

Máster Interuniversitario en Estadística e Investigación Operativa

Título: Modelos mixtos de riesgo a tiempo discreto para la estimación del riesgo de falsos positivos en el cribado mamográfico

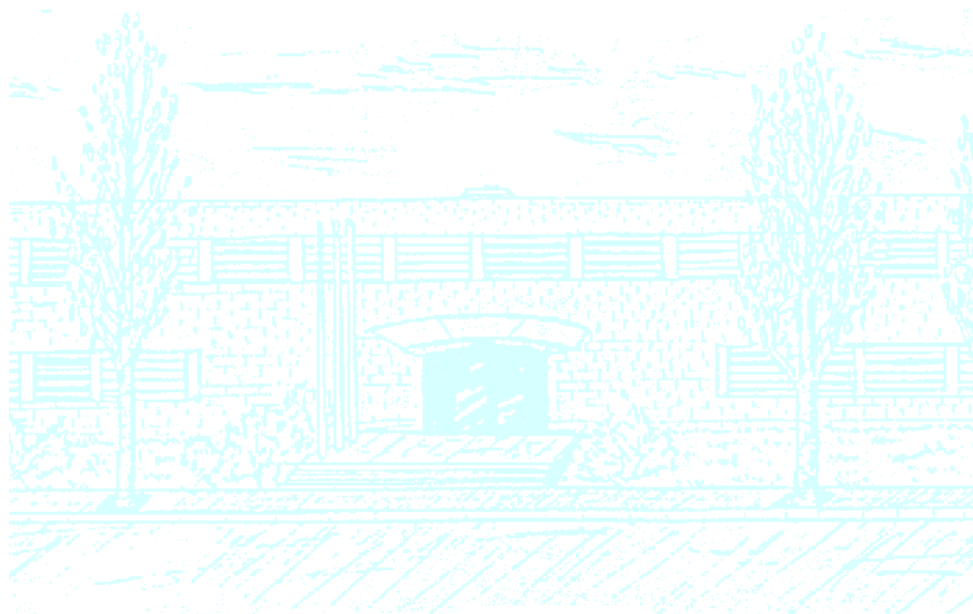
Autor: Rubén Román Expósito

Director: Xavier Castells Oliveres/ Klaus Langohr

Departamento: Estadística e Investigación Operativa

Universidad: Universitat Politècnica de Catalunya

Convocatoria: Febrero 2011



Facultat de Matemàtiques
i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA



UNIVERSITAT DE BARCELONA



TRABAJO DE FIN DE MÁSTER

Facultat de Matemàtiques i Estadística
Universitat Politècnica de Catalunya

Proyecto final del Máster Interuniversitario en
Estadística e Investigación Operativa

**Modelos mixtos de riesgo a tiempo
discreto para la estimación del
riesgo de falsos positivos en el
cribado mamográfico**

Autor

Rubén Román Expósito

Directores

Xavier Castells Oliveres

Klaus Langohr

Departamento de Estadística e Investigación Operativa

Este trabajo está dedicado a mis padres. Por su apoyo incondicional y su cariño inagotable.

Quiero mostrar mi más sincero agradecimiento por su paciencia y dedicación a Xavier Castells, María Sala, Laia Domingo, Esther Martínez, Mercè Comas, Andrea Burón, Jordi Blanch, Anabel Romero, Lola Salas, Raquel Zubizarreta, Nieves Ascunce y Pepa Ibáñez. Y cómo no, a Klaus Langohr, por ser un excelente profesor durante este máster.

RESUMEN

Palabras clave: Cáncer de mama, falso positivo, procedimientos invasivos, factores de riesgo, cribado, variabilidad.

MSC2010:

62J12 Generalized linear models
62N02 Survival analysis and censored data Estimation
62P10 Applications to biology and medical sciences

Antecedentes: El cribado de cáncer de mama reduce la mortalidad por esta enfermedad. Es importante reducir los efectos adversos (falsos positivos) del cribado mamográfico para mejorar la efectividad de esta práctica. El objetivo fue estimar el riesgo de falsos positivos a lo largo de la participación secuencial en sucesivas convocatorias de cribado mamográfico, y se evaluó el impacto de diferentes factores del protocolo de lectura mamográfica y de las características de la mujer sobre el riesgo estimado.

Métodos: Estudio de cohorte retrospectivo con información de 10 programas de cribado del territorio español desde su puesta en marcha en 1990 hasta 2006. Se dispone de información de 1.565.364 mujeres de 45 a 69 años de edad que se realizaron 4.739.498 mamografías de cribado. Se construyeron diferentes modelos mixtos de riesgo con tiempo discreto para estimar el riesgo de falsos positivos asociado a cada participación de la mujer en el cribado. Se estudió de manera independiente el riesgo de falso positivo por cualquier tipo de procedimiento adicional, y el riesgo de falso positivo por procedimientos invasivos. Se estimó el riesgo acumulado a lo largo de 10 mamografías de cribado secuenciales, bajo diferentes escenarios de riesgo, para aquellas mujeres que empezaron el cribado con 50-51 años (máximo número de invitaciones bienales para una mujer de esta edad).

Resultados: Los factores asociados con el riesgo de falso positivo, por cualquier tipo de procedimiento adicional y por procedimientos invasivos fueron la doble lectura de la mamografía (OR=2.06; IC95%: 2.00-2.13 y OR=4.44; IC95%: 4.08-4.84, respectivamente), realización de doble proyección mamográfica (OR=0.77; IC95%: 0.76-0.79 y OR=1.56; IC95%: 1.48-1.64), utilización de mamógrafo digital (OR=0.83; IC95%: 0.72-0.96 para pruebas invasivas), estado premenopáusico (OR=1.31; IC95%: 1.29-1.33; y OR=1.22; IC95%: 1.16-1.29), uso de tratamiento hormonal sustitutivo

(OR=1.03; IC95%: 1.01-1.05; y OR=0.84; IC95%: 0.78-0.90), pruebas invasivas previas (OR=1.52; IC95%: 1.49-1.56 y OR=2.00; IC95%: 1.89-2.12), e historia familiar de cáncer de mama (OR=1.18; IC95%: 1.15-1.20; y OR=1.21; IC95%: 1.13-1.30). El riesgo acumulado de falso positivo para las mujeres que empezaron el cribado con 50-51 años fue del 20.39% (IC95%: 20.02-20.76). Este valor varió entre el 51.43% y el 7.47% para los perfiles de alto y bajo riesgo, respectivamente. El riesgo acumulado para pruebas invasivas fue del 1.76% (IC95%: 1.66-1.87) y varió entre el 12.02% y el 1.58% para los perfiles de alto y bajo riesgo, respectivamente.

Conclusiones: El riesgo acumulado de falso positivo a lo largo de 10 participaciones secuenciales en el cribado varió ampliamente, en función de los factores del protocolo de lectura mamográfica y de las características de la mujer estudiados. Estos resultados son relevantes para mejorar la efectividad de los programas de cribado, ya que aportan información substancial para la evaluación del balance entre riesgos y beneficios de esta práctica. Además permite proporcionar a las mujeres participantes información más precisa sobre los efectos adversos del cribado.

ABSTRACT

Key words: Breast cancer, false positive, invasive procedures, risk factors, screening, variability.

MSC2010:

62J12 Generalized linear models
62N02 Survival analysis and censored data Estimation
62P10 Applications to biology and medical sciences

Background: Breast cancer screening reduces the mortality associated to this disease. It is important to reduce the adverse effects (false-positives) of breast cancer screening to improve its effectiveness. Our aim was to estimate the false-positive risk in a series of sequential participations in breast cancer screening. We examined how the screening protocol and women's characteristics affect the estimated false-positive risk.

Methods: Retrospective cohort study with information of 10 Spanish breast cancer screening programs from 1990 to 2006. We had available information on 1,565,364 women aged 45-69 years who underwent 4,739,498 screening mammograms. Multilevel discrete hazard models were used to estimate the false-positive risk in any of the women's screening participation. We studied independently the false-positive risk for any additional procedure, and the false-positive risk for invasive additional procedures. We estimated the false-positive risk up to the 10th sequential screening participation, under different risk scenarios, for those women starting screening at 50-51 years of age (the maximum number of screening invitations for women aged 50-51 years).

Results: The factors affecting the false-positive risk for any procedure and for invasive procedures were double mammogram reading (OR=2.06; 95%CI: 2.00-2.13 y OR=4.44; 95%CI: 4.08-4.84, respectively), two mammographic views (OR=0.77; 95%CI: 0.76-0.79 y OR=1.56; 95%CI: 1.48-1.64), digital mammography (OR=0.83; 95%CI: 0.72-0.96 for invasive procedures), premenopausal status (OR=1.31; 95%CI: 1.29-1.33; y OR=1.22; 95%CI: 1.16-1.29), use of hormone replacement therapy (OR=1.03; 95%CI: 1.01-1.05; y OR=0.84; 95%CI: 0.78-0.90), previous invasive procedures (OR=1.52; 95%CI: 1.49-1.56 y OR=2.00; 95%CI: 1.89-2.12), and a familial history of breast cancer (OR=1.18; 95%CI: 1.15-1.20; y OR=1.21; 95%CI: 1.13-1.30).

The cumulative false-positive risk for women who started screening at age 50-51 was 20.39% (95%CI: 20.02-20.76), ranging from 51.43% to 7.47% in the highest and lowest risk profiles, respectively. The cumulative risk for invasive procedures was 1.76% (95%CI: 1.66-1.87) ranging from 12.02% to 1.58%, in the highest and lowest risk profiles.

Conclusions: The cumulative false-positive risk over 10 sequential screening participations varied widely, depending on the screening protocol variables, and the women's characteristics studied. These findings are relevant to improve the effectiveness of breast cancer screening programs, as they provide substantial information to evaluate the balance between risk and benefits of breast cancer screening. Besides, it allows the screening programs to provide accurate information on the adverse effects of screening to participating women.

ÍNDICE

1.	INTRODUCCIÓN	13
2.	HIPÓTESIS DEL ESTUDIO	15
3.	OBJETIVOS	16
4.	MATERIAL.....	17
4.1.	Diseño y ámbito del estudio	17
4.2.	Población de estudio	17
4.3.	Definición de falso positivo y exploraciones adicionales invasivas.....	18
5	APROXIMACIÓN A LA PROBLEMÁTICA	20
5.1	Carácter de la Información	20
5.2	Implementación del conocimiento a priori.....	21
5.3	Estructura relacional.....	21
5.4	Estructura multinivel	24
5.5	Tamaño de la muestra	24
6	MÉTODOS	26
6.1	Consistencia y validación lógica.....	26
6.2	Variables de estudio.....	26
6.2.1	Evento de interés.....	26
6.2.2	Variables explicativas	27
6.3	Variabilidad entre Unidades Radiológicas	30
7	ANÁLISIS ESTADÍSTICO	32
7.1	Características del modelo.....	32
7.2	Regresión logística.....	34
7.2.1	Modelos de respuesta binaria	34
7.2.2	Deviance y Criterio de Información de Akaike	36
7.3	Modelos de Riesgo a Tiempo discreto.....	38
7.3.1	Definición de riesgo a tiempo discreto.....	38
7.3.2	Formulación del modelo de riesgo a tiempo discreto.....	39
7.3.3	Estructura de los datos	41
7.3.4	Estimadores de máxima verosimilitud para el DTHM	42
7.3.5	Modelo de riesgo a tiempo discreto con efectos aleatorios	44
7.3.6	Selección de modelos.....	46
7.3.7	Valoración del efecto aleatorio	48
7.4	Estimación del riesgo acumulado	50
8	RESULTADOS.....	52
8.1	Análisis descriptivo.....	52
8.2	Resultados del estudio	55
9	DISCUSIÓN	61
9.1	Implicaciones del estudio	61
9.2	Aproximación estadística	62

9.3	Valoración de los factores de estudio	65
9.4	Limitaciones epidemiológicas.....	66
9.5	Conclusión.....	67
10	BIBLIOGRAFÍA	68
11	ANEXOS.....	71
	Protocolo de las variables de estudio.....	71
	Programa de control de calidad de la base de datos	92
	Valoración de datos faltantes	105

1. INTRODUCCIÓN

El cáncer de mama es el más frecuente entre las mujeres de todo el mundo, (22,7% del total femenino) y el tumor que mayor número de muertes produce en la mujer en nuestro país (1). El cribado de cáncer de mama reduce la mortalidad por esta enfermedad. En el proceso de cribado se realiza una mamografía cada dos años a mujeres asintomáticas, con el objetivo de detectar posibles tumores en fase preclínica, es decir, que aun no se han manifestado como sintomáticos y por lo tanto el individuo desconoce su existencia. Aquellas en las que se confirma la enfermedad reciben tratamiento.

Uno de los efectos adversos más importantes de la mamografía de cribado son los falsos positivos. Un falso positivo se define como la recomendación de realizar exploraciones adicionales (ecografía, resonancia magnética, punción con aguja fina, biopsia excisional, biopsia quirúrgica, etc.) para detectar malignidad a individuos en los que finalmente se descarta la presencia de un cáncer de mama. Los resultados falsos positivos generan preocupación y ansiedad en las mujeres afectadas así como un mayor número de pruebas adicionales con un coste asociado. Sin embargo, los falsos positivos son una parte intrínseca e inevitable de los programas de cribado poblacional de cáncer de mama

Las guías para la práctica del cribado mamográfico (2;3) recomiendan una mamografía bienal a las mujeres entre 50 y 69 años de edad, lo que supondría que la mujer se realizase 10 mamografías de cribado durante este periodo. Los beneficios del cribado se miden en términos de la reducción de mortalidad a lo largo de la participación en sucesivas rondas de cribado. Para poder evaluar el balance beneficio-riesgo del cribado, es importante medir los efectos adversos de manera análoga, calculando el impacto de los mismos en forma de riesgo acumulado.

Actualmente, en España, los programas poblacionales de cribado de cáncer de mama cubren prácticamente la totalidad de la población diana que incluye 4,244,503 mujeres de 50 a 69 años. El primer programa en España se inició en el año 1990 en Navarra y alcanzó el 100% de cobertura en el año 2005. En el territorio español se ha implementado un programa de cribado por cada comunidad, excepto en Cataluña, que debido a las competencias locales dispone de diferentes programas de cribado, no asignados estrictamente a una división política del territorio. En la totalidad de los programas se han adoptado las recomendaciones de las Guías Europeas para el cribado de cáncer de mama (2). Sin embargo, hay diferencias relevantes entre

programas con relación al protocolo de puesta en práctica de la mamografía de cribado. Por ejemplo, la lectura de forma independiente por dos radiólogos se realiza en algunos programas, al igual que la doble proyección no es una práctica generalizada. Por otro lado, la edad de inicio del cribado varía entre los 45 y los 50 años, según el programa (4). Tales diferencias, al igual que otras de orden más organizativo, pueden comportar variaciones en los resultados de los distintos programas de cribado.

En este contexto, se inicia un proyecto para evaluar, de forma conjunta, el riesgo de falsos positivos en el cribado mamográfico y su asociación con el protocolo de lectura mamográfica y las características de la mujer en 10 programas poblacionales de cribado de cáncer de mama de 8 comunidades autónomas del territorio español. Este proyecto está financiado por el *Instituto de Salud Carlos III*, a través del *Fondo de Investigaciones Sanitarias* (PI061230) y lleva por título: "Análisis de la probabilidad acumulada de un falso positivo en el cribado mamográfico en España en la indicación de alguna exploración adicional".

A continuación se exponen las hipótesis de estudio y los objetivos del mismo. Posteriormente se hace se hace hincapié en la fase de pre-análisis, ya que como se comentará posteriormente tuvo una gran relevancia para poder desarrollar este trabajo adecuadamente. A continuación se desarrolla el método estadístico utilizado y posteriormente se presentan los resultados y la discusión.

2. HIPÓTESIS DEL ESTUDIO

Las hipótesis de partida del equipo investigador a la hora de plantear el estudio fueron las siguientes:

1. La probabilidad de falso positivo del cribado mamográfico en la indicación de alguna exploración adicional en la primera mamografía (cribado prevalente) es superior al de las mamografías sucesivas (cribado de incidencia).
2. La probabilidad de falso positivo en la indicación de alguna exploración adicional invasiva en el cribado mamográfico es similar en las mamografías sucesivas (cribado de incidencia).
3. La probabilidad acumulada de falso positivo está asociada con características del protocolo de lectura de la mamografía de cribado:
 - 3.1. Realización o no de la doble lectura: Número de radiólogos lectores para cada mamografía.
 - 3.2. Doble o única proyección: Número de imágenes realizadas en cada mama.
 - 3.3. Tipo de mamografía: Analógica o digital
4. La probabilidad acumulada de falso positivo en el cribado mamográfico está asociada a ciertas características de la mujer (edad, tomar tratamiento hormonal sustitutivo, no tener la menopausia, y tener antecedentes familiares y/o personales de patología mamaria), presentando un rango muy amplio en función de las mismas.

3. OBJETIVOS

El objetivo del estudio es estimar el riesgo, para una mujer participante en el cribado, de presentar un falso positivo en una mamografía de cribado a lo largo de sus 10 posibles participaciones. Se pretende determinar el impacto que ciertas características del protocolo de lectura mamográfica de los programas (lectura simple o doble, mamografía digital o analógica, número de proyecciones, etc.), y ciertas características de la mujer (edad, menopausia, antecedentes familiares o personales de patología mamaria, uso de tratamiento hormonal sustitutivo, etc) tienen sobre el riesgo de falsos positivos.

Objetivos específicos:

1. Estimar la probabilidad acumulada de presentar al menos un falso positivo, con exploraciones adicionales invasivas y/o no invasivas, en diez rondas de cribado con mamografía.
2. Estimar la probabilidad acumulada de presentar al menos un falso positivo, con exploraciones adicionales invasivas (punción y aspiración con aguja fina, core-biopsia y/o biopsia quirúrgica), en diez rondas de cribado con mamografía.
3. Definir un modelo de riesgo de falso positivo en relación al número de rondas de cribado disponibles, para poder predecir/proyectar la probabilidad acumulada a un número determinado de rondas.
4. Evaluar la asociación entre la probabilidad acumulada de falsos positivos y las siguientes características de los protocolos de lectura de la prueba de cribado (mamografía):
 - 4.1. El proceso radiológico asociado a la doble lectura
 - 4.2. El número de proyecciones realizadas
 - 4.3. Tipo de mamografía realizada
5. Evaluar la probabilidad acumulada de falsos positivos y las siguientes características de las mujeres: edad, tratamiento hormonal sustitutivo, menopausia, antecedentes personales y/o- antecedentes familiares.

4. MATERIAL

4.1. DISEÑO Y ÁMBITO DEL ESTUDIO

Estudio retrospectivo sobre una cohorte de mujeres participantes en programas de detección precoz de cáncer de mama en 8 Comunidades Autónomas de España. Los programas están basados en las guías europeas de calidad de la mamografía de cribado y sus resultados se ajustan a los estándares de las recomendaciones de 'Europa Contra el Cáncer' (3). El sistema de salud público cubre los gastos de los programas incluyendo las exploraciones adicionales.

Se invita cada dos años a las mujeres con una edad comprendida entre los 50 y 69 años a realizarse una mamografía, de manera que, una mujer que entra en un programa a la edad de 50 años puede llegar a tener hasta diez mamografías a lo largo de dos décadas.

4.2 POBLACIÓN DE ESTUDIO

La población de estudio está comprendida por las mujeres residentes en las 8 Comunidades Autónomas (CCAA) participantes en el estudio: Navarra, Valencia, Galicia, Canarias, Castilla y León, La Rioja, Asturias y Cataluña. Los programas participantes tienen información estructurada de al menos tres rondas de cribado consecutivas, de manera que la información puede ser analizada como una cohorte (cada mujer debe ser identificada unívocamente mediante un único código en cada participación en el programa). Además, para cada participación de la mujer los programas de cribado tienen información sobre el tipo de exploraciones adicionales (invasivas o no-invasivas) y del resultado histopatológico con el fin de verificar si el positivo es verdadero o falso. Solo se incluyen mujeres participantes al menos una vez en cualquiera de los programas del estudio. Se excluye a las mujeres con historia de cáncer de mama o implante mamario anterior al primer cribado.

Figura 4.1: Programas de cribado participantes en el proyecto RAFP (Riesgo acumulado de falsos positivos), en verde claro.



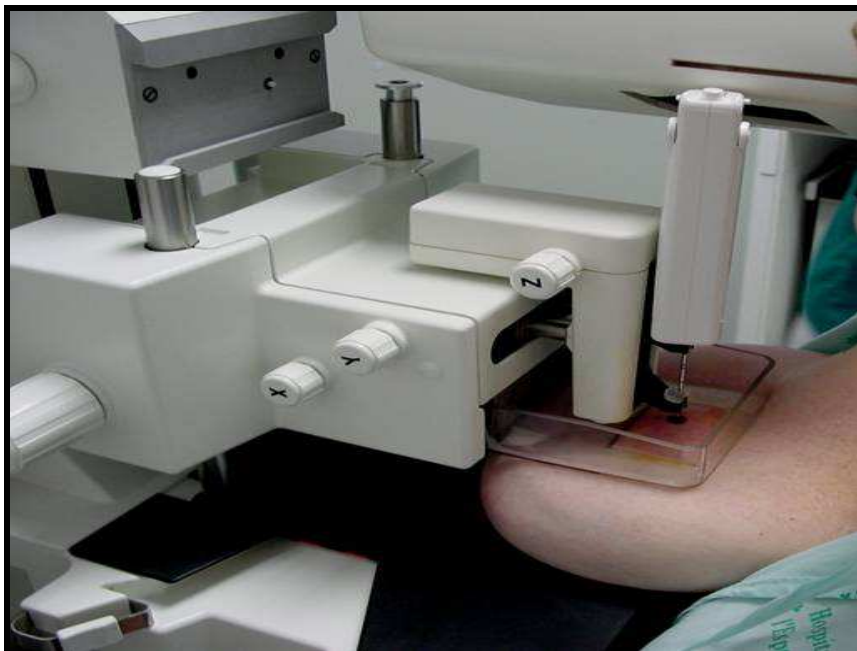
4.3 DEFINICIÓN DE FALSO POSITIVO Y EXPLORACIONES ADICIONALES INVASIVAS

Se consideraran dos posibles resultados de la interpretación mamográfica asociados a los programas: negativo (recomendación de una mamografía en la siguiente ronda, 2 años) y positivo, con realización de exploraciones adicionales en el momento justo después de la mamografía o por reconvocatoria, con el fin de descartar malignidad. Las exploraciones adicionales incluyen pruebas no-invasivas (mamografía adicional y ecografía) e invasivas (punción con aguja fina [PAAF], core-biopsia, biopsia con aguja gruesa y biopsia quirúrgica). La reconvocatoria debida a una mala calidad técnica de la mamografía no se considera un resultado positivo. Se considera un 'verdadero positivo' si, después de las pruebas adicionales se diagnostica cáncer de mama (ductal in situ o invasivo). En caso contrario, el resultado será considerado 'falso positivo'.

Figura 4.2: *Realización de una mamografía de cribado para la detección precoz del cáncer de mama.*



Figura 4.3: *Realización de la prueba invasiva 'biopsia con aguja gruesa' (BAG)*



5 APROXIMACIÓN A LA PROBLEMÁTICA

Ante el planteamiento inicial propuesto surgen diferentes problemas específicos en el contexto de la estadística previos al análisis de los datos.

1) Por un lado, la problemática común a la mayoría de proyectos de investigación que es, el convertir información de carácter administrativo y proveniente de diferentes fuentes de información en información útil, que pueda ser manejable y sirva para ampliar el conocimiento del experto sobre esta materia.

2) Por otra parte, se plantea un problema específico, que es el eje central de este trabajo de fin de máster, y que es identificar y poner en práctica el tipo de metodología que sirva para dar respuesta precisa y fiable a las preguntas planteadas por los investigadores. Se desarrolla brevemente a continuación los principales retos que se han identificado.

A pesar de que nos centraremos en la segunda parte de esta problemática que es la referente a aspectos más metodológicos, repasaremos antes de manera breve como se han resuelto las problemáticas previas al análisis ya que están estrechamente ligadas.

5.1 CARÁCTER DE LA INFORMACIÓN

Participan en el estudio 10 programas poblacionales de cáncer de mama de 8 Comunidades Autónomas. El periodo de estudio del proyecto abarca desde 1990, con la puesta en marcha del primer programa de cribado del territorio español, hasta diciembre de 2006. Se solicitó a cada uno de los programas participantes que aportasen la información completa de todas las mujeres participantes en el cribado al menos una vez desde el inicio de los mismos. Los programas de cribado tienen estructuras administrativas independientes y específicas de cada programa para recoger la información, por lo que se elaboró un detallado protocolo de equivalencias y validación de la información de los mismos. Este protocolo de definición se centró en el tipo de pregunta al que se quería dar respuesta, de manera que la interpretación de las variables sea biunívoca en todos los centros. El total de la información del proyecto cubre el 44% de la población diana española en diciembre de 2005.

En los anexos de este trabajo se encuentran el “**protocolo de definición de variables**” para la recogida de la información.

5.2 IMPLEMENTACIÓN DEL CONOCIMIENTO A PRIORI

En un estudio de estas características en el que diferentes centros son proveedores de información (multicéntrico) y a la vez son centros autónomos e independientes con sistemas de gestión diferentes, se hace especialmente importante el proceso de implementación de un concepto común a partir de la implementación del conocimiento a priori. Es de prever que fuentes de información diferentes aporten definiciones sustancialmente diferentes sobre un mismo concepto. En este sentido se hizo un especial hincapié en la estructura de la base de datos de manera que, sin perder información, se pueda respetar la compleja estructura de información de la manera más sencilla posible.

Se establecieron los niveles de elementales de los que se quería tener información:

- Unidad radiológica en que se realiza la mamografía y programa
- Información identificativa de la mujer,
- Información de cada participación en el cribado (episodio),
- Información de exploraciones adicionales en las mujeres reconvocadas,
- Variables de la mujer cambiantes en el tiempo

En los anexos de este trabajo se encuentran el “**protocolo de definición de variables**” para la recogida de la información.

5.3 ESTRUCTURA RELACIONAL

Las mujeres participantes en el cribado son citadas a participar en el cribado mamográfico cada 2 años mientras están en la edad de la población diana (50-69 años). La información relativa a la mujer como su edad de nacimiento, y fechas de primera mamografía en el programa se genera una única vez.

Cada mujer puede tener un número de participaciones en el cribado que va desde una única participación hasta un máximo de 10, por lo tanto cada mujer puede

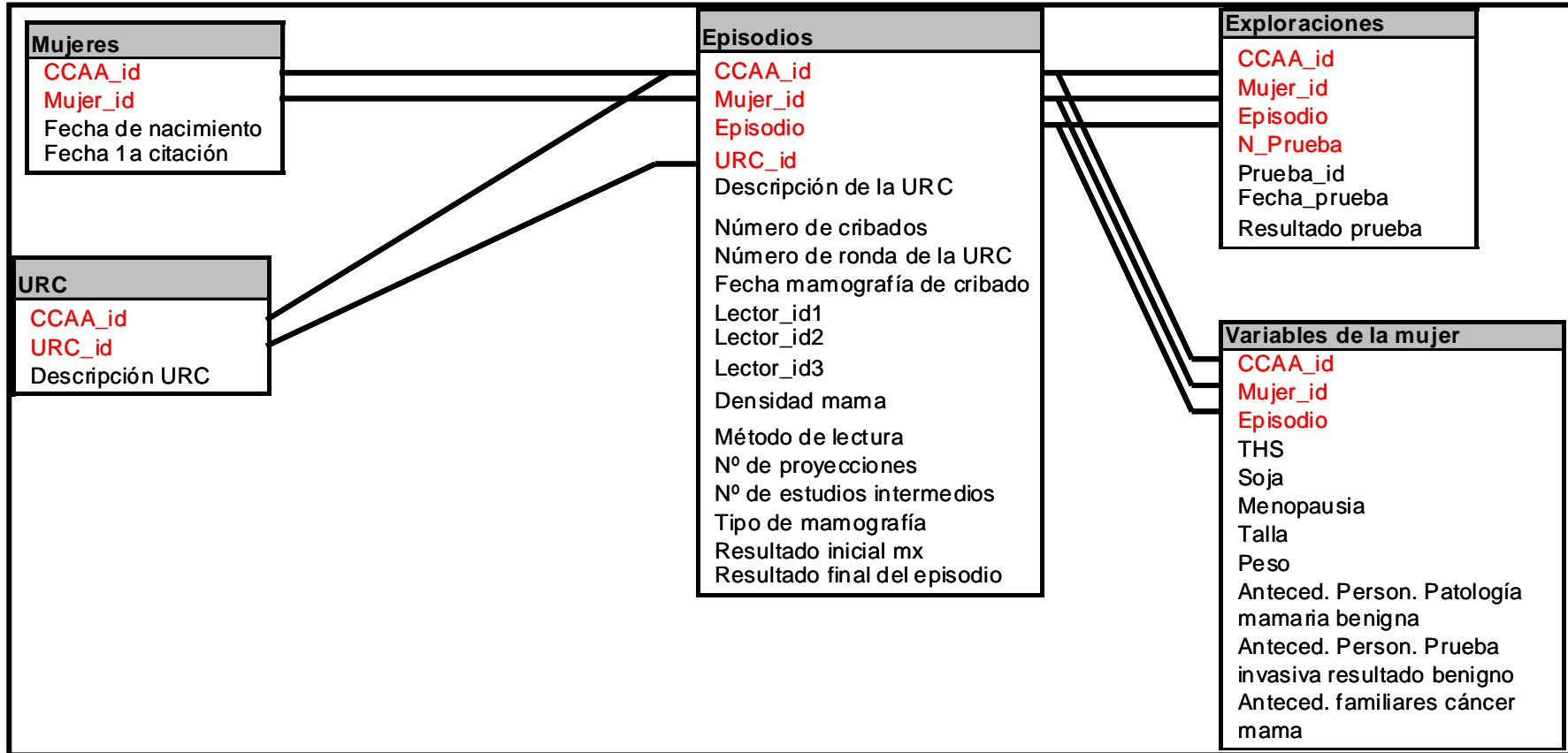
generar un número de registros diferentes de 1 a N , dependiendo de las participaciones que haya tenido.

Por otro lado, las mamografías de cribado con sospecha de malignidad son reconvocadas para realizarse pruebas adicionales para confirmar o descartar malignidad. Sólo algunas mamografías de cribado son reconvocadas para realizar exploraciones adicionales. Además, en estas reconvocatorias puede que se realice más de una prueba a la mujer reconvocada (ecografía y punción con aguja fina, por ejemplo), por lo que una mamografía de cribado puede generar más de un registro de pruebas adicionales realizadas, pero no todas las mamografías ni todas las mujeres son reconvocadas.

Además, existe información relativa a la mujer, como tener antecedentes personales de patologías mamarias, antecedentes familiares de cáncer, o ser menopáusica, que pueden cambiar a lo largo de las diferentes participaciones de la mujer, por lo que las variables de características personales de la mujer pueden ser cambiantes en el tiempo. Estas características se informan en cada una de las participaciones de la mujer, generando un nuevo registro cada vez que la mujer participa en el cribado.

Por otro lado, todas las mamografías realizadas en un mismo programa de cribado y en una misma unidad radiológica están identificadas, por lo que se recoge también la información relativa al programa y unidad radiológica donde se realiza la mamografía. Esta información es muy relevante de cara a proponer un análisis estadístico para los datos ya que es de esperar que mamografías de cribado realizadas en una misma unidad radiológica de cribado estén correlacionadas y por tanto las fuentes de variabilidad sean diferentes. Se optó por recoger la información con una estructura relacional, mediante tablas de dimensiones diferentes, donde las tablas se relacionan entre si con variables de link (**figura 5.1**).

Figura 5.1: Estructura relacional y variables de link entre tablas de información



5.4 ESTRUCTURA MULTINIVEL

Al tratarse de un proyecto multicéntrico en el que participan diferentes programas de cribado, la información tiene una estructura jerárquica. Se identificaron tres niveles. El primer nivel corresponde a las *'mamografías de cribado'*, y contiene toda la información correspondiente a las diferentes participaciones en el cribado. Las mamografías de cribado serán nuestra unidad de análisis de cara a los modelos propuestos. A su vez, diferentes mamografías de cribado están vinculadas a una misma mujer, de manera que tendremos diferentes observaciones de cada mujer. Esta estructura es característica de los estudios longitudinales y se suele denominar "medidas repetidas". Por otro lado, cada mujer está asociada a la *'unidad radiológica'* en el que se realiza las mamografías. Un total de 74 unidades radiológicas participan en el estudio. Cada una de las unidades radiológicas es un proveedor de información autónomo que ha de seguir una estructura lógica de relación en sus datos. Las unidades radiológicas de un mismo programa de cribado comparten protocolos de gestión similares.

Este planteamiento de la estructura de información corresponde con el de los **modelos de regresión multinivel** o de efectos aleatorios, en los que la propia estructura jerárquica de los datos ha de ser tenida en cuenta a la hora de realizar los modelos de regresión.

5.5 TAMAÑO DE LA MUESTRA

Los programas de cribado poblaciones, invitan activamente a realizarse una mamografía de cribado a todas las mujeres de su población diana (50/69 años). Actualmente, en España, más del 99% de las mujeres en edad de cribarse, son invitadas a participar. Como ya se ha comentado, este estudio cubre el 44% de la población diana española y contiene información de 8 comunidades autónomas. El carácter retrospectivo del estudio implica un gran volumen de mamografías de cribado realizadas, desde la puesta en marcha de los programas hasta la actualidad. La base de datos contiene información de 4.739.498 mamografías de cribado de 1.565.364 mujeres diferentes. Además, se informaron 378.060 exploraciones adicionales para confirmar o descartar malignidad. El número de registros que contiene cada una de las tablas en que se dividió la información esta recogido en la tabla 5.1.

Tabla 5.1: *Nº de registros y variables en cada tabla*

Tabla	Registros	Nº variables
URC	74	3
Mujeres	1,565,364	4
Episodios	4,739,498	18
Exploraciones	378,060	7
Variables	4,739,498	11

* URC: Unidad Radiológica de Cribado

Esta es la mayor base de datos sobre efectos adversos del cribado conocida hasta el momento. Ante este volumen de información, la elección de la plataforma para gestionar toda esta información no es trivial, ya que muchos software estadísticos no son capaces de manejar dimensiones de datos tan grandes. El traspaso de información y la puesta en común se hizo a través de texto plano delimitado por tabuladores, que permitió un proceso relativamente ágil de traspaso de información. Finalmente, después de probar diferentes opciones, se decidió trabajar con SPSS 16.0 para la lectura de los datos, fusión de tablas, recodificación, transposición, reestructuración, y descriptivo de la base de datos, previos al análisis. Este software, a pesar de sus limitaciones a nivel de modelización, demostró ser altamente eficiente en el manejo de las tablas, reduciendo considerablemente el tiempo de proceso para estas tareas respecto al software SAS. Para el análisis de los datos se utilizó el software SAS 9.1. que demostró ser el más eficiente en la computación de los diferentes modelos. El software 'R', a pesar de ser el principal candidato por mi preferencia estadística, fue descartado porque es incapaz de cargar un fichero de estas dimensiones. La base de datos de análisis pesa aproximadamente 5 Gb, un modelo de regresión logística con 3 variables explicativas estimado por máxima verosimilitud tarda aproximadamente 30 minutos en converger. Un modelo con efectos aleatorios, y 10 variables explicativas, estimado por pseudo-verosimilitud, tarda en converger 4h aproximadamente. Si tenemos en cuenta la cantidad de modelos que son necesarios hacer antes de elegir un modelo útil (univariados, interacciones, reparametrización, cambios de categoría de referencia), la inversión de tiempo realizada es considerablemente elevada.

6 MÉTODOS

6.1 CONSISTENCIA Y VALIDACIÓN LÓGICA

Una parte fundamental previa al desarrollo de la metodología a plantear es la validación de los datos. Se contrastó la información registrada para comprobar que realmente era representativa de las fuentes de información originales. El objetivo de la validación lógica fue detectar inconsistencias o errores sistemáticos que pudiesen disminuir la validez de los datos o de algunas variables específicas. En este caso, el proceso de validación se llevó a cabo mediante una serie de reglas que establecían la relación lógica entre las variables, de manera que aunque los valores que toman ciertas variables puedan resultar válidos podrían no serlo desde un punto de vista de relación lógica.

Se establecieron tres niveles de validación:

- FORMATO Y VALORES VÁLIDOS (errores de 1^{er} orden)
- RELACIONES LÓGICAS (errores de 2^o orden)
- VALORES POCO PROBABLES (warnings)

En los anexos de este trabajo se encuentran las “**reglas de validación y relaciones lógicas**” para la recogida de la información.

6.2 VARIABLES DE ESTUDIO

6.2.1 EVENTO DE INTERÉS

Se consideró el evento de interés como la presencia de un falso positivo. Se codificó este evento mediante una variable dicotómica que tomaba los valores:

- 1 si en dicha mamografía de cribado la mujer había tenido un resultado falso-positivo
- 0 si el resultado no fue falso-positivo, independientemente de si el resultado final fue negativo (convocatoria la cabo de 2 años) o hubo un diagnóstico de cáncer.

Se utilizaron dos definiciones independientes para la variable respuesta de falso positivo.

‘Falso positivo por cualquier prueba’: se realizó en dicho episodio alguna prueba adicional para descartar malignidad, independientemente de si fueron de carácter invasivo o no invasivo.

‘Falso positivo por pruebas invasivas’: se realizó en dicho episodio alguna prueba adicional de carácter invasivo para descartar malignidad. Forzosamente, todos los “falsos positivos por pruebas invasivas” están contenidos también en la variable “falso positivo por cualquier prueba” ya que ésta última es menos restrictiva e incluye a la anterior.

La unidad de análisis es la mamografía y no la mujer, de tal manera que una misma mujer tendrá un valor para cada una de estas variables en cada participación. Se realizaron modelos de regresión independientes para cada una de las dos variables respuesta definidas.

Además, se utilizó una tercera variable respuesta de **‘detección de cáncer’** que tomaba los valores

- 1 si en dicha mamografía de cribado la mujer tuvo un diagnóstico final de cáncer de mama
- 0 si se descartó la presencia de cáncer, independientemente de si el resultado final fue negativo (convocatoria la cabo de 2 años) o hubo un falso positivo en el proceso.

Los modelos asociados a esta variable respuesta no eran el objetivo principal de este estudio. Estos modelos se hicieron con fines comparativos, para poder evaluar los beneficios (mejora en la tasa de detección) y riesgos (falsos positivos) del cribado mamográfico en la discusión de los resultados. En cualquier caso, este trabajo se centrará en la valoración de los falsos-positivos, a pesar de que en algunas tablas se presentan también datos sobre la detección de cáncer con fines comparativos.

6.2.2 VARIABLES EXPLICATIVAS

▪ Asociadas al protocolo de lectura mamográfica:

‘Método de lectura’: indica si la mamografía de cribado fue leída por un único radiólogo o por varios. Esta variable toma los siguientes valores:

0 ‘lectura simple’: La mamografía es leída por un único radiólogo

1 ‘lectura doble’: La mamografía es leída por dos radiólogos independientes

‘Número de proyecciones’: indica el número de imágenes que se realizan de cada mama. Esta variable toma los siguientes valores:

0 ‘proyección única’: Una única imagen (craniocaudal) en cada mama

1 'doble proyección': Dos imágenes (craniocaudal y mediolateral oblicua) para cada mama

'Tipo de mamografía': Indica el sistema utilizado para realizar la imagen y su lectura. Toma los siguientes valores:

0 'analógica': La imagen es tomada en una placa analógica (*film-screen*) y leída sobre el positivado de la misma.

1 'digital': La imagen es tomada mediante mamógrafo digital y es leída por pantalla, con ayuda de herramientas digitales de medición y diagnóstico.

▪ **Asociadas a las características de la mujer:**

'Edad': Indica la edad de la mujer en el momento de realizarse la mamografía. Esta variable se categorizó en grupos quinquenales con los siguientes valores:

4: edad de la mujer ≥ 44 años y a su vez < 50 años

3. edad de la mujer ≥ 50 años y < 55 años

2. edad de la mujer ≥ 55 años y < 60 años

1. edad de la mujer ≥ 60 años y < 65 años

0. edad de la mujer ≥ 65 años y < 70 años

Se decidió categorizar esta variable en intervalos quinquenales según los intereses de los investigadores. Es habitual en el contexto de los estudios epidemiológicos el tratar la edad como una variable categórica. En particular, el trabajar con intervalos quinquenales se justifica por la comparación con otros estudios internacionales sobre los efectos adversos del cribado que también trabajan con estos grupos de edad.

'Tratamiento hormonal sustitutivo' (THS): Indica si la mujer tomaba tratamiento hormonal sustitutivo en el momento del cribado o en los seis meses anteriores.

0: No

1: Sí

'Menopausia': Indica si la mujer es menopáusica o premenopáusica en el momento del cribado o por el contrario post-menopáusica

0: Menopáusica.

1: Premenopáusica

'Pruebas invasivas previas': Historial de pruebas invasivas previas con resultado benigno fuera del contexto del programa de cribado

0: No

1: Sí

'Antecedentes familiares': Antecedentes familiares de cáncer de mama en familiares de primer grado (padres, hermanos, hijos)

0: No

1: Sí

▪ **Otras variables de ajuste:**

'Periodo': Indica el tiempo en años, desde 1990 a 2006, en el que se realizó la mamografía de cribado. Esta variable se introdujo en los modelos multivariados como variable de ajuste ya que los efectos adversos del cribado se han reducido debido a la introducción de estándares mínimos de calidad, la mejora de las técnicas de lectura mamográfica y el propio aprendizaje de los programas de cribado. Se categorizó en cuatro intervalos:

3. Año de realización de la mamografía ≥ 2003 y < 2007

2. Año de realización de la mamografía ≥ 1999 y < 2003

1. Año de realización de la mamografía ≥ 1995 y < 1999

0. Año de realización de la mamografía ≥ 1990 y < 1995

Se categorizó esta variable en los intervalos mostrados según los intereses de los investigadores. El intervalo de las categorías se estableció arbitrariamente, respetando que fuesen intervalos de amplitudes similares. Al contar con una muestra lo suficiente mente grande no hubo problemas de representatividad dentro de los grupos seleccionados.

'URC' (Unidad Radiológica de Cribado): Indica la unidad radiológica en que se realizó la mamografía de cribado. Esta variable tiene 74 niveles y corresponde a la componente aleatoria que se introducirá en los modelos propuestos

6.3 VARIABILIDAD ENTRE UNIDADES RADIOLÓGICAS

Como ya se explicó en el planteamiento de la problemática inicial del estudio, en este proyecto participan diferentes programas de cribado del territorio español. Cada uno de estos programas tiene diferentes Unidades Radiológicas de Cribado (URC), en las que realizan las mamografías. Cada una de estas URC es independiente del resto. Cada URC se considera una fuente de variabilidad en la que sus observaciones están correlacionadas, ya que cada una de estas unidades utiliza un mamógrafo diferente y tiene un equipo de radiólogos lectores propio.

La representación de las tasas de falsos positivos muestra como éstas variaron en las distintas URC, tanto para falsos positivos por cualquier tipo de prueba (**fig. 6.1**) como para falsos positivos por pruebas invasivas (**fig. 6.2**).

Figura 6.1: Variaciones en las tasas de falsos positivos por cualquier prueba adicional según la 'unidad radiológica de cribado'

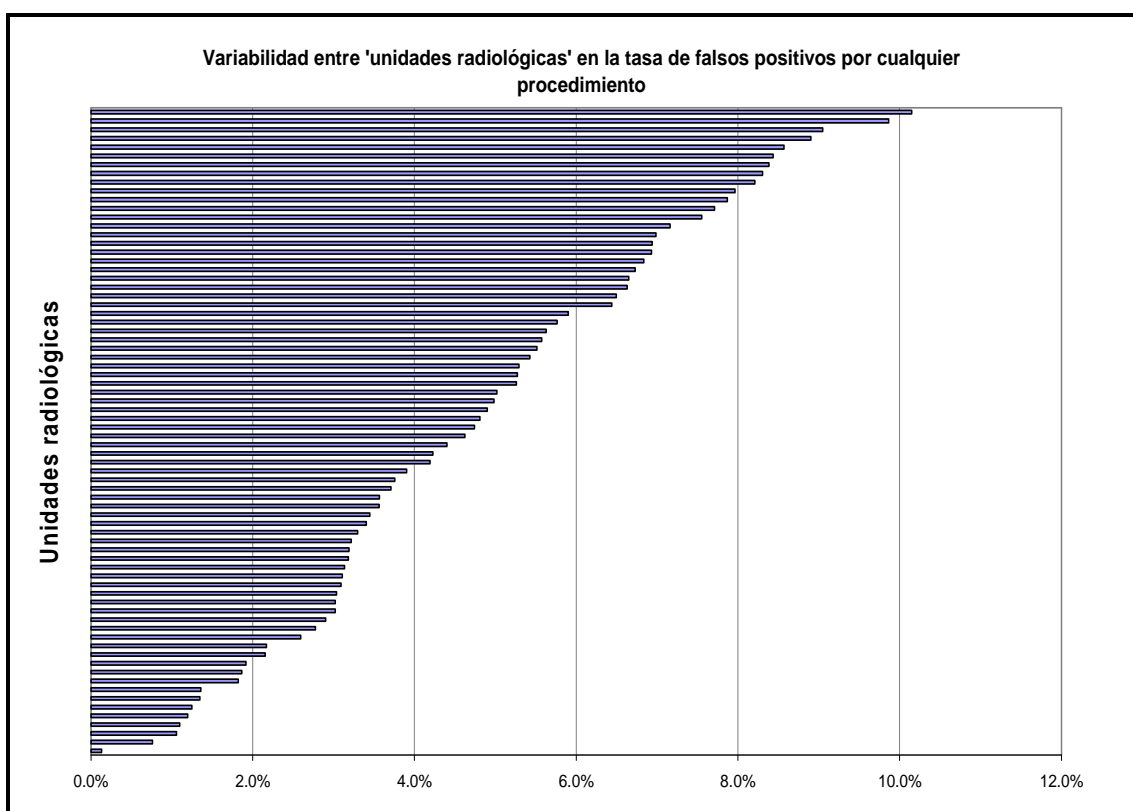
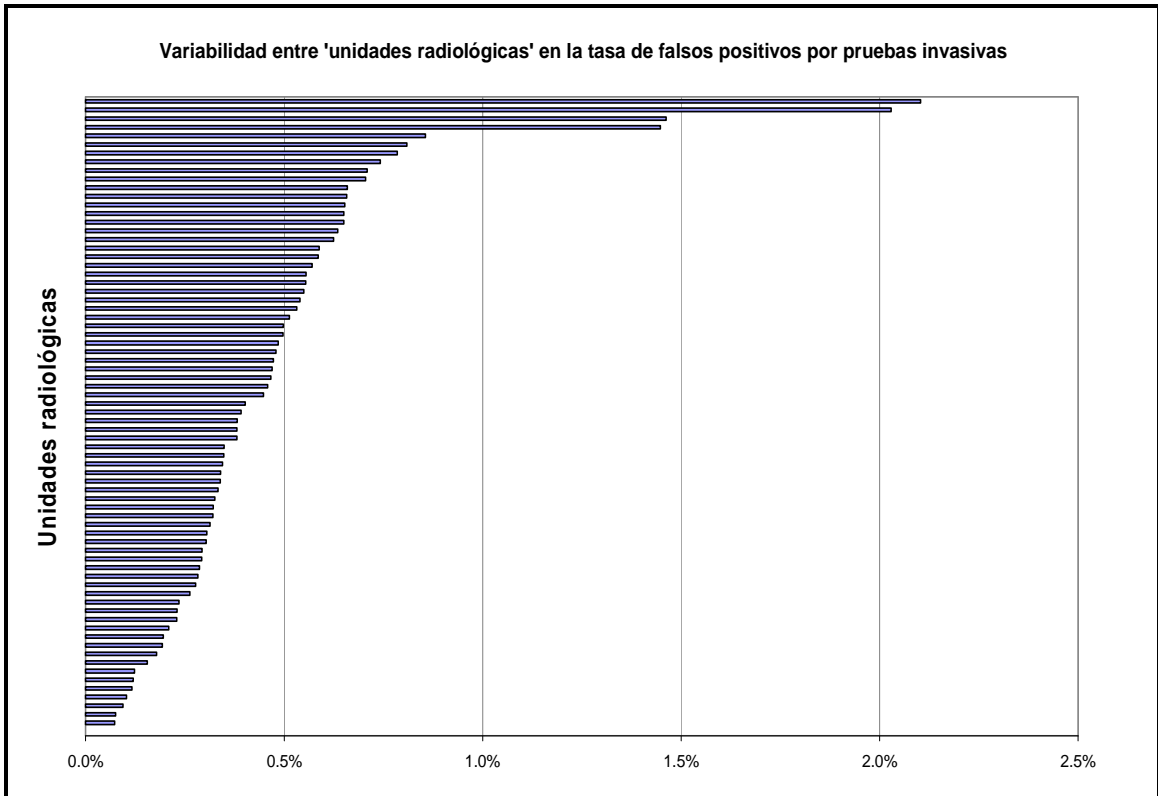


Figura 6.2: Variaciones en las tasas de falsos positivos por pruebas invasivas según la 'unidad radiológica de cribado'



7 ANÁLISIS ESTADÍSTICO

7.1 CARACTERÍSTICAS DEL MODELO

El reto del análisis estadístico se centra en desarrollar un modelo que nos permita evaluar el efecto que tienen las diferentes variables de interés sobre el riesgo de falso-positivo. El modelo propuesto deberá de tener en cuenta que:

- a) La **variable respuesta** (falso-positivo) es una variable **dicotómica**.
- b) **El tiempo es discreto**, y equivale a cada una de las participaciones en el cribado de la mujer. De esta manera, la mujer solo está a riesgo de sufrir un falso-positivo cuando participa en el cribado. Si no participa, o en el tiempo entre participaciones sucesivas, no es un individuo a riesgo. El disponer de una variable de tiempo discreta hace que otras técnicas habituales de regresión para el análisis de la supervivencia, como los *modelos de regresión de Cox*, no sean válidos en este contexto.
- c) Es necesario diferenciar el **riesgo en cada una de las posibles participaciones** de la mujer, es decir, no se considera que la probabilidad de observar el evento de interés (falso-positivo) sea similar a lo largo de las diferentes participaciones de la mujer. Es sabido que el riesgo disminuye cuantas más participaciones haya efectuado la mujer. Nos interesa por lo tanto modelar el riesgo para cada posible participación.
- d) Por la estructura jerárquica de los datos comentada anteriormente, sabemos que las mamografías de una misma unidad radiológica están correlacionadas y deberemos de tener en cuenta este diseño a la hora de realizar los análisis. Será necesario introducir en los modelos una **componente aleatoria** (*random effect*).
- e) Se deberá de tener en cuenta a la hora de formular los modelos las repetidas observaciones de una misma mujer. Las **observaciones** de un mismo individuo se espera que estén **correlacionadas** y deberemos por lo tanto tener en cuenta esta característica de alguna manera.
- f) El modelo deberá de poder considerar **variables cambiantes en el tiempo**. Es decir, modelar los diferentes valores de puedan tomar las variables para una misma mujer a lo largo de sus diferentes participaciones en el cribado (edad, método de lectura, menopausia).

- g) El modelo deberá de proporcionar alguna **medida del efecto** que tienen las variables de estudio sobre el riesgo de falso positivo, de manera que las estimaciones obtenidas puedan ser interpretadas en términos epidemiológicos.

Una aproximación mediante modelos de regresión a la problemática planteada es mediante los **Modelos de Riesgo a Tiempo Discreto** (*Discrete Time Hazard Models*). Este tipo de modelos no han sido estudiados en clase durante el desarrollo de este Máster, por lo que para su implementación y entendimiento se recurrió a bibliografía externa. En particular, estos modelos fueron desarrollados en profundidad por Judith D. Singer y John B. Willet en su libro “*Applied Longitudinal Data Analysis*” del año 2003 (5;6). Los autores modelizan la función de riesgo (*hazard*) de presentar un determinado evento a partir de modelos de supervivencia para intervalos de tiempo discretos, en los que los individuos únicamente están a riesgo cuando son observados. Por lo tanto, los individuos únicamente aportan información a la función de verosimilitud cuando son susceptibles de observar el evento de interés. En nuestro caso, estos intervalos de tiempo se corresponden con las participaciones de la mujer en el cribado. Como extensión a los modelos propuestos por Singer y Willet, en el desarrollo de este trabajo se incorpora una componente aleatoria que tiene en cuenta la estructura jerárquica del muestreo de los datos, en el que mamografías realizadas en una misma unidad radiológica están correlacionadas.

Cabe resaltar que es posible realizar un análisis de la supervivencia (probabilidad de no sufrir el evento de interés hasta un momento determinado) mediante el estimador de Kaplan y Meier para este tipo de datos, ya que este estimador trabaja con intervalos de tiempo discretos y por tanto produciría un estimador apropiado de la supervivencia. En cualquier caso, no permite el ajuste por otros factores de estudio, ni tiene la flexibilidad de modelización que ofrecen los modelos de regresión. Es por esto que nos centramos en el desarrollo de un modelo de regresión apropiado para nuestros datos y se obvia el estimador de Kaplan Meier.

Se desarrolla a continuación el modelo de regresión logística como base para el desarrollo del modelo de riesgo a tiempo discreto.

7.2 REGRESIÓN LOGÍSTICA

En el contexto de los modelos de regresión, los modelos de regresión logística permiten estimar la relación de una variable o conjunto de variables explicativas respecto a una variable de respuesta binaria. En nuestro caso, esta relación será el eje central de los modelos de regresión propuestos. Se explicarán a continuación diferentes características y propiedades de estos modelos de regresión hasta llegar a la justificación de los modelos de regresión elegidos para este trabajo.

Nota: Se obvian los modelos logísticos de respuesta politómica ya que no son de nuestro interés.

7.2.1 MODELOS DE RESPUESTA BINARIA

Definimos, en primer lugar, un conjunto de p variables independientes que denotaremos por el vector $\mathbf{x}' = (x_1, x_2, \dots, x_p)$. Las variables independientes podrán ser de cualquier tipo, sin necesidad de hacer ninguna asunción respecto a la distribución de las mismas ya que la relación de las mismas con la variable respuesta no es una función lineal.

Por otra parte, la variable dependiente de nuestro modelo Y , es dicotómica y tomará los valores 0 y 1, donde 0 representa la ausencia de la característica de interés, y 1 la presencia de la misma. A partir de estas definiciones podemos escribir nuestro modelo de regresión logística a través del nexo logit mediante la siguiente ecuación:

$$\text{logit}(\pi(x)) = \ln \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

Donde $\pi(x) = P(Y = 1 | x')$, o lo que es lo mismo, la probabilidad de que nuestra variable de interés sea igual a 1 condicionado al vector de covariables \mathbf{x}' . Además, podemos expresar esta probabilidad mediante la siguiente ecuación.

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}$$

En función de los valores estimados para los parámetros β_1, \dots, β_p para valores $\beta_i > 0$ tenemos que el factor X_i es un factor de riesgo para Y . Para valores $\beta_i < 0$ tenemos la relación opuesta, es decir, X_i es un factor protector para Y .

Variable explicativa categórica

En el caso de que alguna variable explicativa sea discreta será introducida en el modelo mediante variables artificiales *dummy*. Supongamos que la variable independiente X_p es discreta con k categorías, donde $K \geq 2$. Hemos de definir $k-1$ variables *dummy* de la siguiente forma:

$$X_{p1} = \begin{cases} 1 & X_p = 2 \\ 0 & \text{en caso contrario} \end{cases}; \dots; X_{pk-1} = \begin{cases} 1 & X_p = k \\ 0 & \text{en caso contrario} \end{cases}$$

Donde hemos tomado como categoría de referencia $X_p = 1$

Estimación de parámetros

En los modelos de regresión logística la estimación de parámetros se hace por el método de máxima verosimilitud. El concepto general de la estimación por máxima verosimilitud es el encontrar los valores para los parámetros de las variables de estudio que tienen una mayor probabilidad de reproducir los datos observados.

La función de verosimilitud en un modelo de regresión logística tiene la siguiente forma:

$$L(\beta, X) = \prod_{i=1}^n \left(\frac{\exp(\beta_0 + \beta' X_i)}{1 + \exp(\beta_0 + \beta' X_i)} \right)^{\delta_i} \left(\frac{1}{1 + \exp(\beta_0 + \beta' X_i)} \right)^{1-\delta_i}$$

Donde $\beta = (\beta_1, \dots, \beta_p)'$ es el vector de parámetros asociado a las p covariables del modelo, y δ_i es un indicador que toma el valor 1 en caso de ocurrencia del evento de interés, y 0 en caso contrario.

Interpretación parámetros

La interpretación de los parámetros del modelo de regresión logística puede hacerse en términos de Odds Ratios a través de su transformación logarítmica. Si suponemos que X_k es una variable introducida en el modelo de regresión logística, el odds ratio asociado con $X_k=1$, ajustado por el resto de covariables en el modelo, viene dado por:

$$OR_{X_k} = \frac{\text{odds}(P(D | X_1, \dots, X_k = 1, \dots, X_p))}{\text{odds}(P(D | X_1, \dots, X_k = 0, \dots, X_p))} = \exp(\beta_k)$$

De manera que el odds ratio estimado vendrá dado por:

$$OR_{X_k} = \exp(\hat{\beta}_k)$$

A partir de las estimaciones obtenidas se estudia la significación de las variables incluidas en el modelo. Por otro lado, también será de interés la bondad de ajuste del modelo. La bondad de ajuste y la comparación de modelos se evalúan generalmente mediante la *Deviance* (7).

7.2.2 DEVIANCE Y CRITERIO DE INFORMACIÓN DE AKAIKE (MODELOS LINEALES GENERALIZADOS)

En el caso de los modelos lineales generalizados (GLM) la selección de los modelos se realiza mediante la *Deviance*. Este estadístico cuantifica cuanto se aleja el modelo testado respecto al mejor modelo que pudiésemos estimar. En el caso del modelo de riesgo a tiempo discreto, el modelo saturado reproduciría cada valor observado del 'evento' en nuestro conjunto de datos. El único modelo que podría hacer esto debería de tener tanto parámetros como observaciones tengamos en nuestros datos.

La *Deviance* se define como:

$$D(y) = -2 \left\{ \ln(p(y | \hat{\theta}_0)) - \ln(p(y | \hat{\theta}_s)) \right\}$$

Donde $\hat{\theta}_0$ expresa los valores ajustados de los parámetros en el modelo M_0 , y $\hat{\theta}_s$ define los parámetros ajustados del modelo saturado. O lo que es lo mismo,

Deviance = $-2 \ln$ likelihood, del modelo M_0 que estamos evaluando. Cuanto mejor es el ajuste del modelo menor será el valor de la *Deviance*. Para comparar dos modelos será necesario 1) que ambos modelos estén hechos sobre el mismo conjunto de observaciones, y 2) que un modelo esté anidado dentro del otro. De esta manera comparamos el efecto de introducir una nueva variable (o grupo de variables) en el modelo respecto al mismo modelo sin la inclusión de esa variable (modelo anidado).

Se comparan las verosimilitudes de los dos modelos mediante la *Deviance* de la siguiente manera; $D = -2\ln(\text{verosimilitud modelo 1} / \text{verosimilitud modelo 2})$. Bajo la hipótesis nula H_0 , la *Deviance* sigue una distribución Chi-cuadrado con tantos grados de libertad como la diferencia en número de parámetros entre los dos modelos comparados.

En el caso de que queramos comparar modelos no anidados se puede utilizar el Criterio de Información de Akaike (Akaike's Information Criterion; AIC). En este caso nuestro interés no sería comparar la idoneidad de introducir una variable en el modelo sino comparar entre modelos que incluyen diferentes variables predictoras. Para hacer esta comparación será necesario que ambos modelos estén calculados sobre los mismos datos. Si un modelo tiene p parámetros, $AIC = Deviance + 2p$. El AIC se rige por un razonamiento similar a la *Deviance* en el que cuanto menor es su valor, mejor es el ajuste. Su interpretación se hace en términos absolutos, mediante el valor obtenido del estadístico que nos indicará que modelo tiene un mejor ajuste.

7.3 MODELOS DE RIESGO A TIEMPO DISCRETO

7.3.1 DEFINICIÓN DE RIESGO A TIEMPO DISCRETO

A continuación denotamos T_i a la variable aleatoria que expresa el periodo j (mamografía) en el que el individuo i experimenta el evento de interés.

Para evaluar la probabilidad de ocurrencia del evento de interés en cada periodo de tiempo discreto utilizaremos la probabilidad de ocurrencia del evento, 'hazard', que denotaremos como $h(t_{ij})$. El riesgo en tiempo discreto es exactamente la *probabilidad condicionada de que el individuo i experimente el evento de interés en el periodo j , dado que no lo ha experimentado con anterioridad*. Podemos expresar este riesgo como:

$$h(t_{ij}) = \Pr[T_i = j \mid T_i \geq j] \quad (7.1)$$

El conjunto de probabilidades representadas por la función $h(t_{ij})$ es conocido como la función de riesgo a tiempo discreto de la población.

En esta definición subyace una importante componente de condicionalidad. El individuo i solo puede experimentar el evento en el periodo j si no lo ha experimentado con anterioridad en los $j-1$ periodos anteriores. El 'hazard' representa la probabilidad de ocurrencia del evento de interés entre aquellos individuos a riesgo de experimentar el evento. De esta manera, a medida que los individuos experimentan el evento de interés dejan de estar en el conjunto de individuos a riesgo, de manera que $h(t_{ij})$ expresará la probabilidad de primera ocurrencia del evento.

En un periodo concreto j , podemos expresar el estimador del riesgo, $\hat{h}(t_j)$ como:

$$\hat{h}(t_j) = \frac{d_i}{n_i}$$

Donde d_i expresa el número de individuos que experimentan el evento de interés en el periodo j y n_i el número de individuos a riesgo durante dicho periodo. La magnitud de $\hat{h}(t_j)$ dentro de cada intervalo expresa el riesgo de ocurrencia en dicho intervalo:

- 1) Siempre toma valores entre 0 y 1 (probabilidad)

- 2) Puede tomar cualquier valor dentro de este rango. Cuanto mayor sea el valor de $\hat{h}(t_j)$ mayor riesgo, y viceversa.

De manera similar podemos expresar la probabilidad de que el individuo i no experimente el evento de interés hasta un periodo posterior al j (probabilidad de supervivencia) como:

$$S(t_{ij}) = \Pr[T_i \geq j]$$

Esta probabilidad se puede estimar en la población de interés a partir de la función de riesgo como:

$$\hat{S}(t_j) = \hat{S}(t_{j-1}) [1 - \hat{h}(t_j)]$$

7.3.2 FORMULACIÓN DEL MODELO DE RIESGO A TIEMPO DISCRETO

Tal y como se ha definido, $h(t_{ij})$ expresa la función de riesgo a tiempo discreto para el individuo i en el periodo j . Suponemos ahora que diferentes individuos pueden tener diferentes funciones de riesgo dependiendo de las características que tengan de sus predictores. Este concepto corresponde con el de heterogeneidad observada.

Para introducir esta suposición en la definición de la función de riesgo en primer lugar, identificamos predictores asociados de manera hipotética con la ocurrencia del evento. Estos podrán ser:

- No cambiantes en el tiempo '*time-invariant*' como podrían ser el sexo, o la nacionalidad en otro contexto que no es el de este estudio.
- cambiantes en el tiempo '*time-varying*'. Los predictores cambiantes en el tiempo podrán tomar diferentes valores en cada periodo de tiempo j (aunque asumimos que sus valores permanecen constantes dentro de cada periodo de tiempo. Esto es cierto en el caso de las mamografías de cribado). Para poder expresar este modelo de manera general, de forma que incluya los dos tipos de predictores, denominaremos con $X_{1ij}, X_{2ij}, \dots, X_{p_{ij}}$ las p variables predictoras y $x_{p_{ij}}$ el valor del predictor p -ésimo del individuo i en el periodo j . Si el predictor es no cambiante en el tiempo, tomará el mismo valor en cada periodo j . Si es cambiante en el

tiempo, podrá tomar diferentes valores. Introducimos la heterogeneidad observada en la definición de la función de riesgo mediante:

$$h(t_{ij} | X) = \Pr(T_i = j | T_i \geq j, X_{1ij} = x_{1ij}, X_{2ij} = x_{2ij}, \dots, X_{p_{ij}} = x_{p_{ij}})$$

Que expresa la probabilidad para el individuo i experimentar el evento de interés en el periodo j , condicionado a no haber experimentado el evento con anterioridad y a sus valores particulares de los p predictores en dicho periodo de tiempo.

Se añade a continuación una reformulación del concepto de periodo. Como el evento de interés únicamente puede ocurrir en uno solo de los J periodos de tiempo (mamografías de cribado) en que participa cada individuo, podemos construir un conjunto de variables *dummy*, que llamaremos 'indicadores de tiempo', para expresar cada uno de los posibles valores que toma J . Representamos estos indicadores de tiempo como $D_{1ij}, D_{2ij}, \dots, D_{Jij}$. Cada uno de estos indicadores toma el valor 1 en el periodo de tiempo que representa y 0 en todos los demás. Por ejemplo $D_{1ij}=1$ en el primer periodo de tiempo ($j=1$) y 0 en todos los demás ($j>1$).

La función de riesgo a tiempo discreto tiene dos propiedades muy importantes: 1) Por un lado nos proporciona el riesgo basal; 2) Los cambios en los parámetros del modelo de riesgo muestran el efecto de los predictores en el riesgo basal.

Dado que los $h(t_{ij})$ son probabilidades, podemos expresarlos de manera que tengan una relación logística con los predictores sustantivos y los predictores de tiempo (periodos). Podemos reformular la función de riesgo a tiempo discreto como:

$$h(t_{ij}) = \frac{1}{1 + \exp - [(\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \dots + \alpha_J D_{Jij}) + (\beta_1 X_{1ij} + \beta_2 X_{2ij} + \dots + \beta_p X_{p_{ij}})]} \quad (7.2)$$

De esta manera, aplicando la transformación logística en ambas partes de la ecuación 7.2, utilizando los indicadores de tiempo previamente definidos, así como los p predictores sustantivos podemos representar el modelo de tiempo a riesgo discreto como:

$$\text{logit } (h(t_{ij})) = (\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \dots + \alpha_J D_{Jij}) + (\beta_1 X_{1ij} + \beta_2 X_{2ij} + \dots + \beta_p X_{p_{ij}}) \quad (7.3)$$

- Los parámetros de intercept, $\alpha_1, \alpha_2, \dots, \alpha_j$, representan el valor del *logit hazard* (el log odds de la ocurrencia del evento) en cada periodo de tiempo específico para los individuos en el grupo basal (referencia).

- Los predictores sustantivos, $\beta_1, \beta_2, \dots, \beta_j$, evalúan el efecto de una diferencia unitaria en dicho predictor sobre la ocurrencia del evento, controlando por el efecto de todos los demás predictores del modelo.

7.3.3 ESTRUCTURA DE LOS DATOS

Para poder aplicar la formulación propuesta en nuestro conjunto de datos, es necesario reestructurar la base con la mamografía como unidad de análisis (un registro por cada mamografía), en lugar de un registro por cada mujer. De esta manera se pueden definir en el modelo los indicadores de tiempo para cada participación. La estructura obtenida tiene una forma similar a la mostrada en la figura 7.1.

Figura 7.1: estructura de las variables Dummy indicadoras de tiempo.

Periodo	D ₁	D ₂	...	D _{j-1}	D _j
1	1	0	0	0	0
2	0	1	0	0	0
...	0	0	...	0	0
J-1	0	0	0	1	0
J	0	0	0	0	1

Donde cada uno de los indicadores de tiempo toman el valor 1 en el periodo de tiempo j al que representan, y 0 en todos los demás.

Como se observa en la formulación del modelo de la ecuación (7.3) el modelo no tiene intercept. En su lugar tiene J componentes de tiempo, una para cada periodo que actúan como múltiples intercept. En la especificación del modelo, será necesario forzarlo para no estimar el intercept (constante) para la parte de efectos fijos, de manera que a partir de cada uno de los parámetros asociados a los indicadores de tiempo $\alpha_1, \alpha_2, \dots, \alpha_j$, podremos obtener el valor del riesgo estimado para cada periodo.

Como se comentó anteriormente, es de especial importancia para poder modelizar correctamente la función de riesgo, tener en cuenta independencia entre las observaciones de una misma mujer (medidas repetidas). Para asegurar la independencia entre las observaciones de una misma mujer y obtener un estimador no sesgado de la probabilidad de observar el evento de interés, se modelizan únicamente las participaciones de cada individuo hasta la primera ocurrencia del evento de interés,

sin considerar las apariciones posteriores. Para conseguir esto será necesario sacar de la base de datos de análisis todos los registros posteriores a la ocurrencia del primer evento de interés en cada mujer. Como tenemos un registro por cada participación de la mujer en el cribado, eliminaremos exactamente tantos registros como participaciones posteriores al primer falso positivo de cada individuo tengamos.

La estructura de la información contenida en la base de datos se muestra en la figura 7.2.

Figura 7.2: Estructura de la base de datos de análisis para estimar los modelos de regresión

Individuo	D_1	D_2	D_3	D_4	D_5	...	Evento	X_1	...	X_p
1	1	0	0	0	0		0	X_{111}		X_{11p}
1	0	1	0	0	0		0	X_{121}		X_{12p}
1	0	0	1	0	0		0	X_{131}		X_{13p}
1	0	0	0	1	0		0	X_{141}		X_{14p}
1	0	0	0	0	1		0	X_{151}		X_{15p}
2	1	0	0	0	0		0	X_{211}		X_{21p}
2	0	1	0	0	0		0	X_{221}		X_{22p}
2	0	0	1	0	0		1	X_{231}		X_{23p}
3	1	0	0	0	0		1	X_{311}		X_{31p}
4	1	0	0	0	0		0	X_{411}		X_{41p}
...

Donde los D_j representan las variables Dummy indicadoras de tiempo, la variable 'Evento' indica la ocurrencia del evento de interés, y ninguna mujer genera un registro posterior a la ocurrencia del primer evento, y los predictores X_{ijp} representan el valor que toma el predictor sustantivo p -esimo, para el individuo i -esimo, en su j -esima mamografía.

7.3.4 ESTIMADORES DE MÁXIMA VEROSIMILITUD PARA EL DTHM

Como ya se comentó para el caso de la regresión logística, el método de máxima verosimilitud, busca aquellos estimadores de los parámetros poblacionales (las α 's y las β 's en este caso) que maximizan la verosimilitud de observar los datos en la muestra.

La función de verosimilitud en los modelos de riesgo a tiempo discreto expresa la probabilidad de observar el patrón específico de ocurrencia de los eventos observado en la práctica. – el patrón exacto de 0's y 1's de la variable 'EVENT' en el conjunto de datos persona-periodo.

Como cada individuo tiene J_i observaciones, una por periodo, cada observación contribuye J_i términos a la función de verosimilitud. Como resultado, la función de verosimilitud se compone de tantos términos como registros/observaciones del conjunto de datos persona-periodo.

El valor específico que el individuo i contribuye en el periodo j depende de sus valores observados de riesgo/Hazard en dicho periodo, $h(t_{ij})$, y de cómo el individuo experimentó el evento de interés en ese periodo de tiempo ($EVENT_{ij}$). Determinaremos el valor de cada contribución mediante la siguiente procedimiento/lógica:

- La probabilidad de que el individuo i experimente el evento de interés en el periodo j , dado que el evento no ha ocurrido con anterioridad, es $h(t_{ij})$. En el periodo de tiempo en el que el individuo i experimenta el evento ($EVENT= 1$), el individuo contribuye $h(t_{ij})$ a la función de verosimilitud.
- La probabilidad de que el individuo i NO experimente el evento de interés en el periodo j , dado que el evento no ha ocurrido con anterioridad, es $(1 - h(t_{ij}))$. En todos los periodos de tiempo en los que el individuo i NO experimenta el evento ($EVENT= 0$), el individuo contribuye $(1 - h(t_{ij}))$ a la función de verosimilitud.

Los individuos censurados sólo contribuyen con términos del segundo tipo a la función de verosimilitud. Los individuos que experimentan el evento de interés contribuyen con un término del primer tipo y $J_i - 1$ términos del segundo tipo. A partir de estas características podemos expresar la función de verosimilitud para el modelo de riesgo a tiempo discreto de la siguiente manera:

$$Likelihood = \prod_{i=1}^n \prod_{j=1}^{J_i} h(t_{ij})^{EVENT_{ij}} (1 - h(t_{ij}))^{(1 - EVENT_{ij})} \quad (7.4)$$

Además, podemos reformular la función de verosimilitud en la ecuación 7.4 sustituyendo los valores de $h(t_{ij})$ por su expresión a partir del modelo explícito en la ecuación 7.2.

$$Likelihood = \prod_{i=1}^n \prod_{j=1}^{J_i} \left[\frac{1}{1 + e^{-\{\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \dots + \alpha_J D_{Jij}\} + \{\beta_1 X_{1ij} + \beta_2 X_{2ij} + \dots + \beta_p X_{p ij}\}}} \right]^{EVENT_{ij}} \\ \times \left[1 - \frac{1}{1 + e^{-\{\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \dots + \alpha_J D_{Jij}\} + \{\beta_1 X_{1ij} + \beta_2 X_{2ij} + \dots + \beta_p X_{p ij}\}}} \right]^{(1 - EVENT_{ij})}$$

7.3.5 MODELO RIESGO A TIEMPO DISCRETO CON EFECTOS ALEATORIOS

Los modelos de efectos aleatorios extienden los modelos de efectos fijos incluyendo una componente aleatoria en el modelo que recogerá la heterogeneidad de varianza encontrada entre distintos grupos de observaciones. Añadimos al modelo de riesgo a tiempo discreto una componente aleatoria que vendrá expresada como $Z\gamma_i$ donde:

- El vector $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$, es el vector de coeficientes aleatorios del modelo
- Z es la matriz de diseño de los efectos aleatorios.

Podemos expresar el modelo de riesgo con efectos aleatorios para el k -ésimo grupo como:

$$y_k = \log it(h(t_{ij})) = \alpha D_k + \beta X_k + \gamma_k Z_k \quad (7.5)$$

$$\gamma_k \sim N(0, \Omega)$$

Donde y_k , la respuesta del k -ésimo grupo (k -ésima unidad radiológica), es un vector de longitud igual al número de observaciones (mamografías) en dicho grupo; α es el vector de variables de tiempo (efectos fijos) de longitud j , β es el vector de variables predictoras (efectos fijos) de longitud p , γ_k es el vector de efectos aleatorios de longitud n . D_k , X_k y Z_k son las matrices de efectos fijos y aleatorios, cada una con tantas filas como observaciones (mamografías) en el k -ésimo grupo.

En este caso, como únicamente introducimos una componente aleatoria (también llamado modelo de dos niveles) el vector de coeficientes aleatorios tendrá una única componente γ_1 . Y la matriz de diseño Z tiene tantas filas como registros (mamografías) y una única columna. Podemos generalizar la expresión de la ecuación 7.5 para el modelo de dos niveles (una única componente aleatoria) como:

$$\text{logit}(h(t_{jk})) = \sum_{J=1}^J \alpha_J D_{Jjk} + \sum_{P=1}^P \beta_P X_{Pjk} + \gamma_{ijk}$$

$$\gamma_k \sim NMV(0, \Omega_\gamma)$$

D_{Jjk} representan la matriz de diseño de los J intervalos de tiempo discretos para el individuo i -ésimo en su j -ésima mamografía. X_{Pjk} es la matriz de diseño de los P predictores sustantivos, para el individuo i -ésimo en su j -ésima mamografía. γ_{ijk} representa el efecto aleatorio para el k bloque del efecto (k -ésima unidad radiológica).

Para la estimación de parámetros de este tipo de modelos la aproximación común es mediante la maximización de la función de verosimilitud, la cual solo puede definirse para los efectos aleatorios. Se utilizan en su lugar aproximaciones alternativas como la *quasi-verosimilitud* y la *pseudo-verosimilitud* (8), cada una de ellas con sus ventajas y desventajas, sobre todo en lo respectivo a la minimización de los errores. No se puede recomendar ninguno de estos métodos como el mejor para ajustar todos los modelos lineales generalizados mixtos.

Software estadístico

Para la implementación de los modelos propuestos se utilizó el procedimiento **GLIMMIX** en el software estadístico **SAS 9.1** (SAS Institute, Cary, NC). Este procedimiento es específico para realizar modelos de regresión no lineales con efectos aleatorios.

En nuestro caso particular, tal y como se comenta brevemente en la sección 5.5 de este trabajo y en mayor profundidad en el siguiente capítulo, debido a la problemática del tamaño de la base de datos, se descartaron otros software disponibles para la implementación de este tipo de modelos. El único método que se pudo computar correctamente fue la estimación por pseudo-verosimilitud, teniendo que descartar la aproximación por quasi-verosimilitud (ver siguiente sección).

7.3.6 SELECCIÓN DE MODELOS

En el caso de los modelos lineales generalizados con efectos aleatorios (GLMM) que son los que nos ocupan, la comparación y selección de modelos no es trivial. En la estimación de estos modelos hemos utilizado el procedimiento PROC GLIMMIX de SAS 9.1, que como se explicó anteriormente, fue el único software capaz de trabajar con una base de datos de estas dimensiones. El procedimiento GLIMMIX utiliza pseudo-verosimilitud para estimar los parámetros. Mediante pseudo-verosimilitud no podemos obtener estadísticos de ajuste para la comparación de modelos. En la salida del modelo obtenemos una serie de estadísticos de ajuste basados en la pseudo-verosimilitud, pero que únicamente tienen un carácter informativo, en ningún caso nos servirán para establecer una valoración formal. La siguiente figura muestra un ejemplo de los estadísticos obtenidos.

Figura 7.3: 'Output' del procedimiento GLIMMIX de SAS basado en una aproximación por pseudo-verosimilitud

Fit Statistics	
-2 Res Log Pseudo-Likelihood	25780244
Pseudo-AIC	25780274
Pseudo-AICC	25780274
Pseudo-BIC	25780473
Pseudo-CAIC	25780488
Pseudo-HQIC	25780326
Generalized Chi-Square	4265092
Gener. Chi-Square / DF	1.00

REML information criteria are adjusted for fixed effects and covariance parameters. Fit statistics based on pseudo-likelihoods are not useful for comparing models that differ in their pseudo-data.

Además, si consultamos el manual de referencia de SAS nos advierte nuevamente de que no es posible utilizar estos estimadores para comparar modelos: ***"Note that the (residual) log pseudo-likelihood in a GLMM is the (residual) log likelihood of a linearized model. You should not compare these values across different statistical models, even if the models are nested with respect to fixed and/or G-side random effects. It is possible that between two nested models the larger model has a smaller pseudo-likelihood."***

Una aproximación para poder comparar modelos no lineales con efectos aleatorios podría ser utilizar el procedimiento PROC NLMIXED de SAS. Este procedimiento se basa en una aproximación integral (cuadratura gaussiana) para la

estimación de parámetros que permite realizar un test de razón de verosimilitud. No se utilizó este procedimiento para la estimación de los modelos propuestos en este trabajo ya que es un modelo más restrictivo en cuanto a su computación y menos eficiente que el GLIMMIX y no se pudo resolver ninguno de los modelos propuestos mediante NLMIXED y fue descartado.

La nueva versión de SAS 9.2 ya tiene una opción implementada basada en la aproximación integral que permite la comparación de modelos, pero debe ser aplicada con cautela ya que es experimental. Este trabajo se desarrolló con la versión 9.1 de SAS que era la única disponible ya que es un software de pago.

La selección y valoración de los modelos presentados en este trabajo se hizo mediante la significación estadística de los parámetros introducidos en los modelos, y la relevancia epidemiológica de los diferentes factores de estudio.

Por un lado, se trabajó con un número reducido de variables de estudio. Todos los factores de riesgo estudiados son relevantes en el contexto del cribado de cáncer de mama, y han sido recogidas expresamente para su valoración. El propósito de este estudio fue valorar el impacto de estas variables en el riesgo de falsos positivos. No se consideró como cuestión de interés el comparar la eficiencia entre diferentes modelos, ya que todas las variables de estudio se consideraban relevantes.

Por otro lado, el gran tamaño de muestra de la base de datos de análisis nos permite poder introducir parámetros en el modelo sin que esto represente una disminución sensible de la potencia estadística del modelo.

A raíz de los problemas encontrados para la selección y comparación de modelos en el contexto de los modelos lineales generalizados con efectos aleatorios, queda planteado el debate sobre las posibles aproximaciones para resolver esta problemática. Las soluciones encontradas en la actualidad son de carácter provisional, y dejan patente la necesidad de profundizar y avanzar en este campo en futuro más inmediato.

7.3.7 VALORACIÓN DEL EFECTO ALEATORIO

Se muestra en las figuras 7.4 y 7.5 el estimador del efecto aleatorio para cada una de sus componentes (unidad radiológica). El conjunto de estos estimadores tiene media $\mu=0$ y una componente de error σ_i para cada una de las unidades. Se construyeron los siguientes gráficos a partir de la estimación del efecto en un modelo basal que únicamente incluía las componentes de tiempo discreto.

Figura 7.4: *Estimadores del efecto aleatorio para cada una de sus componentes en el modelo de falsos positivos por cualquier tipo de prueba*

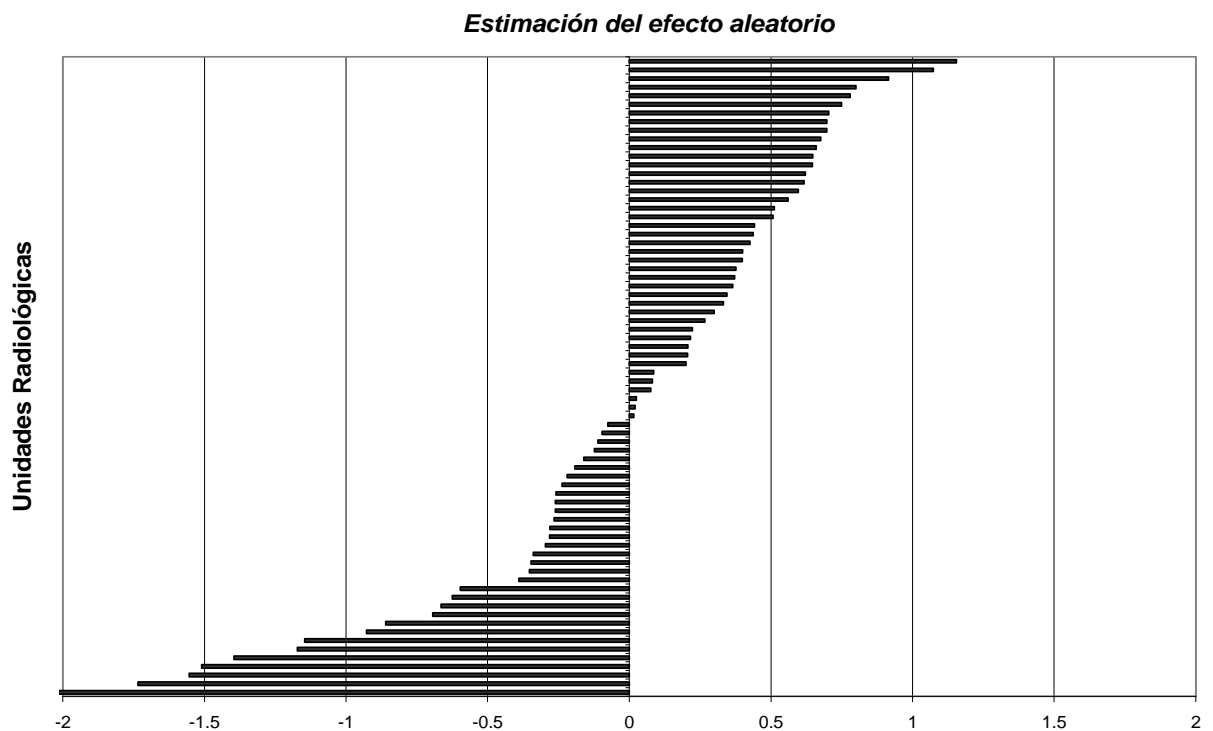
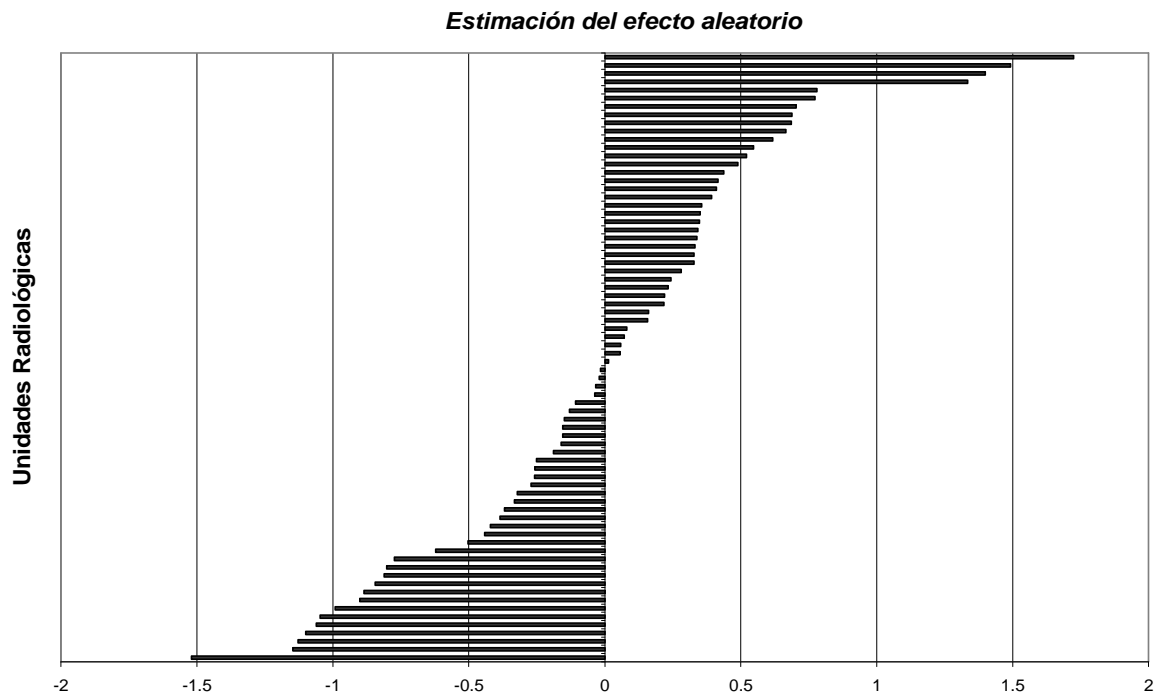


Figura 7.5: Estimadores del efecto aleatorio para cada una de sus componentes en el modelo de falsos positivos por pruebas invasivas



7.4 ESTIMACIÓN DEL RIESGO ACUMULADO

El riesgo acumulado de falso positivo a lo largo de las 10 posibles participaciones de la mujer en el cribado mamográfico se calculó como una secuencia de probabilidades. De manera análoga al estimador de Kaplan–Meier en el análisis de supervivencia, obtenemos un estimador para el riesgo acumulado de falso-positivo. En particular, tal y como se expresa en la ecuación 7.1, para la mujer *i*-ésima en su mamografía *j*-ésima, h_{ij} expresa la Pr(falso-positivo en la *j*-ésima mamografía dado que no ha tenido un falso-positivo en las previas *j*-1 mamografías).

h_{ij} es una función dependiente del vector de predictores asociados a cada mamografía. En el modelo de estimación del riesgo basal, que no tiene predictores sustantivos, dependerá únicamente de los valores del α_j , predictor de tiempo, asociado a cada participación.

Obtenemos la probabilidad de ocurrencia del evento, asociado a la *j*-ésima mamografía como:

$$\hat{h}_j = 1 / (1 + \exp(-\alpha_j))$$

donde α_j es el parámetro estimado en el modelo de regresión para dicha *mamografía j-ésima*.

Obtenemos la probabilidad de ocurrencia del evento, asociado a la *j*-ésima mamografía para el modelo con predictores sustantivos como:

$$\hat{h}_j = 1 / (1 + \exp(-\alpha_j + \sum X_{ij}))$$

Donde el conjunto de los X_{ij} representan los predictores sustantivos de las variables estudiadas para la proyección.

Por otro lado, para cada mamografía de cribado conocemos la probabilidad de no haber experimentado el evento de interés, S_j ; que equivale a Pr(No falso-positivo en la mamografía *j*-ésima dado que no ha tenido un falso-positivo en las previas *j*-1 mamografías). Podemos calcular esta probabilidad, que representará la proporción de mujeres que no han tenido un falso positivo hasta la *j*-ésima mamografía, ($j=2, \dots, 10$) como:

$$\hat{S}_j = \hat{S}_{j-1} * (1 - \hat{h}_j) \quad (7.6)$$

De esta manera el Riesgo Acumulado de Falso Positivo (RAFP) hasta la mamografía *j-esima*, expresa la probabilidad acumulada de ocurrencia del evento, y será equivalente a:

$$RAFP_j = \hat{h}_j * (\hat{S}_{j-1}) + RAFP_{j-1} = 1 - \hat{S}_j \quad (j = 2, \dots, 10) \quad (7.7)$$

Las expresiones 7.6 y 7.7 son validas tanto para el modelo sin predictores sustantivos, como para el modelo con predictores sustantivos.

* NOTA: El Riesgo acumulado de falso positivo (RAFP), no expresa el “riesgo acumulado” de la función de riesgo h_{ij} , tal y como se suele expresarse en términos del análisis de supervivencia. En este contexto, denomina la “Probabilidad acumulada de ocurrencia de un primer falso positivo), o lo que es lo mismo, $RAFP_j = 1 - \hat{S}_j$.

En particular, para la primera participación en el cribado (j=1) tenemos:

$$RAFP_1 = \hat{h}_1$$

$$\hat{S}_1 = 1 - \hat{h}_1$$

Intervalos de confianza

Para calcular el error estándar de la probabilidad de supervivencia se utilizó la aproximación de Greenwood. En el año 1926 Greenwood demostró que el error estándar de la probabilidad de supervivencia en el periodo j podía ser aproximado por:

$$s.e(\hat{S}(t_j)) = \hat{S}(t_j) \sqrt{\frac{\hat{h}(t_1)}{n_1(1-\hat{h}(t_1))} + \frac{\hat{h}(t_2)}{n_2(1-\hat{h}(t_2))} + \dots + \frac{\hat{h}(t_j)}{n_j(1-\hat{h}(t_j))}}$$

A partir de esta aproximación, se calcularon los intervalos de confianza para la probabilidad de supervivencia \hat{S}_j ; y para el $RAFP_j$

8 RESULTADOS

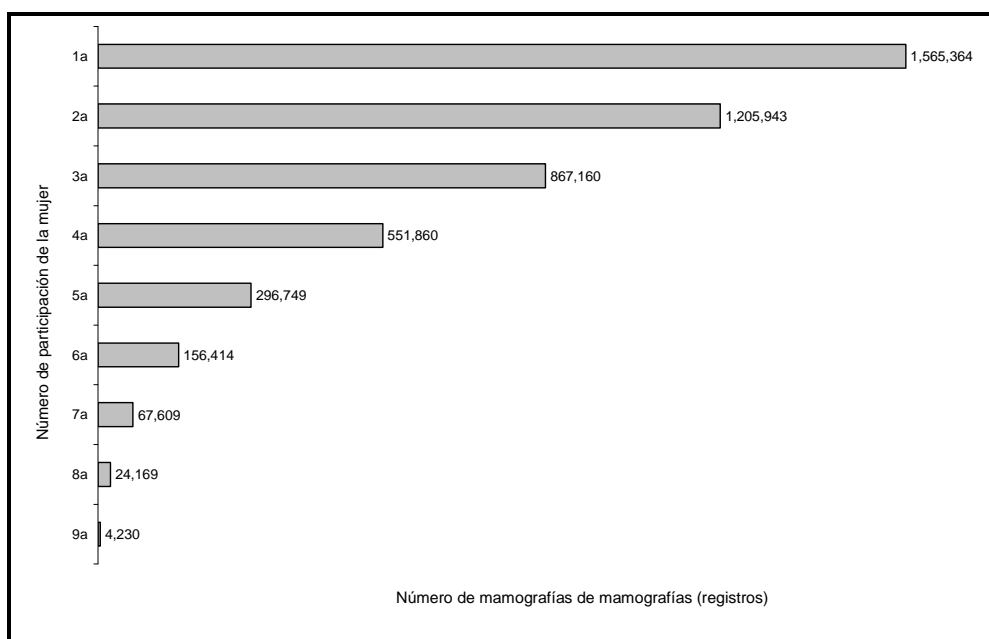
Los resultados de este estudio han sido aceptados recientemente para su publicación en la revista “*Annals of Oncology*”

Román R; Sala M; Salas D; Ascunce N; Zubizarreta R; Castells X; and the Spanish Breast Cancer Screening Research Group. “*Effect of protocol-related variables and women’s characteristics on the cumulative false-positive risk in breast cancer screening*”. *Annals of Oncology* 2011 [en prensa].

8.1 ANÁLISIS DESCRIPTIVO

Se analizaron un total de 4.739.498 mamografías de cribado realizadas a 1.565.364 mujeres. De estas mujeres que participaron al menos una vez 1.205.943 (77.04%) tuvieron al menos una segunda mamografía, 867.160 (55.40%) una tercera y 156.414 (9.99%) llegaron hasta la sexta. La **figura 8.1** muestra el número de mujeres participantes según el número de cribado desde el 1º hasta el 9º, que corresponde al máximo número de participaciones posibles en nuestro periodo de estudio.

Figura 8.1: Número de mamografías de cribado según el número de participaciones de la mujer.



* En la primera participación el número de mamografías equivale al número de mujeres participantes.

En la **tabla 8.1** se muestra la información de la población de estudio en relación a los 16 años del periodo de estudio. La fecha de puesta en marcha de las unidades radiológicas es diferente, siendo la puesta en marcha de la primera unidad en 1990. Debido a esta incorporación gradual, el número de unidades radiológicas en funcionamiento aumenta a medida que avanzamos cronológicamente. Como se aprecia en esta primera tabla, se realizaron mamografías de cribado en 74 unidades radiológicas diferentes, con un promedio de 64.047 mamografías de cribado (percentil 10^o-90^a: 9.159–117.988) y 21.154 mujeres cribadas (percentil 10^o -90^o: 3.424–38.268) por unidad radiológica.

Tabla 8.1. Información sobre las mamografías de cribado realizadas según periodos bienales

	1990 - 1992	1993 - 1994	1995 - 1996	1997 - 1998	1999 - 2000	2001 - 2002	2003 - 2004	2005 - 2006	Total
	Num. (%)	Num. (%)	Num. (%)	Num. (%)	Num. (%)	Num. (%)	Num. (%)	Num. (%)	Num.
Mamografías de cribado	67,806 (1.4)	233,407 (4.9)	371,033 (7.8)	485,800 (10.3)	714,981 (15.1)	849,415 (17.9)	932,861 (19.7)	1,084,195 (22.9)	4.739.498
Mujeres participantes (1er cribado)	61,746 (3.9)	178,245 (11.4)	198,190 (12.7)	198,721 (12.7)	298,747 (19.1)	240,817 (15.4)	196,470 (12.6)	192,428 (12.3)	1.565.364
Mamografías (cribado sucesivo)	6,060 (0.2)	55,162 (1.7)	172,843 (5.4)	287,079 (9.0)	416,234 (13.1)	608,598 (19.2)	736,391 (23.2)	891,767 (28.1)	3.174.134
Exploraciones adicionales	13,037 (3.4)	24,013 (6.4)	35,070 (9.3)	41,886 (11.1)	68,603 (18.1)	64,991 (17.2)	63,945 (16.9)	66,515 (17.6)	378.060
Mujeres con un FP *	10,175 (3.9)	18,992 (7.2)	27,727 (10.5)	30,077 (11.4)	46,024 (17.5)	43,707 (16.6)	42,278 (16.0)	44,627 (16.9)	263.607
Mujeres con un FP (invasivas) †	566 (2.3)	2,532 (10.4)	2,471 (10.1)	3,075 (12.6)	4,511 (18.5)	4,259 (17.4)	3,687 (15.1)	3,306 (13.5)	24.407
Unidades radiológicas ‡	9	21	33	41	63	68	71	74	74

* : Falso positivo por cualquier exploración adicional (invasiva o no invasiva)

† : Falso positivo por exploración adicional invasiva

‡ : Expresado como número de unidades radiológicas en funcionamiento en dicho periodo

Del total de 1.565.364 mujeres que participaron en al menos un cribado, 467.910 se cribaron por primera vez con 44-49 años, 477.177 con 50-54 años, 300.901 con 55-59 años, 260.223 con 60-64 años, y 59.153 con 65-69 años. En la **tabla 8.2** se muestran los descriptivos de algunos indicadores básicos en la evaluación del cribado mamográfico. Entre ellos la tasa de falsos-positivos por cualquier prueba y por pruebas invasivas, y la tasa de detección de cáncer para cribado inicial y cribados sucesivos.

Tabla 8.2. Número y porcentaje de falsos positivos y cáncer detectado según cribado inicial o sucesivo

	Cribado inicial		Cribado sucesivo		Total	
	Núm.	Porcentaje [IC 95%]	Núm.	Porcentaje [IC 95%]	Núm.	Porcentaje [IC 95%]
Falso positivo	134.757	8.6 (8.56-8.65)	130.044	4.10 (4.08-4.12)	264.801	5.59 (5.57-5.61)
Falso positivo (invasivas)	15.894	1.02 (1.00-1.03)	8.542	0.27 (0.26-0.28)	24.436	0.52 (0.51-0.52)
Detección cáncer	7.065	0.45 (0.44-0.46)	9.464	0.30 (0.29-0.30)	16.529	0.35 (0.34-0.35)

8.2 RESULTADOS DEL ESTUDIO

Los Odds Ratios ajustados para el riesgo de falso-positivo por cualquier prueba, falso-positivo por pruebas invasivas y para las tasas de detección en función de las variables de las variables de protocolo (método de lectura, número de proyecciones, y tipo de mamografía) se muestran en la **tabla 8.3**. Las mamografías leídas mediante el método de doble lectura tuvieron un mayor riesgo de presentar un falso-positivo (OR=2.06; IC95%: 2.00-2.13) que las leídas mediante lectura simple. Este riesgo fue aun mayor para el falso positivo por pruebas invasivas (OR=4.44; IC95%: 4.08-4.84). La doble proyección mamográfica tuvo un efecto protector para el falso-positivo por cualquier tipo de prueba (OR=0.77; IC95%: 0.76-0.79), pero fue un factor de riesgo para el falso-positivo por pruebas invasivas (OR=1.56; IC95%: 1.48-1.64). La mamografía digital tuvo un efecto protector en el falso-positivo por pruebas invasivas (OR=0.83; IC95%: 0.72-0.96), pero este efecto no fue estadísticamente significativo para el falso-positivo por cualquier tipo de prueba. Se detectaron más cánceres con doble lectura y con mamografía digital (OR=1.08; IC95%: 1.04-1.12 y OR=1.26; IC95%: 1.10-1.45, respectivamente).

Tabla 8.3. Riesgo de falso positivo y detección de cáncer para las variables asociadas al protocolo de lectura mamográfica (N= 4,739,498)

	Mamografías cribado	Falso positivo (cualquier prueba)	Falso positivo (pruebas invasivas)	Detección cáncer
		Análisis Multivariado [OR, I.C.95%] †	Análisis Multivariado [OR, I.C.95%] †	Análisis Multivariado [OR, I.C.95%] †
Método de lectura				
Lectura simple	1.734.930	Ref.	Ref.	Ref.
Lectura doble	3.004.568	2.06 (2.00, 2.13) *	4.44 (4.08, 4.84) *	1.08 (1.04, 1.12) *
Número proyecciones				
Una	1.482.503	Ref.	Ref.	Ref.
Dos	3.256.995	0.77 (0.76, 0.79) *	1.56 (1.48, 1.64) *	1.02 (0.97, 1.06) *
Tipo mamografía				
Analógica	4.676.138	Ref.	Ref.	Ref.
Digital	63.360	0.96 (0.92, 1.01)	0.83 (0.72, 0.96) *	1.26 (1.10, 1.45) *

* Estadísticamente significativo con un nivel de confianza del 95%
† Análisis multivariado ajustado por número de cribado de la mujer, unidad radiológica (efecto aleatorio), periodo y edad

El modelo incluyendo las variables de características de la mujer se muestra en la **tabla 8.4**. Se observó un mayor riesgo de falso-positivo por cualquier tipo de prueba y por pruebas invasivas en las mujeres más jóvenes (OR=1.50; IC95%: 1.46-1.54; y OR=1.44; IC95%: 1.30-1.58), mujeres con pruebas invasivas previas (OR=1.52; IC95%: 1.49-1.56 y OR=2.00; IC95%: 1.89-2.12), mujeres con historia familiar de cáncer de mama (OR=1.18; IC95%: 1.15-1.20; y OR=1.21; IC95%: 1.13-1.30,) y mujeres premenopáusicas (OR=1.31; IC95%: 1.29-1.33; y OR=1.22; IC95%: 1.16-1.29). El THS se asoció con un menor riesgo de falso-positivo por pruebas invasivas (OR=0.84; IC95%: 0.78-0.90). Se detectaron más cánceres en las mujeres premenopáusicas (OR=1.16; IC95%: 1.07-1.25), mujeres con pruebas invasivas previas (OR=1.31; IC95%: 1.20-1.42) y en aquellas con historia familiar de cáncer de mama (OR=1.66; IC95%: 1.55-1.79). Las mujeres más jóvenes y aquellas que utilizaban THS tuvieron tasas de detección más bajas (OR=0.39; IC95%: 0.35-0.43; y OR=0.86; IC95%: 0.80-0.94, respectivamente).

Tabla 8.4. Riesgo de falso positivo y detección de cáncer para las variables asociadas a las características de la mujer (ajustado por las variables de lectura mamográfica) (N= 2,777,429)

	Mamografías cribado	Falso positivo (cualquier prueba)	Falso positivo (pruebas invasivas)	Detección cáncer
		Análisis Multivariado [OR, I.C.95%] [†]	Análisis Multivariado [OR, I.C.95%] [†]	Análisis Multivariado [OR, I.C.95%] [†]
Edad en el cribado (años)				
44-49	469.047	1.50 (1.46, 1.54) *	1.44 (1.30, 1.58) *	0.39 (0.35, 0.43) *
50-54	699.256	1.26 (1.23, 1.29) *	1.26 (1.15, 1.37) *	0.48 (0.44, 0.52) *
55-59	695.921	1.13 (1.10, 1.16) *	1.06 (0.97, 1.16)	0.67 (0.62, 0.73) *
60-64	633.845	1.06 (1.03, 1.09) *	0.96 (0.88, 1.06)	0.84 (0.77, 0.90) *
65-69	279.360	Ref.	Ref.	Ref.
THS				
No	2.485.550	Ref.	Ref.	Ref.
Sí	291.879	1.03 (1.01, 1.05) *	0.84 (0.78, 0.90) *	0.86 (0.80, 0.94) *
Menopausia				
Menopáusica	2.157.627	Ref.	Ref.	Ref.
Premenopáusica	619.802	1.31 (1.29, 1.33) *	1.22 (1.16, 1.29) *	1.16 (1.07, 1.25) *
Pruebas invasivas previas				
No	2.585.871	Ref.	Ref.	Ref.
Sí	191.558	1.52 (1.49, 1.56) *	2.00 (1.89, 2.12) *	1.31 (1.20, 1.42) *
Historia familiar cáncer mama				
No	2.581.981	Ref.	Ref.	Ref.
Sí	195.448	1.18 (1.15, 1.20) *	1.21 (1.13, 1.30) *	1.66 (1.55, 1.79) *

* Estadísticamente significativo con un nivel de confianza del 95%

† Análisis multivariado ajustado por número de cribado de la mujer, unidad radiológica (efecto aleatorio), periodo, y variables de lectura mamográfica (método de lectura, número de proyecciones, tipo de mamografía)

THS: tratamiento hormonal sustitutivo; Menopausia: estado pre/perí-menopáusico o menopáusico

El riesgo acumulado de falso-positivo basal por cualquier tipo de prueba y por pruebas invasivas en mujeres con 50-51 años de edad en su primer cribado cuando se proyectó a 10 rondas fue 20.39% (IC95%: 20.02-20.76) y 1.76% (IC95%: 1.66-1.87), respectivamente. Las **figuras 8.2** y **8.3** muestran el riesgo acumulado estimado para mujeres de 50-51 años, con los perfiles de riesgo mayor y menor. El riesgo acumulado en 10 participaciones en mujeres con doble lectura, una proyección, mamografía analógica, estado premenopáusico, pruebas invasivas previas, e historia familiar de cáncer de mama se estimó en 51.43% (IC95%: 51.02–51.84) mientras que mujeres sin todos estos factores de riesgo tuvieron un riesgo estimado de 7.47% (IC95%: 7.23–7.72) (**Figura 8.2**). El riesgo diferencial entre el mayor y el menor perfil de riesgo fue 43.96%. Las características de protocolo explicaron un 54.2% de este riesgo diferencial entre perfiles de riesgo, mientras que las características personales de la mujer explicaron el 45.8% restante. El riesgo acumulado de falso-positivo por pruebas invasivas en las mujeres de alto riesgo (doble lectura, dos proyecciones, sin THS, estado premenopáusico, pruebas invasivas previas, e historia familiar de cáncer de mama) fue de 12.02% (IC95%: 11.75–12.30) mientras que aquellas que estaban en el

perfil de bajo riesgo (categorías opuestas) fue 1.58% (IC95%: 1.48–1.69) (**Figura 8.3**). El riesgo diferencial entre el perfil de alto y bajo riesgo fue de 10.44%. Las características personales de la mujer explicaron el 73.3% de este riesgo diferencia. El ratio máximo-mínimo entre el perfil de alto riesgo y el de bajo riesgo se estimó en 6.88 y 7.59 para el falso-positivo por cualquier prueba y por pruebas invasivas, respectivamente.

Figura 8.2: Riesgo acumulado (líneas) y riesgo de falso positivo (barras) según el número de participación para los perfiles de alto y bajo riesgo. Falso positivo por cualquier tipo de procedimiento

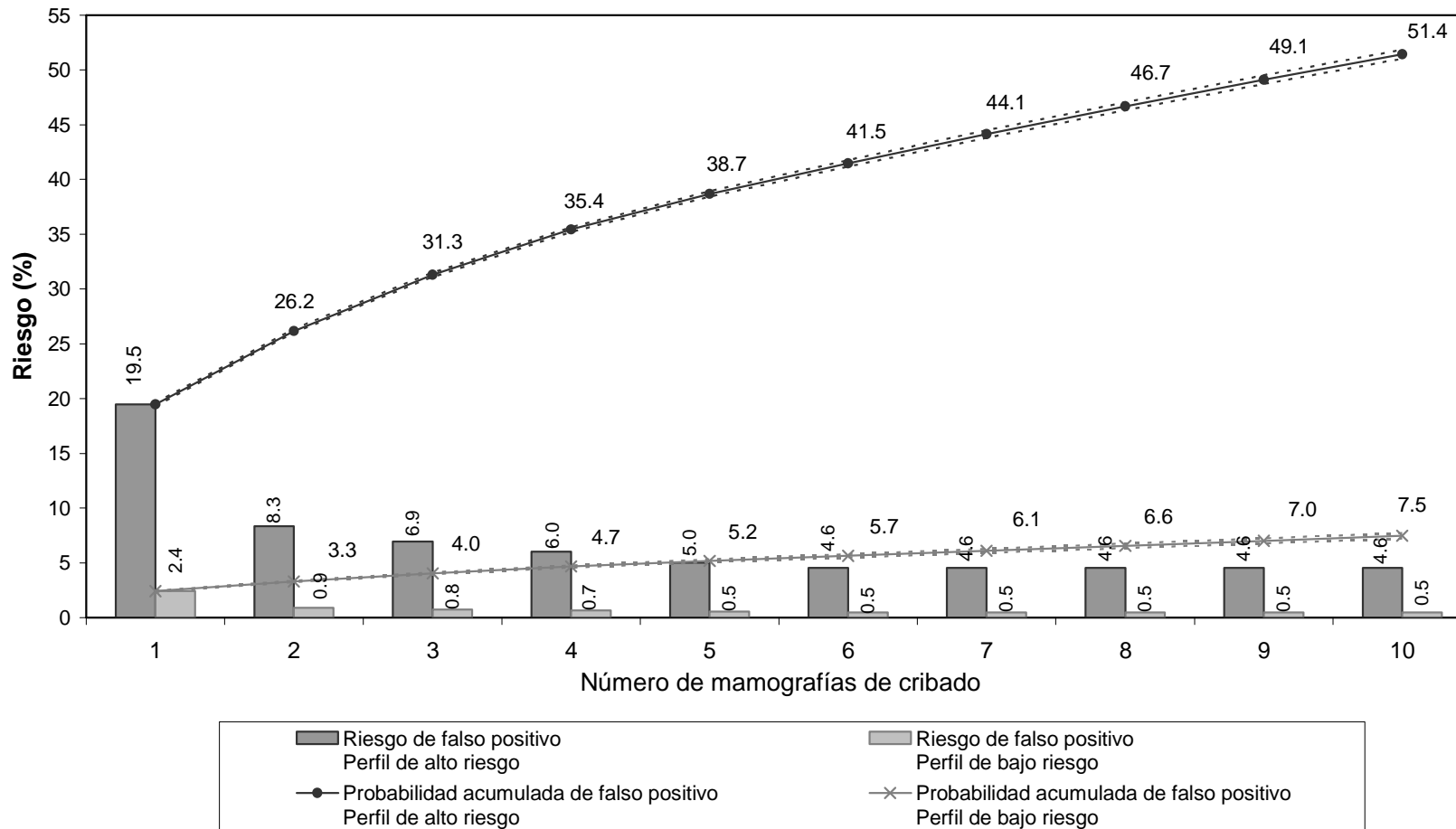
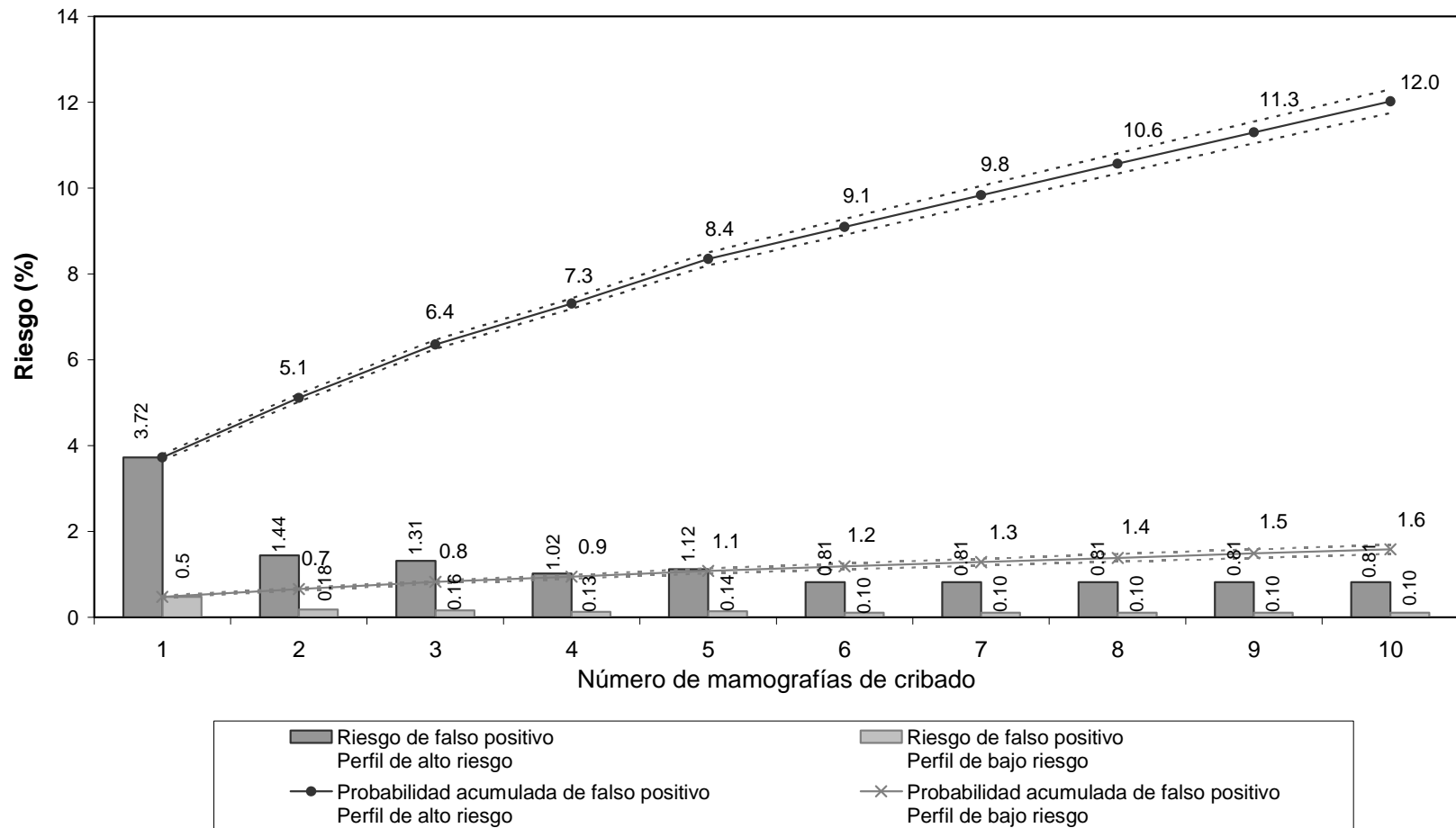


Figura 8.3: Probabilidad acumulada (líneas) y riesgo de falso positivo (barras) según el número de participación para los perfiles de alto y bajo riesgo. Falso positivo por pruebas invasivas



9 DISCUSIÓN

9.1 *IMPLICACIONES DEL ESTUDIO*

La estimación del riesgo acumulado de falso-positivo tiene el propósito de aportar la máxima información disponible a las mujeres invitadas a participar en el cribado poblacional de cáncer de mama. Hoy en día, los falsos-positivos son un efecto adverso indeseable del cribado. Si la reducción de la mortalidad por cáncer de mama como beneficio del cribado se analiza en términos de una secuencia de múltiples participaciones, los efectos adversos deberían de ser estudiados de la misma manera.

Hemos estimado que 1 de cada 5 mujeres que participan a lo largo de 10 convocatorias en un programa poblacional de cribado de cáncer de mama tendrán un falso positivo. Estos resultados son consistentes con hallazgos de otros estudios realizados en Noruega (9) y el Reino Unido (10), donde la organización de los programas de cribado es similar. Estos resultados fueron mucho menores que el 49.1% encontrado en EEUU (11;12). Estas diferencias se encontraron también en una comparación entre EEUU y Reino Unido (10). Una explicación para estas diferencias podría ser que el cribado de cáncer de mama en EEUU no está subvencionado ni gestionado por el gobierno, mientras que en Europa los programas de cribado han de cumplir una serie de estándares de calidad que implican menores tasas de falsos-positivos (10;13-15).

Merece la pena destacar que el riesgo acumulado de falso-positivo para pruebas invasivas fue 10 veces menor que el falso-positivo por cualquier tipo de prueba. A pesar de su menor riesgo, los efectos adversos de un falso-positivo debido a pruebas invasivas son mayores en términos de impacto físico y psicológico para la mujer, además de implicar un mayor coste que los procedimientos de imagen (no invasivos) y un retraso en el proceso de información hacia la mujer, debido al mayor tiempo de proceso y análisis de las muestras.

Estudios anteriores encontraron un mayor riesgo acumulado de falso-positivo por pruebas invasivas (9;10) en un contexto europeo, y un riesgo aun mayor en EEUU (10;12). Sin embargo, serian necesarios más estudios para analizar y determinar la gran variabilidad encontrada en el riesgo acumulado por pruebas invasivas en el contexto europeo.

9.2 APROXIMACIÓN ESTADÍSTICA

Se han propuesto diferentes aproximaciones en la bibliografía para la valoración del impacto de los efectos adversos del cribado. Destacan dos citas que por su relevancia y por ser los precursores en la estimación del riesgo de falsos positivos en el cribado mamográfico son tomados como artículos de referencia.

En 1994 Joann Elmore et al. (12) publicaron un artículo en '*New England Journal of Medicine*' en el que estimaban el riesgo acumulado de falsos positivos utilizando una versión *Bayesiana* del estimador de Kaplan-Meier. En estos cálculos se utilizaba el número de mamografías de cribado (valor discreto) en lugar de una variable de tiempo (continuo). Este artículo fue pionero en la estimación del riesgo acumulado de falso positivo y ha sido muy referenciado. Los estimadores calculados no se obtuvieron a partir de un modelo de regresión y presentan varias limitaciones. 1) No se pueden tener en cuenta posibles factores de confusión como la edad de las mujeres, el uso de tratamiento hormonal, variabilidad entre las fuentes de información, etc., que pudiesen estar sesgando estas estimaciones; 2) No permiten evaluar el impacto de ningún factor de riesgo sobre la probabilidad de presentar un falso positivo. Todos los datos mostrados en el artículo sobre las variables de estudio son meramente descriptivos.

Por su parte, en el año 2004 Solveig Hofvind et al. (9) publicaron un estudio para la estimación del riesgo acumulado de falsos positivos basado en otra aproximación. El cálculo del riesgo acumulado en este estudio está basado en el cálculo de probabilidades. Los autores seleccionan únicamente las mujeres con un primer cribado realizado con 50-51 años y que además hayan participado en tres cribados consecutivos. La probabilidad se calculó sumando la probabilidad de falso positivo en cada cribado, restando la intersección de los falsos positivos en dos cribados y añadiendo las intersecciones a tres rondas.

Este método ha sido muy criticado por no tener en cuenta la independencia entre los falsos positivos observados en las repetidas participaciones de una misma mujer (16).

Posteriormente otros autores como Castells et al. (17), a partir de la aproximación propuesta por Hofvind, reproducen el modelo en el contexto español con datos de un único programa de cribado. Estos autores proponen además un modelo de regresión logística para evaluar el impacto de los factores de estudio sobre el riesgo de falso positivo. En este modelo fijan las variables cambiantes en el tiempo

(antecedentes personales, IMC, edad, etc) al valor que toman en la primera participación de la mujer, de manera que el análisis se plantea de forma transversal, sin tener en cuenta la temporalidad, ni las diferencias en riesgo asociadas al número de participación.

Los modelos planteados en este trabajo permiten una aproximación más precisa a una problemática común a los estudios para la evaluación de los efectos adversos del cribado. Tal y como se ha comentado en la sección de métodos, estos modelos presentan una serie de ventajas sobre los propuestos hasta la fecha:

- a) Tiene en cuenta el tiempo como variable principal en la estimación del riesgo. El tiempo es discreto, y equivale a cada una de las participaciones en el cribado de la mujer, ya que la mujer solo está a riesgo de sufrir un falso-positivo cuando participa en el cribado.
- b) Estima un riesgo asociado a cada una de las posibles participaciones de la mujer.
- c) Considera las variables cambiantes en el tiempo, teniendo en cuenta los diferentes valores que puedan tomar las variables para una misma mujer a lo largo de sus diferentes participaciones.
- c) Tiene en cuenta la estructura jerárquica de los datos, ya que las mamografías de una misma unidad radiológica están correlacionadas.
- e) Tiene en cuenta en la formulación del modelo las repetidas observaciones de una misma mujer, de manera que la unidad de análisis es la propia mamografía y no la mujer.
- g) Proporciona una medida del efecto de las variables introducidas en el modelo sobre el riesgo de falso positivo.

Limitaciones estadísticas

A su vez, nos encontramos con varias limitaciones a la hora de aplicar esta metodología.

1) La comparación entre los diferentes modelos propuestos se hace complicada. Al trabajar con una Pseudo-verosimilitud, las comparaciones basadas en

Deviance, propuestas generalmente para la comparación de modelos no lineales, no son aconsejadas en el caso de modelos no lineales con efectos aleatorios, ya que podrían llevar a valores de la Log pseudo-verosimilitud sesgados. Este mismo problema surge para la comparación mediante el criterio de información de *Akaike*, o el *BIC* que no pueden ser obtenidos mediante la estimación por pseudo-verosimilitud

2) La categorización de las variables 'Edad' y 'Periodo' se hizo en función de los criterios epidemiológicos propuestos por el grupo de investigadores, de manera que los resultados fuesen comparables con los de otros estudios. La categorización de la edad en grupos quinquenales (45-49; 50-54; 55-59; 60-64; 65-70) está sujeta a la categorización encontrada en la bibliografía. La categorización de la variable periodo (1990-1994; 1995-1998; 1999-2002; 2003-2006) se hizo de forma arbitraria en función de las preferencias de los investigadores, respetando que fuesen intervalos de amplitudes similares. En cualquier caso, al disponer de un tamaño de muestra lo suficientemente grande no hubo problemas de representatividad dentro de los grupos seleccionados.

3) La información de las características personales de la mujer (menopausia, tratamiento hormonal, antecedentes personales, antecedentes familiares) tiene una proporción importante de valores perdidos. La información referente a estas variables no se recogía de forma rutinaria en todas las unidades radiológicas, y en otras, la manera en que estaba recogida no se ajustaba al protocolo del estudio. Se valoró el impacto de estos valores perdidos comparando la distribución de edad entre los registros con valores ausentes y los registros con información completa. También se compararon las tasas de falsos positivos entre estos grupos (ver anexos).

No se aplicaron técnicas de imputación de valores faltantes. El equipo investigador valoró la representatividad de la información disponible sobre variables de la mujer y se consideró que la submuestra con información disponible de estas variables (N=2,777,429, 58.6% de la muestra) era suficientemente representativa y permitía valorar conjuntamente en un mismo modelo las variables del protocolo de lectura y de características de la mujer.

9.3 VALORACIÓN DE LOS FACTORES DE ESTUDIO.

Se han descrito con anterioridad varios factores que afectan el riesgo de falsos positivos, incluyendo el método de lectura, el número de proyecciones mamográficas, la calidad de la mamografía y la experiencia del radiólogo (18-20). En concordancia con los resultados de varios estudio previos (14;21;22) encontramos que la doble lectura se asoció con un mayor riesgo de falso-positivo (OR=2.06) y mayor tasas de detección de cáncer (OR=1.08) que la lectura simple. Este aumento en la tasa de detección justifica por si mismo la recomendación de utilizar doble lectura de las '*Guías europeas para el cribado de cáncer de mama*'. Sin embargo, hay una gran variabilidad en estudios previos en el balance entre riesgos y beneficios de la doble lectura frente a lectura simple (13-15;23). Algunos estudios argumentan que el aumento en la tasa de falsos positivos asociado con la doble lectura se reducía cuando se utilizaba doble lectura con consenso o con arbitraje en lugar de doble lectura sin consenso (23-25). En nuestro estudio, a pesar de que el tipo de doble lectura (consenso, arbitraje, o sin consenso) no constituía una variable de estudio, el 84.8% de las dobles lecturas realizadas fueron con consenso o arbitraje, mientras que únicamente el 15.2% fueron dobles lecturas sin consenso.

A pesar de la recomendación de las guías europeas para realizar doble proyección, en nuestro estudio algunas unidades radiológicas realizaban únicamente una proyección, especialmente en el cribado inicial (primer cribado). Nuestros resultados concuerdan con estudios previos en los que la doble proyección se asoció con un menor riesgo de falso-positivo por cualquier tipo de prueba (13), pero también encontramos que el uso de doble proyección aumentaba el riesgo de falso-positivo por pruebas invasivas. Una explicación para este hallazgo podría ser que las mujeres con una mamografía de cribado no negativa a las que se les hubiese realizado doble proyección podrían ser reconvocadas para realizarse pruebas adicionales invasivas para descartar malignidad sin la necesidad de realizar pruebas de imagen adicionales, lo que explicaría el aumento en la tasa de falsos positivos por pruebas invasivas. En nuestro estudio, no se encontraron diferencias en la detección de cáncer con el uso de doble o proyección única, pero estudios previos encontraron que la sensibilidad del test de cribado aumentaba cuando se utilizaba doble proyección en lugar de una (26). Observamos una mayor detección de cáncer y un menor riesgo de falso-positivo con mamografía digital. Mayores tasas de detección mediante mamografía digital han sido descritas en las mujeres más jóvenes (27;28), mientras que una reducción en el riesgo de falso-positivo por cualquier prueba ha sido encontrada en algunos estudios (29;30) pero no en otros (31).

Nuestros resultados sobre la influencia de las características de la mujer están en consonancia con las de estudios previos. El riesgo de falso-positivo es mayor en las mujeres más jóvenes, después de ajustar por el número de cribado (participación), lo que posiblemente refleja ciertas características relacionadas con la edad como la densidad mamaria, que no pudo ser estudiada ya que no se disponía de esta información. El THS no se asoció con un mayor riesgo de falso-positivo, lo que parece contradictorio dada la relación existente de este tratamiento con la densidad mamaria y el riesgo de cáncer. Sin embargo, este hallazgo podría ser explicado por la menor utilización de la combinación de estrógenos con progesterona – que está asociada con la densidad mamaria (32;33) – en España, en comparación con las recomendaciones realizadas en otros países europeos (34). Como era de esperar, las pruebas invasivas previas y los antecedentes familiares de cáncer de mama estuvieron asociados con un mayor riesgo de falso-positivo por cualquier tipo de prueba y por pruebas invasivas.

Se observó un amplio rango en el riesgo acumulado de falso-positivo estimado en función de los diferentes perfiles de riesgo, dependiendo de las características de la mujer y de las características del protocolo de lectura mamográfica. El riesgo de falso-positivo a lo largo de 10 participaciones para los perfiles de mayor y menor riesgo varió entre 51.4% y 7.5%, respectivamente (ratio máximo-mínimo: 6.8). Las variables de protocolo de lectura mamográfica, explicaban más de la mitad de este rango entre máximo y mínimo. Una proporción similar en rango (1.6% hasta 12.0%) se observó para pruebas invasivas (ratio máximo-mínimo: 7.6). Sin embargo, las características de la mujer jugaban aquí un papel más importante y explicaban el 73.3% de esta variabilidad. Es importante mencionar que las características asociadas a la mujer, excepto el uso de THS, no pueden modificarse, pero la evaluación de su impacto proporciona información esencial sobre el balance riesgo-beneficio del cribado mamográfico. Diferencias en el riesgo acumulado de falso-positivo entre diferentes perfiles de riesgo han sido estudiadas con anterioridad (11;17).

9.4 LIMITACIONES EPIDEMIOLÓGICAS

Como se ha comentado anteriormente en el apartado de limitaciones estadísticas, la información sobre las características personales de la mujer tiene datos faltantes ya que no siempre se disponía de esta información en algunas unidades radiológicas o no se recogía de forma rutinaria en otras. La distribución de edad entre los datos con y sin valores faltantes respecto a las variables de la mujer fue

razonablemente similar (ver anexos), sin embargo encontramos una tasa de falsos positivos por cualquier prueba moderadamente más baja y una tasa moderadamente más alta para falso-positivo por pruebas invasivas en los datos faltantes. Se analizó una submuestra con la máxima información disponible (N=2,777,429), lo que permitió evaluar las variables del protocolo de lectura y de características de la mujer de forma conjunta.

La información del método específico de doble lectura utilizado no se incluyó como una variable de estudio y no pudo ser evaluada, pero la mayoría de las unidades radiológicas realizaban doble lectura con consenso o arbitraje. No se disponía de información de la experiencia de los radiólogos dentro y fuera del programa de cribado. La experiencia del radiólogo es un factor importante para la precisión de la interpretación mamográfica (19) y podría afectar las tasas de falsos positivos. Las guías europeas recomiendan que los radiólogos lean al menos 5.000 mamografías al año y la mayoría de los radiólogos que leen en los programas de cribado participantes alcanzaban este número.

9.5 CONCLUSIÓN

Como conclusión, este estudio utiliza información de programas de cribado con diferentes protocolos de lectura mamográfica y diferentes etapas de desarrollo y experiencia de los mismos, siendo esta la mayor cohorte de mujeres cribadas estudiada hasta la fecha. Encontramos que las características del protocolo de lectura y propias de la mujer tienen un fuerte impacto en el riesgo acumulado a 10 rondas de falso-positivo por cualquier tipo de prueba y por pruebas invasivas. Entender las fuentes de variabilidad ayuda a mejorar la efectividad de los programas de cribado, pudiendo reducir los efectos adversos y proporcionando a las mujeres participantes información más precisa.

La utilización de modelos mixtos de riesgo a tiempo discreto fue de gran utilidad para atajar gran parte de las limitaciones encontradas hasta la fecha por los estudios que valoraban los efectos adversos del cribado. Su aplicabilidad en este contexto pone de manifiesto la utilidad de la estadística para sintetizar y resolver problemas complejos de otras disciplinas, como la sanitaria en este caso.

10 BIBLIOGRAFÍA

1. Ferlay, J., Parkin, D. M., and Steliarova-Foucher, E. Estimates of Cancer Incidence and Mortality in Europe in 2008. *Eur J Cancer*.2010.Mar.;46.(4):765.-81.Epub.2010.Jan.29. 2010;46:765-81.
2. Perry, N; Broeders, M; de Wolf, C; Törnberg, S; Holland, R; von Karsa, L. European guidelines for quality assurance in breast cancer screening and diagnosis. 2006. Luxemburg, Office for Official Publications of the European Communities.
3. Assembly of Members. CONSTITUTION AND BYLAWS of the European Breast Cancer Network. 2001.
4. Ascunce, N., Salas, D., Zubizarreta, R., Almazan, R., Ibanez, J., and Ederra, M. Cancer Screening in Spain. *Ann.Oncol*.2010.May.;21 Suppl 3:iii43.-51. 2010;21 Suppl 3:iii43-iii51.
5. Singer J; Willett JB. Fitting basic discrete-time hazard models. *Applied longitudinal data analysis: modelling change and event occurrence*. New York: Oxford University Press; 2003. pp.357-467.
6. Singer J; Willett JB. Describing Discrete-Time Event Occurrence Data. *Applied longitudinal data analysis: modelling change and event occurrence*. New York: Oxford University Press; 2003. pp.325-56.
7. Hosmer, DW and Lemeshow, S, *Applied logistic regression*. John Wiley and Sons; 2011.
8. Jorgensen, B, *The Theory of Dispersion Models (Monograph on Statistics & Applied Probability)*. Chapman & Hall; 2007.
9. Hofvind, S, Thoresen, S, and Tretli, S. The Cumulative Risk of a False-Positive Recall in the Norwegian Breast Cancer Screening Program. *Cancer*.2004.Oct.1;101.(7.):1501.-7. 1-10-2004;101:1501-7.
10. Smith-Bindman, R., Ballard-Barbash, R., Miglioretti, D. L., Patnick, J., and Kerlikowske, K. Comparing the Performance of Mammography Screening in the USA and the UK. *J Med.Screen*.2005.;12.(1):50.-4. 2005;12:50-4.
11. Christiansen, C. L., Wang, F., Barton, M. B., Kreuter, W., Elmore, J. G., Gelfand, A. E., and Fletcher, S. W. Predicting the Cumulative Risk of False-Positive Mammograms. *J Natl.Cancer Inst*.2000.Oct.18;92.(20.):1657.-66. 18-10-2000;92:1657-66.
12. Elmore, J. G., Barton, M. B., Mocerri, V. M., Polk, S., Arena, P. J., and Fletcher, S. W. Ten-Year Risk of False Positive Screening Mammograms and Clinical Breast Examinations. *N.Engl.J Med*. 16-4-1998;338:1089-96.
13. Elmore, J. G., Nakano, C. Y., Koepsell, T. D., Desnick, L. M., D'Orsi, C. J., and Ransohoff, D. F. International Variation in Screening Mammography Interpretations in Community-Based Programs. *J Natl.Cancer Inst*.2003.Sep.17;95.(18):1384.-93. 17-9-2003;95:1384-93.
14. Hofvind, S., Vacek, P. M., Skelly, J., Weaver, D. L., and Geller, B. M. Comparing Screening Mammography for Early Breast Cancer Detection in Vermont and Norway. *J Natl.Cancer Inst*.2008 Aug.6.;100.(15.):1082.-91.Epub.2008 Jul.29. 6-8-2008;100:1082-91.

15. Yankaskas, B. C., Klabunde, C. N., Ancelle-Park, R., Renner, G., Wang, H., Fracheboud, J., Pou, G., and Bulliard, J. L. International Comparison of Performance Measures for Screening Mammography: Can It Be Done? *J Med.Screen.*2004.;11(4):187.-93. 2004;11:187-93.
16. Xu, J. L., Fagerstrom, R. M., Prorok, P. C., and Kramer, B. S. Estimating the Cumulative Risk of a False-Positive Test in a Repeated Screening Program. *Biometrics.*2004.Sep.;60.(3):651.-60. 2004;60:651-60.
17. Castells, X., Molins, E., and Macia, F. Cumulative False Positive Recall Rate and Association With Participant Related Factors in a Population Based Breast Cancer Screening Programme. *J Epidemiol.Community Health.*2006.Apr;60.(4):316.-21. 2006;60:316-21.
18. Beam, C. A., Conant, E. F., and Sickles, E. A. Association of Volume and Volume-Independent Factors With Accuracy in Screening Mammogram Interpretation. *J Natl.Cancer Inst.*2003.Feb. 19-2-2003;95:282-90.
19. Esserman, L., Cowley, H., Eberle, C., Kirkpatrick, A., Chang, S., Berbaum, K., and Gale, A. Improving the Accuracy of Mammography: Volume and Outcome Relationships. *J Natl.Cancer Inst.*2002.Mar.6.;94.(5):369.-75. 6-3-2002;94:369-75.
20. Kan, L., Olivotto, I. A., Warren Burhenne, L. J., Sickles, E. A., and Coldman, A. J. Standardized Abnormal Interpretation and Cancer Detection Ratios to Assess Reading Volume and Reader Performance in a Breast Screening Program. *Radiology.*2000.May.;215.(2):563.-7. 2000;215:563-7.
21. Harvey, S. C., Geller, B., Oppenheimer, R. G., Pinet, M., Riddell, L., and Garra, B. Increase in Cancer Detection and Recall Rates With Independent Double Interpretation of Screening Mammography. *AJR Am.J Roentgenol.*2003.May.;180.(5):1461.-7. 2003;180:1461-7.
22. Kopans, D. B. Double Reading. *Radiol.Clin.North Am.*2000.Jul.;38.(4):719.-24. 2000;38:719-24.
23. Ciatto, S., Ambrogetti, D., Risso, G., Catarzi, S., Morrone, D., Mantellini, P., and Rosselli, Del Turco. The Role of Arbitration of Discordant Reports at Double Reading of Screening Mammograms. *J Med.Screen.*2005.;12.(3):125.-7. 2005;12:125-7.
24. Duijm, L. E., Groenewoud, J. H., Fracheboud, J., van Ineveld, B. M., Roumen, R. M., and de Koning, H. J. Introduction of Additional Double Reading of Mammograms by Radiographers: Effects on a Biennial Screening Programme Outcome. *Eur J Cancer.*2008 Jun.;44.(9):1223.-8.Epub.2008 Apr 8. 2008;44:1223-8.
25. Shaw, C. M., Flanagan, F. L., Fenlon, H. M., and McNicholas, M. M. Consensus Review of Discordant Findings Maximizes Cancer Detection Rate in Double-Reader Screening Mammography: Irish National Breast Screening Program Experience. *Radiology.*2009.Feb.;250.(2):354.-62. 2009;250:354-62.
26. Blanks, R. G., Given-Wilson, R. M., and Moss, S. M. Efficiency of Cancer Detection During Routine Repeat (Incident) Mammographic Screening: Two Versus One View Mammography. *J Med.Screen.* 1998;5:141-5.
27. Pisano, E. D., Gatsonis, C., Hendrick, E., Yaffe, M., Baum, J. K., Acharyya, S., Conant, E. F., Fajardo, L. L., Bassett, L., D'Orsi, C., Jong, R., and Rebner, M. Diagnostic Performance of Digital Versus Film Mammography for Breast-Cancer Screening. *N.Engl.J Med.*2005.Oct.27.;353.(17):1773.-83.Epub.2005.Sep.16. 27-10-2005;353:1773-83.

28. Pisano, E. D., Hendrick, R. E., Yaffe, M. J., Baum, J. K., Acharyya, S., Cormack, J. B., Hanna, L. A., Conant, E. F., Fajardo, L. L., Bassett, L. W., D'Orsi, C. J., Jong, R. A., Rebner, M., Tosteson, A. N., and Gatsonis, C. A. Diagnostic Accuracy of Digital Versus Film Mammography: Exploratory Analysis of Selected Population Subgroups in DMIST. *Radiology*.2008 Feb.;246.(2):376.-83. 2008;246:376-83.
29. Del Turco, M. R., Mantellini, P., Ciatto, S., Bonardi, R., Martinelli, F., Lazzari, B., and Houssami, N. Full-Field Digital Versus Screen-Film Mammography: Comparative Accuracy in Concurrent Screening Cohorts. *AJR Am.J Roentgenol*.2007.Oct.;189.(4):860.-6. 2007;189:860-6.
30. Sala, M., Comas, M., Macia, F., Martinez, J., Casamitjana, M., and Castells, X. Implementation of Digital Mammography in a Population-Based Breast Cancer Screening Program: Effect of Screening Round on Recall Rate and Cancer Detection. *Radiology*.2009.Jul.;252.(1):31.-9.Epub.2009.May.6. 2009;252:31-9.
31. Vigeland, E., Klaasen, H., Klingen, T. A., Hofvind, S., and Skaane, P. Full-Field Digital Mammography Compared to Screen Film Mammography in the Prevalent Round of a Population-Based Screening Programme: the Vestfold County Study. *Eur Radiol*.2008 Jan.;18(1):183.-91.Epub.2007.Aug.7. 2008;18:183-91.
32. Noh, J. J., Maskarinec, G., Pagano, I., Cheung, L. W., and Stanczyk, F. Z. Mammographic Densities and Circulating Hormones: a Cross-Sectional Study in Premenopausal Women. *Breast*.2006.Feb.;15.(1):20.-8.Epub.2005.Jul.5. 2006;15:20-8.
33. Topal, N. B., Ayhan, S., Topal, U., and Bilgin, T. Effects of Hormone Replacement Therapy Regimens on Mammographic Breast Density: the Role of Progestins. *J Obstet.Gynaecol.Res*.2006.Jun.;32.(3):305.-8. 2006;32:305-8.
34. Benet, Rodriguez M., Carvajal Garcia-Pando, A., Garcia, Del Pozo, Alvarez, Requejo A., and Vega, Alonso T. [Hormonal Replacement Therapy in Spain]. *Tratamiento Hormonal Sustitutivo En Espana. Un Estudio De Utilizacion. Med.Clin.(Barc.)*. 2002.Jun.8.;119.(1):4-8. 8-6-2002;119:4-8.

11 ANEXOS

ANÁLISIS DE LA PROBABILIDAD ACUMULADA DE UN FALSO POSITIVO EN EL CRIBADO MAMOGRÁFICO EN ESPAÑA EN LA INDICACIÓN DE ALGUNA EXPLORACIÓN ADICIONAL

PROTOCOLO DE LAS VARIABLES DE ESTUDIO

CRITERIOS DE INCLUSIÓN

1. Los programas disponen de información de al menos tres rondas de cribado consecutivas (e incluye la primera ronda).
2. Los programas disponen de un identificador único por mujer que permite enlazar (linkar) a la misma mujer a través de las distintas rondas de cribado (como si de una cohorte se tratase). En estas condiciones, para cada mujer y ronda, disponen de información sobre el resultado de la mamografía de cribado y se dispone de información respecto a las pruebas adicionales tanto invasivas como no invasivas.
3. Se incluirá en la base de datos a todas las mujeres incluidas en el programa (de cualquier edad en la fecha de exploración) con al menos 1 cribado (al menos una vez participante), pudiendo tener su primer cribado (cribado inicial) en cualquier ronda del programa.
4. Se incluirá a las mujeres con fecha de mamografía de cribado desde el inicio del programa hasta el 31 de Diciembre de 2006. Se registrarán las pruebas adicionales realizadas desde el inicio del programa hasta 30 de Junio de 2007
5. El estudio de los *cánceres de intervalo / falsos negativos* está fuera del marco de análisis de este estudio y no se contemplarán.

DESCRIPCIÓN DE LAS VARIABLES (VER CATEGORÍAS EN EL ANEXO 1)

Con el fin de mejorar la comprensión de las variables que vamos a recoger para el *análisis de la probabilidad acumulada de al menos un falso positivo en el cribado mamográfico en España en la indicación de alguna exploración adicional invasiva (PAAF, core-biopsia o biopsia quirúrgica)* proponemos el siguiente ejemplo:

IMAGINEMOS UN PROGRAMA DE CATALUNYA (CIUTAT VELLA) CON CINCO RONDAS (VUELTAS O CAMPAÑAS) Y UNA MUJER NACIDA 15/01/1948 Y CONVOCADA A PARTIR DE LA SEGUNDA (FECHA PRIMERA CONVOCATORIA 01/03/1998) Y HASTA LA QUINTA RONDA. SUPONGAMOS QUE LA MUJER HA PARTICIPADO EN LA SEGUNDA, CUARTA Y QUINTA RONDA DEL PROGRAMA.

SUPONGAMOS QUE SU TERCERA MAMOGRAFÍA DE CONTROL EN EL PROGRAMA (RONDA 5 DEL PROGRAMA) SE REALIZA EL 23/04/04. TRAS LA LECTURA DE LOS RADÍOLOGOS NO ES POSIBLE DESCARTAR MALIGNIDAD (MAMOGRAFÍA UNA PROYECCIÓN Y DOBLE LECTURA). EN FECHAS POSTERIORES A LA MAMOGRAFÍA SE LE REALIZAN CONSECUTIVAMENTE LAS SIGUIENTES PRUEBAS: ECOGRAFÍA, PAAF Y BIOPSIA QUIRÚRGICA. FINALMENTE NO SE DIAGNOSTICA CÁNCER DE MAMA PERO SE LE RECOMIENDA A LA MUJER LA REALIZACIÓN DE UNA MAMOGRAFÍA INTERMEDIA A LOS 6 MESES. PASADOS LOS 6 MESES (15/11/04), SE LE REALIZA LA MAMOGRAFÍA Y JUSTO DESPUÉS DE LA MAMOGRAFÍA (EN ESA MISMA FECHA), SE LE REALIZA UNA ECOGRAFÍA Y UNA BIOPSIA ESCISIONAL PERCUTÁNEA, CON UN RESULTADO DEFINITIVO HISTOLÓGICO DE CÁNCER DE MAMA.

1) CCAA (1 registro por CCAA)

- *CCAA_id:*

Código que se asignará a cada una de las Comunidades Autónomas participantes en el proyecto. Se enumerarán sucesivamente.

- *Nombre de la CCAA:*

Descripción (etiqueta) de la Comunidad Autónoma.

EJEMPLO:

CATALUNYA

2) Unidad Radiológica de Cribado (1- N registros por CCAA / 1 registro por episodio)

- *URC_id:*

Código de cada una de las Unidades de Cribado donde se realizan las exploraciones. Se entenderá como URCs, unidades organizativas independientes dentro de una misma CCAA, sean estas fijas o móviles. Se van a enumerar las URC de forma sucesiva a partir del número 1, sin ningún orden específico. En el caso de CCAA sin estructura de URC bastará con poner 1 en esta variable y el nombre de la CCAA en la descripción de la URC.

- *Descripción de la URC:*

Nombre y/o descripción (etiqueta) de la Unidad de Cribado donde se realizan las exploraciones. Las CCAA sin estructura de URC la etiquetarán con el nombre de la propia CCAA. Aquellas CCAA que no identifiquen las URC, las etiquetaran con el nombre de la Comunidad seguido del número de URC_id que le hayan asignado (Comunidad1, Comunidad2, ..., ComunidadN)

3) Mujeres (1 registro por mujer)

- *Mujer_id:*

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

- *Fecha de nacimiento:*

Fecha de nacimiento de la mujer.

EJEMPLO:

15/01/1948

- *Fecha de la 1ª citación en el programa:*

Fecha en la cual la mujer es invitada a participar por 1ª vez en la URC (independientemente de si participa o no).

09/09/9999 si desconocido

EJEMPLO:

01/03/1998

4) Episodios (1 episodio por ronda del programa / 1-N registros por mujer)

- *Episodio:*

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

EJEMPLO:

LA BASE DE DATOS DE EPISODIOS CONTENDRÍA TRES REGISTROS PARA ESTA MUJER QUE SERÍAN RELLENADOS CON LOS NÚMEROS 1,3 Y 4 CONSECUTIVAMENTE.

- *Número de cribados:*

Número ordinal que indica el número de exploraciones de cribado que una mujer lleva realizadas hasta el episodio (incluyendo la del episodio).

EJEMPLO:

EL CAMPO NÚMERO DE CRIBADOS SERÍA RELLENADO CON LOS NÚMEROS 1, 2 Y 3 CONSECUTIVAMENTE PARA CADA UNO DE LOS TRES EPISODIOS.

- *Número de ronda de la URC:*

Identifica el número de ronda (vuelta, campaña) del programa en cada episodio.

EJEMPLO:

EL CAMPO NÚMERO DE RONDA DE LA URC SERÍA RELLENADO CON LOS NÚMEROS 2, 4 Y 5 CONSECUTIVAMENTE PARA CADA UNO DE LOS TRES EPISODIOS.

- *Fecha mamografía de cribado:*

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

EJEMPLO:

23/04/04.(EN EL EPISODIO CODIFICADO COMO 4).

- *Lector_id1:*

Identificador del radiólogo lector número 1 de la mamografía de cribado.

- *Lector_id2:*

Identificador del radiólogo lector número 2 de la mamografía de cribado.

- *Lector_id3:*

Identificador del radiólogo lector número 3 de la mamografía de cribado (Ej: el de arbitraje).

- *Método de lectura:*

Indicación del tipo de lectura en el episodio para una mujer.

EJEMPLO:

DOBLE LECTURA (EN EL EPISODIO CODIFICADO COMO 4).

- *Nº de proyecciones:*

Indicación del número de proyecciones en el episodio para una mujer.

EJEMPLO:

UNA PROYECCIÓN (EN EL EPISODIO CODIFICADO COMO 4).

- *Número de estudios intermedios:*

Número total de estudios con mamografía, u otras pruebas, intermedias dentro de un episodio, realizados por indicación del programa independientemente de cuando y dónde se realicen. Si dentro del episodio no se ha realizado ningún estudio intermedio, el valor de la variable será 0 (cero).

EJEMPLO:

1 ESTUDIO INTERMEDIO (EN EL EPISODIO CODIFICADO COMO 4).

- *Tipo de mamografía:*

Permitirá recoger la información sobre si la mamografía realizada es convencional o digital (variable ligada al proyecto de la mamografía digital).

- *Resultado inicial de la mamografía de cribado*

Esta variable, juntamente con el resultado final del episodio, va a permitir evaluar el riesgo acumulado de falsos positivos de la mamografía como test de cribado.

Se indicará el resultado inicial de la mamografía de cribado para ese episodio, aunque posteriormente, y durante el mismo episodio, haya otros resultados de otros tests (mamografías intermedias, pruebas de imagen o exploraciones adicionales invasivas). En el caso de repetición técnica, se tomará como inicial la técnica.

Se codificará de la siguiente manera:

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

EL CAMPO RESULTADO (INICIAL) DE LA MAMOGRAFÍA SERÍA RELLENADO CON LA ETIQUETA: EXPLORACIONES ADICIONALES (EN EL EPISODIO CODIFICADO COMO 4).

- *Resultado final del episodio:*

El resultado final del episodio es el diagnóstico definitivo del mismo, tenga o no tenga exploraciones adicionales, tenga o no tenga mamografías intermedias, sea cual sea la casuística durante el episodio. Se codificará de la siguiente manera:

- *1: Cáncer: el diagnóstico definitivo (histológico) del episodio es de cáncer de mama. No se incluirán los lobulares in situ.*

- *2: No cáncer: el resultado del episodio es de no cáncer de mama y se recomienda a la mujer un control rutinario.*

- *3: Pendiente de exploraciones adicionales: mujeres con episodios incompletos. Como se incluirá a las mujeres desde el inicio del programa hasta el 31 de Diciembre de 2006 es posible que durante el último período haya mujeres que están todavía pendientes de la realización (o de los resultados) de pruebas adicionales y no se pueda establecer un resultado final del episodio.*

- *4: Pendiente estudios intermedios: mujeres con episodios incompletos. Como se incluirá a las mujeres desde el inicio del programa hasta el 31 de Diciembre de 2006 es posible que durante el último período haya mujeres que están todavía pendientes de la realización de un estudio intermedio y no se pueda establecer un resultado final del episodio.*

- *5: Seguimiento incompleto: mujeres a las que no se puede completar las exploraciones recomendadas en el episodio (pruebas adicionales, exploraciones intermedias, etc) por ejemplo por rechazo a las mismas por parte de la mujer, fallecimiento antes de completar proceso diagnóstico etc.*

- 99: Desconocido: mujeres de las que no se conoce el resultado final del episodio (cáncer o no cáncer).

En los programas de detección precoz con un circuito que no permita tener la información de todas la pruebas adicionales (por ejemplo, en el caso de un programa que tiene la información de algunas pruebas, pero la mujer puede ser derivada en algún momento a seguimiento hospitalario, momento a partir del cual la información sobre las pruebas que se realizan se desconocen), se procederá de la siguiente manera:

- a. Si se conoce el resultado final del episodio (Cáncer / No cáncer),***
 - *Se indicará dicho resultado.*
 - *Se informará en la base de datos de todas las pruebas adicionales que sean conocidas para este episodio si existen (invasivas y no invasivas).*
 - *Se considerará el proceso de seguimiento hospitalario como un estudio intermedio*
 - *Puesto que no todas las pruebas adicionales son conocidas (por derivación a seguimiento hospitalario):*
 - i. Si el resultado final del episodio es negativo (No cáncer): Se contabilizará una exploración adicional que se codificará como “prueba desconocida realizada durante el seguimiento hospitalario”. Se reserva la codificación 88 para estos casos particulares. (Prueba_id = 88). El resultado de esta prueba desconocida será siempre 99 (desconocido).*
 - ii. Si el resultado final del episodio es positivo (Cáncer) se contabilizará una “prueba invasiva desconocida” (2.9) con resultado positivo.*
- b. Si no se conoce el resultado final del episodio éste será clasificado como 99 (Desconocido). En dicho caso, si existen, se informará en la base de datos de todas las pruebas adicionales conocidas.***

En el caso de mujeres pendientes de alguna exploración adicional derivada de una mamografía intermedia se codificará como “3: Pendiente de exploraciones adicionales”.

EL CAMPO RESULTADO FINAL DEL EPISODIO NÚMERO 4 SERÍA RELLENADO CON LA ETIQUETA: CÁNCER.

5) Exploraciones adicionales (0-N registros por episodio)

- *N prueba:*

Es el contador del número de pruebas realizadas a una misma mujer dentro de cada episodio. Para cada prueba realizada a la mujer se le asignará el número de prueba que le corresponde.

- *Prueba_id:*

Se recogerán todas las pruebas realizadas en cada episodio aunque sean del mismo tipo (Ej: dos PAAF). En el caso de una mujer-episodio con pruebas en ambas mamas en la misma fecha, se registrará únicamente la más maligna.

Se codificará de la siguiente manera:

1. Pruebas no Invasivas

1.1. Otras proyecciones mamográficas

1.2. Ecografía

1.3. Resonancia magnética

1.4. Otras pruebas no invasivas

1.9. Prueba no invasiva desconocida

2. Pruebas Invasivas

2.1. PAAF

2.2. Biopsia aguja gruesa

2.3. Biopsia asistida por vacío

2.4. Biopsia escisional percutánea

2.5. Biopsia quirúrgica

2.6. Otras pruebas invasivas

2.9. Prueba invasiva desconocida

99. Desconocido

Además, se incluye la siguiente codificación para aquellos programas en que la mujer puede ser derivada a seguimiento hospitalario y se desconocen todas las pruebas realizadas.

88. Prueba desconocida realizada durante el seguimiento hospitalario

DENTRO DEL EPISODIO NÚMERO 4 (TERCER CRIBADO DE LA MUJER, RONDA 5 DEL PROGRAMA), EN LA BASE DE DATOS DE EXPLORACIONES ADICIONALES CONSTARÍAN LAS SIGUIENTES 5 PRUEBAS: ECOGRAFÍA, PAAF, BIOPSIA QUIRÚRGICA, ECOGRAFÍA Y BIOPSIA ESCISIONAL PERCUTÁNEA.

• *Fecha_prueba:*

09/09/9999 si desconocido.

• *Resultado prueba:*

Se incluirá el resultado de cada prueba.

La categoría no concluyente se reserva únicamente para pruebas invasivas no valorables y/o resultado no concluyente.

Se codificará de la siguiente manera:

1. Positiva, sospecha o certeza de malignidad

2. Negativa

3. No concluyente

4. Pendiente

99. Desconocido: sólo en el caso de que se confirme que la prueba se ha realizado pero no se conoce el resultado

SI SUPONEMOS QUE SE REALIZARON LAS 5 PRUEBAS ANTERIORES, ÉSTAS SERÍAN CLASIFICADAS CONSECUTIVAMENTE COMO: SOSPECHOSA, NEGATIVA (PODRÍA SER NO CONCLUYENTE), NEGATIVA (PODRÍA SER NO CONCLUYENTE), SOSPECHOSA, POSITIVA.

6) Variables de las mujeres (1 registros por episodio)

El objetivo es obtener información orientativa sobre el perfil de la mujer que se realiza la exploración de cribado. Puesto que algunas de las características de las mujeres pueden variar en cada ronda de cribado, las siguientes variables se recogerán para cada episodio:

- *THS*

Se evaluará el uso del THS en el momento de hacerse la exploración de cribado. Se considerará que la mujer usaba THS si lo tomaba en el momento de la mamografía de cribado o en los 6 meses anteriores.

- *Soja:*

Alimentación rica en Soja. Se considerará que la mujer usaba Soja si lo tomaba en el momento de la mamografía de cribado o en los 6 meses anteriores

- *Menopausia:*

Se considerarán dos niveles: posmenopáusica y, premenopáusica o perimenopáusica.

- *Talla*

Talla en centímetros.

- *Peso*

Peso en kilogramos.

- *Antecedentes personales de patología mamaria benigna (inespecífica):*

Se considerará que una mujer tiene antecedentes personales de patología mamaria benigna, cuando se tenga conocimiento de que ha sufrido alguna de las patologías mamarias benignas posibles, de forma previa a la exploración de cribado.

- *Antecedentes personales de prueba invasiva con resultado benigno:*

Cuando se tenga conocimiento de que la mujer ha sufrido alguna prueba invasiva con resultado benigno, previa a la exploración de cribado.

- *Antecedentes familiares de cáncer de mama:*

Se considerará que una mujer tiene antecedentes familiares de cáncer de mama cuando tiene algún familiar de primer grado afectado (madre, hermanas o hijas) con cáncer de mama.

CCAA (1 registro por CCAA)	Codificación
CCAA_id	Código asignado a cada CCAA
Nombre de la CCAA	Nombre de la CCAA
URC (1- N registros por CCAA, 1 registro por episodio)	Codificación
URC_id	Código asignado a cada Unidad Radiológica de Cribado donde se realizan las exploraciones
Descripción URC	Descripción identificativa de la Unidad Radiológica de Cribado
Mujeres (1 registros por mujer)	Codificación
CCAA_id	
Mujer_id	Número identificador interno del programa
Fecha de nacimiento	09/09/9999 Desconocido
Fecha 1a citación en el programa	09/09/9999 Desconocido
Episodios (1 registro por cribado de la mujer, 1-N registros por mujer)	Codificación
CCAA_id	
Mujer_id	
Episodio	1,2,3,4,...
URC_id	Código asignado a cada Unidad Radiológica de Cribado donde se realizan las exploraciones

Descripción de la URC	Nombre de la URC
Número de cribados	1,2,3,4,...
Número de ronda de la URC	1,2,3,4...
Fecha mamografía de cribado	09/09/9999 Desconocido
Lector_id1	Identificador del radiólogo; 99: Desconocido
Lector_id2	Identificador del radiólogo; 99: Desconocido; [Vacío si no es pertinente]
Lector_id3	Identificador del radiólogo; 99: Desconocido; [Vacío si no es pertinente]
Método de lectura	1: Lectura simple; 2: Doble lectura sin consenso (resultado más desfavorable); 3: Doble lectura con consenso (deciden ambos radiólogos); 4: Doble lectura con arbitraje; 5: Doble lectura con arbitraje a ciegas; 99: Desconocido
Nº de proyecciones	1,2
Nº de estudios intermedios	0,1,2,3,4,...
Tipo de mamografía	1: Mamografía convencional; 2: Digital indirecta / lectura placa; 3: Digital indirecta / lectura monitor; 4: Digital directa / lectura placa; 5: Digital directa / lectura monitor; 6: Otros; 99: Desconocido
Resultado inicial de la mamografía de cribado	1: Negativo; 2: Exploraciones adicionales; 3: Mamografía intermedia (Estudios intermedios)

	99:Desconocido
Resultado final del episodio	1: Cáncer; 2: No cáncer; 3: Pendiente de exploraciones adicionales; 4: Pendiente estudios intermedios; 5: Seguimiento incompleto; 99: Desconocido
Exploraciones adicionales (0-N registros por episodio)	Codificación
CCAA_id	
Mujer_id	
Episodio	1,2,3,4,...
N_Prueba	1,2,3,4,...
Prueba_id	1.1. Otras proyecciones mamográficas; 1.2. Ecografía; 1.3. Resonancia magnética; 1.4. Otras pruebas no invasivas; 1.9. Prueba no invasiva desconocida; 2.1. PAAF; 2.2. Biopsia aguja gruesa; 2.3. Biopsia asistida por vacío; 2.4. Biopsia escisional percutánea; 2.5. Biopsia quirúrgica; 2.6. Otras pruebas invasivas 2.9. Prueba invasiva desconocida 99. Desconocida prueba 88: Prueba desconocida seguimiento hospitalario
Fecha_prueba	09/09/9999 si desconocido
Resultado prueba	1:Sospecha o certeza de malignidad; 2: Negativo; 3: No concluyente; 4: Pendiente; 99: Desconocido

Variables de las mujeres (1 registro por episodio)	Codificación
<p>CCAA_id</p> <p>Mujer_id</p>	
Episodio	1,2,3,4,...
THS	1: Sí, en el momento del cribado o en los 6 meses anteriores; 2: No; 99: Desconocido
Soja	1: Sí, en el momento del cribado o en los 6 meses anteriores; 2: No; 99: Desconocido
Menopausia	1: Posmenopáusica; 2: Premenopáusica o perimenopáusica; 99: Desconocido
Talla	Talla en cm; 99: Desconocido
Peso	Peso en Kg; 999: Desconocido
Antecedentes personales de patología mamaria benigna (inespecífica)	1: Sí; 2: No; 99: Desconocido
Antecedentes personales de prueba invasiva con resultado benigno	1: Sí; 2: No; 99: Desconocido
Antecedentes familiares de cáncer de mama (madre, hermanas e hijas)	1: Sí; 2: No; 99: Desconocido

ANÁLISIS DE LA PROBABILIDAD ACUMULADA DE UN FALSO POSITIVO EN EL CRIBADO MAMOGRÁFICO EN ESPAÑA EN LA INDICACIÓN DE ALGUNA EXPLORACIÓN ADICIONAL

PROGRAMA DE CONTROL DE CALIDAD DE LA BASE DE DATOS

RANGO DE VALORES

Se considera que un valor está fuera de rango, y por lo tanto es un valor imposible, cuando no está dentro del conjunto de valores predeterminados definidos para esa variable.

A continuación definimos el rango de valores para cada variable.

- *CCAA_id*

Campo definido como texto de longitud 2. Se codificará con un identificador del 1 hasta el 9 según la CCAA:

<u>CCAA_id</u>	<u>Nombre CCAA</u>
1	Asturias
2	Canarias
3	Castilla y León
4	Catalunya
5	Galicia
6	La Rioja
7	Navarra
8	Valencia

- *Nombre de la CCAA*

Campo definido como texto de longitud 25. Ver equivalencias en la definición de la anterior variable.

- *URC_id*

Campo definido como numérico de longitud 2. Se asignará un único valor 1 a esta variable si no existe estructura de URCs (En este caso se asignará el nombre de la CCAA en la descripción de la URC).

- *Descripción de la URC*

Campo definido como texto de longitud 25. Puede ser introducido como un valor numérico, como un valor alfanumérico, o como una combinación de ambos. En el caso de que únicamente exista una URC_id para toda la CCAA, rellenar este campo con el nombre de la CCAA.

- *Mujer_id*

Campo definido como texto de longitud 15. Puede ser introducido como un valor numérico, un valor alfanumérico, o una combinación de ambos.

- *Fecha de nacimiento*

Campo definido como fecha de longitud 10, con formato 00/00/0000 correspondiente a día/mes/año (09/09/9999 si desconocido). No se añadirán restricciones de valor para esta variable, pero sí se incluirán valores poco probables.

- *Fecha de la 1ª citación en el programa*

Campo definido como fecha de longitud 10, con formato 00/00/0000 correspondiente a día/mes/año. Esta variable toma valores comprendidos entre 01/01/1989 y 31/12/2006 ó bien 09/09/9999 si es desconocido.

- *Episodio*

Campo definido como numérico de longitud 2. Será menor de 20

- *Número de cribados*

Campo definido como numérico de longitud 2. Será menor de 20

- *Número de ronda de la URC*

Campo definido como numérico de longitud 2. Esta variable toma valores en el rango 1 - 9 ó 99 si es desconocido.

- *Fecha mamografía de cribado*

Campo definido como fecha de longitud 10, con formato 00/00/0000 correspondiente a día/mes/año. Su rango de valores está comprendido entre 01/01/1989 y 31/12/2006 ó bien 09/09/9999 si es desconocido.

- *Lector_id1*

Campo definido como texto de longitud 50. Puede ser introducido con un valor numérico, con un valor alfanumérico, o con una combinación de ambos.

- *Lector_id2*

Campo definido como texto de longitud 50. Puede ser introducido con un valor numérico, con un valor alfanumérico, o con una combinación de ambos. Si no es pertinente, no se informará y se dejará su correspondiente campo 'vacío'.

- *Lector_id3*

Campo definido como texto de longitud 50. Puede ser introducido con un valor numérico, con un valor alfanumérico, o con una combinación de ambos. Si no es pertinente, no se informará y se dejará su correspondiente campo 'vacío'.

- *Método de lectura*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Lectura simple;
- 2: Doble lectura sin consenso (resultado más desfavorable);
- 3: Doble lectura con consenso (deciden ambos radiólogos);
- 4: Doble lectura con arbitraje;
- 5: Doble lectura con arbitraje a ciegas;
- 99: Desconocido

- *Nº de proyecciones*

Campo definido como numérico de longitud 2. Los únicos valores posibles son el 1, el 2, y el 99 si es desconocido.

- *Número de estudios intermedios*

Campo definido como numérico de longitud 2.

- *Tipo de mamografía*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Mamografía convencional;
- 2: Digital indirecta / lectura placa;
- 3: Digital indirecta / lectura monitor;
- 4: Digital directa / lectura placa;
- 5: Digital directa / lectura monitor;
- 6: Otros;
- 99: Desconocido

- *Resultado inicial de la mamografía de cribado*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Negativo;
- 2: Exploraciones adicionales;
- 3: Mamografía intermedia;
- 99: Desconocido

- *Resultado final del episodio*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Cáncer;
- 2: No cáncer;
- 3: Pendiente de exploraciones adicionales;
- 4: Pendiente mamografía intermedia;
- 5: Seguimiento incompleto;
- 99: Desconocido

- *N prueba*

Campo definido como numérico de longitud 2.

- *Prueba_id*

Campo definido como texto de longitud 2. Ver equivalencias:

- 11: Otras proyecciones mamográficas;
- 12: Ecografía;
- 13: Resonancia magnética;
- 14: Otras pruebas no invasivas;
- 19: Prueba no invasiva desconocida;
- 21: PAAF;
- 22: Biopsia aguja gruesa;
- 23: Biopsia asistida por vacío;
- 24: Biopsia escisional percutánea;
- 25: Biopsia quirúrgica;
- 26: Otras pruebas invasivas;
- 29: Prueba invasiva desconocida;
- 99: Desconocido
- 88: Prueba desconocida realizada durante el seguimiento hospitalario

- *Fecha_prueba*

Campo definido como fecha de longitud 10, con formato 00/00/0000 correspondiente a día/mes/año (09/09/9999 si desconocido). Su rango de valores está comprendido entre 01/01/1989 y 30/06/2007 o bien 09/09/9999 si es desconocido.

Para aquellas excepciones en que por un error en la información la *'Fecha de prueba'* sea anterior a la *'Fecha de mamografía de cribado'* la "Fecha de prueba" deberá tener el valor 09/09/9999 "Desconocido".

- *Resultado prueba*

Campo definido como texto de longitud 2. Ver equivalencias:

En el caso de realización de prueba invasiva, al menos una de las pruebas no invasivas del episodio debe estar clasificada como: positiva, sospecha de malignidad o certeza de malignidad.

- 1: Sospecha o certeza de malignidad;
- 2: Negativo;
- 3: No concluyente;
- 4: Pendiente;
- 99: Desconocido

- *THS*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Sí, en el momento de la mamografía o en los 6 meses previos;
- 2: No;
- 99: Desconocido

- *Soja:*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Sí, en el momento del cribado o en los 6 meses previos;
- 2: No;
- 99: Desconocido

- *Menopausia:*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Posmenopáusica;
- 2: Premenopáusica o perimenopáusica;
- 99: Desconocido

- *Antecedentes personales de patología mamaria benigna (inespecífica).*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Sí;
- 2: No;
- 99: Desconocido

- *Antecedentes personales de prueba invasiva con resultado benigno.*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Sí;
- 2: No;
- 99: Desconocido

- *Antecedentes familiares de cáncer de mama*

Campo definido como texto de longitud 2. Ver equivalencias:

- 1: Sí;
- 2: No;
- 99: Desconocido

INCOHERENCIAS

1. Comprobar que los códigos para la variable “CCAA_id” se corresponden con los de la variable “Nombre de la CCAA” tal y como sigue:

1	Asturias
2	Canarias
3	Castilla y León
4	Catalunya
5	Galicia
6	La Rioja
7	Navarra
8	Valencia

2. No pueden existir dos mujeres o más con el mismo código de “Mujer_id” para una determinada CCAA.

3. La variable “Fecha de nacimiento” tendrá formato 00/00/0000 en el orden día / mes / año y NO mes / día / año. Así mismo, el año se especificará con 4 dígitos.

4. La “Fecha de la 1ª citación en el programa” no puede ser un campo vacío. Ha de cumplir con la especificación correcta de fecha, tal y como se definió anteriormente, y además estará dentro del rango de valores predefinidos. En el caso de una mujer participante sin invitación previa (por los motivos que sea), este campo se rellenará con la fecha de realización de la primera mamografía.

5. La variable “Episodio” será un número natural estrictamente mayor o igual que 1. Es ordinal pero puede tener saltos (1,3,4...). Este campo no puede estar vacío para un registro existente. Comprobar que para una misma CCAA_id y Mujer_id no existan dos episodios iguales.

6. La variable “Número de cribados” es un contador ordinal, representado por un número natural. Empieza con el 1 y no tiene saltos (1,2,3,4...). Este campo no puede estar vacío para un episodio existente. Es un número igual o menor al número de “Episodio”. Cuando se introduce un nuevo episodio con mamografía reliazada, la variable “Número de cribados” incrementa necesariamente en uno el contador.

7. El “Número de ronda de la URC” es un número natural estrictamente mayor o igual que 1. En el caso de programas para el cual la ronda de cribado se asigna por municipios (no por Unidades de Exploración), se asignará el número de ronda del municipio en que reside la mujer.

8. La variable “Fecha mamografía de cribado” ha de ser igual o posterior a la “Fecha de 1ª citación en el programa” y además igual o inferior a “Fecha_prueba”. Tiene formato 00/00/0000 en el siguiente orden: día / mes / año. Así mismo, el año se especificará con 4 dígitos.

9. Para un mismo episodio, el “lector_id1” es distinto al “lector_id2”. Si para un mismo episodio, las variables “lector_id1” y “lector_id2” no son nulos, la variable “Método de lectura” no puede ser codificada con un 1 (lectura simple).

10. La variable “Tipo de mamografía” no puede estar vacía para un episodio concreto.

11. Si la variable “Resultado inicial de la mamografía de cribado” es igual a 1 (Negativo) para un episodio concreto, la variable “Prueba_id” quedará en blanco, es decir, no pueden existir pruebas para una mujer que tenga un resultado inicial de la mamografía negativo.

12. El “Resultado inicial de la mamografía de cribado” no puede ser un campo vacío para un episodio existente. Si el resultado inicial de la mamografía de cribado para un episodio concreto es igual a 1 (Negativo), la variable “Resultado final del episodio” será codificada necesariamente con el 2 (No cáncer).

13. Para un Mujer_id y episodio concreto, si la variable “Resultado final del episodio” se codifica como cáncer, no puede existir para esa misma mujer un episodio posterior. En el caso excepcional dónde se conozca el resultado final de cáncer una vez que se ha vuelto a invitar a la mujer y se ha realizado un nuevo cribado, esta mamografía se eliminará de la base de datos.

14. Si TODAS las pruebas no invasivas (códigos: 11, 12, 13, 14, o 19 de la variable Prueba_id) correspondientes a un episodio, tienen como valor en la variable “Resultado prueba” 2 (negativo), no debe haber ninguna prueba invasiva en el episodio.

Se puede dar el caso de episodios con ‘mamografía de cribado’ no negativa, resultado de TODAS las pruebas no invasivas 2 (negativo) y que posteriormente tengan una prueba invasiva. A pesar de que esta regla de validación se mantendrá como alerta de control para posibles errores de la base de datos queda registrado que esta casuística es factible

15. Si la variable "Prueba_id" está codificada como una prueba de imagen (códigos: 11, 12, 13, 14, o 19) y la variable "Resultado prueba" es 1 (sospecha o certeza de malignidad), debe cumplirse al menos una de las condiciones siguientes:

- a)** Existe alguna prueba adicional invasiva en el episodio.
- b)** El número de controles avanzados en ese episodio es 1 o más.
- c)** El resultado final del episodio es 5 (Seguimiento incompleto) o 3 (Pendiente de exploraciones adicionales).
- d)** Existe alguna otra prueba adicional no invasiva (de cualquier tipo) en el episodio, con resultado negativo.

16. Si la variable “Resultado prueba” es codificada con un 3 (No concluyente) implica que la variable “Prueba_id” tiene que haber sido codificada necesariamente con un 21, ó 22, ó 23, ó 24, ó 25, ó 26, ó 29 (PAAF, BAG, BAV, BEP, Otra prueba invasiva). No puede haber un resultado no concluyente en una prueba no invasiva.

*Dado que no se admite el resultado “no concluyente” en **pruebas no invasivas**, el resultado de las **pruebas no invasiva** en el caso en que remiten a la mujer a un control avanzado, y estuviese codificada como no concluyente, se codificará como **2. Negativo***

17. Si la variable “Resultado final del episodio” es codificada con un 3 (pendiente de exploraciones adicionales) implica que habrá por lo menos una prueba con “Resultado prueba” codificado como 4 (pendiente).

18. La variable “Menopausia” no puede ser codificada como 1 (posmenopáusica) en un episodio y en cualquier otro episodio posterior como 2 (premenopáusica o perimenopáusica).

19. La variable “Antecedentes personales de patología mamaria benigna (inespecífica)” si en un episodio se clasifica como 1 (Sí), en los subsiguientes no puede clasificarse como 2 (No). Debe seguir siendo clasificada como 1.

20. La variable “Antecedentes personales de prueba invasiva con resultado benigno” si en un episodio se clasifica como 1 (Sí), en los subsiguientes no puede clasificarse como 2 (No). Debe seguir siendo clasificada como 1.

21. La variable “Antecedentes familiares de cáncer de mama” si en un episodio se clasifica como 1 (Sí), en los subsiguientes no puede clasificarse como 2 (No). Debe seguir siendo clasificada como 1.

VALORES POCO PROBABLES

Los valores poco probables se definen para que la base de datos dé una señal de alerta indicando que probablemente existe una incoherencia.

1. Fecha de nacimiento

Se consideran valores poco probables de esta variable los que estén fuera del intervalo comprendido entre 01/01/1919 y 31/12/1962.

2. El número de cribados de la mujer no debería ser un número superior al número de ronda de la URC.

3. Número de estudios intermedios ≥ 4

4. N prueba ≥ 10 por episodio

5. La talla se recogerá en centímetros, así pues, es de esperar que sea un número que esté aproximadamente entre 120 y 200.

6. El peso se recogerá en kilogramos, así pues, es de esperar que sea un número que esté aproximadamente entre 40 y 120.

Para validar definitivamente la base de datos se sugiere la realización de "frecuencias" de cada variable. Los resultados de las frecuencias dependerán de los protocolos de cada CCAA, por lo que, de momento, no se propone una validación específica.

VALORACIÓN DE DATOS FALTANTES

Para evaluar las posibles diferencias entre la población inicial del estudio y el subconjunto de registros con información completa de variables de la mujer, se compararon las tasas de falso positivo y la distribución de edad entre los registros con datos faltantes y los registros con información completa para cada una de las variables de la mujer.

Apéndices Tabla 1. Tasa de falsos positivos entre población con datos faltantes y población con información completa sobre las variables de la mujer

		falsos positivos (cualquier prueba) %	falsos positivos (pruebas invasivas) %
THS	Datos completos	5.84%	0.525%
	Datos faltantes	5.01%	0.493%
Menopausia	Datos completos	5.89%	0.511%
	Datos faltantes	5.08%	0.523%
Pruebas invasivas Previas	Datos completos	5.96%	0.441%
	Datos faltantes	4.99%	0.634%
Antecedentes familiares Cáncer	Datos completos	5.88%	0.529%
	Datos faltantes	4.95%	0.487%

Apéndices Tabla 2. Riesgo de falso positivo para la base de datos completa y para la submuestra con información completa sobre variables de la mujer

	falsos positivos (cualquier prueba) %	falsos positivos (pruebas invasivas) %
Base datos completa (N= 4,739,498)	5.59%	0.5156%
Base datos submuestra (N= 2,777,429)	5.08%	0.629%

Apéndices Tabla 3. Distribución de edad entre datos faltantes y datos informados para las variables de la mujer

	THS				Menopausia				Pruebas invasivas Previas				Antecedentes familiares Cáncer				
	Datos informados		Datos faltantes		Datos informados		Datos faltantes		Datos informados		Datos faltantes		Datos informados		Datos faltantes		
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
Edad en el cribado (años)																	
50-54	853,658	30.7%	372,804	31.4%	751,747	30.2%	474,715	32.2%	751,001	31.1%	475,461	30.6%	834,315	30.5%	392,147	31.7%	
55-59	827,415	29.7%	372,785	31.4%	749,268	30.1%	450,932	30.6%	719,469	29.8%	480,731	30.9%	811,556	29.7%	388,644	31.4%	
60-64	777,810	28.0%	315,441	26.6%	684,142	27.5%	409,109	27.7%	655,125	27.1%	438,126	28.2%	764,920	28.0%	328,331	26.6%	
65-69	322,911	11.6%	124,685	10.5%	306,346	12.3%	141,250	9.6%	287,981	11.9%	159,615	10.3%	320,628	11.7%	126,968	10.3%	