30	12.86	31.43	1.36
Recte	3.38	228	5.86	1.93	2.42	0.20
Ronyó	1.53	103	2.65	1.03	1.14	0.10
Sistema Nerviós Central	4.28	289	7.43	4.01	6.21	0.45
Tiroides	0.41	28	0.72	0.26	0.35	0.03
Úter	3.20	216	5.55	2.26	3.56	0.25
Totes les localitzacions	NA	6777	174.12	64.78	90.79	6.85
Totes les localitzacions excepte pell no Melanoma	100.00	6748	173.38	64.63	90.67	6.84

Taula A.4. SART Descriptive Results. Catalonia Cancer Mortality descriptive. Period: 2015 - 2015 . Age from 0 to + 85 years. Men.

	%	N	CR	ASR(W)	TR(W)	CumR
Bufeta de l'orina	8.70	1094	28.55	11.72	10.42	1.19
Cavitat Oral i Faringe	2.61	328	8.56	5.07	11.25	0.60
Còlon	10.47	1317	34.37	14.76	15.69	1.53
Colorrectal	14.11	1774	46.30	20.11	21.86	2.13
Esòfag	2.23	281	7.33	4.19	8.46	0.50
Estómac	4.24	533	13.91	6.38	8.07	0.71
Fetge	3.43	431	11.25	5.40	6.18	0.67
Laringe	1.70	214	5.58	3.02	5.55	0.36
Leucèmies	3.91	492	12.84	6.36	5.95	0.61
Limfoma de Hodgkin	0.17	22	0.57	0.37	0.54	0.03
Limfoma no Hodgkin	3.10	390	10.18	5.26	7.35	0.57
Malaltia immunoproliferativa o Mieloma múltiple	1.13	142	3.71	1.58	1.61	0.17
Melanoma de pell	0.83	104	2.71	1.49	2.55	0.16
Pàncreas	4.62	581	15.16	7.54	10.70	0.87
Pulmó	27.85	3501	91.37	45.60	68.36	5.34
Recte	3.68	463	12.08	5.49	6.41	0.61
Ronyó	2.43	306	7.99	3.86	4.83	0.42
Sistema Nerviós Central	2.58	324	8.46	5.18	7.97	0.57
Tiroides	0.19	24	0.63	0.34	0.41	0.04
Pròstata	7.25	911	23.78	7.75	2.79	0.55
Testicle	0.05	6	0.16	0.08	0.23	0.01
Totes les localitzacions	NA	12646	330.04	150.41	186.62	16.16
Totes les localitzacions excepte pell no Melanoma	100.00	12573	328.13	149.83	186.18	16.12

%: Percentage, N: Number of cases , CR: Crude Rate x 100000 person-years, ASR: Age Standardized Rate x 100000 person-years, TR: Truncated Rate x 100000 person-years by the 35-64 age group, CumR: Cumulative Rate (percent) by the 0-74 age group, (-): Reference Population.

ANNEX 2.13. Estimació del nombre de casos morts per càncer a la comarca de l'Alt Penedès en el 2007.

Model selection by group and sex

Models fitted

	group	sex	model.chosen	pv.test	significant	method.AIC.GOF
1	Bufeta de l'orina	2	Poisson age-drift	1.00	No	Minimum AIC
2	Bufeta de l'orina	1	Poisson age-drift	0.99	No	Minimum AIC
3	Cavitat Oral i Faringe	2	Poisson age-drift	1.00	No	Minimum AIC
4	Cavitat Oral i Faringe	1	Poisson age-drift	0.96	No	Minimum AIC
5	Cell Uter	2	Poisson age-drift	0.99	No	Minimum AIC
6	Colon	2	Poisson age-drift	0.99	No	Minimum AIC
7	Colon	1	Poisson age-drift	1.00	No	Minimum AIC
8	Colorrectal	2	Poisson age-drift	0.99	No	Minimum AIC
9	Colorrectal	1	Poisson age-drift	0.99	No	Minimum AIC
10	Endometri	2	Poisson age-drift	1.00	No	Minimum AIC
11	Esofag	2	Poisson age-drift	1.00	No	Minimum AIC
12	Esofag	1	Poisson age-drift	0.99	No	Minimum AIC
13	Estomac	2	Poisson age-drift	1.00	No	Minimum AIC
14	Estomac	1	Poisson age-drift	0.93	No	Minimum AIC
15	Fatge	2	Poisson age-drift	1.00	No	Minimum AIC
16	Fatge	1	Poisson age-drift	0.36	No	Minimum AIC
17	Laringe	2	Poisson age-drift	1.00	No	Minimum AIC
18	Laringe	1	Negative Binomial age-drift	1.00	No	Minimum AIC
19	Leucèmies	2	Poisson age-drift	0.35	No	Minimum AIC
20	Leucèmies	1	Poisson age-drift	0.73	No	Minimum AIC
21	Linfoma de Hodgkin	2	Poisson age-drift	1.00	No	Minimum AIC
22	Linfoma de Hodgkin	1	Poisson age-drift	1.00	No	Minimum AIC
23	Linfoma no Hodgkin	2	Poisson age-drift	0.99	No	Minimum AIC
24	Linfoma no Hodgkin	1	Poisson age-drift	0.82	No	Minimum AIC
25	Malaltia immunoproliferativa o Mieloma múltiple	2	Poisson age-drift	1.00	No	Minimum AIC
26	Malaltia immunoproliferativa o Mieloma múltiple	1	Poisson age-drift	1.00	No	Minimum AIC
27	Mama	2	Poisson age-specific slope	0.85	No	Minimum AIC
28	Melanoma de pell	2	Poisson age-drift	0.99	No	Minimum AIC
29	Melanoma de pell	1	Poisson age-drift	1.00	No	Minimum AIC
30	Ovari	2	Poisson age-drift	0.88	No	Minimum AIC
31	Pancreas	2	Poisson age-drift	0.99	No	Minimum AIC
32	Pancreas	1	Poisson age-drift	0.97	No	Minimum AIC
33	Pulmo	2	Poisson age-specific slope	0.97	No	Minimum AIC
34	Pulmo	1	Poisson age-specific slope	0.74	No	Minimum AIC
35	Recte	2	Poisson age-drift	1.00	No	Minimum AIC
36	Recte	1	Poisson age-drift	0.99	No	Minimum AIC
37	Ronyó	2	Poisson age-drift	0.65	No	Minimum AIC
38	Ronyó	1	Poisson age-drift	0.99	No	Minimum AIC
39	Sistema Nerviós Central	2	Poisson age-drift	1.00	No	Minimum AIC
40	Sistema Nerviós Central	1	Poisson age-drift	0.51	No	Minimum AIC
41	Tiroides	2	Poisson age-drift	1.00	No	Minimum AIC
42	Tiroides	1	Poisson age-drift	1.00	No	Minimum AIC
43	Totes les localitzacions	2	Poisson age-drift	0.11	No	Minimum AIC
44	Totes les localitzacions	1	Poisson age-specific slope	0.19	No	Minimum AIC
45	Totes les localitzacions excepte pell no Melanoma	2	Poisson age-drift	0.12	No	Minimum AIC
46	Totes les localitzacions excepte pell no Melanoma	1	Poisson age-specific slope	0.20	No	Minimum AIC
47	Uter	2	Poisson age-drift	1.00	No	Minimum AIC
48	Pròstata	1	Poisson age-drift	1.00	No	Minimum AIC
49	Testicle	1	Poisson age-drift	1.00	No	Minimum AIC

group: Group of interest; sex: 0= Both sexes, 1= Men, 2= Women; model.chosen: The fitted model chosen to predict; pv.test: The p-value of the model goodness of fit test; significant: Is the preceding p.value significant?; method.AIC.GOF: The method used for model selection (AIC: Akaike Information Criterion; GOF: chi-square goodness of fit test)

MOSTRA DE LES ESTIMACIONS DE MORTS PER CÀNCER A L'ALT PENEDEÈS EN EL 2007:

```
sex;age.group;year;group;cases;population
2;1;2007;Bufeta de l'orina;0;3097
2;2;2007;Bufeta de l'orina;0;2595
2;3;2007;Bufeta de l'orina;0;2143
...
1;1;2007;Colorrectal;0;3269
1;2;2007;Colorrectal;0;2608
1;3;2007;Colorrectal;0;2270
1;4;2007;Colorrectal;0;2212
1;5;2007;Colorrectal;0;2660
1;6;2007;Colorrectal;0;3881
1;7;2007;Colorrectal;0;4808
1;8;2007;Colorrectal;0;4537
1;9;2007;Colorrectal;0;4031
1;10;2007;Colorrectal;0;3573
1;11;2007;Colorrectal;1;3011
1;12;2007;Colorrectal;1;2687
1;13;2007;Colorrectal;1;2318
1;14;2007;Colorrectal;2;1716
1;15;2007;Colorrectal;3;1679
1;16;2007;Colorrectal;4;1533
1;17;2007;Colorrectal;4;1058
1;18;2007;Colorrectal;3;724
...
2;1;2007;Totes les localitzacions excepte pell no Melanoma;0;3097
2;2;2007;Totes les localitzacions excepte pell no Melanoma;0;2595
2;3;2007;Totes les localitzacions excepte pell no Melanoma;0;2143
2;4;2007;Totes les localitzacions excepte pell no Melanoma;0;2191
2;5;2007;Totes les localitzacions excepte pell no Melanoma;0;2473
2;6;2007;Totes les localitzacions excepte pell no Melanoma;0;3623
2;7;2007;Totes les localitzacions excepte pell no Melanoma;1;4365
2;8;2007;Totes les localitzacions excepte pell no Melanoma;1;3952
2;9;2007;Totes les localitzacions excepte pell no Melanoma;2;3702
2;10;2007;Totes les localitzacions excepte pell no Melanoma;3;3358
2;11;2007;Totes les localitzacions excepte pell no Melanoma;3;2937
2;12;2007;Totes les localitzacions excepte pell no Melanoma;4;2593
2;13;2007;Totes les localitzacions excepte pell no Melanoma;5;2347
2;14;2007;Totes les localitzacions excepte pell no Melanoma;5;1666
2;15;2007;Totes les localitzacions excepte pell no Melanoma;9;1974
2;16;2007;Totes les localitzacions excepte pell no Melanoma;13;1952
2;17;2007;Totes les localitzacions excepte pell no Melanoma;14;1519
2;18;2007;Totes les localitzacions excepte pell no Melanoma;17;1342
...
1;16;2007;Testicle;0;1533
1;17;2007;Testicle;0;1058
1;18;2007;Testicle;0;724
```

ANNEX 2.14. Càlcul de les SIMR per a la Mortalitat per càncer de Catalunya, prenent una confiança del 90% per als intervals de confiança.

SIMR Results

Standardized Incidence or Mortality Ratio by sex and group

Men

Group	CASES	SIMR	LCI	UCI
All sites excluding skin non-melanoma	47171	0.97	0.96	0.97
All sites including skin non-melanoma	47314	0.97	0.96	0.97
Bladder	2782	1.09	1.06	1.13
Brain Central Nerv Sys	1067	1.06	1.01	1.12
Colon	4252	1.03	1.01	1.06
Esophagus	1274	0.93	0.89	0.97
Hodgkin Lymphoma	120	0.74	0.63	0.86
Kidney	931	1.06	1.01	1.12
Larynx	1163	0.88	0.84	0.93
Leukemia	1468	1.08	1.03	1.13
Liver	2395	0.86	0.84	0.89
Lung	13090	0.98	0.96	0.99
Multiple myeloma and immunoproliferative neoplasms	679	0.85	0.80	0.91
Non Hodgkin Lymphoma	1141	1.14	1.08	1.20
Oral Cavity and Pharynx	1403	0.91	0.87	0.95
Pancreas	1849	1.03	0.99	1.07
Prostate	4094	0.84	0.82	0.86
Rectum	1492	1.06	1.02	1.11
Skin Melanoma	336	1.10	1.00	1.20
Stomach	2662	0.88	0.85	0.91
Testis	20	0.93	0.62	1.36
Tyroid	66	1.08	0.87	1.32

Women

Group	CASES	SIMR	LCI	UCI
All sites excluding skin non-melanoma	28517	0.95	0.94	0.96
All sites including skin non-melanoma	28618	0.95	0.95	0.96
Bladder	609	0.93	0.87	1.00
Brain Central Nerv Sys	850	1.11	1.04	1.17
Breast	5003	0.93	0.91	0.95
Cervix	1336	0.83	0.79	0.87
Colon	3400	0.94	0.92	0.97
Corpus Uteri	395	0.79	0.73	0.86
Endometrium	665	0.94	0.88	1.00
Esophagus	203	0.94	0.83	1.06
Hodgkin Lymphoma	90	0.84	0.70	1.00
Kidney	471	0.94	0.87	1.02
Larynx	53	1.22	0.96	1.53
Leukemia	1166	1.02	0.97	1.07
Liver	1275	0.86	0.82	0.90
Lung	1778	1.14	1.10	1.18
Multiple myeloma and immunoproliferative neoplasms	696	0.85	0.80	0.90
Non Hodgkin Lymphoma	1051	1.13	1.07	1.18
Oral Cavity and Pharynx	319	1.00	0.91	1.10
Ovary	1502	1.05	1.00	1.09
Pancreas	1717	1.17	1.12	1.21
Rectum	968	0.94	0.89	0.99
Skin Melanoma	283	1.11	1.00	1.22
Stomach	1652	0.83	0.80	0.86
Tyroid	123	0.84	0.72	0.98

SIMR:Standardized Incidence or Mortality Ratio, LCI:Lower Confidence Interval, UCI: Upper Confidence Interval

Aquest projecte s'ha desenvolupat en el marc del Grup d'Investigació Emmergent. Registre de Càncer de Catalunya. SGR 2009-2013 de la AGAUR (Agència de Gestió d'Ajuts Universitaris i de Recerca)