

**Desarrollo de una herramienta de simulación  
Monte Carlo de los programas y ensayos  
aleatorios controlados de cribado en cáncer de  
mama**

Luis I. Zamora Ardoy

Tesis de doctorado

Departamento de Física Atómica, Molecular y Nuclear  
Universidad de Granada

Programa de doctorado en Física y Ciencias del Espacio



Granada, julio de 2019

Editor: Universidad de Granada. Tesis Doctorales  
Autor: Luis Ignacio Zamora Ardoy  
ISBN: 978-84-1306-355-3  
URI: <http://hdl.handle.net/10481/57804>



## Agradecimientos

No existe fuerza motriz más poderosa que la que nace del interior más profundo y desconocido de uno mismo. A menudo buscamos darle nombre, por esa idea que supone que todo aquello que puede etiquetarse es conocido: amor propio, orgullo, responsabilidad,... quién sabe, lo único que realmente conocemos de ella es que sólo puede calmarse cuando se ha cerrado la puerta que un día, sin saber cómo, se abrió. Sea como fuera, agradezco que la llama que mantenía vivo el fuego, una veces intensa pero otras muchas débil, nunca se apagara por completo.

La vida está llena de paradojas, de tal forma que las mayores dificultades no suelen ser las que se encuentran a la vista. Se trata de aquellas que se ocultan hábilmente tras las que consideramos más inexpugnables las que acaban resultando más complicadas de vencer. Podría resultar paradójico, y en cierto modo lo es, que este sea uno de los apartados más complicados de acometer en este proceloso camino que ha supuesto para mí la realización de este trabajo. No porque aquellas personas que refiero más abajo no sean merecedoras de unas simples líneas, sino porque dudo que pueda mostrar la enorme importancia que todas y cada una de ellas han tenido, de una u otra forma, en que este trabajo salga a la luz. No soy además persona hábil en estas tareas, sabéis aquellos que me conocéis mejor cuánto me cuesta mostrar mi agradecimiento.

A mi mujer Cristina, parafraseando a Groucho..., yo no querría formar parte de un club que me admitiera como socio, pero soy el primer socio del club que ella preside. Ella ha sido la persona que más me ha animado a continuar, que más ha creído en mi y en este trabajo, incluso cuando ni yo mismo creía en él. Espero ser capaz de devolverle desde aquí solo una pequeña parte de lo mucho que ella me ha dado.

A mis hijos Luis y Andrés. Siempre digo que “los vivo” con mucha intensidad, algo de esa intensidad que quito a otras cosas..., no sé hacerlo de otra manera. Parte de esta tesis se escribió bajo sus pelotazos, lloros, risas y besos, ahora siento que no podría haber tenido un entorno mejor.

A mis hermanas Sofía y María Aurora, por aquello que vivimos juntos y que nos unirá para siempre.

A Damián, apoyo desinteresado, valía, conocimiento y opinión siempre preferente. Reconozco en este punto mi individualismo que no busca referentes..., hasta que redescubrí a Damián.

A Antonio Lallena, quien fuera mi profesor en aquellos ya lejanos años de la facultad. La vida nos ha vuelto a juntar para mostrarme que sin duda es el Maestro.

A todos aquellos compañeros de trabajo que me permitieron robar tiempo al trabajo para dedicárselo a este, en no pocos momentos, incierto proyecto, en especial a Jose Carlos.

A mi padre Juan, con el que seguí el extraño camino que lleva a alejarte, para descubrir tras los años y no sin asombro, lo cerca que siempre hemos estado.

*A mi Madre Sofía*

# Índice general

<b>Prólogo</b>	<b>1</b>
<b>Introducción</b>	<b>5</b>
<b>1. MODELOS NUMÉRICOS DE CRIBADO MAMOGRÁFICO</b>	<b>15</b>
1.1. REVISIÓN DE MODELOS NUMÉRICOS PREVIOS . . . . .	15
1.1.1. Microsimulaciones . . . . .	16
1.1.2. Modelos de Markov . . . . .	24
1.1.3. Simulaciones deterministas y otros modelos semi-analíticos . . . . .	30
1.2. MODELO DESARROLLADO EN ESTE TRABAJO . . . . .	33
<b>2. ESTRUCTURA GENERAL DEL CÓDIGO Y ALGORITMOS DE SIMULACIÓN</b>	<b>37</b>
2.1. ESTRUCTURA GENERAL . . . . .	39
2.2. MODELO DE DETECCIÓN . . . . .	39
2.2.1. Muestreo de la edad de la candidata . . . . .	39
2.2.2. Modelo de crecimiento tumoral . . . . .	40
2.2.3. Incidencia y prevalencia: muestreo de la aparición del cáncer de mama	42
2.2.4. Mortalidad por causas diferentes al cáncer de mama . . . . .	44
2.2.5. Probabilidad de detección. Densidad de la mama, tipos histológicos y transiciones entre ellos . . . . .	45
2.2.6. Modelo de detección clínica. Cánceres de intervalo . . . . .	51
2.2.7. Modelo de cánceres radioinducidos . . . . .	52
2.2.8. Mujeres sin cáncer. Falsos positivos . . . . .	54
2.3. MODIFICACIONES PARA ESTUDIAR LA REDUCCIÓN DE MORTALIDAD . . . . .	55

2.3.1.	Modificaciones para tratar el problema de la supervivencia . . . . .	55
2.3.2.	Muestreo del año de nacimiento. Corrección de nacimiento uniforme	56
2.3.3.	Sensibilidad global y curva de adelanto . . . . .	57
2.3.4.	Supervivencia tras detección y tratamiento . . . . .	60
2.3.5.	Evaluación de la mortalidad por cáncer de mama. Incertidumbre . .	63
2.4.	MODIFICACIONES PARA EL ESTUDIO Y LA SIMULACIÓN DE LOS ENSAYOS ALEATORIOS CONTROLADOS . . . . .	64
<b>3.</b>	<b>APLICACIONES DEL MODELO</b>	<b>69</b>
3.1.	RESULTADOS RELACIONADOS CON LA DETECCIÓN. SIMULACIÓN DE PROGRAMAS REALES, ANÁLISIS DE LA SENSIBILIDAD Y DEL SOBREDIAGNÓSTICO . . . . .	70
3.1.1.	Simulación de los resultados de detección en programas reales. Es- tudio de las diferentes configuraciones . . . . .	71
3.1.2.	Análisis de sensibilidad . . . . .	73
3.1.3.	Estudio del sobrediagnóstico . . . . .	79
3.2.	RESULTADOS RELATIVOS A LA REDUCCIÓN DE MORTALIDAD EN LOS PROGRAMAS DE CRIBADO . . . . .	83
3.2.1.	Validez de la sensibilidad global y curva de adelanto . . . . .	85
3.2.2.	Estudio de las diferentes configuraciones y de la participación en los programas de cribado en términos de reducción de mortalidad . . . .	87
3.2.3.	Análisis Coste-Utilidad . . . . .	90
3.3.	SIMULACIÓN DE LOS PRINCIPALES ENSAYOS ALEATORIOS CONTROLADOS DE CRIBADO MAMOGRAFICO . . . . .	94
3.3.1.	Simulación de la reducción de mortalidad e intervalos de confianza de los EAC . . . . .	95
<b>4.</b>	<b>CONCLUSIONES Y PERSPECTIVAS</b>	<b>103</b>
4.1.	CONSIDERACIONES FINALES . . . . .	103
4.2.	CONCLUSIONES . . . . .	105
4.3.	PERSPECTIVAS . . . . .	106
	<b>Bibliografía</b>	<b>107</b>

<i>ÍNDICE GENERAL</i>	v
<b>Lista de figuras</b>	<b>119</b>
<b>Lista de tablas</b>	<b>121</b>
<b>Publicaciones, comunicaciones y citas destacadas</b>	<b>123</b>
Publicaciones . . . . .	124
Comunicaciones . . . . .	124
Otras citas . . . . .	125



# Prólogo

El comienzo de este trabajo no es sólo el de una tesis doctoral, sino el de un proyecto de vida. Digo esto porque el origen de lo que aquí se culmina está en el trabajo de investigación “Optimización de un programa de cribado mamográfico”, de Cristina Forastero Rodríguez, que hoy es mi mujer y que se llevó a cabo allá por septiembre de 2008. Damián Guirado Llorente tenía este asunto del cribado mamográfico en su infinito repertorio de futuros trabajos por hacer, inoculando en Cristina y en mí el “veneno” de esta controvertida disciplina que aúna aspectos diagnósticos, radiobiológicos, clínicos e incluso económicos y psicológicos. De esta forma comenzamos a trabajar en ello a principios del 2008. Durante estos primeros meses generamos los códigos FORTRAN que darían lugar al primer modelo y al trabajo de Cristina. Todavía recuerdo aquellos primeros resultados del modelo que mostraban tasas de detección oscilantes, más aún, recuerdo el momento en que entendimos, prácticamente de forma simultánea, el motivo de este supuesto artefacto. Podría ser este uno de esos casos de sincronidad de la que hablaba Jung. Tras una primera publicación relativamente sencilla en *Physics in Medicine and Biology* en el 2010, vino la primera travesía del desierto.

Este primer trabajo no abordaba de lleno las cuestiones más polémicas que el cribado mamográfico ha generado y todavía provoca en la comunidad científico-profesional. Prueba de esto fue la ponencia que con el título “Riesgos al paciente en Radiodiagnóstico. Apartado de programas de cribado mamográfico”, tuve oportunidad de presentar en junio de 2009 en el Congreso Nacional de las sociedades españolas de Física Médica y Protección Radiológica de Alicante, que generó entre algunos compañeros cierta incomodidad cuando no un indisimulado enfado.

Tras unos meses de parada fuimos conscientes de que era necesario ir más allá de los resultados ligados exclusivamente a la detección, había que introducir la supervivencia tras detección y tratamiento para obtener resultados de reducción de mortalidad, pues esta es la principal métrica que mide la bondad de los programas de cribado. Era este un proyecto ambicioso, pues implicaba conceptos radiobiológicos y clínicos más difícilmente parametrizables en un modelo, al fin y al cabo, matemático. Una vez más, el modelo desarrollado volvía a arrojar resultados sorprendentemente coherentes, al menos a mí parecer siempre algo pesimista, publicando el resultado de este trabajo en la revista *Medical Physics* en el año 2012.

Nuestro proyecto parecía haber llegado a su fin, pero una vez más Damián quiso ir más allá, seguramente porque era el único que veía este trabajo no sólo como una serie de publicaciones más o menos exitosas sobre un asunto de interés, sino como una futura tesis doctoral. Así, propuso que estudiáramos los ensayos aleatorios controlados que habían

sido la base de la evidencia sobre la materia hasta la fecha. Tras realizar las modificaciones oportunas en el código para adaptarlo a este nuevo tipo de simulaciones, publicamos el trabajo, no sin dificultad, en la revista *Physica Medica* entrado ya el 2016.

Se completaba de este modo el número de tres artículos publicados sobre la materia, momento este en el que la idea de una tesis pasó en mi cabeza de ser un rumor lejano a una poderosa voz interior. Se hacía necesario todavía algo más: una tesis que versara sobre cribado mamográfico debería contener un apartado especial dedicado al sobrediagnóstico/sobretatamiento, pues es este un punto crucial y de difícil abordaje, pero que puede determinar la justificación, en cualquiera de sus formas, de un programa de cribado. Este ha sido el trabajo que nos ha ocupado durante aproximadamente el último año y que, tras dura batalla, hemos conseguido publicar en *BMJ open* durante este 2019 en curso.

Debo en este punto destacar la tremenda labor llevada a cabo por el catedrático Antonio M. Lallena Rojo, que ha ocupado un papel central en todo este trabajo que aquí resumo, siendo de especial relevancia en la publicación de este último artículo, devolviendo con precisión y maestría cada una de las “bolas” que nos enviaron los revisores.

Espero que la redacción de esta tesis haya sido capaz de captar el “alma” de todo el trabajo realizado durante este periodo, en numerosos momentos interrumpido, de más de 10 años.

*Para cada enfermedad que los doctores curen con medicinas, producen diez mil en las personas saludables, inoculándolas con ese virus que es mil veces más poderoso que cualquier microbio..., la idea de que uno está enfermo.*

Marcel Proust.



# INTRODUCCIÓN

Se entiende por programa de cribado aquella actividad orientada a la detección precoz de la enfermedad, su diagnóstico y tratamiento temprano, que se ofrece activamente al conjunto de la población susceptible de padecer la enfermedad, aunque no tengan síntomas ni hayan demandado ayuda médica (ley 33/2011, General de Salud Pública). En particular, el cribado mamográfico entendido como la invitación a un grupo de población de un intervalo de edades prefijadas y asintomático por cáncer de mama, para someterse a una técnica de detección basada en mamografía, con el fin de detectar la enfermedad en un estadio previo al que se presentaría de manera sintomática, tiene su origen en el desarrollo de la técnica de mamografía por rayos X (RX). Aquí cabe destacar como pioneros a Gershon-Cohen (EE. UU.), Leborgne (Uruguay) y Gross (Francia).

La idea de utilizar esta técnica de detección en programas masivos o poblacionales para reducir la mortalidad originada por el cáncer de mama, pronto empieza a fraguarse. El primer ensayo aleatorio controlado encaminado a probar la reducción de mortalidad se lleva a cabo entre los años 1963 y 1969 y es conocido como New York HIP [Sha85]. Los resultados preliminares de estos ensayos animaron a la realización de muchos más, cada vez más depurados, a la vez que se ponían en marcha los primeros programas poblacionales, es decir, intervenciones en grupos diana con fines preventivos (prevención secundaria) cuya finalidad básica es disminuir la incidencia de esta patología, la mortalidad de la misma y aumentar la calidad de vida de esta población. En España el primer programa de cribado de cáncer de mama se acomete en la comunidad Foral de Navarra en el año 1990 [Asc04].

Rápidamente empiezan a surgir las primeras controversias alrededor de la propia existencia de los programas, de sus resultados en términos de reducción de la mortalidad y de su configuración (grupo etario al que se dirigen, tiempo de intervalo entre las diferentes pruebas, técnica de la mamografía y diferentes proyecciones a emplear, gestión del diagnóstico, etc.). En este sentido cabe plantearse la justificación de estos programas en términos de un análisis coste-beneficio, para lo que se definen algunos conceptos que son de utilidad [Eco07]:

- **Eficacia.** Define una intervención en salud que pueda provocar las metas preestablecidas empleada en condiciones óptimas.
- **Eficiencia.** Se define como la propiedad de alcanzar el máximo beneficio en salud en función de los recursos disponibles.
- **Efectividad.** Se dice que una intervención tiene esta propiedad cuando produce más bienestar que daño si se emplea en las circunstancias habituales.

- **Coste directo.** Valor monetario de los recursos utilizados en la prevención, tratamiento o rehabilitación de las personas con un determinado padecimiento. Incluye el coste de los fármacos y tratamientos y otras necesidades como el apoyo psicológico, transporte, pérdida económica asociada a una reducción del rendimiento o la productividad por la morbilidad asociada al tratamiento, etc.
- **Coste intangible.** Constituye el estigma social, el dolor y el sufrimiento que un individuo presenta por ser portador de una enfermedad. Este es un aspecto crucial pero su difícil, cuando no imposible, cuantificación dificulta la justificación global de cualquier intervención en materia de salud pública.
- **Análisis coste-efectividad.** Considera el costo de una nueva intervención respecto a la antigua en relación con el número de años de vida ganados.
- **Análisis coste-utilidad.** Considera el costo de una nueva intervención respecto a la antigua en relación con el número de años de vida ganados ajustados por calidad de vida (AVAC o *QALY Quality Adjusted Life Years* en inglés).
- **Análisis coste-beneficio.** Describe y cuantifica las ventajas (ingresos sociales) y desventajas (costes y gastos sociales) de un proyecto de inversión o de una política, ambos en términos monetarios, por tanto, directamente comparables.

La medida de efectividad (años de vida) se traduce en una medida de utilidad (*QALY*) multiplicando por un factor entre 0 y 1, donde 0 representa estar muerto y 1 el estado de salud perfecto. En efecto, la utilidad de un estado de salud  $Q$ , mantenido durante  $Y$  años viene dada por  $U(Q, Y) = U(Q) \times Y$ . Existen fundamentalmente 2 métodos para determinarlo:

1. Método de compensación temporal. El valor de la utilidad para un estado de salud,  $U(Q)$ , se obtendría reponiendo a la pregunta ¿qué cantidad de vida estaría dispuesto a ceder a cambio de una mejora del estado de salud?. Si una persona sorda con esperanza de vida 20 años estaría dispuesta a “ceder” 3 años por no estarlo, se podría escribir  $U(20 \text{ años, sordo}) = U(17 \text{ años, buena salud})$  y por tanto  $20 \times U(\text{sordo}) = 17 \times U(\text{buena salud})$ , es decir  $U(\text{sordo}) = 17/20 = 0.85$
2. Método de la “lotería estándar”. Se basa en el riesgo de muerte que una persona estaría dispuesta a asumir a cambio de recuperar la buena salud. Así, para el ejemplo anterior, si esta persona asume un 5 % de riesgo de fallecimiento, podríamos escribir:  $U(20 \text{ años, sordo}) = 0.05 \times U(\text{muerte}) + 0.95 \times U(20 \text{ años, buena salud})$ , que supondría  $20 \times U(\text{sordo}) = 0.05 \times 0 + 0.95 \times 20 \times 1$  y, por tanto,  $U(\text{sordo}) = 0.95$ .

La utilidad, a su vez, se convierte en unidades monetarias para estimar el beneficio.

Junto a estos conceptos empleados en el ámbito de ciencias de la salud, es interesante definir otros puramente económicos, de los que sólo destacaremos:

- **Coste marginal.** Mide la tasa de variación del coste dividida por la variación del producto. Representa, por tanto, en términos de salud, el coste adicional que supone la ganancia de un año de vida más. Permite estudiar el punto de inversión en recursos más allá del cual los costes se disparan sin aportar grandes beneficios, es decir, el gasto se vuelve muy poco eficiente.

- **Coste de oportunidad.** Designa el coste de la inversión en la mejor alternativa posible. Habría que descontar de los beneficios aquellos estimados de esta alternativa para una evaluación adecuada.
- **Tasa de descuento social.** Procedimiento para calcular los costes o beneficios monetarios que ocurren en distintos periodos de tiempo con objeto de actualizarlos al presente. Es la de más difícil aplicación en el ámbito de ciencias de la salud, pues tiene en cuenta que tanto los costes como los beneficios de una intervención pueden sobrepasar la escala generacional pudiendo, incluso, tener coste y beneficio escalas de tiempo diferentes. Existen diferentes modelos económicos que intentan traducir estos costes y beneficios a un “valor equivalente” en el presente.

A la luz de estas definiciones se desprende que las intervenciones en materia de salud pública raramente resultan “beneficiosas” en términos de salud exclusivamente, supuesta una adecuada, o cuando menos razonable, conversión de los *QALY* en unidades monetarias. La clave por tanto parece estar en el análisis coste-utilidad, es decir un análisis en términos de *QALY* que tenga en cuenta el coste de oportunidad mediante la mejor alternativa posible y que debe encontrar su configuración óptima de funcionamiento en función del coste marginal. Llevado al caso concreto que nos ocupa, el cribado poblacional de cáncer de mama mediante mamografía, el análisis coste-utilidad debería incluir entre los costes directos:

1. **Coste de la mamografía.** Incluye personal técnico y facultativo, medios técnicos, amortización de recursos no fungibles y costes fungibles.
2. **Coste de los tratamientos** en pacientes diagnosticados correctamente y de las pruebas alternativas diagnósticas (biopsias, mamografías adicionales, etc.), tanto en los diagnósticos correctos (verdaderos positivos, VP) como en los incorrectos (falsos positivos, FP y falsos negativos, FN).
3. **Estimaciones de los costes asociados a pérdida de días de trabajo y productividad** que suponen tanto la asistencia al programa como un resultado positivo en algún momento del programa de cribado.
4. **Costes de apoyo psicológico y rehabilitación.**
5. **Coste asociado a la radioinducción de cánceres**, entendido como costes monetarios relacionados con el diagnóstico y tratamiento.

En términos cualitativos y más allá del análisis en cualquiera de las formas anteriores, antes de continuar habría que responder a dos preguntas fundamentales sobre los programas de cribado mamográfico en relación con los conceptos definidos previamente de eficacia, eficiencia y efectividad:

1. ¿Puede la detección precoz interrumpir el crecimiento del cáncer de mama (historia natural) antes de su diseminación metastásica?
2. ¿Cuál sería la edad de inicio y la de finalización de los programas y con qué frecuencia se realizaría el cribado?

Actualmente se piensa que la detección más temprana presenta ventajas en los resultados como consecuencia de la aparición de tratamientos adyuvantes como la quimioterapia (QT) y la hormonoterapia, junto a los tratamientos tradicionales como la radioterapia (RT) y la cirugía [UKP12]. Es más, en algunos casos se presupone una ventaja debida a la modalidad de detección, en cierto modo asociada no tanto al tamaño como al estadio y al pronóstico, que también suelen ser mejores en etapas tempranas [San03]. En cualquier caso, es también conocido que hay tumores incurables aún siendo pequeños o incluso microscópicos [Kop98].

Se suele dividir la historia natural del cáncer de mama en tres etapas:

1. **Fase indetectable**, en la cual el tumor no tiene el tamaño necesario para su detección por mamografía, ni muestra síntomas clínicos.
2. **Fase preclínica** (también llamado tiempo de intervalo, estancia o *sojourn time*). En ella el tumor es detectable por mamografía pero no clínicamente.
3. **Fase clínica**, a partir de cuyo inicio el tumor es detectable por sus síntomas.

La duración de esta fase preclínica es clave para el beneficio del cribado desde el punto de vista de la configuración de los programas, produciéndose la transición entre la fase indetectable y la preclínica cuando el tumor supera el umbral de 2 mm [Kop98]. Este valor será más elevado en mujeres jóvenes con mamas, en general, más densas que presentan un peor contraste radiológico. Diferentes estudios sostienen que esta fase preclínica es de menor duración en las mujeres jóvenes [Mos86, Fei94, Tab95, Dij97], lo que reduce la ventana de detección por mamografía en este grupo etario y, por consiguiente, el cribado proporciona un menor tiempo de adelanto a la detección clínica.

Desde el punto de vista de esta clasificación, el tiempo de intervalo entre las diferentes mamografías debe ser del orden de la mitad de la duración de la fase preclínica. Tiempos mayores no ofrecerían ganancia respecto a la detección clínica, mientras que tiempos mucho menores no generarían detecciones adicionales al estar muchos tumores en fase indetectable. La edad máxima para someterse al cribado estaría limitada por el beneficio obtenido que, a su vez, estaría condicionado, según algunos estudios, por una mayor duración de la fase preclínica en mujeres mayores o, al menos, por un valor más elevado del producto (sensibilidad)  $\times$  (duración fase preclínica), entendida la sensibilidad como la probabilidad de detectar un tumor por mamografía. Por su parte, la edad mínima estaría asociada a un tiempo de intervención suficientemente temprano antes de que aparezca la diseminación metastásica.

Un análisis puramente aritmético, en términos de tumores detectados mediante cribado frente a detección clínica, se muestra claramente insuficiente. La imagen mamográfica no es capaz de distinguir algunas lesiones benignas de tumores más agresivos. El mismo cribado lleva asociado un “sesgo de duración” que implica una mayor probabilidad de detectar aquellos tumores de crecimiento lento y, por tanto, menos agresivos en principio. La supervivencia sin más tampoco es suficiente pues el adelanto en el tiempo de diagnóstico proporciona unos mayores tiempos de supervivencia sin que, necesariamente, se retrase el momento de la muerte (sesgo *lead time*, o tiempo de adelanto).

Este problema del cribado mamográfico, como indicamos anteriormente, debe abordarse en términos de coste de oportunidad, es decir, comparado con la mejor opción alternativa posible. Habría por tanto también que decidir cuál es esta: sistema general de salud sin más, sistema general de salud con especial atención a grupos de riesgo (antecedentes familiares, mutaciones BRCA1/2, etc.), alternativas diagnósticas como autoexploración tras el adecuado adiestramiento, ecografía, resonancia magnética, etc.

Llegados a este punto, si el resultado del análisis coste-utilidad fuera favorable al cribado pese a un mayor coste económico, habría que decidir cuál sería el total de recursos a dedicar y cuál la configuración elegida, en términos de coste marginal. Será esta una decisión político-administrativa fuera ya de los parámetros estrictamente económicos, pues podría darse el caso de que el análisis arrojara un mayor número de *QALY* del cribado frente a la mejor opción alternativa, que podría ser no llevarlo a cabo, pero a costa de un mayor número de recursos. Se tendría que decidir entre una u otra alternativa en función de los recursos disponibles. Finalmente, en función del coste marginal, tendría que establecerse dónde detener la inversión. Podría llegarse a una situación ciertamente paradójica en el ámbito de la salud pública: el cribado podría presentar un peor coste-utilidad en términos de €/QALY pero su implementación supondría un montante total de *QALY* mayor que la no realización del mismo, pese a su evidente mayor coste.

Este enfoque global es sin duda muy ambicioso y, como hemos señalado, requiere de la estimación de numerosos costes relacionados con la evolución de tumores en ausencia/presencia del cribado (sobretretamiento, reducción de mortalidad o tumores radio-inducidos), junto con la evaluación de la forma más precisa posible de los detrimentos inherentes al cribado y el coste de los tratamientos. En este sentido se han llevado a cabo estudios más o menos sofisticados, de entre los cuales es una referencia fundamental el artículo de Lindfords y Rosenquist [Lin95].

Dentro de los costes intangibles, cobran una especial importancia aquellos asociados al sobrediagnóstico/sobretretamiento [UKP12, Got09], entendido como el detrimento asociado a aquellos casos diagnosticados como positivos en el programa que no causarían ningún perjuicio en ausencia del mismo. A efectos de la curva de incidencia, el sobrediagnóstico se manifestaría como un incremento en la misma que comienza alrededor de 5 años después de la edad de inicio de cribado, y se extendería hasta la edad de finalización más el *lead time*, momento en que regresaría al nivel basal sin cribado. El sobrediagnóstico comprende dos posibles escenarios:

1. Tumores que sin el cribado nunca llegarían a ser sintomáticos.
2. Muerte por causas diferentes al cáncer de mama.

En el primer caso nos referimos a los tumores que, de seguir su historia natural, nunca serían detectados ni causarían detrimento alguno. Nunca pasarían a invadir estructuras fuera de su origen debido a sus propiedades histológicas o a una velocidad de crecimiento extremadamente lenta. El segundo punto es ciertamente inevitable como origen de sobrediagnóstico.

Este es, sin duda, un apartado crucial en el estudio de los programas de cribado, tanto por la dificultad de su evaluación como por el elevado peso que se estima puede

tener en el análisis; hasta el punto de poder hacer estos programas ineficientes frente a otras alternativas. Se trataría, por tanto, de tener una evaluación precisa del llamado sobrediagnóstico/sobret ratamiento, lo que a su vez, supone un doble inconveniente: por un lado implica una comparación entre la evolución de un caso diagnosticado en el programa y la evolución que tendría sin esta “interrupción”; por otro, quedaría la dificultad de evaluar de manera adecuada el detrimento que supone el diagnóstico precoz en estos casos.

El tratamiento de los dos principales asuntos a determinar en este ámbito de conocimiento, a saber, la reducción de la mortalidad que supone la implementación de un programa y la evaluación del sobrediagnóstico/sobret ratamiento, han sido abordados por la mejor herramienta de la que dispone la medicina basada en la evidencia: el diseño y desarrollo de ensayos aleatorios controlados (EAC). Actualmente se consideran 11 EAC como los de mayor evidencia en la estimación de la reducción de mortalidad por cáncer de mama. Estos EAC se basan en la evaluación de la relación entre el riesgo de mortalidad por cáncer de mama de las mujeres invitadas al cribado y el riesgo basal (RR). Del correspondiente metanálisis de todos ellos, se ha derivado una reducción de mortalidad por cáncer de mama del 20 % (11 %-27 %; IC 95 %) [Got09].

Otro tipo de estudios se deben considerar de menor evidencia. Entre estos están: (i) los estudios observacionales, de entre los cuales destacamos los estudios ecológicos o de tendencia temporal, que aportan muy poca evidencia al estar basados en comparaciones entre poblaciones con diferentes riesgos basales y niveles socioeconómicos, (ii) los estudios de casos y controles y (iii) los estudios de mortalidad basados en la incidencia. Estos dos últimos tipos conllevan una sobreestimación de los resultados [UKP12].

En lo referente a la segunda cuestión, el sobrediagnóstico, a día de hoy no es posible determinar si un tumor detectado será o no un caso de sobrediagnóstico y tan solo nos es posible estimar, con una elevada incertidumbre, la frecuencia de aparición de ese sobrediagnóstico. La mayor fuente de evidencia para estas estimaciones vuelven a ser los EAC, pero presentan el problema del escaso tiempo de seguimiento. Además, la mayoría de estos ensayos incluyen un cribado al grupo de control al final del estudio. En cuanto al primer problema, debe extenderse la duración del seguimiento, al menos, un periodo igual al *lead time* más allá de la edad límite estudiada, como vimos anteriormente, para poder distinguir de manera adecuada entre detección temprana y sobrediagnóstico. Este *lead time* se estima entre 5 y 10 años tras la finalización del ensayo [Bie07, Pul11].

Un tratamiento aparte merecería el detrimento debido a la radioinducción de tumores asociada a la radiación. Si bien la implantación de un programa de cribado tiene un carácter “masivo” y, por tanto, deben evaluarse los riesgos asociados a la exposición, no parece, a la vista de las estimaciones que se han realizado hasta la fecha, que esta juegue un papel decisivo en la justificación de los mismos. La utilización de los EAC para este extremo se vuelve improductiva: serían necesarios estudios con  $10^8$  mujeres para poder estimar un resultado con significación estadística [Kop98], por lo que se hace necesario acudir a modelos de proyección de riesgo aditivos o multiplicativos [ICRP91]. Según la *Health Protection Agency* (EE.UU.), por cada  $10^5$  mujeres sometidas a una mamografía se inducen entre 1 y 10 cánceres. Estimaciones más recientes del *National Health System* (Reino Unido) [UKP12] afirman que por cada  $10^4$  mujeres cribadas entre los 47 y 73 años cada 3 años se inducirían entre 3 y 6 cánceres.

Otros daños asociados al cribado mamográfico serían:

- Dolor durante la mamografía.
- Falsos positivos. En el caso del programa inglés NHSBP se estimó en un 3.4 % la tasa de rellamadas, es decir, mujeres que vuelven a ser convocadas tras la mamografía para pruebas más específicas y que resultaron no tener cáncer.
- Falsos negativos. En el caso inglés se estimó en 0.2 por cada  $10^3$  mujeres cribadas.

La morbilidad asociada al tratamiento podría resumirse en [UKP12]:

1. No se ha demostrado mortalidad asociada a la hormonoterapia.
2. No se ha probado un efecto en la mortalidad de la radioterapia.
3. Un máximo de 0.2 muertes por cada  $10^3$  mujeres debido a la cirugía.
4. Un máximo de 1.3 muertes por cada  $10^3$  mujeres debido a la quimioterapia.

Por tanto, si hablamos exclusivamente en términos de mortalidad, tendríamos un máximo del 0.15 % asociada a los tratamientos. Si suponemos 122 sobrediagnósticos/sobretratamientos por cada 40 muertes evitadas, podríamos atribuir 0.2 muertes debidas al sobrediagnóstico/sobretratamiento. Así pues, se evitarían 200 muertes mediante el cribado por cada muerte inducida por tratamiento de mujeres que no deberían haberlo recibido en ausencia del programa. Con los datos del informe Cochrane [Got09] el balance es menos favorable, situándose en una proporción de 70 a 1.

Existen otros riesgos no probados asociados a la mamografía o el cribado, de entre los cuales citamos [Kop98]:

- Riesgo de ruptura del encapsulamiento asociado a la presión de la mamografía, que podría favorecer la diseminación metastásica [Hol94].
- La posible relación del crecimiento de las lesiones metastásicas con la extirpación del tumor primario, aunque parece que este efecto sólo sería significativo durante algunas horas [Sil95].

Quedan por tanto numerosas cuestiones a resolver en lo referente a los programas masivos de cribado mamográfico. Si bien parece probada una reducción de mortalidad global, que no sería inferior al 15 %, quedaría por establecer qué parte es debida al cribado y cuál atribuible al avance en las terapias, fundamentalmente las adyuvantes. En este sentido cabe destacar el aumento de la incidencia en el cáncer de mama experimentado desde los años 70, asociado entre otras causas a la generalización de las terapias hormonales, el retraso en la edad de nacimiento, el menor número de hijos, la obesidad, el consumo de alcohol, el aumento de la esperanza de vida y la propia generalización de los programas de cribado. Frente a este aumento, existe una tendencia a disminuir la mortalidad atribuida al cribado, al desarrollo de las terapias y a algunas mejoras en la especialización y organización del cuidado del cáncer.

El carácter masivo o poblacional inherente a este tipo de intervenciones en salud pública lleva asociado resultados, cuando menos, polémicos. En este sentido, el informe WHO

2002 [IAC02] establece que el cribado se traduce en un aumento de 1 día de vida en promedio para cada mujer invitada al mismo. Por el contrario, actuaciones específicas sobre “blancos” más delimitados, como puede ser la terapia hormonal con tamoxifeno aplicada a mujeres con ganglios positivos, producirían un aumento en la vida media de 6 meses [Got09] (200 veces más). El rango de edades a los que debería aplicarse el programa tampoco está exento de polémica. Si bien parece clara una reducción de mortalidad efectiva para mujeres de 50 años, no está claro lo que sucede en el rango de 40 a 49. Es conocida la menor incidencia del cáncer de mama a estas edades junto a una menor eficiencia de la técnica mamográfica en mujeres más jóvenes con mamas densas. Actualmente se cree que los tumores que surgen a edades más tempranas muestran también una menor duración de la fase preclínica, disminuyendo aún más la eficacia de cribado. Pero la propia segregación brusca en mujeres menores de 50 y mayores de 50 años asociada, sin duda, a la edad media de la menopausia es, a todas luces, artificial. La menopausia lleva asociados cambios hormonales significativos y, con ellos, puede que cambios en el comportamiento natural de los posibles nuevos tumores que aparecieran junto a estos cambios. Es conocida la relación que existe entre este tipo de tumores y cierto tipo de hormonas, así como cambios progresivos en la composición del tejido mamario con la edad pero, en ningún caso, estos cambios ocurrirían de forma abrupta ni serían generalizables a toda la población.

Las recomendaciones internacionales no son muy uniformes. La *International Agency for Research on Cancer* (IARC) [IAC02] revisó la evidencia sobre la cuestión, concluyendo que había una reducción de la mortalidad por cáncer de mama entre mujeres de 50 a 69 años del 35 %, con una evidencia limitada para el grupo de 40 a 49 años. Por su parte, la *US Preventive Services Task Force* [Woo10] recomendó excluir de los programas a mujeres menores de 50 años y pasar a tiempos de intervalo de 2 años, mientras que la *Canadian Task Force* [CTF11] encontró que el beneficio para mujeres menores de 50 años era menor, siendo mayor por el contrario el sobrediagnóstico. Finalmente, los diferentes informes de la fundación Cochrane, en particular [Got09], son muy críticos con el balance riesgo-beneficio de estos programas.

El problema es, por tanto, extremadamente complejo, pues en él confluyen (*i*) aspectos biológicos, derivados del comportamiento natural de estos tumores, su velocidad de crecimiento, capacidad metastásica y la dificultad de inferir el grado de malignidad de los mismos en fases tempranas; (*ii*) aspectos técnicos derivados de la técnica mamográfica, su sensibilidad, especificidad y la dependencia de estos parámetros con la edad y el tipo de tumor; (*iii*) vertientes puramente terapéuticas, que tienen que ver con la eficiencia de los tratamientos tempranos; (*iv*) la evaluación del detrimento y la morbilidad que suponen estos programas, así como la estimación del riesgo de radioiducción de nuevos tumores y, finalmente, (*v*) la dificultad de diseñar EAC que resuelvan de manera definitiva. Todas estas cuestiones hace aún más difícil aportar certezas en esta materia.

Es necesario, por tanto, introducir nuevas herramientas que ayuden a responder de forma precisa todas estas preguntas. Es en este punto donde la simulación matemática de estos programas puede aportar luz por su indudable capacidad para modelar aspectos biológicos, poblacionales, técnicos y epidemiológicos mediante distribuciones aleatorias de los parámetros fundamentales. Si bien un modelo matemático de simulación no puede nunca, por sí solo, ser una evidencia en el ámbito de las ciencias de la salud, sí que puede, una vez que se han ajustado los parámetros del modelo y probado su eficiencia fuera del ámbito de este ajuste, ser una herramienta de gran valor para la justificación o no de los

programas de cribado, dar una respuesta a cuestiones radiobiológicas y epidemiológicas esenciales, indicar cuáles son las configuraciones óptimas de los programas e incluso ser una medida de la bondad metodológica y de validez externa de ciertos ensayos aleatorios controlados sobre esta cuestión.

Este trabajo intenta resolver las dudas que los programas de cribado ofrecen en términos de reducción de mortalidad, volumen de sobrediagnóstico que inducen y configuración más adecuada de edades y tiempo de intervalo para una mejor justificación en términos de coste-utilidad. Para ello, empleamos la herramienta de simulación Monte Carlo, capaz de reproducir escenarios reales en los que determinar estos valores, a la vez que permite obtener conclusiones en materia de salud pública mediante las simulaciones de diferentes poblaciones y configuraciones que los programas o ensayos aleatorios no podrían evaluar, bien por motivos éticos o simplemente estadísticos asociados al volumen de población necesario para obtener conclusiones significativas.

Para desarrollar esta herramienta Monte Carlo se ha empleado el principio de máxima simplicidad, de forma que no se asumen ideas previas que no hayan sido suficientemente probadas, se parametrizan las distribuciones empleadas mediante cantidades observables o de escasa incertidumbre, de cara a obtener resultados que no estén condicionados por estas ideas, o no dependan de estados definidos de manera artificial. Así, se asume un modelo de detección por mamografía sencillo, que introduce de forma directa la dependencia de la sensibilidad con la densidad de la mama, tipo histológico del tumor y tamaño del mismo. Por su parte, la detección clínica sigue un modelo exponencial, que puede justificarse de forma intuitiva y que depende únicamente de dos parámetros suficientemente conocidos. En cuanto al crecimiento del tumor, se asume también un modelo plausible y realista que no presupone velocidades de crecimiento diferentes en función de la edad de la mujer. Para estimar la supervivencia tras detección y tratamiento estándar de los tumores invasivos, se ha hecho uso de la evidencia que indica que la supervivencia es únicamente dependiente del tamaño del tumor en la detección, de forma que no se introducen sesgos en la supervivencia asociados al tipo de detección.

El modelo así obtenido, se emplea para reproducir los resultados asociados a la detección generados por los programas reales, suficientemente documentados y de reconocida solvencia, previo análisis de sensibilidad, que permita comprobar su estabilidad y consistencia en función de los parámetros de entrada empleados. Una vez completada esta primera etapa, pasamos a estudiar aspectos críticos de la justificación de los programas como son la reducción de mortalidad por cáncer de mama que supuestamente producen, el sobrediagnóstico y la justificación de las diferentes configuraciones de rango de edad y tiempo entre mamografías empleadas, así como la influencia del porcentaje de participación en los resultados. Finalmente, se aplica el modelo para reproducir los resultados reportados por los ensayos aleatorios principales realizados hasta la fecha, a fin de estudiar su validez interna (corrección metodológica) y externa (aplicabilidad a otras poblaciones).

Este hilo conceptual se implementa en los 4 capítulos que se desarrollan a continuación. El capítulo 1 está dedicado a la descripción de los principales modelos matemáticos desarrollados hasta la fecha sobre la materia. En el capítulo 2 se detallan y justifican los diferentes algoritmos empleados en las simulaciones. En el capítulo 3 se aplica el modelo para reproducir programas de cribado reales y ensayos aleatorios, así como para evaluar las configuraciones de los programas más eficientes y el sobrediagnóstico. Por último, el capítulo 4 se dedica a las conclusiones y perspectivas.

*...En arte, nada que merezca la pena se puede hacer sin genio; en ciencia, incluso una capacidad muy modesta puede contribuir a un logro supremo.*

Bertrand Russell.

# Capítulo 1

## MODELOS NUMÉRICOS DE CRIBADO MAMOGRÁFICO

En esta capítulo se hará una revisión de los principales modelos numéricos que han abordado el asunto del cribado mamográfico desde diferentes perspectivas, de cara a situar nuestro modelo de manera precisa en el “mapa” de este ámbito de conocimiento y poder tener una idea clara de sus aspectos diferenciales respecto al resto de trabajos llevados a cabo hasta la fecha.

### 1.1. REVISIÓN DE MODELOS NUMÉRICOS PREVIOS

Existen multitud de códigos que simulan los programas de cribado mamográfico. Atendiendo a su naturaleza los podemos agrupar en *microsimulaciones*, simulaciones basadas en *procesos de Markov* y simulaciones *deterministas*, a continuación se describen los aspectos más relevantes de cada uno de ellos, mencionando los ejemplos más significativos:

1. **Microsimulaciones.** Se trata de simulaciones Monte Carlo de la historia de las mujeres sometidas al cribado que utilizan distribuciones aleatorias de los diferentes parámetros que caracterizan el proceso: incidencia, detección clínica, detección mamográfica, crecimiento tumoral, supervivencia, etc. Emplean una simulación detallada de la historia natural del tumor, es decir, de su evolución biológica debida a su naturaleza y al entorno cuando no se interviene externamente sobre él.
2. **Códigos basados en procesos de Markov.** En estos cálculos la historia natural del tumor se modela mediante estados más o menos artificiales para los que se calculan las distribuciones de tiempo de permanencia en los mismos, utilizando la matriz de intensidad de las transiciones entre estados y resolviendo las ecuaciones de Chapman-Kolmogorov. Estas distribuciones se parametrizan y son necesarios métodos del tipo máxima verosimilitud o mínimos cuadrados generalizados para estimar los parámetros con ayuda de resultados conocidos: EAC, recopilaciones epidemiológicas, etc. Debe hacerse notar que estos modelos se emplean únicamente para reproducir la historia natural del tumor; posteriormente se realizan simulaciones detalladas empleando las distribuciones obtenidas mediante el procedimiento anterior o, en otros casos, se llevan a cabo estimaciones analíticas con las mismas.

3. **Simulaciones deterministas.** En estas simulaciones se generan cohortes de mujeres para las que se sigue su historia utilizando parámetros con valores fijos, no sujetos a distribuciones aleatorias.

### 1.1.1. Microsimulaciones

Las microsimulaciones son simulaciones Monte Carlo de la historia de las mujeres sometidas al cribado que utilizan distribuciones aleatorias de los diferentes parámetros que caracterizan el proceso: incidencia, detección clínica, detección mamográfica, crecimiento tumoral, supervivencia, etc. Además, emplean una simulación detallada de la historia natural del tumor. Se denominan microsimulaciones porque representan la manera más detallada de simular estos procesos, sin hipótesis previas sobre estados o transiciones entre ellos.

El primer ejemplo de este tipo de cálculos se debe a Jansen y Zoetelief [Jan95], que presentan el modelo MBS (*Model risk benefit analysis of Breast cancer Screening*) de simulación Monte Carlo de los programas de cribado mamográfico. Este paquete supone un crecimiento gompertziano de los tumores [Pee93] y una función de supervivencia a 10 años exclusivamente dependiente del diámetro de detección,  $d$ , ajustada mediante la siguiente expresión:

$$S(d) = \exp(c_1 d - c_2 d^2), \quad (1.1)$$

donde  $c_1$  y  $c_2$  son parámetros libres. Se utiliza la distribución log-normal tanto para el muestreo del tiempo de duplicación inicial del tumor, como para los umbrales de detección espontánea y por mamografía, una para mujeres menores de 48 años y otra para mayores de 48, de forma que se puede modelar así el efecto en la detección en las mamas más densas. El muestreo de estas distribuciones se realiza mediante el método de Box-Muller [Box58] y no se introducen beneficios adicionales por modalidad de detección. En este modelo se incluye también la simulación de la radioinducción de tumores usando factores derivados de modelos de proyección aditivos/multiplicativos [Mil89]. Las distribuciones de edad de incidencia (Gompertz), tiempo de duplicación inicial (log-normal) y umbrales de detección (log-normal) se calculan a partir de datos ajustados del ensayo *Swedish Two-county* [Tab81].

El modelo reproduce razonablemente bien los resultados del ensayo en cuanto a prevalencia, tasas de cánceres de intervalo y distribución de tumores detectados en cribado. Sin embargo, para reproducir los datos en mujeres entre 40-49 años, se requiere un umbral (media de las distribuciones log-normal) un 40 % más alto que en mujeres mayores. Finalmente, predice un tiempo de duplicación que crece considerablemente con la edad y muestra un beneficio debido al cribado para mujeres mayores de 40 años.

Nuevamente, Jansen y Zoetelief [Jan97], utilizando el *software* desarrollado anteriormente, abordan un estudio riesgo-beneficio desde un punto de vista no económico. En este artículo se entiende por beneficio la reducción de mortalidad por cáncer de mama y por riesgo la radioinducción de tumores, por lo que no se trata de un estudio coste-beneficio *sensu stricto*. Calculan la diferencia en los tamaños tumorales detectados clínicamente y mediante cribado, empleando las tablas de supervivencia de Tabar *et al.* [Tab92] para el análisis. Para modelar la radioinducción de tumores utilizan modelos de proyección aditivos y multiplicativos, obteniendo los factores necesarios de un estudio canadiense sobre

pacientes con tuberculosis expuestos a fluoroscopia [Mil89]. El análisis es, quizás, demasiado simplista desde el punto de vista de la justificación pues, como hemos señalado anteriormente, la principal dificultad proviene del sobrediagnóstico/sobretratamiento y la morbilidad que ocasiona en la población, aspecto no cuantificado en este análisis.

Otros estudios coste-utilidad basados en microsimulaciones emplean el *software* MICROLIFE. Es el caso del estudio de Szeto y Devlin [Sze96], que utiliza distribuciones de estadios y mortalidad de Van Ootmarssen *et al.* [Oor90] y datos de incidencia del periodo 1985-89 en Nueva Zelanda (periodo previo a la implantación del cribado). Las tablas de mortalidad para causas diferentes del cáncer de mama son del periodo 1985-1987, también en Nueva Zelanda. La sensibilidad se fija a valores de 0.70 o 0.95 para tumores con diámetros inferiores o superiores a 10 mm, respectivamente, mientras que la especificidad se toma como 0.92 (teniendo en cuenta el 8% de falsos positivos que implican biopsias quirúrgicas). A los años de vida ganados mediante el cribado se les aplica una reducción del 5% (y hasta el 10%) atribuido a la tasa de descuento social, comúnmente empleada en estudios de carácter económico. Por su parte, los costes incluyen el de prueba/mujer y el del tratamiento, este último basado en protocolos desarrollados por Salked y Gerard [Sal94], que incluyen 10 modalidades de tratamiento (2 para CDis y 8 para tumores invasivos). En las simulaciones se asume que todos los CDis son tratados con la opción más barata y todos los tumores invasivos con la más cara.

Si bien los costes totales no muestran grandes diferencias entre las diferentes opciones, en términos del coste marginal sí que son importantes: 1 año de vida adicional en la configuración 50-64 años y tiempo de intervalo 2 años cuesta alrededor de 22264 \$ más que en la configuración 50-64 años y tiempo de intervalo 3 años. En cualquier caso, el análisis coste-utilidad no muestra un ahorro económico frente a la opción alternativa que es no hacer cribado. Este análisis se puede considerar una primera aproximación al problema de la justificación real de los programas de cribado, pero adolece de una falta de precisión en la estimación del detrimento que supone el sobrediagnóstico/sobretratamiento, más allá del coste que implican los tratamientos asociados al mismo. En este sentido, los años de vida de estas mujeres deberían ponderarse utilizando, por ejemplo, el concepto de *QALY*, como hemos indicado en otros casos.

El análisis coste-utilidad de diferentes estrategias combinando la mamografía y el examen físico en términos de años de vida ganados es abordado por Shen y Parmigiani [Shen05]. En este modelo, el momento del inicio del cáncer se obtiene deconvolucionando la curva de incidencia de detección clínica con la distribución de duración de la fase preclínica. Esta fase preclínica se modela mediante una distribución exponencial, explorándose también modelos log-normales en un análisis de sensibilidad. Para el ajuste se utilizan datos de los EAC canadienses y el ensayo Nijmegen [Col84, Ver84]. La sensibilidad se asume dependiente del tamaño y la edad a la que se realiza el test mediante un modelo de regresión logit, introduciéndose la variabilidad en la población a través de la función beta [Haz01]. Tras el diagnóstico se calculan la supervivencia y los años de vida ajustados por calidad (factores correctores que tienen en cuenta el detrimento). Esta supervivencia se calcula en función de la estrategia de tratamiento de acuerdo con el NHI (*Consensus Conference on Early Breast Cancer, 1991*). Los resultados de las simulaciones muestran que la mamografía bienal puede ser efectiva si se ayuda de exploración física anual.

Volviendo a los estudios coste-beneficio, Jacobi *et al.* [Jac06] parten de un programa de referencia que incluye mujeres entre 50 y 75 años, con tiempo de intervalo de 2 años,

que se supone efectivo desde el punto de vista coste-beneficio. Se simulan diferentes configuraciones y cohortes de edad en mujeres con antecedentes familiares de cáncer de mama sin mutaciones BRCA1/BRCA2. La incidencia se simula mediante el modelo de Jonker *et al.* [Jon03], que presenta un ajuste basado en la función normal acumulada para evaluar el riesgo en función de la edad de ser portador de cualquiera de las mutaciones BRCA1, BRCA2 o BRCAu en mujeres con antecedentes familiares de cáncer de mama u ovario. La sensibilidad se muestrea con distribuciones discretas en función de la edad, con valores de 64 % (menores de 50 años) y 85 % (mayores de 50 años), mientras que para la especificidad se utilizan valores entre 0.967 y 0.996 (menores de 50 años) y 0.997 (mayores de 50 años). Para los tiempos de duplicación del tumor se emplean valores medios de 80 días (menores de 50 años) y 157 días (mayores de 50 años). Se supone doble mamografía con un riesgo de radioinducción de  $1.65 \times 10^{-3}$  % en mujeres menores de 50 años y de  $1.14 \times 10^{-3}$  % si son mayores de 50 años. En cuanto a la supervivencia de las pacientes con cáncer de mama se distingue entre las que no se curan, para las que se utiliza una expectativa de vida media de 5.1 años, y las que si lo hacen, para las que se asume una expectativa de vida lineal con la edad de detección de  $(80.52 - \text{edad detección}) \times 0.996$ .

En el análisis coste-beneficio se utiliza el concepto *QALY*, asumiendo una reducción del 10 % en ese factor para mujeres que se curan y un 50 % para las que no. En este contexto, un programa se dice coste-efectivo para mujeres por debajo de 50 años si el coste del cribado por año de vida ganado por cada  $10^3$  mujeres cribadas es menor o igual que el del cribado de mujeres entre 50-52 años de la población general.

Se realizan simulaciones de  $10^5$  mujeres en cohortes que, o bien presentan mutaciones BRCA1, BRCA2 y BRCAu o bien no presentan ninguna de ellas, muestreando si tendrán cáncer a lo largo de su vida según la incidencia antes referenciada. Si tienen cáncer, se muestrea la edad de inicio (*onset*) que se distribuye según una normal, y el tiempo de duplicación, distribuido según una log-normal de valores medios dependientes de la edad, como indicamos anteriormente. La duración de la fase preclínica está relacionada con el tiempo de duplicación, permitiendo conocer la edad de detección clínica (conocidas la edad de *onset* y la duración de la fase preclínica).

Como salida del modelo se tiene el número de tumores detectados/no detectados, edad de detección y tamaño tumoral para cada grupo de riesgo. A estas salidas se le aplican las supervivencias anteriormente descritas para calcular los años ganados y corregirlos por un factor de calidad para generar valores de *QALY* según el modelo de Koning *et al.* [Kon91]. Se encuentra que el cribado anual es coste-efectivo si el riesgo acumulado hasta los 80 años de edad de padecer cáncer de mama es superior al 27 %, 32 %, 38 % y 47 % para mujeres mayores de 45, 40, 35 y 30 años, respectivamente. El análisis se muestra coste-efectivo respecto al programa de referencia, el cribado anual de mujeres menores de 50 años, cuando tienen al menos dos familiares afectados, uno de ellos en primer grado, diagnosticado con menos de 50 años (riesgo acumulado 33 %). El modelo es, sin duda, completo, si bien el cálculo de los años de vida ganados se realiza fuera de la simulación (estimación determinista), y la estimación de *QALY* según el modelo de Koning *et al.* [Kon91] es demasiado simplista.

Plevritis *et al.* [Ple06] implementan simulaciones Monte Carlo para evaluar el coste-efectividad de añadir resonancia magnética (RM) al cribado mamográfico anual de mujeres entre 25 y 69 años portadoras de mutaciones BRCA1/BRCA2. La RM es una técnica sensible a la detección de cáncer de mama pero, a cambio, incrementa los falsos positivos

y es 10 veces más cara que la mamografía. Los beneficios en salud son estimados en términos de *QALY*. En este trabajo, la probabilidad acumulada de desarrollar cáncer de mama a la edad  $a$  viene dada por:

$$P(a_{\text{clin}} \leq a) = 1 - \prod_{i=25}^a \left(1 - \frac{h_i}{100000}\right), \quad (1.2)$$

donde  $h_i$  son las incidencias clínicas por  $10^5$  mujeres a la edad  $i$  y  $a_{\text{clin}}$  es la edad de detección clínica. Con esta función se genera la edad de detección clínica en cada cohorte. Puesto que se toman conjuntamente tumores invasivos y CDis, se actúa con un factor arbitrario que determina la proporción de uno y otro para reproducir los datos del *Surveillance, Epidemiology and End Results program* (SEER) [Hol06]. El tiempo necesario para la detección clínica se modela con una distribución que depende linealmente del volumen del tumor, para el que se supone un crecimiento exponencial con un tiempo de duplicación distribuido según una función gamma de dos parámetros. Se introduce también la distribución de probabilidad del paso del tumor local a regional. De forma similar, se modela el riesgo de metástasis a distancia, condicionada a la diseminación regional. Para la estimación de los parámetros se emplea el método de máxima verosimilitud. En cuanto al modelo de detección por mamografía se propone una distribución de umbrales de detección truncada para diámetros tumorales  $d \leq 2$  mm. Los resultados de la aplicación de este modelo muestran que, con la adición de la RM, el cribado mamográfico anual es más coste-efectivo en portadoras de la mutación BRCA1 que en BRCA2 y, además, muy dependiente de la edad.

A partir del año 2006 se realiza una serie de trabajos enfocados a reproducir los resultados del cribado y terapia adyuvante en EE.UU. relativos al periodo 1975-2000 mediante modelos numéricos. Esta serie se inaugura con el modelo anteriormente descrito [Plev06b] aplicado ahora a este nuevo escenario, modelando los beneficios asociados a la terapia adyuvante mediante la introducción de factores en la supervivencia para las diferentes modalidades de tratamiento en función de la edad a la que se aplican: 0.72 ( $\leq 40$  años), 0.73 (40-49 años), 0.84 (50-59 años) y 0.91 ( $\geq 60$  años). En el caso de tamoxifeno se asume un factor de 0.69. Estas curvas de supervivencia son dependientes de la edad, tamaño tumoral y estadio SEER, introduciendo factores que tienen en cuenta la modalidad de detección y, como hemos descrito anteriormente, el tipo de terapia adyuvante. Con estas entradas, el modelo reproduce con bastante precisión la tendencia de mortalidad entre 1975 y 1995, pero no la posterior a este periodo. Este desacuerdo se atribuye a variaciones temporales en algunos de los parámetros de entrada. De forma similar, muestra una incidencia algo mayor que la reportada (alrededor de un 4% más) posiblemente debida a una sobreestimación de la incidencia. Las contribuciones de las terapias adyuvantes y el cribado en la reducción de mortalidad resultan similares a la luz de este artículo. El trabajo es elegante metodológicamente, pero carece de un modelado de la historia natural de CDis, aspecto este importante, dado que aproximadamente el 20% de los tumores detectados en cribado son CDis.

Esta serie se continúa con el trabajo de Berry *et al.* [Ber06], con un modelo de simulación bayesiano que utiliza distribuciones *a priori* y, en virtud de un algoritmo de rechazo, genera las distribuciones *a posteriori*, que son los verdaderos *inputs* de la simulación. El objetivo de este trabajo es simular el aumento de incidencia de cáncer de mama experimentado en los últimos 25 años en EE.UU. junto con la disminución de la mortalidad

(19.6 % entre los años 1990 y 2000). Esta última se atribuye a una combinación entre la mayor difusión de los programas de cribado y una mejora y mayor implantación de las terapias adyuvantes (QT y tamoxifeno en tumores estrógeno receptores, ER+).

Se simulan cohortes por año de nacimiento de 2 millones de mujeres seguidas entre los años 1975 y 2000. Para generar los casos prevalentes (mujeres con cáncer en 1975) se muestrea según datos SEER en función de la edad y los correspondientes al año simulado. Además, se tiene en cuenta la supervivencia por muerte natural (no debida al cáncer de mama) y la debida a cáncer de mama, asociada a la edad de detección y a la terapia aplicada según el año de detección. Las causas de muerte compiten de forma que, si la supervivencia por cáncer de mama generada para una historia resulta menor que la debida a muerte por otra causa, se clasifica como muerte por cáncer de mama. Respecto a la detección clínica, se muestrea según el modelo APC [Hol06], que incorpora tendencias en la incidencia y en el que la distribución depende de la edad de la mujer y año de nacimiento y procede de los datos SEER del año 1975. En cuanto a la detección por mamografía en un año determinado, la probabilidad depende de la edad de la mujer, utilizando datos del *National Breast and Cervical Cancer Early Detection Program 2002* (datos no publicados). Esta probabilidad depende también de si se trata de la primera mamografía y, de no ser así, del tiempo transcurrido desde esta. Si este tiempo supera los 3 años, la probabilidad será igual a la del primer test. Los tumores de intervalo dependen de la edad, tiempo desde la última mamografía y año. Su estimación se basa en diversas fuentes, entre las cuales están los resultados de dos estudios escandinavos [Hak95, Fri87] que estiman la probabilidad de tener cáncer de intervalo de un determinado estadio. Finalmente, para modelar el tratamiento, la supervivencia basal se modifica para incluir la QT adyuvante y la terapia con tamoxifeno para tumores estrógeno receptores.

El algoritmo empleado para generar las distribuciones *a posteriori* se basa en reproducir los valores anuales de mortalidad en el periodo 1975-2000, con una ventana de aceptación anual, junto con la condición de reproducir las tendencias en los periodos 1985-1990, 1990-1995 y 1995-2000. Si el resultado de esta simulación previa verifica estas condiciones, el conjunto de parámetros se acepta como una muestra de la distribución. De esta forma se aceptan hasta 176 muestras de los parámetros con los que se construyen las distribuciones mediante un método de *kernel density*.

Los resultados muestran que el descenso en la mortalidad se debe tanto al cribado como a la terapia adyuvante. Además, la contribución a la reducción de mortalidad se estima en  $(10.6 \pm 5.7)$  % para el cribado y de  $(19.5 \pm 5.4)$  % de la terapia adyuvante.

Este método de simulación bayesiano presenta la ventaja de adaptar las distribuciones utilizadas para reproducir las tendencias en la mortalidad observada e incorpora la información previa mediante distribuciones *a priori*, que son modificadas mediante simulaciones de ajuste. Se incorpora, además, el efecto de mejora en la supervivencia por modalidad de detección. No utiliza un modelo biológico de crecimiento del tumor ni modela las sensibilidades de detección desde un punto de vista físico, limitándose a incorporar la información exhaustiva anual obtenida de SEER para el periodo 1975-2000. Esto limita el modelo para su utilización fuera de este escenario.

Fryback *et al.* [Fry06] continúa esta serie de artículos monográficos simulando 2.95 millones de mujeres en cohortes de nacimiento entre los 20 y los 100 años seguidas durante el periodo 1950-2000, utilizando la distribución poblacional de Wisconsin. En este modelo se

asume que el CDis es un estadio inicial del invasivo, utilizando un modelo de crecimiento tumoral gompertziano asintótico con límite de crecimiento de 8 cm y umbral de la fase preclínica de 2 mm. La tasa de crecimiento del tumor se muestrea según la distribución gamma, sin introducir dependencias con la edad y el tipo de tumor (todos surgen como CDis). Con pasos de 6 meses se muestrea la diseminación metastásica mediante una distribución de Poisson descrita por Shwartz *et al.* [Shw75, Shw78, Shw81, Shw92] que es función del tamaño tumoral y de la tasa de crecimiento instantánea. El modelo de crecimiento tumoral se ajusta a los 4 estadios SEER caracterizados por el tamaño y el número de ganglios afectados. Por debajo de cierto tamaño el tumor se clasifica automáticamente como CDis. Este modelo se calibra utilizando las incidencias por edad dadas por SEER en el periodo 1975-1999. Cuando el tumor alcanza la fase de diseminación distal (supera los 5 ganglios positivos) se muestrea la muerte mediante una distribución de supervivencia basada en datos SEER 1975-1982. Esta distribución tiene una media de 1.95 años y mediana de 5.22 años.

La tasa de tumores que entran en la fase preclínica está dada por un modelo APC en ausencia de cribado. Este modelo no predice la incidencia sino el *onset* o momento de entrada del tumor en la fase preclínica. No todos los tumores pasan a esta fase, por lo que el cociente entre los tumores incidentes y aquellos que llegan a la fase preclínica va desde 1 hasta valores en el entorno de 1.2.

Respecto a la detección del cáncer de mama, el modelo utiliza datos de cribado correspondientes a los años 1975-2000. La probabilidad de detección en cribado es una función del tamaño tumoral y se introducen diferentes valores para mujeres menores de 50 años. Además, se incorporan diferencias en la sensibilidad y detección clínica en función del año correspondiente.

Para el modelado del tratamiento, se utiliza dos estados: curación y no curación. En el primer caso se interrumpe la progresión del tumor, mientras que en el segundo progresa hasta enfermedad distal. En cuanto al tratamiento aplicado, sigue la estadística correspondiente al año y edad de la mujer en el momento de la detección. La evolución del tratamiento va desde la era de la mastectomía con radioterapia hasta la inclusión de terapias adyuvantes de quimioterapia y tamoxifeno.

Para modelar estos efectos se utilizan resultados de metanálisis. En concreto se consideran datos de mortalidad por causas diferentes al cáncer de mama, incidencia e implantación del cribado y la terapia obtenidos del *Cancer Intervention and Surveillance Modelling Network* (CISNET) y otra información sobre los efectos de la terapia adyuvante y ER obtenidos de la bibliografía. Los parámetros de la historia natural del tumor, en particular las probabilidades de detección y la distribución de estadios sin cribado se ajustan para reproducir los resultados CISNET, que incorporan el efecto del cribado, mediante un procedimiento heurístico iterativo.

Los resultados de la simulación indican que el cribado produce una reducción de mortalidad por cáncer de mama del 20.3% (con terapias correspondientes a 1975) y las terapias adyuvantes del 20.8%.

Este artículo presenta un modelo de crecimiento tumoral y de sensibilidad de detección por cribado, lo que le da un carácter más fundamental y una mayor validez externa que otros modelos similares. Además, introduce el concepto de LMP (tumores con potencial

maligno limitado) para simular aquellos tumores que no progresan y que están directamente relacionados con el sobrediagnóstico. En lo que respecta a la supervivencia, no se introducen mejoras por modalidad de detección. Un aspecto discutible es la introducción del tipo CDis como una fase inicial de todos los tumores.

Greuter *et al.* [Gre10] simulan programas de cribado mamográfico en mujeres de alto riesgo (mutaciones BRCA1/BRCA2 o con antecedentes familiares) con el *software* SIM-RISC. Se cree que estas mujeres tienen una mayor susceptibilidad genética a la radioinducción del cáncer, por lo que el sometimiento a mamografías desde edades muy tempranas es controvertido.

Para introducir la radioinducción del cáncer se utiliza un modelo de exceso relativo de riesgo [Pres02], con una relación cuadrática con la edad de la mujer y sin dependencia con la edad en el momento de la exposición. La distribución de duración de la etapa preclínica se muestrea de acuerdo con una log-normal, cuyos parámetros se derivan de Peer *et al.* [Pee96] con tiempo medio de duplicación de 1.9 años y la probabilidad de padecer cáncer a una determinada edad, distribuida según una normal. Finalmente, la sensibilidad en la fase preclínica se modela como:

$$S(t) = S_0 \left( \frac{S_1}{S_0} \right)^{\frac{t}{T}}, \quad (1.3)$$

donde  $S_0$  la sensibilidad en el *onset* del tumor,  $S_1$  es la sensibilidad en la fase clínica,  $T$  es la duración del periodo preclínico y  $t$  el momento de evaluación. Los resultados se comparan con tres estudios de la bibliografía [Cor06, Kri06, Lea05] en términos de porcentaje de tumores detectados en cribado y de intervalo en estos estudios predicho por la simulación, llegando hasta un  $71 \pm 4\%$  o incluso el  $81 \pm 6\%$  si se excluye la ronda de prevalencia. Estos resultados se consideran una validación del modelo al objeto de realizar estudios riesgo/beneficio en términos de tumores radioinducidos/detectados. Sin embargo, puesto que la comparación con los valores de la bibliografía se hace en base a tumores totales detectados, no hay forma de asegurar que la proporción radioinducidos/detectados es la predicha por el modelo, aspecto crucial en la justificación que se pretende.

Bayley *et al.* [Bai10] emplean el *software Breast Cancer Screening Simulator* (BCSS) con objeto de averiguar qué factores pesan más en los peores resultados de mortalidad asociados al cribado en mujeres de entre 40 y 49 años frente al grupo de 50 a 69 años; compiten, por un lado, el aspecto tecnológico de una menor sensibilidad en la detección debido a mamas más densas y, por otro, el factor biológico de tumores más agresivos y de crecimiento más rápido a edades tempranas. Este *software* implementa la simulación Monte Carlo de historias de mujeres con cáncer de mama, generando la historia natural de los tumores invasivos (no CDis) en términos de crecimiento y progresión del estadio clínico, edad en el momento del cribado, umbral de detección por mamografía en el momento del examen, edad de diagnóstico clínico, tamaño y tipo en el momento de la detección, tipo de tratamiento adyuvante recibido, edad de la muerte y causa (cáncer de mama u otras causas).

La distribución del tiempo de duplicación depende de la edad de detección clínica de manera lineal y el umbral de detección es función de la densidad de la mama, que a su vez se modela también linealmente con la edad de la mujer. Los ajustes del modelo se realizan de forma que se reproduzcan los datos de incidencia por edad del estudio SEER y la

distribución de tumores detectados en cribado del *Breast Cancer Surveillance Consortium* (BCSC). Para evaluar la bondad del ajuste se utiliza un test  $\chi^2$  con una función de mérito dada por:

$$f = \frac{1}{10} \sum_{i=1}^{10} (\chi_{\text{incidencia}}^2)_i + \frac{1}{4} \sum_{j=1}^4 (\chi_{\text{tamaño}}^2)_j, \quad (1.4)$$

donde  $i$  representa los 10 grupos de edad (35-39, 40-44, ..., 75-79) y  $j$  los cuatro grupos (30-49, 50-59, 60-69 y >70), mientras que  $\chi_{\text{incidencia}}$  se calcula para los resultados BCSS de incidencia frente a los datos SEER y  $\chi_{\text{tamaño}}$  para los resultados BCSS y datos obtenidos de BCSC referidos a tamaños tumorales en la detección. Esta función es optimizada por el método de Nelder-Mead [Nel65].

Los resultados muestran una mayor influencia de la peor detectabilidad asociada a mamas densas frente a la de las tasas de crecimiento: el 79 % de la reducción de mortalidad se reproduce mediante cambios en la sensibilidad, mientras que el 21 % restante es reproducido mediante cambios en el tiempo de duplicación. Los detalles de este modelo en lo referente a las distribuciones concretas utilizadas son escasos en el artículo original y presenta la limitación adicional de la no inclusión de los CDis.

En 2012 Sigal *et al.* [Sig12] utilizan el *software* descrito en los dos artículos de Plevritis *et al.* [Ple06, Plev06b] para comparar diferentes estrategias (mastectomía radical bilateral, ooforectomía radical y cribado con mamografía anual desde los 25 años más RM) en mujeres con alto riesgo de padecer cáncer de mama/ovarios por mutaciones BRCA1/BRCA2.

Modela, como indicamos antes, el crecimiento del tumor y el riesgo de invasión local/regional/distal en función del estadio en el momento de la detección sintomática. A partir de la detección, utiliza curvas de supervivencia basadas en SEER. Para la simulación de la incidencia se tiene en cuenta el mayor riesgo que suponen las mutaciones BRCA1/BRCA2.

Los resultados de las simulaciones muestran una mayor eficacia en expectativas de vida de la ooforectomía, seguida de la mastectomía profiláctica en portadoras de la mutación BRCA1, si bien no ocurre así con las portadoras de la mutación BRCA2. El cribado muestra ganancias significativas, pero inferiores a las de otras estrategias.

Este modelo carece de la implementación del riesgo de cáncer radioinducido que, para estas mujeres con una supuesta mayor susceptibilidad y comienzo temprano del cribado, podría tener trascendencia en los resultados finales.

Tsunematsu y Kakehashi [Tsu15] emplean un modelo que divide la historia natural del tumor en 10 estados y evalúan las tasas de transición entre estos mediante métodos de máxima verosimilitud, utilizando registros del cáncer de Japón y EE.UU. Las tasas de supervivencia para mujeres con cáncer se estiman de las supervivencias a 10 años procedentes de datos publicados por diferentes fuentes, ajustados logarítmicamente. Las simulaciones arrojan una ganancia de 13 días de vida para las mujeres japonesas mayores de 40 años debida al cribado y de 25 días para las estadounidenses, con un número de falsos positivos del 65 % y 53 % respectivamente.

Este modelo carece de una descripción analítica de las distribuciones empleadas, trabajando únicamente con tasas de transición entre los estados que conforman la historia

natural del tumor. Se fundamenta, únicamente, en valores estadísticos, sin un sustento biológico importante, lo que limita su aplicabilidad en un ámbito más general.

Huang *et al.* [Hua18] utilizan el modelo de Berry *et al.* anteriormente descrito [Ber06] para evaluar diferentes estrategias de cribado en términos de beneficio (años de vida ganados) frente a daños (falsos positivos y sobrediagnóstico) en diferentes configuraciones del programa de cribado: tiempos de intervalo de 1, 2 y 3 años (y combinaciones de estos), edades de comienzo en el rango de 40 a 50 años y de finalización entre 70 y 80 años. Obtienen una reducción de mortalidad debida al cribado del 16% y del 23% debido a las terapias adyuvantes.

Citaremos por último el modelo de Ishende y Humphreys [Ish19], que modela un crecimiento continuo del tumor mediante una función exponencial y un parámetro de crecimiento distribuido según una función gamma. La detección clínica se introduce mediante un modelo lineal con el tamaño del tumor y la detección en cribado por medio de una función logística, dependiente del tamaño del tumor, la densidad de la mama y un término de interacción entre ambas. Los ajustes se fundamentan en 1901 casos procedentes de las regiones de Suecia y, en particular, Estocolmo, referidos a mujeres entre 50 y 74 años. El modelo pretende únicamente reproducir esta cohorte de mujeres mediante microsimulaciones, por contraposición a los modelos de Markov, de los que estos autores señalan la limitación de presentar tasas de transición entre estados constantes.

### 1.1.2. Modelos de Markov

En estos modelos la historia natural del tumor se describe mediante un proceso de Markov [Uhr10], con diferentes estados posibles, más o menos artificiales, para los que se calculan las distribuciones de tiempos de permanencia en los mismos, utilizando una matriz de probabilidades de transición entre estados y resolviendo las ecuaciones de Chapman-Kolmogorov. Usando métodos de máxima verosimilitud o mínimos cuadrados generalizados se ajustan las distribuciones obtenidas para reproducir resultados conocidos de EAC, recopilaciones epidemiológicas, etc. Debe hacerse notar que estos modelos se emplean únicamente para reproducir la historia natural del tumor para, posteriormente, llevar a cabo simulaciones detalladas o bien estimaciones analíticas empleando estas distribuciones obtenidas.

Uno de los primeros trabajos en este ámbito es el de Habbema *et al.* [Hab85], basado en el paquete MISCAN (*Microsimulation of Screening Analysis*). Como en todo proceso de Markov, las transiciones se consideran independientes de la edad de la mujer, pero se puede introducir esta dependencia *ad hoc* mediante un polinomio cuadrático. Utiliza funciones de Erlang, Weibull y escalón para modelar los tiempos de permanencia en los diferentes estados e incluye tablas de incidencia y mortalidad por causas diferentes al cáncer de mama. Resulta un modelo muy versátil en cuanto a su configuración, mostrándose, sin embargo, un tanto artificial en los estados definidos. Además, considera el CDis como un estado previo al invasivo, aspecto este un tanto controvertido.

Otros autores como Van Oortmarsen *et al.* [Oor90] emplean este modelo con parametrizaciones basadas en datos de casos y controles del estudio Utrecht y Nijmegen [Col84, Ver84] que, por no ser un ensayo controlado, resulta sesgado. Estos autores asumen el efecto pronóstico en la detección por cribado, entendido como un sesgo a favor de

este y que asume que los tumores que se detectan en cribado tienen un mejor pronóstico que los detectados clínicamente, aunque esta mejora no está asociada únicamente al tamaño del tumor.

Paci *et al.* [Pac95] utilizan también el paquete MISCAN para simular y ajustar los datos del programa de cribado llevado a cabo en el distrito de Florencia (FDP) entre los años 1975-1986, aunque la distribución de tumores detectados clínicamente introducida no corresponde exactamente con la requerida por MISCAN. Emplean, además, una sensibilidad dependiente del tipo histológico y del estadio, pero independiente de la edad. Con estos valores se estima el tiempo de duración de la fase preclínica o *sojourn time*:

$$t_{\text{sojourn}} = \frac{1}{S} \frac{\text{TDP}}{I}, \quad (1.5)$$

donde  $S$  representa la sensibilidad, TDP la tasa de detección obtenida en la ronda de prevalencia (primera ronda) e  $I$  la tasa de incidencia (rondas posteriores).

Las discrepancias observadas entre los resultados del programa FDP y las predicciones del modelo, en lo que hace referencia a la proporción de tumores en las rondas de incidencia, son atribuidas a que el tiempo de intervalo empleado en dicho programa de cribado no corresponde con el valor medio de 30 meses utilizado en los ajustes.

Uno de los artículos de referencia acerca de los estudios coste-efectividad, en el marco de los modelos de Markov, es el de Lindfors *et al.* [Lin95], que está basado en el *software* Smeltree 2.9 y que utiliza datos del *National Cancer Institute* (NCI) para los valores de incidencia y prevalencia, así como para modelar las muertes por causas diferentes al cáncer de mama. El análisis coste-efectividad en este modelo se realiza en términos de coste marginal por año de vida salvado, calculado como el cociente entre el coste marginal del cribado (diferencia entre el coste total del cribado y el coste total del grupo de observación en el periodo observado) y la efectividad marginal (diferencia en los años de vida acumulados en ambos grupos). Con estas premisas se analizan diferentes configuraciones de los programas de cribado y resulta más ventajosa, en lo que al coste-efectividad se refiere, la mamografía anual en mujeres entre 40 y 49 años, pasando a bienal en el rango 50 a 79 años. En términos de efectividad marginal, la mejor estrategia es la mamografía anual en el grupo de 40 a 79 años.

Si bien este es uno de los estudios de coste-efectividad más completos, no introduce el detrimento asociado al cribado. Si utilizamos el concepto de *QALY* podemos aplicar un factor de 0.83, derivado de Previtis *et al.* [Ple06] para cáncer de mama a aquellos años de vida de mujeres sobrediagnosticadas/sobretreatadas. Suponemos 3 sobrediagnósticos por vida salvada y 30 años por cada una de estas. Con estos ingredientes resultaría que una efectividad marginal de 0.1 años como las reportadas en este artículo, se traduciría en  $(0.01 \times 30) - (0.01 \times 30 \times 0.83) = 0.05$  años menos en el grupo de cribado respecto al observado, lo que significa hasta un 50% menos en términos de *QALY*. Aunque se trata de una estimación muy grosera, muestra cómo quedan comprometidos los resultados si se introduce el detrimento del sobrediagnóstico en los cálculos.

Basado también en modelos de Markov de progresión tumoral de 3 y 6 estados encontramos el trabajo de Duffy [Duf97]. Los datos para el ajuste del mismo se obtienen del EAC *Swedish Two-County* y el modelo proporciona la distribución subyacente de los diferentes estados, la duración media de la fase preclínica, las sensibilidades y las especificidades.

Los resultados en este modelo arrojan unos tiempos medios de duración de la fase preclínica de 2.44 años (40-49 años), 3.70 años (50-59 años) y 4.17 años (60-69 años) y una sensibilidad de 83 %, 100 % y 100 % para esos rangos de edad, respectivamente.

En una segunda fase, se trabaja con un modelo de Markov de 5 estados, con lo que la resolución del sistema de ecuaciones de Chapman-Kolmogorov se vuelve muy complicada, por lo que se emplean métodos de regresión no lineal para estimar los parámetros característicos del modelo [Duf95]. Con estos modelos y la utilización de datos de supervivencia ajustados mediante regresión de Cox [Bag10], calculan la mortalidad para diferentes tiempos de intervalo (1, 2 y 3 años) y la comparan con la de los diferentes EAC. Los resultados muestran una sensibilidad de la mamografía menor en mujeres entre 40 y 49 años, a la vez que una progresión de los tumores más rápida en estas edades. También el grado de malignidad progresa más rápidamente en mujeres de este rango de edad. Se concluye que el tiempo de intervalo óptimo para mujeres entre 40 y 49 años debe ser menor de 2 años. El modelo puede hacerse más completo incluyendo además el tamaño tumoral en la definición de los diferentes estados.

Boer *et al.* [Boe98] emplean el paquete MISCAN para un estudio coste-beneficio de la actual configuración del programa de cribado inglés [UKP12]. Retrasan la edad de finalización de 64 a 69 años y cambian el tiempo de intervalo de 3 a 2 años. El análisis se realiza en términos de coste en libras por año de vida ganado y de coste marginal. Los tiempos de permanencia en las diferentes fases preclínicas se supone que siguen distribuciones exponenciales, determinadas a partir de los valores medios de duración de la fase preclínica, mientras que las tasas de transición entre la fase preclínica y clínica, son inferidas de la distribución de tumores detectados clínicamente. Además se introduce una mejora en la supervivencia asociada a la modalidad de la detección derivada de los ensayos suecos. Las distribuciones de edades, estadios y supervivencia se obtienen del ensayo de Utrecht y Nijmegen, mientras que para la estimación del número de años ganados por muerte prevenida se utilizan tablas de mortalidad por causas diferentes al cáncer de mama. Este estudio carece de una evaluación de *QALY* para incluir el detrimento asociado a estos programas (análisis coste-utilidad) debido al tratamiento y el sobrediagnóstico, junto a otras limitaciones importantes.

Mandelblatt *et al.* [Man06], continuando con la serie de trabajos encaminados a estudiar el cribado y las terapias adyuvantes en EE.UU. en el periodo 1975-2000, desarrollan su modelo utilizando el paquete SPECTRUM para generar cohortes de mujeres según su fecha de nacimiento y evaluar la mortalidad en el periodo de interés. En el caso de mujeres con cáncer de mama, se muestrea el estadio de detección clínica, el estado de receptor de estrógenos del tumor, que se supone invariante y la duración de la fase preclínica según la distribución exponencial, tomando como *lead time* en caso de detección por cribado, el umbral de este periodo.

Para tratar la incidencia en ausencia de cribado se emplean los datos del NCI, mediante la inclusión de tendencias temporales para cada cohorte según el año de nacimiento. Esta distribución depende del estadio SEER, ER y de la edad de la mujer en el momento de la detección clínica. El cribado se introduce en términos de sensibilidad y especificidad del test, produciendo un desplazamiento hacia estadios más tempranos del tumor. Se utilizan para este propósito distribuciones del estadio en función de la edad de detección por cribado y *lead time*. No se introduce, sin embargo, efecto por modalidad de detección

(*length bias*). Finalmente, la supervivencia se modela usando datos SEER dependientes de la edad, estadio, el estado de receptor de estrógenos del tumor y modalidad de tratamiento.

Respecto a la historia natural del tumor, se hacen asunciones mínimas que suponen distribuciones de tiempo de permanencia exponenciales en cada estadio. Los tiempos medios de permanencia en cada estado se calculan como la inversa de la tasas de transición entre ellos, con datos procedentes de ensayos aleatorios controlados, si bien los resultados del pre-ajuste muestran una proporción de CDis más baja que la encontrada en los datos SEER. En cuanto a la construcción de las distribuciones de estadios de detección por cribado condicionada por *lead time*, se parte de las conocidas para detección clínica y se emplea la información de tiempos de permanencia en cada uno de los estadios (CDis, local, regional y distal) mediante el teorema de Bayes. Las distribuciones de duración de la fase preclínica y la sensibilidad son ajustadas mediante simulaciones previas intentando reproducir los datos de incidencia y distribución de estadios conocidas.

En este modelo, la sensibilidad de la detección por cribado se define como el cociente de los verdaderos positivos (VP) y el número de mujeres con tumor, entendidas como aquellas con tumor en la fase preclínica. Esta sensibilidad se asume dependiente de la edad de la mujer en el primer test, aunque es independiente de esta en exámenes posteriores. No se introduce dependencia de la detección con el tamaño tumoral.

Como resultado principal, el modelo muestra una reducción de la mortalidad por cáncer de mama atribuible al cribado del 12.4%, mientras que la asignable a las terapias (tamoxifeno y QT) es del 14.6% en el año 2000. Al igual que el resto de publicaciones monográficas descritas referentes a la reducción de mortalidad en EE.UU. en el periodo 1975-2000, el modelo resulta de escasa validez externa más allá de la población y periodo estudiados. La historia natural del tumor no incluye un modelo de crecimiento tumoral, sino una transición de estadios SEER basada en un modelo de Markov. Estos estadios implican, como señalamos anteriormente, que el CDis es el origen de todo tumor invasivo. Por otra parte, el modelo de sensibilidad carece de realismo, al ser independiente del tamaño tumoral en la detección.

El último de los modelos de esta serie es el de Lee *et al.* [Lee06]. En esta simulación se clasifican los estados de salud en  $S_0$  (sin enfermedad o indetectable),  $S_P$  (preclínica),  $S_C$  (clínica) y  $S_D$  (muerte). Se asume que la transición sigue esta secuencia y que la ganancia del cribado procede del adelanto que supone la detección en un estado más temprano. Para las simulaciones se emplean distribuciones condicionadas, procedentes de diferentes fuentes y que corresponden a la supervivencia, dependiente del estadio, la sensibilidad de la mamografía, la duración de la fase preclínica, los estadios con y sin cribado, según la clasificación del *American Joint Committee on Cancer* (AJCC), la implantación del cribado y la terapia en el periodo de interés (1975-2000) y el estado de receptor de estrógenos del tumor.

El origen de las distribuciones de incidencia son los datos SEER del periodo 1975-2000 en función del estadio. De igual forma se procede con la supervivencia. En cuanto a la mortalidad por causas diferentes al cáncer de mama, se utilizan datos procedentes del CISNET.

El modelo no trata explícitamente el CDis, suponiendo que la fracción de estos tumores que pasa a invasivos es pequeña y su efecto queda incluido en el modelado de estos últimos. La distribución de duración de la fase preclínica sigue un modelo exponencial, con media

dependiente de la edad, obtenida de datos de los EAC. En cuanto a la sensibilidad, se define un modelo dependiente de la edad basado en datos BCSC del periodo 1996-1998, modelo que se modifica en función del periodo o año en el que se detecta el tumor por cribado. Por su parte, los patrones de diseminación de cribado y tratamiento se obtienen de datos CISNET y NCI. Para validar el modelo se simulan 8 EAC, y los resultados de mortalidad quedan dentro de los intervalos de confianza que proporcionan estos ensayos.

Este modelo atribuye un efecto en la reducción de mortalidad por cáncer de mama del 23 % al cribado y del 11 % a la terapia, aunque presenta, como todos los anteriores de esta serie, problemas de validez externa. La ausencia de un modelo biológico del tumor y un modelado de la sensibilidad algo más realista son también objeciones que ponen en duda el carácter fundamental de estas simulaciones.

En el año 2011 aparecen nuevos trabajos de simulación [Rav11] que emplean una versión diferente del código MISCAN, para comparar dos estrategias en la población del Reino Unido: cribado anual frente a cribado cada 3 años, comparando los resultados con los dos EAC disponibles hasta la fecha. En esta aproximación, MISCAN utiliza el concepto de “universo paralelo”, consistente en simular cómo cambia la historia individual de una mujer con cribado frente a la ausencia del mismo. Además, incluye un submodelo para el CDis. Simulan la historia de mujeres sin cáncer y se ajusta el modelo a los datos demográficos, incluyendo posteriormente la incidencia del cáncer. La historia natural del tumor se modela con un crecimiento continuo, bajo la hipótesis de que todo tumor tiene un diámetro (denominado “diámetro fatal”) alcanzado el cual, el diagnóstico y tratamiento no se traducen en un aumento de años de vida. La variabilidad entre tumores es modelada con distribuciones de tasas de crecimiento de los mismos, umbral de detección en mamografía, diámetro de detección clínica, diámetro fatal y supervivencia desde diámetro de detección fatal. En esta aproximación, se supone que aquellas mujeres a las que se les detecta el tumor antes de que se alcance el diámetro fatal se curarán, muestreándose su mortalidad según distribuciones de muerte por causas diferentes al cáncer de mama. Así, el beneficio del cribado se traduciría en un mayor número de tumores con diámetro pre-fatal.

En cuanto a la sensibilidad, se supone 0 por debajo del umbral de detección por mamografía y 100 % por encima. Esta distribución depende de la edad de la mujer y del año. Se entiende que esta última dependencia introduce una relación entre la sensibilidad y el momento del cribado. En esta aproximación, el umbral decrece tanto con el año como con la edad de la mujer. El tratamiento adyuvante se simula mediante un aumento en el diámetro fatal. Con estos condicionamientos, el modelo proporciona un riesgo relativo de 0.83 en el grupo de cribado anual frente al de cribado cada 3 años. El resultado de *UK Frequency Trial* es de 0.93 (0.63,1.37) con una confianza del 95 %.

Esta aproximación de diámetro fatal es una simplificación excesiva, pues la idea de que tumores detectados con tamaños menores que este se “curen”, y la mortalidad se deba a causas diferentes al cáncer de mama, y aquellas con diámetro de detección justo mayor no obtengan beneficio del tratamiento no parece realista. Junto a esta aproximación, el modelado de la sensibilidad de la detección también resulta discutible.

Cabe destacar el trabajo de Gunsoy *et al.* [Gun12], que emplea los datos derivados del EAC *UK Age Trial* para obtener la probabilidad de detectar un tumor (CDis o invasivo) en una fase preclínica, la duración media de la fase preclínica (para tumores CDis no progresivos, progresivos e invasivos), la proporción de CDis detectados en cribado que son

no progresivos y la proporción de cánceres detectados que no se presentaría clínicamente en ausencia de cribado (sobrediagnóstico).

El modelo de Markov definido presenta 6 estados, a saber: sano, CDis no progresivo detectable por cribado, CDis no progresivo detectable clínicamente, CDis progresivo detectable por cribado, tumor invasivo preclínico detectable en cribado y tumor invasivo detectado clínicamente.

Para simular el sobrediagnóstico se pasa a un modelo de Markov de 10 estados y se obtienen las intensidades de transición de muy diversas fuentes. Se simulan 10 millones de mujeres seguidas durante 15 años desde los 40 años de edad.

Los resultados muestran una sensibilidad de la mamografía para cánceres invasivos del 90 % y del 82 % para CDis. En cuanto a la duración media de la fase preclínica, sería de 1.3 años y 0.11 años para CDis no progresivos y progresivos respectivamente. Para invasivos es de 0.8 años. La proporción de CDis detectados en cribado que son no progresivos es del 55 % en la primera ronda y el 40 % en rondas sucesivas. Finalmente, se concluye que el 0.7 % de los tumores detectados en cribado sería sobrediagnóstico, pudiendo ser hasta el 2.9 % si se modifican los *inputs* (análisis de sensibilidad del modelo). En cualquier caso, este valor es muy bajo, lejos de las estimaciones más realistas obtenidas de los EAC y estudios poblacionales, que lo sitúan en el entorno del 19 % [UKP12].

Estos modelos de Markov se revelan, por tanto, útiles para simular procesos como la evolución tumoral, proporcionando además información muy valiosa referida a diferentes configuraciones posibles de los programas de cribado. Sin embargo, el modelo no contempla una dependencia con la edad en la probabilidad de transiciones entre estados, pese a lo cual, se encuentra una menor duración de la fase preclínica en mujeres jóvenes. Además, presenta algunas limitaciones metodológicas derivadas de la estimación de los parámetros de tasas instantáneas de transición obtenidas mediante métodos de regresión no lineal. Por otro lado, realiza la asunción de incidencia de la fase preclínica constante a lo largo de la vida.

En 2015, Arrospeide *et al.* [Arr15] utilizan un *software* comercial genérico de simulaciones (*Arena Simulation Software*) para modelar la población sometida a cribado en el País Vasco entre los años 1996 y 2011, utilizando registros propios para esta estimación. Estiman la sensibilidad del cribado y detección clínica basándose en los tumores detectados de una u otra forma respecto a los resultados de las pruebas diagnósticas y utilizan el modelo tumoral de Lee *et al.* [Lee06], descrito por un modelo de Markov de 4 estados. Para estimar la supervivencia emplean datos debidos a VilaprinYO *et al.* [Vil09] y asignan a cada mujer dos edades de muerte, una por cáncer de mama y otra por causas diferentes al mismo, asociadas a la edad y estadio del tumor detectado, prevalenciando la menor de ellas.

Con estos ingredientes, se estima una reducción de la mortalidad por cáncer de mama del 16 % en esta población y resulta un sobrediagnóstico del 4 % de los tumores de detectados en cribado. El modelo resulta poco extrapolable, pues se basa en registros de escaso tiempo de seguimiento y no tiene en cuenta las mejoras en la supervivencia producidas por los nuevos tratamientos adyuvantes, por estar basado en registros del año 2011.

Gocgun *et al.* [Goc15] emplean un modelo de Markov de la historia natural del tumor de 8 estados para evaluar el coste-beneficio del cribado. Este modelo implementa un estado

saludable, 4 estados preclínicos, 1 estado sintomático, 1 estado de muerte por cáncer de mama y, finalmente, otro estado de muerte por causas diferentes al cáncer de mama. Los parámetros de las distribuciones exponenciales se ajustan mediante métodos de máxima verosimilitud, usando datos de los ensayos *Canadian National Breast Screening Studies* (CNBSS) [Mil02] correspondientes al periodo 1980-1989. La sensibilidad, tanto del cribado como sintomática, se introduce a partir de los datos de Taghipour *et al.* [Tag13], agrupada en intervalos de edad de 10 años. Finalmente, no utiliza un modelo explícito de supervivencia, pues esta viene dada por las tasas de transición a los estados de muerte por cáncer u otras causas del propio modelo de Markov.

A partir de este modelo se deduce que las mujeres de entre 40 y 49 años no ofrecen beneficio en años de vida/\$. La mejor configuración resulta ser la que incluye mujeres de entre 50 y 69 años con un tiempo de intervalo de 5 años, siendo descartable, en todos los casos, el cribado anual.

Este modelo considera únicamente los costes asociados al cribado, tratamiento y diagnóstico y no realiza una conversión a *QALY*. Introduce el sobrediagnóstico como una *input* más de la simulación, fijándolo en 0.20 casos por cada cáncer detectado en el cribado.

El modelo de Taghipour *et al.* [Tag17] se basa en cadenas de Markov de 4 estados, donde no se incluye muerte por cáncer de mama, pues se ocupa únicamente de resultados relacionados con la detección. Intenta reproducir los resultados de los ensayos canadienses I y II, tanto para el grupo de intervención, que utilizan para obtener los parámetros del modelo mediante métodos de máxima verosimilitud, como del grupo de control, con el que consiguen un segundo ajuste por mínimos cuadrados, variando los parámetros en los intervalos de confianza generados con la primera simulación del grupo de cribado.

### 1.1.3. Simulaciones deterministas y otros modelos semi-analíticos

El trabajo de Fett *et al.* [Fet01] debe enmarcarse en lo que se ha dado en llamar simulaciones deterministas, entendidas como aquellas en las que los parámetros de la simulación no se generan de acuerdo con distribuciones aleatorias, sino que son *inputs* dependientes de la edad obtenidos por interpolación lineal, tanto para la sensibilidad como para el *sojourn time*. Utilizan la supervivencia tras detección clínica, independientemente de la modalidad de detección. En el caso de Fett *et al.*, los autores obtienen esa información de datos del sur de Australia. Estas curvas son función de la edad de detección y del *sojourn time* y permiten deducir la probabilidad de muerte mediante la interpolación en función del momento de la fase preclínica en que se detecta el tumor por cribado. Este trabajo reproduce bastante fielmente la mortalidad acumulada por año de seguimiento en los grupos de cribado y control simulados.

Otros trabajos desarrollan modelos semi-analíticos, generalmente encaminados a obtener información de algún aspecto relacionado con el cribado mamográfico. En este sentido, Michaelson *et al.* [Mic99], parten de los datos de Tabar *et al.* [Tab92b, Tab95b, Tab96] para derivar la supervivencia en función únicamente del tamaño tumoral en la detección. En este análisis, características de los tumores como el estadio, malignidad, tipo histológico, etc., juegan el papel de variables de confusión correlacionadas con el tamaño tumoral. Es, por tanto, posible calcular la fracción de pacientes que presentarán metástasis a partir de los datos de tamaño tumoral y tiempo medio que una célula tarda en producir una

metástasis a distancia. En este contexto se llevan a cabo las simulaciones explorando los tiempos de intervalo óptimos. Se encuentra que en ausencia de cribado se tiene un 45 % de incidencia de metástasis. Para el cribado con tiempo de intervalo de 3 años hay una reducción en las metástasis del 14 %, 22 % para tiempo de intervalo 2 años y 51 % para tiempo de 1 año. Estos resultados se muestran razonablemente estables frente a variaciones en las distribuciones radiobiológicas.

El análisis coste-beneficio realizado para mujeres mayores de 50 años suponiendo un coste de 100\$ por mamografía, muestra un elevado beneficio en la reducción de mortalidad, con un coste por año de vida salvado para tiempos de intervalo entre 1 y 3 años muy similar y costes marginales también ajustados. A partir de tiempos de intervalo de 3 meses la reducción de mortalidad es casi insignificante y el coste medio por año y el marginal se disparan. Se concluye así que es más eficiente aumentar la frecuencia del cribado que mejorar la sensibilidad de la detección.

Una vez más, el estudio carece de una estimación de los costes del tratamiento y del detrimento del sobrediagnóstico/sobretretamiento. Incluye sólo los costes derivados de la mamografía, pero es, por el contrario, una buena justificación teórica del uso de curvas de supervivencia únicamente dependientes del tamaño tumoral en la detección, sin incluir beneficios adicionales derivados de la modalidad de detección, siempre controvertidos.

Existen otros modelos que pretenden estimar parámetros críticos del cribado, como la duración de la fase preclínica, basados en métodos de máxima verosimilitud o mínimos cuadrados generalizados, como es el caso del de Shen y Zelen [She05]. En este trabajo se utilizan los valores del EAC HIP, iniciado en 1960 con  $62 \times 10^3$  mujeres de entre 40 y 64 años, para estimar los parámetros de la función de distribución de la fase preclínica y la sensibilidad tras tres rondas. La incidencia anual es de 2 casos por cada  $10^3$  mujeres y se obtiene una sensibilidad de la mamografía de 0.71 y una duración media de la fase preclínica de 2 años (inferior a una año en el 20 % de los casos). El modelo se aplica también al ensayo CNBSS [Mil02] y resulta una sensibilidad global de 0.84 para el grupo de mujeres entre 40-49 años y 0.74 para aquellas entre 50-59 años, con duración de la fase preclínica de 2 y 3.3 años, respectivamente.

Este trabajo supone una alternativa al modelado de la duración de la fase preclínica mediante distribuciones analíticas paramétricas, y se trata de un método robusto para obtener distribuciones realistas, además de estimaciones de la sensibilidad. Por contra, pueden presentarse problemas de convergencia, asociados con la aparición de máximos locales, cuando el número de parámetros es escaso (en nuestro casos, pocas rondas de cribado), por lo que se utilizan varios métodos alternativos de estimación de los parámetros.

En la misma línea se encuentra el trabajo de Cong *et al.* [Con05] que profundiza más en estos modelos mediante la introducción de una sensibilidad específica del cribado,  $\beta$ , dependiente de la edad,  $x$ , y de la edad mínima de la cohorte de mujeres,  $x_0$ , y utiliza para ello un modelo logit con dos parámetros libres  $a$  y  $b$ , de la forma:

$$\beta(x) = \frac{\exp[a + b(x - x_0)]}{1 + \exp[a + b(x - x_0)]}, \quad (1.6)$$

Esto permite llegar a una función condicional de verosimilitud que debe resolverse mediante un método no-lineal de optimización propuesto por Byrd *et al.* [Byr95] para estimar  $a$  y

b. De forma similar se introduce una dependencia lineal de la duración media de la fase preclínica con la edad de detección,  $x$ :

$$\mu(x) = \mu_0 + \xi(x - x_0), \quad (1.7)$$

El trabajo contempla tres situaciones o modelos: el Modelo 0 supone que ni el *sojourn time* ni la sensibilidad dependen de la edad; en el Modelo 1 sólo la sensibilidad depende de la edad, mientras que el Modelo 2 sólo afecta al *sojourn time*. Con estas asunciones, la sensibilidad estimada para el estudio HIP varía entre 0.47 y 0.85 según el rango de edad en el Modelo 1 y el *sojourn time* medio sería  $2.57 \pm 0.46$  años ( $k=1$ ). Estas estimaciones son similares a las obtenidas con el Modelo 0. Para el estudio CNBSS, la sensibilidad con el Modelo 1 iría entre 0.58 y 0.99 y un *sojourn time* medio de  $2.42 \pm 1.22$  años, también similares a los valores obtenidos con el Modelo 0. Con el Modelo 2, las estimaciones de sensibilidad en HIP son de  $0.59 \pm 0.15$  y del *sojourn time* 3.04 años, mientras que para CNBSS la sensibilidad es  $0.80 \pm 0.20$  y 2.9 años para el *sojourn time*.

Los resultados muestran que cuando se introduce la dependencia con la edad en la distribución de la fase preclínica o en la sensibilidad, hacerlo también en el otro parámetro no mejora significativamente los resultados del ajuste. Esto es debido, seguramente, a la fuerte correlación que existe entre esos dos parámetros: cuando la sensibilidad es muy alta los tumores son detectables antes, así que la duración de la fase preclínica es menor. Se puede considerar, por tanto, que la introducción de una dependencia con la edad en el *sojourn time* “sobreparametriza” los modelos y no aporta un mejor ajuste de los datos.

Incluimos en este apartado el modelo de Seigneurin *et al.* [Sei16], puesto que no realiza una simulación detallada de la historia natural del tumor, sino que utiliza un modelo de 13 parámetros, en el que el tumor presenta 5 estados preclínicos. Para su estimación se emplean métodos bayesianos de rechazo, basándose en los registros de Isère y de la *Office De Lutte contre le Cancer* para reproducir los resultados relativos al sobrediagnóstico para mujeres entre 50 y 74 años en el periodo 2007-2010. Como resultado se concluye que un 17% de todos los tumores *in situ* (entre el 2.5% y el 35%, intervalo de confianza del 95%) y un 5.5% (entre 0.8% y 9.8%) de todos los invasivos detectados serían sobrediagnósticos.

Incluimos a continuación dos trabajos que, si bien utilizan una metodología basada en parte en procesos de Markov, la parte más original e importante de los mismos no puede encuadrarse en este apartado. Así, Hsu *et al.* [Hsu18] implementan un proceso estocástico de tiempo no continuo basado en la familia de distribuciones binomiales negativas, que pretende responder a las preguntas: ¿cuántas rondas de cribado son necesarias para detectar un tumor y cuándo se puede detener el cribado tras varios resultados negativos del mismo en pacientes de bajo riesgo?.

Estos autores suponen un proceso Markov de tres estados, donde cada uno de ellos sigue una distribución multinomial que determina las probabilidades de transición. El modelo incorpora un tratamiento de los datos censurados, así como de la sensibilidad y especificidad. Basándose en este formalismo, emplean una función de enlace logística (modelo geométrico), binomial negativa o binomial negativa generalizada. Los parámetros se obtienen por el método de máxima verosimilitud, con datos de los programas de cribado de Finlandia y emplean, para la resolución de las ecuaciones, el método Monte Carlo bayesiano de cadenas de Markov mediante un muestreo de Gibbs.

La aplicación de este formalismo supone que son necesarias entre 2.79 y 2.91 rondas para detectar un cáncer de mama (con una confianza del 95 %), dependiendo de la sensibilidad de la mamografía. En mujeres entre 50 y 60 años, el número de cribados necesarios para descartar seguir en el programa sería de alrededor de 8.

Se trata de un método estadístico elegante, aunque no trata con los aspectos biológicos o físico-médicos del problema. Carece, además, de una estimación del sobrediagnóstico.

Existen otros trabajos [Rav15] que emplean códigos tipo Markov junto con microsimulaciones del proyecto CISNET, para realizar un balance riesgo-beneficio del cribado en mujeres mayores de 74 años, así como una estimación del sobrediagnóstico. Este estudio ofrece resultados de entre 7.8 y 11.4 años de vida ganados por cada 1000 cribados en mujeres de 74 años, entre 4.8 y 7.8 años en mujeres mayores de 80 años y entre 1.4 y 2.4 años en mujeres de 90 años, con rangos de sobrediagnóstico de 1.2 a 5, 1.8 a 6 y 3.7 a 7.5 sobrediagnósticos por cada 1000 cribados respectivamente (los intervalos están dados al 95 % de confianza).

Finalmente, incluimos en este apartado el modelo desarrollado por Wu *et al.* [Wu18] que, si bien no realizan una simulación detalla de la historia natural del tumor, tampoco se basa en modelos de Markov. Se trata de un modelo híbrido que emplea probabilidades de transición entre 4 estados del tumor preclínicos y 3 estados clínicos (observables). Para simular los tiempos de transición se utilizan funciones exponenciales definidas por tramos, a fin de introducir diferentes tasas de transición en función del tiempo. Respecto a la estimación de parámetros, se utiliza un algoritmo quasi-Newton [Kal85], basados en resultados derivados de la población inglesa. Se simula un EAC de mujeres entre 51 y 69 años con cribado bienal mediante 100 repeticiones de 1 millón de historias y se estima un valor para el sobrediagnóstico de alrededor del 12.5 %.

Con la información disponible sobre estos códigos podemos realizar el resumen mostrado en las Tablas 1.1 y 1.2, donde se realiza un recopilación de los modelos más importantes, con una breve descripción del tipo al que pertenecen, tipo de simulación de la historia natural del tumor, modelo de sensibilidad de la detección y modelo de supervivencia que implementan.

## 1.2. MODELO DESARROLLADO EN ESTE TRABAJO

Dentro de la terminología utilizada, nuestra herramienta Monte Carlo puede ser encuadrada en el apartado de microsimulaciones y se caracteriza por emplear un modelo de crecimiento tumoral logístico, determinado por un único parámetro distribuido según una distribución log-normal. La evolución del tipo histológico la reproduce mediante un modelo sencillo, ajustado en virtud de los resultados de detección por ronda de los programas de cribado. Se implementa una sensibilidad de detección por mamografía que es función del tipo histológico, el tamaño tumoral y la densidad de la mama, junto con la evolución de esta última con la edad. Utiliza un modelo de detección clínica exponencial, determinado por dos únicos parámetros libres. Finalmente, modela la supervivencia tras detección y tratamiento loco-regional en función únicamente del tamaño tumoral en la detección, incluyendo la mortalidad por causas diferentes al cáncer de mama.

En cuanto a las características de la población simulada, hemos buscado distribuciones de incidencia libres del artefacto que introducen los propios programas de cribado mediante la utilización de incidencias previas a la implantación de estos, sustrayendo de las mismas el retraso debido a la detección para obtener la incidencia pura de la enfermedad. Se emplean, finalmente, distribuciones de edades estándar, propias de sociedades occidentales típicas.

El modelo se desarrolla en tres partes:

1. Una primera que supone la simulación realista cuantitativa en términos de tumores detectados y cánceres de intervalo. Esta primera parte genera, además, las distribuciones de tiempos de adelanto (*lead time*) y sensibilidad global para diferentes configuraciones de programas que se utilizan como *inputs* de las otras dos partes. Esta primera iteración permite abordar el problema de la estimación del sobrediagnóstico, aspecto central de la justificación de los programas de cribado. Para verificar el modelo se lleva a cabo un análisis de sensibilidad del mismo frente a las diferentes entradas que determinan los resultados.
2. Una segunda parte que calcula la reducción de mortalidad que supone la implantación de los programas de cribado mamográfico en la población mediante la introducción de la supervivencia tras detección y tratamiento loco-regional.
3. Una tercera parte que intenta modelar los EAC, principal fuente de evidencia sobre el cribado mamográfico.

Comprende, por tanto, todos los aspectos relevantes en la cuestión del cribado mamográfico, a saber: desarrolla un modelo que es testeado numéricamente a efectos de los resultados asociados a la detección. Tras comprobar la consistencia del mismo y su estabilidad frente a variaciones de los *inputs* se emplea para abordar el asunto del sobrediagnóstico pues, como se ha dicho, es consecuencia de una falta de balance entre tumores detectados en el programa de cribado y aquellos detectados en ausencia del mismo. Posteriormente, se introduce un modelo de supervivencia tras detección y tratamiento, lo que permite abordar el no menos crucial aspecto de la reducción de mortalidad asociada a estos programas. En este punto se pasa a evaluar el aspecto de las configuraciones óptimas, entendidas como rango de edades y tiempos de intervalo a emplear para obtener los mejores resultados. Finalmente, se emplea el modelo para evaluar la consistencia de los principales ensayos aleatorios llevados a cabo hasta la fecha, fuente principal de la evidencia científica sobre la cuestión que nos ocupa.

Como resultado de todo lo anterior contaremos con una herramienta fundamental para el abordaje integral de todos los aspectos y controversias que rodean el cribado poblacional del cáncer de mama. Esta herramienta estará basada en parámetros objetivos observables, pudiendo ser adaptada a poblaciones particulares, técnicas de diagnóstico con diferentes sensibilidades de detección, tratamientos más o menos efectivos y cualquiera de las configuraciones deseadas. Puede, además, ser empleada para realizar estudios de coste-efectividad, utilidad o beneficio de manera sencilla.

Código	Tipo	Historia tumor (fase preclínica)	Sensibilidad	Supervivencia
<b>MISCAN</b> <sup>1</sup> [Hab85]	Markov (permite ajuste poblométrico edad)	Configurable (diferentes distribuciones)	Función del tipo de tumor (valores constantes). Modela errores sistemáticos	Función clasificación del tumor en detección con mejora por modalidad detección
<b>MBS</b> [Jan95]	Microsimulación	Crecimiento modelo Gompertz distrib. tiempos duplicación log-normal	Distrib. log-normal umbral ( $\leq 48$ años y $> 48$ años)	Función del tamaño tumoral en detección
<b>Smltree</b> [Lin95]	Markov	No disponible	No disponible	No disponible
<b>Duffy et al.</b> [Duf95]	Markov (3 y 6 estados)	Duración fase preclínica exponencial	Se obtiene comparación tumores en fase preclínica/ ensayo <i>Swedish Two County</i>	Supervivencia ajustada mediante modelo de Cox
<b>MICROLIFE</b> [Sze96]	Microsimulación	Duración exponencial (duración media $m_a = \frac{0.3m_{50}}{2} (1 - 0.9a)$ donde $m_{50}$ parámetro que va entre 0.58 y 1.00 y $a = edad/100$ )	Función del tamaño tumoral	Dependiente del tamaño (ajuste log-normal datos Nueva Zelanda)
<b>POWERSIM</b> [Fet01]	Simulación determinista	Interpolación lineal	Interpolación lineal	Relación lineal mortalidad-momento fase preclínica
<b>Plevritis et al.</b> [Ple06]	Microsimulación	Tiempo detección clínica lineal con el volumen del tumor. Tiempo duplicación según distribución gamma	Distribución de umbral de detección $\nu^3$	Dependiente de la edad, tamaño y estadio del tumor
<b>Lee et al.</b> [Lee06]	Markov (4 estados)	Diferentes estados: sin enfermedad, preclínico, clínico y mortal, modelados por distribuciones exponenciales de media dependiente de la edad	Dependiente de la edad (datos BCSC <sup>2</sup> ) factores globales dependientes del periodo	Basada en datos SEER (1975-1979) modificados para introducir el beneficio de la QT reportada en EBCTCG <sup>3</sup> por grupos de edad
<b>SPECTRUM</b> [Man06]	Markov(4 estados)	Estadíos SEER. Distribuciones de permanencia y de duración fase preclínica exponencial (medias dependientes de la edad)	La sensibilidad varía entre la primera mamografía y las siguientes y en función del tramo de edad	Supervivencia específica por estadio y ER sin tratamientos adyuvantes <sup>4</sup> y reducciones de probabilidad de muerte asociadas a terapias adyuvantes

Tabla 1.1: Resumen de los códigos de simulación. (1) Versión Fadia incluye “estado fatal de detección”; (2) *The Breast Cancer Surveillance Consortium*; (3) *Early Breast Cancer Trialists’ Collaborative Group*; (4) Datos procedentes de la población USA posteriores a 1975.

Código	Tipo	Historia tumor (fase preclínica)	Sensibilidad	Supervivencia
Berry <i>et al.</i> [Ber06]	Microsimulación	Basada en frecuencia de estadios, ganglios afectados y ER obtenido de distribuciones de frecuencia dependientes de la edad, modo de detección y tiempo tras última mamografía	Dependiente de la edad según datos NBCCEDP <sup>5</sup> y de si es primera mamografía o del tiempo que transcurre desde esta	Riesgo basal modelado por $S_{ij}(t)$ con $t$ la supervivencia a tiempo $t$ con $i = 1, \dots, 4$ estadio AJCC y $j$ grupo de edad. Se introduce exponencialmente $\lambda$ muestreado con distribución uniforme [0,80,1] para incertidumbre. El efecto del tratamiento adyuvante se modela con exponente $1 - \gamma$ muestreado $\gamma$ mediante distribuciones beta o uniforme
Wisconsin Breast Cancer Epidemiology Simulation Model [Fry06]	Microsimulación	Crecimiento modelo Gompertz distrib. tasas de crecimiento función gamma. Modelado de extensión regional (ganglios+ y distal) distribución de Poisson	Probabilidad de detección en función del tamaño, dependiente de la edad y el periodo	Distribuciones semiempíricas basadas en datos SEER (1975-1982) de supervivencia tras diagnóstico enfermedad distal (media 1,95 años y mediana 5,22 años)
SIMIRISC [Gre10]	Microsimulación	Log-normal	$S(t) = S_0 \left( \frac{S_1}{S_0} \right)^{\frac{t}{T}}$	No se evalúa
BCSS [Bai10]	Microsimulación	Dependiente edad de detección	Función densidad mama (dependiente edad)	No disponible
Gunsoy <i>et al.</i> [Gun12]	Markov (6 y 10 estados)	Duración exponencial	Se obtiene comparación tumores en fase preclínica/ensayo UK Age trial	No se evalúa
Gocgun <i>et al.</i> [Goc15]	Markov (8 estados)	3 estados preclínicos	Se obtiene de Taghipour <i>et al.</i> [Tag13]	Mediante las transiciones a estados de muerte por cáncer de mama y otras causas
Seigneurin <i>et al.</i> [Sei16]	Simulación con estimación de parámetros bayesiana	5 estados preclínicos	Se obtiene del registro de Isère	Isère 2010
Isheden y Humphreys [Ish19]	Microsimulación	Exponencial (parámetro de crecimiento según gamma)	Modelo logístico (dependiente del tamaño y densidad)	No se implementa

Tabla 1.2: Resumen de los códigos de simulación. 1:2006-2019. (5) National Breast and Cervical Cancer Early Detection Program.

## Capítulo 2

# ESTRUCTURA GENERAL DEL CÓDIGO Y ALGORITMOS DE SIMULACIÓN

En este capítulo se describe la estructura general del código de simulación que se ha utilizado en las diferentes partes que componen este trabajo, junto con los algoritmos empleados en cada una de ellas.

En la primera parte se busca reproducir los resultados relativos a la detección tumoral en un programa de cribado de cáncer de mama, caracterizado por el rango de edades de las mujeres que se someten al mismo y la frecuencia de las exploraciones (tiempo de intervalo entre mamografías). Para ello es necesario modelar la sensibilidad de la mamografía en función de la densidad de la mama, el tamaño tumoral y el tipo histológico del mismo. Se hace necesario, por tanto, describir la evolución de la densidad de la mama de las mujeres con la edad, de forma que el modelo pueda reproducir la dependencia de la sensibilidad de la mamografía con la edad de las mujeres a través de esta, pues es conocida la peor detectabilidad en mamas de mujeres jóvenes [Kop98]. De igual manera, hay también que describir el crecimiento tumoral, así como la evolución de los tipos histológicos, aspectos ambos cruciales tanto en la detección por cribado como en la detección clínica de los tumores.

Por esta razón es imprescindible introducir un modelo de detección clínica, que dé cuenta de aquellos tumores que aparecen por sus síntomas durante el desarrollo del programa (cánceres de intervalo), así como de los resultados en ausencia de programas de cribado, lo que permite analizar cuáles son los efectos reales de estos programas.

Otro aspecto crítico en esta primera parte es la proporción inicial subyacente de tumores invasivos/*in situ* y su evolución con el tiempo asumiendo transiciones entre los distintos tipos histológicos. Esta proporción no es observable y no se corresponde necesariamente con las proporciones observadas en la detección, debido a la compleja relación entre la probabilidad de detección y las características del tumor.

Las características de las distribuciones poblacionales, así como su evolución, también han de tenerse en cuenta. En particular, es necesario incluir en las simulaciones las distribuciones de edades de las mujeres, las de incidencia de cáncer de mama corregidas por

tiempo medio de detección y las probabilidades de mortalidad. Debe señalarse en este apartado que, en general, se han empleado distribuciones genéricas, si bien, en algunos, y con el fin de reproducir resultados proporcionados por programas concretos, se ha hecho uso de distribuciones específicas.

Los resultados de este primer modelo están encaminados a verificar si es capaz de reproducir los resultados de un programa real, teniendo en cuenta que estos se llevan a cabo durante periodos de tiempo relativamente largos en los que las incidencias del cáncer de mama, los aspectos tecnológicos asociados a la detección o, incluso los detalles logísticos relativos a la propia estructura y características del programa de cribado, pueden haber sufrido modificaciones.

Dadas sus características, el algoritmo desarrollado en esta primera parte se puede adaptar para el estudio del sobrediagnóstico. Para ello es necesario simular un programa de cribado activo durante un periodo determinado y comparar el número de tumores presentes en la población con el correspondiente a la misma población no sometida a cribado alguno. Este es, sin duda, el aspecto más interesante de esta primera parte, pues es donde los EAC muestran más limitaciones a la hora de proporcionar una estimación precisa.

En la segunda parte del trabajo damos cuenta de la mortalidad tras detección y tratamiento loco-regional del cáncer de mama. En este apartado se produce un cambio radical en la filosofía de las simulaciones, que pasan de describir de manera detallada la historia de las mujeres sometidas al cribado a utilizar un algoritmo “condensado” en el que se hace uso directo de la distribución de adelanto de detección y la sensibilidad global. Ambas se simulan con ayuda del código original para cada configuración del programa de cribado, entendida como rango de edades y tiempo de intervalo entre mamografías. Este cambio de filosofía implica la simulación de la historia de las mujeres desde su nacimiento hasta el momento en el que entran en el programa de cribado, siendo necesaria, por tanto, una adaptación de la distribución de edades de las mismas para tener simulaciones realistas. Junto a estos cambios en el código, es necesario realizar algunas consideraciones teóricas relativas a la supervivencia tras tratamiento, asociadas principalmente al tamaño del tumor en el momento de la detección.

Estas modificaciones nos permiten estudiar la reducción de mortalidad asociada a las diferentes configuraciones de los programas, aspecto crucial en los estudios coste-utilidad y coste-beneficio de los programas de cribado.

Finalmente, la tercera parte de este capítulo se dedica a describir las modificaciones realizadas en el código para modelar los EAC y los resultados que los diferentes ensayos han proporcionado en términos de reducción de mortalidad por cáncer de mama y sus intervalos de confianza. Comparando los resultados del modelo con los proporcionados por los ensayos se pueden sacar ciertas conclusiones sobre las características y la validez externa de los mismos. Las características principales de estos ensayos que deben modelarse son el número de mujeres que participan, la duración de los mismos (incluyendo el tiempo de seguimiento), el número de rondas que comprenden, las edades y los tiempos de intervalo entre mamografías, los porcentajes de participación por ronda y otros aspectos más complejos y difícilmente parametrizables como son la autoexploración que se introduce en algunos ensayos en los grupos de control o, incluso también, en el grupo sometido al cribado [Mil00, Mil02].

De cara a la generación de las curvas de adelanto y sensibilidad global para esta tercera parte, se tiene en cuenta el número limitado de rondas que comprenden estos ensayos, aspecto que influye principalmente en la segunda. La modificación de esta sensibilidad global se utiliza para modelar la diferente participación por ronda en estos ensayos.

## 2.1. ESTRUCTURA GENERAL

A continuación se describen los algoritmos empleados en el modelo y se realiza una breve descripción teórica de los mismos, así como del proceso heurístico que se ha seguido para determinar los parámetros que caracterizan las distribuciones empleadas. Además, cuando procede, se llevan a cabo contrastes de los *inputs* empleados con valores de referencia publicados. De esta manera, se ha pretendido evitar la utilización de distribuciones artificiales o excesivamente parametrizadas que podrían dar lugar a un modelo de escasa o nula validez externa.

## 2.2. MODELO DE DETECCIÓN

Se describen los algoritmos empleados en la primera parte del trabajo, donde se trabaja únicamente con los resultados asociados a la detección de tumores, tanto clínica como mediante cribado. Además, con este primer modelo, se generan los *inputs* utilizados en la segunda y tercera parte (distribución de adelanto y sensibilidad global) y se lleva a cabo la estimación del sobrediagnóstico.

### 2.2.1. Muestreo de la edad de la candidata

Se muestrea la edad de cada mujer en el momento del inicio del programa de cribado,  $e_0$ , de acuerdo con la distribución de edades de las mujeres españolas en el año 2007, proporcionada por el Instituto Nacional de Estadística [INE07] (ver Figura 2.1).

Para el muestreo según este histograma, se divide el intervalo  $[0,1]$  en 18 sub-intervalos de anchura proporcional al porcentaje relativo de mujeres en cada uno de los 18 intervalos de edad, desde 0 a 87 años, que se incluyen en el histograma. El intervalo de edad al que pertenece la mujer que entra en el cribado se obtiene muestreando la distribución  $U[0,1]$ . Finalmente, la edad  $e_0$  de la mujer se determina muestreando nuevamente la distribución uniforme correspondiente al intervalo seleccionado.

Si  $e_0 \geq E_{\text{MIN}}$ , la edad mínima de la configuración del programa, la mujer entraría en el programa al comienzo del mismo, con  $e = e_0$ . En caso contrario, la mujer se someterá al programa cuando alcance la edad mínima  $E_{\text{MIN}}$  y lo hará en la ronda  $R_{\text{in}} = E_{\text{MIN}} - e_0$ .

Con esta manera de proceder simulamos el comienzo y desarrollo de un programa de cribado aplicado a las mujeres de un grupo de edad prefijado. Está, por tanto, implícita la idea de un programa de cribado que se inicia en un momento determinado del tiempo, aspecto este que tendrá su trascendencia en los resultados, como veremos más adelante.

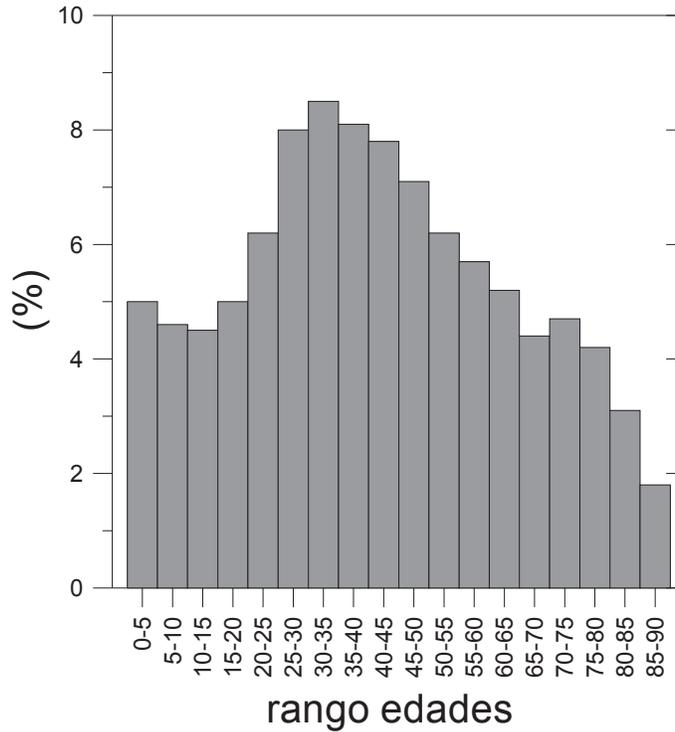


Figura 2.1: Histograma que representa la distribución de edad de las mujeres españolas en el año 2007 (información obtenida del Instituto Nacional de Estadística, INE)

### 2.2.2. Modelo de crecimiento tumoral

Asumimos un crecimiento tumoral según un modelo logístico dado por Spratt *et al.* [Spr93]. La ecuación general del modelo se escribe como:

$$V_{\text{tum}}(t) = V_{\text{tum}}(\infty) \left[ 1 + \frac{1}{\exp N(bt + c)} \right]^{-\frac{1}{N}}, \quad (2.1)$$

donde  $V_{\text{tum}}(t)$  es el volumen tumoral,  $b$  la tasa de crecimiento relativo,  $t$  el tiempo de crecimiento del tumor y  $N$  y  $c$  son parámetros ajustables del modelo.

$N$  y  $c$  se relacionan imponiendo la condición inicial  $V_{\text{tum}}(t = 0) = V_0$ . Se acepta que el volumen inicial es el volumen correspondiente a una célula y el tamaño máximo,  $V_{\text{tum}}(\infty)$ , corresponde a  $2^{40}$  células.

Del ajuste del crecimiento de 448 tumores seguidos en el tiempo realizado por Spratt *et al.* [Spr93b] se obtiene el valor de  $N$  y, por lo tanto, de  $c$ . La ecuación resultante del modelo será:

$$V_{\text{tum}}(t) = 1.1 \cdot 10^6 \left[ 1 + \frac{1023}{\exp\left(\frac{bt}{4}\right)} \right]^{-4}, \quad (2.2)$$

La distribución de  $b$  (Tabla 2.1) se ajusta a una log-normal, de forma que para su muestreo se utiliza el método Box-Müller [Box58]. En esta distribución asumimos que no existen tasas de crecimiento tales que el tumor alcance 1 cm en menos de 5 años, lo que

se traduce en una cota superior para  $b$  de  $0.01 \text{ días}^{-1}$ . Esta restricción pudiera parecer excesiva, pues se estima que algunos tumores de mama pueden tener tasas de crecimiento que les haga alcanzar diámetros de 1 cm en 2 años [Gen81]. En cualquier caso, el porcentaje de estos tumores especialmente agresivos es mínimo (ver siguiente apartado) y se introduce la restricción para evitar artefactos en las simulaciones.

Tabla 2.1: distribución del parámetro de crecimiento  $b$  de Spratt *et al.* [Spr93b].

$10^{-3} \text{días}^{-1}$	percentil (%)
33.72	1
11.77	5
5.17	25
3.22	50
1.73	75
0.56	95
0.20	99

Como puede verse en la Figura 2.2, el crecimiento presenta una tasa muy baja hasta los 10-20 años desde su inicio, tiempo necesario para alcanzar un volumen de detección clínico medio ( $500 \text{ mm}^3$ ), acelerándose entonces. Sigue luego una fase de crecimiento lineal hasta que vuelve a estabilizarse en tamaños elevados, deteniendo finalmente su evolución. Con este modelo, el 90% de los tumores (entre el percentil 5% y el 95%) presentan un tiempo de duplicación en la fase preclínica (entendida como aquella en la que el tumor alcanza un diámetro de 5 mm) entre 0.18 años ( $b = 0.01177 \text{ días}^{-1}$ ) y 3.8 años ( $b = 0.00056 \text{ días}^{-1}$ ), con un valor para la mediana ( $b = 0.0032 \text{ días}^{-1}$ ) de 0.66 años. Estos valores son consistentes con los obtenidos por otros autores (ver por ejemplo Peer *et al.* [Pee93]).

Calculando la distribución de tiempos de duplicación, TD, según nuestro modelo, en el momento en que el tumor alcanza los 10 mm (valor representativo del tamaño medio de detección), obtenemos una distribución con media TD=0.95 años, muy similar a la aportada en el estudio de referencia de Anerlöv *et al.* [Jan95, Ane92] sobre 158 casos de tumores, que es de 0.91 años (incluyendo para el cálculo de la media 11 tumores que no mostraron crecimiento), si bien esta distribución de referencia presenta una mayor variabilidad que la nuestra (ver Figura 2.3). A la hora de valorar la compatibilidad de ambas distribuciones, debe tenerse en cuenta la aproximación empleada para la comparación, que fija el tamaño de detección en 10 mm, pues la distribución de Anerlöv *et al.* incluye tumores detectados tanto por mamografía de cribado como en intervalo.

El modelo que hemos usado reproduce, por tanto, el tipo de crecimiento de los tumores de mama con bastante fidelidad, a juzgar por los resultados obtenidos de la bibliografía disponible, y se caracteriza por un único parámetro libre distribuido según una log-normal. Esta distribución es conocida, fácilmente muestreable y se limita para evitar artefactos indeseables en las simulaciones.

Finalmente debemos señalar que en este estudio de referencia no se encuentra diferencia significativa entre los tiempos de duplicación para mujeres menores de 50 años respecto a los de las mujeres mayores de 50, resultado que avala nuestra hipótesis de mayor simplicidad de acuerdo con los resultados experimentales probados. En este sentido, no se consideran

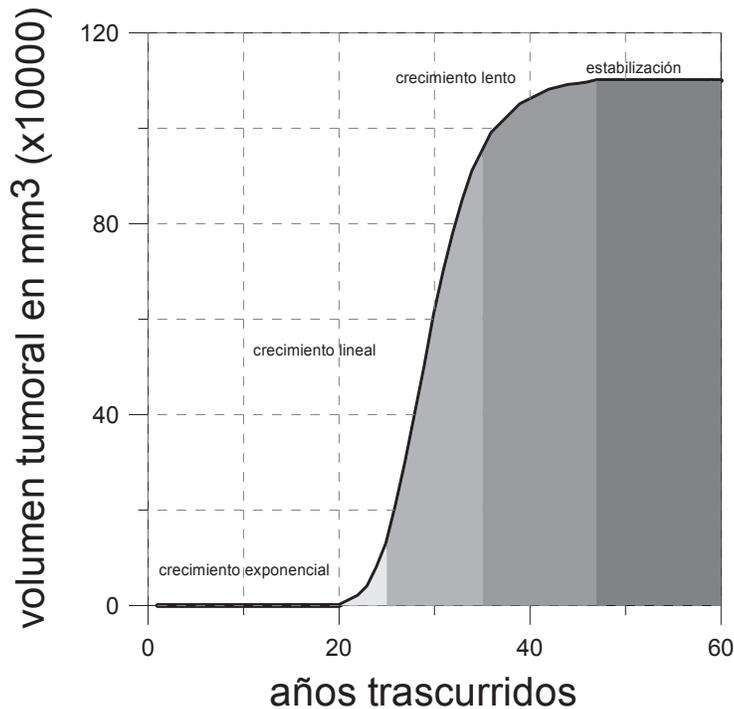


Figura 2.2: Modelo de crecimiento de Spratt ajustado para el seguimiento de 448 tumores con un valor de  $b$  promedio de  $0.0032 \text{ días}^{-1}$ . Se aprecian las diferentes velocidades de crecimiento en la historia natural del tumor: una primera fase de tumor latente prácticamente sin crecimiento, inicio del crecimiento exponencial seguida de un crecimiento lineal, para llegar a una zona de crecimiento lento y estabilización.

dependencias con la edad de la tasa de crecimiento, hipótesis que en numerosas ocasiones es introducida *ad hoc*, condicionando los resultados obtenidos.

### 2.2.3. Incidencia y prevalencia: muestreo de la aparición del cáncer de mama

Tras generar la edad de la mujer al inicio del programa de cribado, se muestrea si padecerá cáncer a lo largo de su vida y, si es así, el momento en el que aparece (*onset*). Para ello es necesario acudir a las definiciones de prevalencia e incidencia:

- La **prevalencia** cuantifica el número de individuos que padecen la enfermedad en un momento dado y se estima a partir de estudios de carácter transversal.
- La **incidencia** es la proporción de individuos sanos que desarrollan la enfermedad a lo largo de un periodo concreto. Su valor indica la probabilidad o riesgo de padecer la enfermedad durante ese periodo concreto. Para su estimación se utilizan estudios de tipo longitudinal.

En la Figura 2.4 se muestra la curva de incidencia típica para cáncer de mama, que muestra un aumento brusco en el entorno de los 30 años, llegando a una meseta alrededor

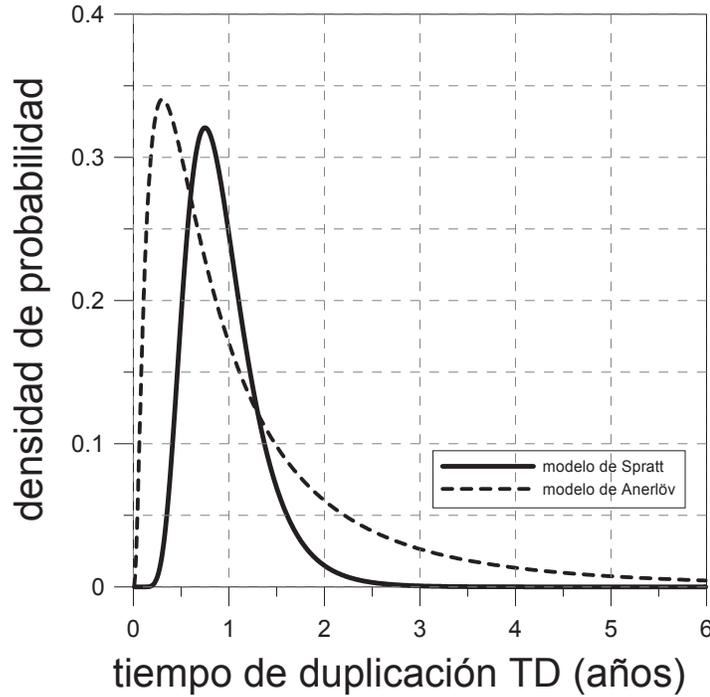


Figura 2.3: Comparación de las curvas de tiempo de duplicación (TD) de nuestro modelo con el de Anerlöv, asumiendo en nuestras simulaciones tamaños de detección de 10 mm. Ambas distribuciones son compatibles teniendo en cuenta que la de Anerlöv presenta mayor heterogeneidad en su construcción respecto a nuestro modelo teórico.

de los 50 años, para luego volver a crecer rápidamente hasta los 75 años. Esta curva se obtiene a partir del número de tumores detectados en función de la edad, por lo que no es una incidencia pura; además está afectada por los propios programas de cribado, que producen un aumento aparente de la incidencia a edades tempranas, asociado con el adelanto en el tiempo de diagnóstico. Por esta razón empleamos una curva de incidencia correspondiente a una zona y periodo no afectado por los programas de cribado. Así sucede con la debida al Registro Poblacional de Cáncer de las Islas Canarias del año 1996 [SCS96], que proporciona datos previos a la implantación del programa de cribado mamográfico.

Para obtener una curva de incidencia pura utilizamos el modelo de crecimiento descrito anteriormente con un valor de  $b = 0.0032 \text{ días}^{-1}$ . Suponiendo un diámetro medio de detección clínica de 17 mm [Ber06], se calcula el tiempo que tardaría el tumor en alcanzar ese tamaño, resultando un valor de 19.2 años. Este valor medio es sustraído de la curva de incidencia  $I(t)$ , que se normaliza de acuerdo con la prevalencia, tomada como 0.12 hasta los 85 años [Kop98].

$$0.12 = \text{cte} \int_0^{85} p(\tau) [I(\tau) - 19.2] d\tau, \quad (2.3)$$

donde  $p(\tau)$  es la distribución de edades de la población que, en esta aproximación, se supone constante. De esta forma obtenemos la probabilidad de padecer cáncer en un intervalo de tiempo determinado como:

$$P(t, t + \Delta t) = c \int_t^{t+\Delta t} [I(\tau) - 19.2] d\tau, \quad (2.4)$$

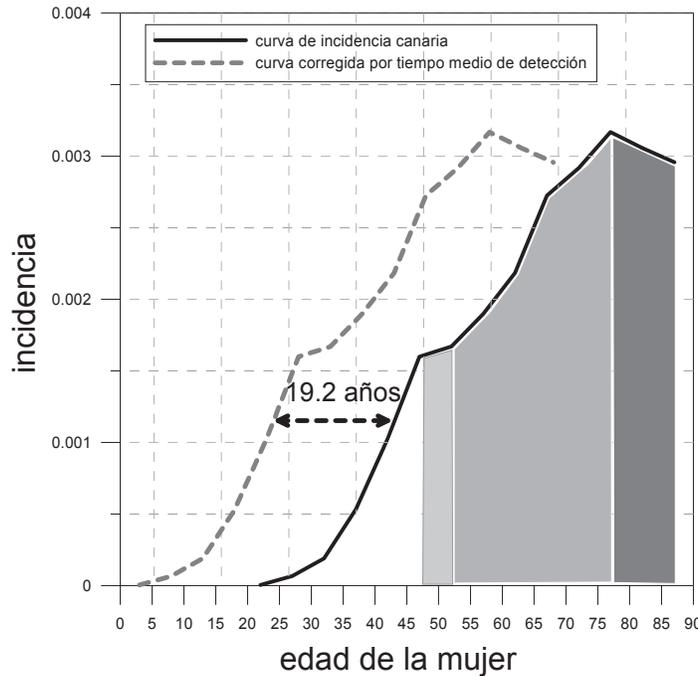


Figura 2.4: Curva de incidencia de cáncer de mama de la población canaria (1996) y curva corregida por tiempo medio de detección (19.2 años). Las zonas en diferentes tonos de gris indican los tramos indicados en el texto. Se aprecia como la curva corregida llega sólo a los 70 años.

Con este resultado, estimamos la probabilidad de padecer cáncer en cualquier intervalo de tiempo de la vida de la mujer. Una vez que se calcula esta, se muestrearán si padecerá o no cáncer con ayuda de la distribución uniforme.

#### 2.2.4. Mortalidad por causas diferentes al cáncer de mama

Para estimar la probabilidad de muerte por causas diferentes al cáncer de mama en un intervalo de edades entre  $t$  y  $t + \Delta t$  utilizamos datos del INE del año 2006 [INE07], restando las muertes por cáncer de mama a la mortalidad general de las mujeres. Se tiene la distribución mostrada en la Tabla 2.2.

Esta probabilidad se asume constante en los 5 años del intervalo, de forma que la probabilidad de muerte por causas diferentes al cáncer de mama en un periodo entre  $t$  y  $t + \Delta t$  se calcula como:

$$P_{\text{mnc}}(t; t + \Delta t) = 1 - \prod_{i=t}^{t+\Delta t} (1 - P_i), \quad (2.5)$$

donde  $P_i$ , con  $i = t, \dots, t + \Delta t$ , representa la probabilidad de muerte por causas diferentes al cáncer de mama en los diferentes intervalos de edad. Debe notarse que la probabilidad complementaria de “no morir por causas diferentes al cáncer de mama” sería continuar viva o morir por cáncer de mama, esta segunda opción se trata de manera segregada y como veremos más adelante, solo se considera tras la detección del tumor. Una vez estimada la probabilidad se muestrea según la distribución uniforme.

Tabla 2.2: Valores de mortalidad por causas diferentes al cáncer de mama correspondiente al año 2006 proporcionada por el INE. Los valores de la tercera columna se obtienen restando las muertes por cáncer de mama a la mortalidad general de las mujeres durante ese año y dividiendo por el número total de mujeres vivas al principio del 2006.

	nº muertes por cualquier causa menos cáncer de mama	probabilidad de muerte/año por causas distintas al cáncer de mama
De 30 a 34 años	597	0.03 %
De 35 a 39 años	960	0.05 %
De 40 a 44 años	1432	0.08 %
De 45 a 49 años	1885	0.12 %
De 50 a 54 años	2419	0.18 %
De 55 a 59 años	3140	0.25 %
De 60 a 64 años	4310	0.38 %
De 65 a 69 años	6341	0.63 %
De 70 a 74 años	12724	1.18 %
De 75 a 79 años	22120	2.36 %
De 80 a 84 años	33941	4.89 %
de 85 años y mas	79502	13.35 %

### 2.2.5. Probabilidad de detección. Densidad de la mama, tipos histológicos y transiciones entre ellos

Para la estimación de la probabilidad de detección en mamografía, asumimos un modelo factorial que supone el producto de dos probabilidades: una dependiente del tamaño del tumor y otra función del tipo histológico y la densidad de la mama.

$$P_{\text{det}} = P_{\text{tam}} \cdot P_{\text{hist, dens}}, \quad (2.6)$$

La densidad de la mama es consecuencia de los diferentes tipo de tejidos que la componen, principalmente adiposo, glandular y conjuntivo. El Colegio Americano de Radiología en su publicación [ACR] clasifica las mamas como: predominantemente grasas (I), grasas con algunas regiones fibroglandulares (II), densas heterogéneas (III) y extremadamente densas (IV).

Este es un aspecto crucial desde el punto de vista de la sensibilidad de la mamografía pues los tejidos más densos muestran un menor contraste radiológico con las lesiones, dificultando la detección de las mismas. Además, hay una clara relación entre la edad de la mujer y el carácter más o menos graso de la mama: mujeres más jóvenes tienden a tener mamas más densas [Kop98]. El efecto se explica por los cambios hormonales que tienen lugar en las mujeres asociados a la edad, favoreciendo que el tejido fibroglandular se transforme en grasa (ver Figura 2.5).

En nuestras simulaciones consideramos la distribución de la densidad de las mamas con la edad de acuerdo con la clasificación anterior y estimamos la probabilidad de transición desde los subtipos menos grasos hasta los más grasos. Para la simulación de estas transiciones se asume que (*i*) las transiciones se producen cuando la edad de la mujer cambia

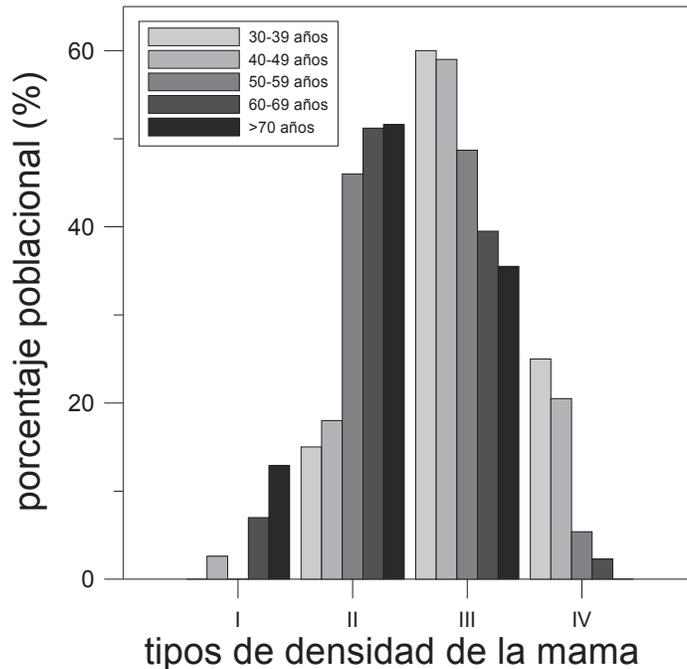


Figura 2.5: Distribución de los tipos de densidad de la mama según la clasificación BI-RADS. Los valores se han obtenido de Kerlikowske *et al.* [Ker96]

de década, y (ii) solo están permitidas las transiciones desde un tipo más denso hasta el inmediatamente menos denso en cada iteración.

Ninguna de las dos restricciones resulta ser muy limitante y, además, la primera simplifica las simulaciones y la segunda responde a las transiciones suaves que, parece, tienen lugar en la naturaleza. Si bien se argumenta en algunas ocasiones que la menopausia supone cambios bastante radicales, en cualquier caso esta transición se refleja en las transiciones del tipo IV al III y del III al II que tienen lugar alrededor de los 49 años. Las probabilidades de mantener el tipo de densidad al final de cada década se resumen en la Tabla 2.3. Con este modelo, la probabilidad de que una mama del tipo IV a los 30 años de edad se mantenga como tal a los 70 años sería  $(1 - 0.82) \times (1 - 0.26) \times (1 - 0.43)$ , valor que ronda el 10%. En este contexto, se utiliza la probabilidad de detección (sensibilidad de la mamografía) en función de los tipos histológicos y de la clasificación de densidad de la mama anterior, mediante los valores mostrados en la Tabla 2.4 [Beg04].

Los tumores de la mama se clasifican según el lugar anatómico donde se originan y su capacidad para invadir los tejidos adyacentes. Así, los carcinomas ductales son aquellos originados en los ductos y se conocen como *in situ* o intraductales (CDi). Puede ocurrir que estos tumores, en su evolución, rompan la membrana basal que rodea el conducto e invadan tejidos vecinos, dando lugar a un carcinoma ductal invasivo o infiltrante. Con menor proporción (menos del 10%) se detectan cánceres con origen en el epitelio lobular (carcinomas lobulillares) que, si se encuentran confinados dentro del lóbulo, se clasifican del mismo modo como *in situ*, pudiendo progresar también hacia invasivos. Como puede verse en la Tabla 2.5 estos son los tipos más frecuentes de tumores y los considerados en nuestras simulaciones.

Tabla 2.3: Probabilidad de mantener el tipo de mama según la clasificación BI-RADS al final de cada década. Los cálculos se realizan a partir de las distribuciones de la Figura 2.8 asumiendo transiciones sólo entre tipos próximos.

densidad de la mama	% de población por grupo de edad				
	30-39	40-49	50-59	60-69	≥ 70
<b>I Predominantemente grasa</b>	100	100	100	100	100
<b>II Grasa con regiones fibroglandulares</b>	83	100	85	100	100
<b>III Densa heterogénea</b>	91	57	75	84	100
<b>IV Extremadamente densa</b>	82	26	43	0	100

Tabla 2.4: Probabilidad de detección en función del tipo histológico del tumor y de la densidad de la mama [Beg04]. Se aprecia la clara dependencia del tipo histológico para las diferentes densidades, salvo para el tipo de mama grasa. Los valores de probabilidad de detección para estas mamas son del 100 %, y son mucho más bajos para mamas del tipo (4), extremadamente densas.

densidad de la mama	Tipo histológico			
	CD infiltrante	CL infiltrante	CD <i>in situ</i>	otros
<b>I Predominantemente grasa</b>	100	100	100	100
<b>II Grasa con regiones fibroglandulares</b>	89	60	60	25
<b>III Densa heterogénea</b>	86	36	64	21
<b>IV Extremadamente densa</b>	60	11	25	15

Estas proporciones detectadas no necesariamente corresponden a las subyacentes pues, como hemos señalado, se pueden producir transiciones entre unos estados y otros a lo largo de la historia natural del tumor y no todos estos estados tienen la misma probabilidad de detección. No hay una evidencia consistente sobre cómo evolucionan los tumores ni en qué proporción lo hacen. Algunas hipótesis apuntan a que el tipo intraductal no es más que una fase precoz del tumor invasivo; otras, sin embargo, afirman que los tumores invasivos presentan características metastásicas desde su primer clon.

Para estimar la proporción de tumores subyacente y su evolución con el crecimiento de los mismos, hemos realizado un proceso heurístico donde se ha tenido en cuenta que la proporción detectada en los programas de cribado oscila entre un 70 % y 80 % de tumores invasivos y un 20 %-30 % de CDis [Ros06]. Menos del 10 % de los tumores tienen su origen en el epitelio lobular, por lo que no consideraremos estos tipos de tumores en nuestras simulaciones. Comenzamos por simular la distribución de tumores en la población en un instante determinado. La información sobre la proporción de cada tipo y su evolución aún no ha sido introducida en la simulación. En este apartado se utiliza básicamente el modelo de crecimiento tumoral ya descrito con la distribución del parámetro de crecimiento  $b$  para generar la distribución de tumores en función del diámetro  $d$ . Esta distribución es ajustada a una función potencial fácilmente muestreable en las simulaciones posteriores.

$$P(d) = 0.547d^{-1.3747}, \quad (2.7)$$

Con esta distribución como *input* y las probabilidades de detección simplicadas de la

Tabla 2.5: Tipos histológicos de cáncer de mama y frecuencia de aparición. Los valores se obtiene de [Kop98].

Tipo histológico	Frecuencia (%)
Carcinoma ductal infiltrante	74.0
Carcinoma lobulillar infiltrante	6.0
Carcinoma intraductal	3.5
Carcinoma lobulillar <i>in situ</i>	3.0
Carcinoma medular	2.7
Cistosarcoma	2.5
Enfermedad de Paget	2.3
Carcinoma mucinoso	2.0
Carcinoma papilar	1.5
Carcinoma inflamatorio	1.5
Sarcomas y otros	1.0
<b>Total</b>	<b>100</b>

manera que explicamos más adelante, buscamos encontrar una distribución los de tumores subyacente que reproduzca, de manera realista, los porcentajes de detectados de cada tipo. Realizamos las siguientes hipótesis para este cálculo:

1. Buscamos reproducir los datos de proporción de tumores invasivos detectados en cribado proporcionada por el Instituto Nacional del Cancer Norteamericano (INC) [Ros06].
2. Como se comprueba en las distribuciones de densidad de la mama, las más frecuentes son las tipo II y III. Promediando las probabilidades de detección para estas dos densidades en la Tabla 2.4 obtenemos los valores:  $P_{\text{hist,dens}}(\text{invasivo}) = 88\%$ ;  $P_{\text{hist,dens}}(\text{in situ}) = 62\%$ .
3. Utilizamos una probabilidad de detección en función del tamaño simplificada, modelada por una distribución normal que ajustamos con las condiciones:  $P_{\text{det}}(5 \text{ mm}) = 26\%$  y  $P_{\text{det}}(10 \text{ mm}) = 91\%$  [Wee08]. Suponemos además que todos los tumores de 10 mm son del tipo invasivos y todos los de 5 mm son CDis, con lo que obtenemos despejando en la ecuación (2.6) que  $P_{\text{tam}}(5 \text{ mm}) = 42\%$  y  $P_{\text{tam}}(10 \text{ mm}) = 100\%$ .
4. Acotamos las distribución de tumores generada entre 1.5 mm y 50 mm para limitar efectos indeseables de borde: no es probable encontrar tumores de menos de 1.5 mm por la resolución de la mamografía, ni de más de 50 mm, pues serían palpables previamente.

A partir de estas hipótesis empleamos una simulación sencilla de forma que vamos variando la proporción de tumores invasivos/*in situ* en cada intervalo de tamaños y comparando los resultados de la simulación con los de referencia INC [Ros06]. Así, obtenemos la distribución de tumores subyacentes en función del tamaño que se muestra en la Tabla 2.6.

Tabla 2.6: Distribución de tumores suyacente obtenida de las simulaciones buscando reproducir los datos de tumores detectados en función del tamaño del INC [Ros06]. En estas simulaciones se han empleado las simplificaciones especificadas en el texto.

Tipo histológico	Diámetro (mm)				
	1-6	6-11	11-15	16-20	>20
<i>in situ</i>	80 %	50 %	15 %	10 %	10 %
invasivo	20 %	50 %	85 %	90 %	90 %

Con esta distribución de tumores subyacentes podemos estimar la probabilidad de que los tumores *in situ* se mantengan como tales cuando sobrepasan los 6, 11, 16 y 20 mm (ver Tabla 2.7). En esta estimación se asume que sólo los tumores del tipo *in situ* pueden pasar a invasivos en algún momento. Para la estimación de las probabilidades de la Tabla 2.7 se ha utilizado exclusivamente la proporción de tumores subyacente para los diferentes tamaños, por lo que un cambio en esta proporción se traducirá en cambios en las probabilidades de mantener el tipo histológico *in situ* a lo largo de la historia natural del tumor. En las simulaciones, esta probabilidad se muestrea mediante la distribución uniforme cada vez que se sobrepasa uno de los 4 umbrales establecidos. De esta tabla se deduce de forma intuitiva que existe una cierta predisposición de los tumores *in situ* a cambiar a invasivos cuando sobrepasan el umbral de los 11 mm.

Tabla 2.7: Probabilidad de mantener el tipo CDis cuando el tamaño del tumor sobrepasa 6, 11, 16 y 20 mm.

Tipo histológico	Tamaño tumoral (mm)			
	>6	>11	>16	>20
<i>in situ</i>	0.62	0.30	0.67	1

Una vez obtenida la distribución subyacente óptima la utilizamos para realizar un ajuste más fino de las probabilidades de detección en función del tamaño tumoral, para ello volvemos a hacer uso del trabajo de Weedon-Fekjaer *et al.* [Wee08] que afirma que la probabilidad de detección crece rápidamente para tumores pequeños hasta alcanzar un *plateau*. Por otro lado, Barclay *et al.* [Bar96] sostienen que la probabilidad de detección decrece para tumores de elevado tamaño. A la luz de estos resultados, proponemos para la probabilidad de detección en función del tamaño tumoral una suma de distribuciones normales centradas en 18 mm y 25 mm y pesadas de la forma:

$$P_{\text{tam}} = aN(18 \text{ mm}; \sigma_1) + (1 - a)N(25 \text{ mm}; \sigma_2), \quad (2.8)$$

A partir de esta suma de gaussianas de pesos relativos desde 1:0 (para  $a = 1$ ) hasta 1:1 (para  $a = 0.5$ ) es posible generar una forma simétrica para estos dos casos, o introducir cierta asimetría para valores de  $a$  mayores que 0.5. Las probabilidades de detección para tumores pequeños o muy grandes se ajustan mediante las colas de la distribución. Para este ajuste, las probabilidades de detección se normalizan al máximo.

Realizamos simulaciones de 100 millones de tumores variando los 3 parámetros libres  $\sigma_1$  entre [5,35],  $\sigma_2$  entre [5,70] y  $a$  entre [1.0,0.5] y calculando  $\chi^2$  respecto a los valores reportados por el INC anteriormente referidos:

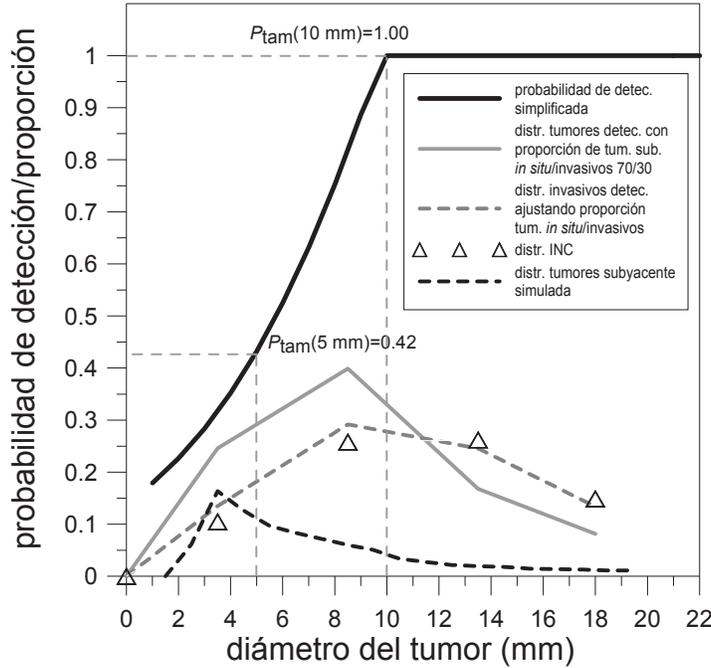


Figura 2.6: La curva negra continua representa la probabilidad de detección simplificada propuesta que pasa por los puntos  $P_{\text{tam}}(5 \text{ mm})$  y  $P_{\text{tam}}(10 \text{ mm})$ , obtenidos de [Wee08]; la curva gris continua representa la distribución de tumores invasivos detectados para una proporción subyacente de tumores invasivos/*in situ* constante de 70/30; la curva gris discontinua representa la distribución de tumores invasivos detectados que obtenemos aplicando las proporciones de tumores invasivos propuesta; la curva negra discontinua es la distribución de tamaños tumorales presentes en la población sometida al cribado, obtenida de nuestra simulación.

$$\chi^2 = \sum_{i=1}^5 \left( \frac{o_i - e_i}{e_i} \right)^2, \quad (2.9)$$

donde  $i = 1, 6, 11, 16$  y  $20 \text{ mm}$ ,  $o_i$  son las proporciones simuladas y  $e_i$  las proporcionadas por el INC para cada tamaño  $i$ .

Esta función se minimiza para  $\sigma_1 = 6 \text{ mm}$  y  $a = 1$ . Finalmente, la distribución se reescala relativa al valor de  $P_{\text{tam}}(5 \text{ mm})$  del 42 %, lo que da lugar a un factor de normalización de 62 y se asume que tamaños superiores a 10 mm tendrán un probabilidad de detección del 100 %, como indicamos anteriormente, resultado la siguiente probabilidad de detección en función del tamaño:

$$P_{\text{tam}} = \begin{cases} 62 \cdot N(\mu = 18 \text{ mm}; \sigma = 6 \text{ mm}) & \text{si } d \leq 10 \text{ mm} \\ 1 & \text{si } d > 10 \text{ mm} \end{cases} \quad (2.10)$$

Con esta probabilidad de detección realizamos una simulación de 100 millones de tumores, que arroja el acuerdo con la distribución de invasivos INC mostrado en la Figura 2.7.

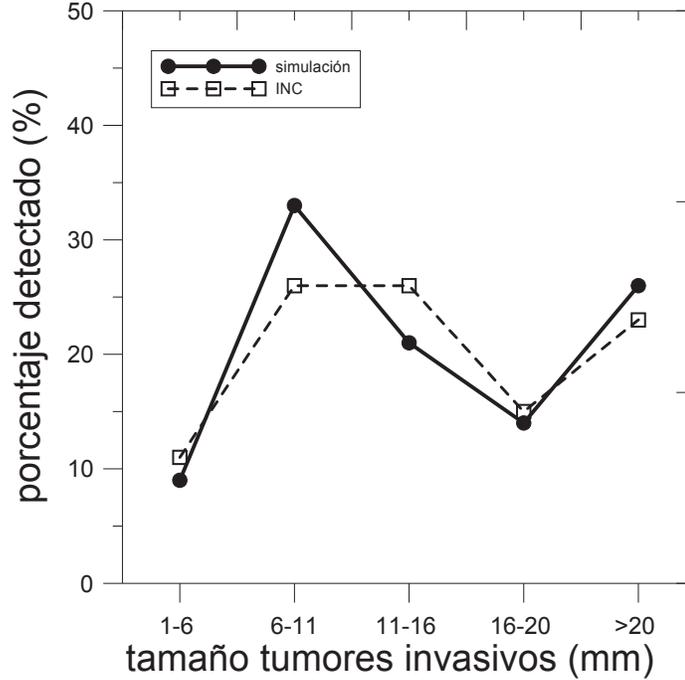


Figura 2.7: Resultado de la comparación de la proporción de tumores invasivos detectados en función del tamaño entre nuestras simulaciones con 100 millones de tumores y los resultados proporcionados por INC [Ros06]. Las incertidumbres de la simulación van desde  $6 \times 10^{-5}$  hasta  $9 \times 10^{-5}$ , por lo que no se han incluido las barras de error.

Mediante este proceso hemos podido obtener una evolución realista de la proporción de tumores invasivos/*in situ* que permite reproducir las distribuciones típicas de tumores invasivos detectados mediante mamografía junto a una parametrización de esta probabilidad de detección en función del tamaño del tumor.

### 2.2.6. Modelo de detección clínica. Cánceres de intervalo

Se considera detección clínica o sintomática aquella que se origina por la manifestación clínica de la enfermedad. Se denominan cánceres de intervalo aquellos que se detectan sintomáticamente entre las diferentes rondas de cribado, bien porque no se detectaron en rodas de cribado previas o porque presentan un crecimiento muy rápido, alcanzando el umbral de detección clínica en el periodo entre cribados consecutivos.

Nuestro modelo incorpora la detección clínica teniendo en cuenta que el principal parámetro implicado en la lamisma es el tamaño tumoral. Si llamamos  $\lambda$  al tiempo medio de detección clínica por cualquier tipo de prueba,  $\lambda^{-1}$  será la probabilidad de detección por unidad de tiempo. Si  $p(\tau)$  es la densidad de probabilidad de detección clínica, la probabilidad de que trascurra un tiempo  $t$  sin que la paciente sea diagnosticada será:

$$F(t) = 1 - \int_0^t p(\tau) d\tau = \int_t^\infty p(\tau) d\tau, \quad (2.11)$$

La probabilidad de ser diagnosticada en  $(t, t + dt)$ ,  $p(t)dt$ , será el producto de  $F(t)$  por  $\lambda^{-1}dt$ :

$$p(t)dt = \lambda^{-1}dt \int_t^{\infty} p(\tau)d\tau, \quad (2.12)$$

que en forma diferencial se escribe como:

$$\frac{dp}{dt} = -\frac{1}{\lambda}p(t), \quad (2.13)$$

Si imponemos la condición de contorno  $p(\infty) = 0$  obtenemos una función de distribución de la detección clínica:

$$p(t) = A \exp\left(-\frac{t}{\lambda}\right), \quad (2.14)$$

Si asumimos que cuando no hay ningún programa de cribado en funcionamiento, todos los tumores acaban siendo detectados clínicamente en algún momento, podemos imponer la condición de normalización:

$$\int_0^{\infty} p(\tau)d\tau = 1, \quad (2.15)$$

$$A = \frac{1}{\lambda}, \quad (2.16)$$

El mismo razonamiento puede ser aplicado a los tamaños de detección, de forma que  $p(d)$  sería la función de distribución de tamaños de detección y  $\lambda$  el tamaño medio de detección clínica. Este tamaño puede ser fijado a 20 mm [Ber06], aunque podrá ser ajustado en virtud del tamaño medio detectado de cánceres de intervalo en el caso de cribado o de detección clínica en ausencia del mismo. En el modelo debe incorporarse *ad hoc* un umbral de detección clínica, a fin de limitar artefactos no deseados debidos a detección sintomática de tumores muy pequeños. Este umbral se fija inicialmente en 15 mm [Ber06].

La distribución de probabilidad de detección clínica en función del tamaño tumoral queda, por tanto, determinada por un único parámetro con significado clínico y fácilmente accesible en la bibliografía, junto con un umbral asociado a la propia detección. Por otra parte, la forma exponencial de la distribución responde al modelado habitual de duración de la fase preclínica cuando se utilizan modelos de Markov para la simulación de la historia natural del tumor. En nuestro modelo, la justificación teórica desarrollada permite pasar de una representación de tiempos a una de tamaños de detección, que a su vez requiere de un tamaño medio de detección frente a una duración media de la fase preclínica, parámetro este mucho más difícil de conocer.

### 2.2.7. Modelo de cánceres radioinducidos

Para introducir en nuestras simulaciones el fenómeno de los cánceres radioinducidos partimos de un modelo de Exceso Absoluto de Riesgo (EAR) basado en el trabajo de Preston *et al.* [Pre02]. El EAR se define en términos de la diferencia entre la incidencia debida a la radiación y la incidencia natural. Si se supone linealidad con la dosis, el número de casos de cáncer de mama por cada  $10^5$  mujeres se puede escribir como:

$$\text{EAR}(e, a, D) = 94 D \exp[-0.05(e - 30)] \left(\frac{a}{60}\right)^\eta, \quad (2.17)$$

con  $\eta = 3.5$  para  $a < 50$  años,  $\eta = 1$  para  $a \geq 50$  años,  $a$  la edad alcanzada,  $e$  la edad de exposición, ambas en años y  $D$  la dosis.

Para transformar la incidencia natural en probabilidad hacemos uso de las definiciones de prevalencia e incidencia vistas en el apartado 2.2.3, así la prevalencia,  $P$ , se relaciona con la incidencia,  $I$ , mediante:

$$P = A \int_0^{e_{\max}} p(\tau)I(e)de \simeq A \sum_i I_i \Delta e, \quad (2.18)$$

$$A = \frac{P}{5 \sum_i I_i}, \quad (2.19)$$

En este apartado debemos señalar que sólo pretendemos obtener un orden de magnitud. No se ha introducido, por tanto, la distribución de edades de la población (se supone constante) y se asume una misma relación entre la incidencia y la prevalencia para los tumores radioinducidos que para los totales. Mediante la utilización de la curva de incidencia y la prevalencia descritas anteriormente se obtiene  $A = 7.57923 \times 10^{-6}$  y, por tanto, la probabilidad de desarrollar un cáncer en  $\Delta t$ , se podrá escribir como:

$$p = AI\Delta t, \quad (2.20)$$

El incremento de la probabilidad debido a incrementos pequeños de la incidencia como es el caso de la radioinducción, se puede escribir:

$$dp = AdI\Delta t, \quad (2.21)$$

Si tomamos,  $\Delta t = 1$  año (la probabilidad de desarrollar un cáncer a una edad  $a$  se entiende como la probabilidad de desarrollarlo durante ese año), el incremento de la probabilidad anual de desarrollar un cáncer radioinducido se relaciona con el exceso absoluto de riesgo mediante:

$$\Delta p_{\text{anual}} = A\Delta I_{\text{anual}}, \quad (2.22)$$

Este tipo de modelos de estimación de riesgo debido a la radiación presentan un carácter semiempírico y carecen de la precisión deseable. Puesto que en este apartado buscamos encontrar únicamente una cota superior de este efecto, hemos optado por introducir en las simulaciones un factor 5, de forma que se sobreestime el número de cánceres radioinducidos. Así, la constante que relaciona el incremento de probabilidad de desarrollar un cáncer por radioinducción debido al exceso absoluto de riesgo será  $A' = 3.78962 \times 10^{-5}$ . De esta forma, la probabilidad de tener un cáncer radioinducido a una edad determinada con un número  $n$  de exposiciones previas será:

$$P_{\text{radio}}(a) = A' \left[ \sum_{i=1}^n \text{EAR}(a, e_i, D_i) \right], \quad (2.23)$$

$$P_{\text{no radio}}(a) = 1 - P_{\text{radio}}(a), \quad (2.24)$$

Por lo que la probabilidad de no tener un cáncer radioinducido a la edad  $a$  será la probabilidad de no tenerlo hasta  $a - t_{\text{int}}$ , por la probabilidad de no desarrollarlo durante  $t_{\text{int}}$ :

$$P_{\text{no radio}}(a) = P_{\text{no radio}}(a - t_{\text{int}}) \times P_{\text{no radio}}(t_{\text{int}}), \quad (2.25)$$

Despejando en (2.24) y utilizando (2.23), se tiene:

$$P_{\text{no radio}}(t_{\text{int}}) = \frac{1 - P_{\text{radio}}(a)}{P_{\text{no radio}}(a - t_{\text{int}})}, \quad (2.26)$$

Y utilizando las probabilidades complementarias queda, finalmente:

$$P_{\text{radio}}(t_{\text{int}}) = 1 - \frac{1 - P_{\text{radio}}(a)}{1 - P_{\text{radio}}(a - t_{\text{int}})}, \quad (2.27)$$

que permite estimar la probabilidad de desarrollar cánceres radioinducidos en cada tiempo de intervalo entre mamografías.

En nuestro modelo se ha supuesto doble exposición de cada mama, con una dosis por mama de 4.5 mGy. Por tanto, de cara a la estimación de riesgos, tomaremos un valor de dosis de 9 mGy para cada cribado. Con estos ingredientes, el número de cánceres radioinducidos por cada  $10^4$  mujeres sometidas a cribado entre los 47 y los 73 años cada 3 años, evaluados a los 73 años, es de 1.2, lo que supondría 1.3 cánceres por cada  $10^5$  mujeres y mamografía (9 mamografías por mujer). Como vimos anteriormente, las estimaciones más recientes hablan de  $10^3$  entre 3 y 6 cánceres radioinducidos, mientras que la *Health Protection Agency* (HPA) del año 2001 sugiere valores de entre 1 y 10. Teniendo en cuenta la amplitud e incertidumbre de las estimaciones, nuestro modelo muestra resultados compatibles con las diferentes estimaciones de la bibliografía en el orden de magnitud del efecto.

En la simulación se acumula el número de exposiciones y la edad de la mujer en cada mamografía, estimándose, en cada tiempo de intervalo, la probabilidad de desarrollar un cáncer radioinducido en virtud de las exposiciones anteriores. Para cada probabilidad estimada se muestrea, de acuerdo con distribución uniforme, si se genera o no el cáncer radioinducido, que se asume como ductal invasivo [NCR06].

### 2.2.8. Mujeres sin cáncer. Falsos positivos

Dentro de las mujeres que se someten al cribado están aquellas que no tienen cáncer, a pesar de lo cual, pueden ser diagnosticadas erróneamente como positivas en virtud de la especificidad de la prueba diagnóstica. Se define esta especificidad, E, como la probabilidad de que el resultado de la prueba sea negativo no estando la enfermedad presente.

$$E = \frac{VN}{VN + FP}, \quad (2.28)$$

donde VN es la proporción de verdaderos negativos y FP de falsos positivos.

Esta especificidad es una característica de la prueba diagnóstica en la que no influye la prevalencia de la enfermedad. Para tener en cuenta esta prevalencia, un indicador más adecuado es el valor predictivo positivo (VPP) de la prueba, entendido como la probabilidad condicionada de tener la enfermedad cuando el resultado de la prueba diagnóstica es positivo:

$$VPP = \frac{VP}{VP + FP}, \quad (2.29)$$

Si se utiliza el teorema de Bayes, se puede relacionar el valor predictivo positivo, (VPP), con la sensibilidad, (S) y la especificidad, (E) de las pruebas, mediante la prevalencia, (P), de la enfermedad en la población (información *a priori*):

$$\text{VPP} = \frac{\text{VP}}{\text{VP} + \text{FP}} = \frac{P \times S}{P \times S + (1 - P)(1 - E)}, \quad (2.30)$$

Este valor predictivo se evalúa a partir de los estudios poblacionales de cribado, pues va más allá de las características intrínsecas de la prueba diagnóstica, e incluye propiedades de la población sometida al cribado a través de la prevalencia. El programa inglés [Ber06] proporciona la evolución de este VPP en los últimos años, que se mueve en el entorno del 15%. El porcentaje de mujeres rellamadas que efectivamente presentan cáncer de mama tras la realización de pruebas complementarias es del 18.8% en el informe inglés del año 2012 [UKP12].

Nuestras simulaciones incluyen como *input* la especificidad. Para su ajuste buscamos reproducir un VPP en el entorno del 15%, simulando un programa con la configuración del inglés: 50-70 años y tiempo de intervalo 3 años. El valor de la especificidad que genera VPP compatibles con el 15% es de 96.6%.

## 2.3. MODIFICACIONES PARA ESTUDIAR LA REDUCCIÓN DE MORTALIDAD

Como se indicó anteriormente al comienzo del apartado 2.2, la primera parte del trabajo se utiliza también para generar los *inputs* de entrada empleados en la segunda y tercera, es decir, la sensibilidad global y la curva de adelanto en la detección. En este apartado se describen todas las modificaciones que se realizan para introducir la reducción de mortalidad en las simulaciones de los programas reales de cribado (ver Figura 2.8).

### 2.3.1. Modificaciones para tratar el problema de la supervivencia

Hasta el momento, nuestro código es capaz de generar únicamente valores de tasa de detección por ronda, número de cánceres de intervalo, distribución de tamaños detectados, edad de detección y otros parámetros cuantitativos asociados, todos ellos, al momento de la detección. En esta segunda versión pretendemos abordar el problema de la simulación de la mortalidad por cáncer de mama, con y sin programa de cribado. Trataremos este problema desde dos grandes ejes:

1. Generación de las curvas de adelanto en la detección y la sensibilidad global mediante el código anterior. Estas serán usadas como entradas en la segunda parte de nuestro trabajo.
2. Utilización de un modelo de supervivencia tras detección y tratamiento loco-regional. Aspecto central que permitirá comprobar si los programas de cribado y los tratamientos asociados producen una reducción de la mortalidad por cáncer de mama o, simplemente, un adelanto en la detección sin alterar el momento de la muerte.

De acuerdo con esta filosofía, dada una mujer con cáncer de mama, se muestreará, de acuerdo con la sensibilidad global, si se detecta o no en el programa de cribado. Este valor de sensibilidad global, calculado para cada configuración, contendrá toda la información condensada relativa a la misma, así como de la evolución de los tumores. Si el resultado es la detección, se muestreará, de acuerdo con la distribución de adelanto, cuánto tiempo se ganará respecto a la detección clínica. En esta segunda iteración del modelo generamos el momento de nacimiento de las mujeres mediante el muestreo de un año entre 0 y 100 antes del año de evaluación de la mortalidad. Las mujeres entrarán en el programa una vez que, alcanzando la edad de inicio sin síntomas clínicos de la enfermedad, acepten entrar en el mismo según el porcentaje de participación, de manera que no se generen efectos de borde asociados al comienzo del programa de cribado. Este programa llevará así un tiempo indefinido en funcionamiento en el año en que se evalúa la mortalidad anual por cáncer de mama.

Las simulaciones de cada configuración de un programa de cribado estudiada (40-70 años, 50-70 años, 50-65 años con tiempos de intervalo de 1, 2 y 3 años) se realizan con  $5 \times 10^8$  historias y se expresa la mortalidad por año y  $10^5$  mujeres (IARC [IAR.12]). Las incertidumbres se calcularon con un factor de cobertura  $k=3$ .

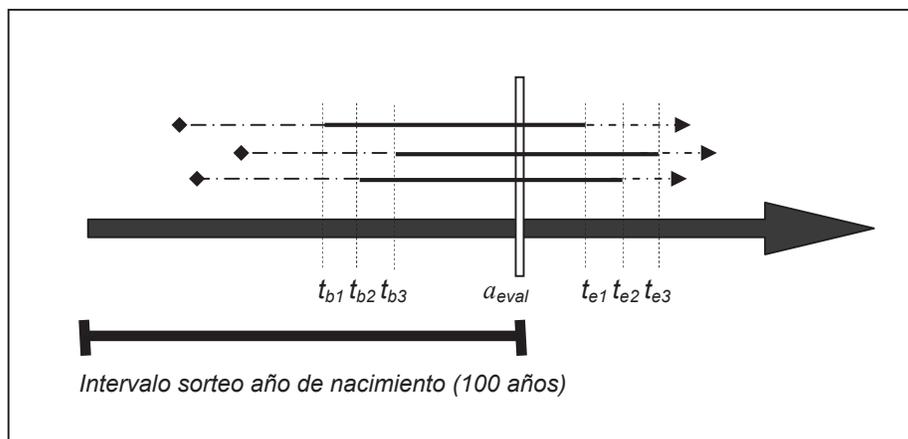


Figura 2.8: Se representa de forma esquemática la estructura de las simulaciones de la primera iteración del código. Se muestrean los momentos de nacimiento y se sigue la historia de la mujer hasta que alcanza la edad de comienzo de cribado en un instante determinado,  $t_{bi}$ . Se muestrea la participación y posteriormente se evalúa en un tiempo determinado,  $a_{eval}$ , la mortalidad por cáncer de mama. Las mujeres saldrán del programa alcanzada la edad de finalización en el instante  $t_{ei}$ . Con esta estructura se simula un programa de cribado que funciona desde un tiempo indeterminado, evitando efectos de borde o de sincronización de las poblaciones que se someten al cribado (rondas de prevalencia e incidencia).

Los algoritmos empleados en esta segunda parte se describen a continuación.

### 2.3.2. Muestreo del año de nacimiento. Corrección de nacimiento uniforme

A diferencia de la iteración anterior, donde se muestreaba la edad de la mujer al inicio del programa de cribado según una distribución conocida de referencia (ver apartado 2.2.1),

en esta segunda versión se muestrea el año de nacimiento de manera uniforme entre 0 y 100 años antes del momento de la evaluación de mortalidad. Una vez obtenido este año de nacimiento, se simula la historia de la mujer hasta el momento en que alcanza la edad de inicio en el programa ( $E_{MIN}$ ). La distribución de edad de las mujeres en el año de evaluación de la mortalidad no se ajustará, por tanto, necesariamente a la de referencia. Para corregir este artefacto se modifican la anchura de los intervalos de edad de nacimiento (inicialmente iguales) de manera iterativa hasta encontrar una distribución que se ajuste a la de referencia. En la Figura 2.9 se muestran estas distribuciones y en la 2.10 los pesos relativos aplicados a cada intervalo de edades para reproducir esta distribución. Se comprueba como son desfavorecidos los intervalos de nacimiento más próximos al momento de la evaluación de mortalidad (entre los años 80 y 100), que darían lugar a mujeres entre 0 y 20 años. Se pesan más los intervalos de nacimiento entre 75 y 55 que generan mujeres con edades entre 45 y 25 años en el año de evaluación y, finalmente, se pasan especialmente los intervalos entre 10 y 20, que dan lugar a mujeres de mayor edad (80-90 años), cuya abundancia mostraba un claro déficit con la generación uniforme de años de nacimiento.

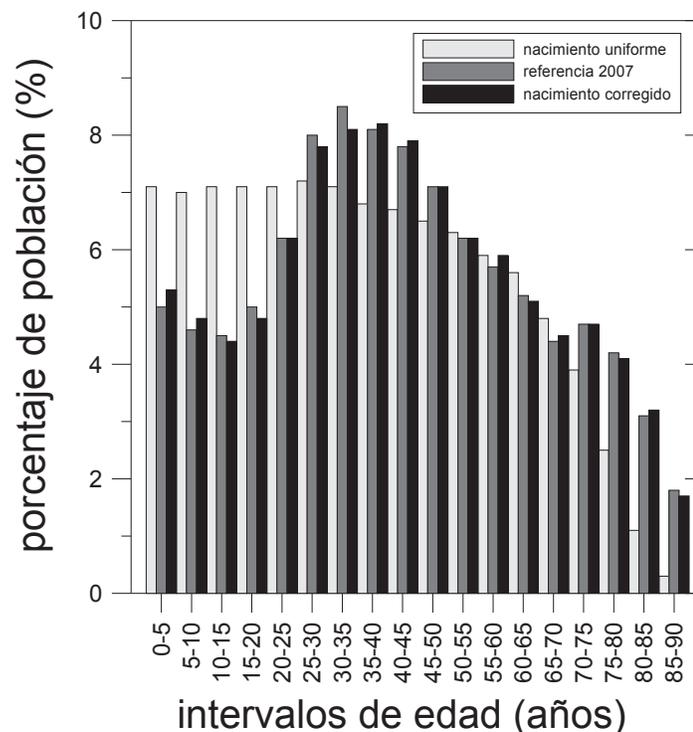


Figura 2.9: Gráfico que representa el proceso de corrección del nacimiento uniforme en el momento de evaluación de la mortalidad (distribución gris clara). El histograma gris oscuro representa la distribución de referencia (2007) y el negro la distribución de nacimiento uniforme corregida mediante los pesos adecuados.

### 2.3.3. Sensibilidad global y curva de adelanto

Definimos la sensibilidad global como el porcentaje de tumores que son detectados en cribado relativo al número de ellos que alcanzan el umbral de detectabilidad por mamografía.

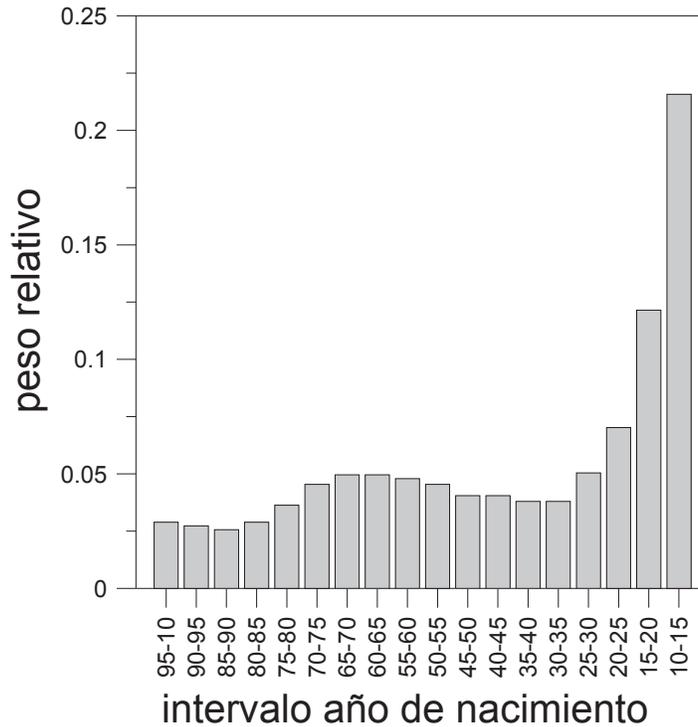


Figura 2.10: Pesos relativos aplicados a la distribución de nacimiento uniforme.

grafía en algún momento entre la edad de comienzo del cribado y la edad de finalización. Para el cálculo de esta sensibilidad global se utiliza el programa principal realizando simulaciones de  $5 \times 10^5$  mujeres sin considerar la distribución de edades, esto es, todas las mujeres se colocan a la edad  $E_{MIN}$ , de forma que se eliminen efectos asociados a esta distribución. Esta definición caracteriza el programa de manera global, si bien no debe aplicarse a mujeres individuales fuera del contexto de la simulación, pues la sensibilidad de la mamografía como prueba diagnóstica depende de numeroso factores como la densidad de la mama, tamaño del tumor, terapia hormonal, número de proyecciones, calidad de la mamografía e incluso del la experiencia y sesgos en el diagnóstico asociados al radiólogo. En nuestras simulaciones, la sensibilidad global juega el papel de algoritmo “condensado” que puede aplicarse en virtud del sorteo estocástico y no es comparable de manera inmediata con la sensibilidad de la prueba diagnóstica. Para el muestreo de la detección según la sensibilidad global se utiliza la distribución uniforme.

La distribución del tiempo de adelanto que supone el cribado respecto a la detección clínica es el otro *input* que se introduce en esta segunda iteración. Las curvas se calculan utilizando el programa original y el resultado para las configuraciones estudiadas en esta segunda parte son las curvas mostradas en la Figura 2.11.

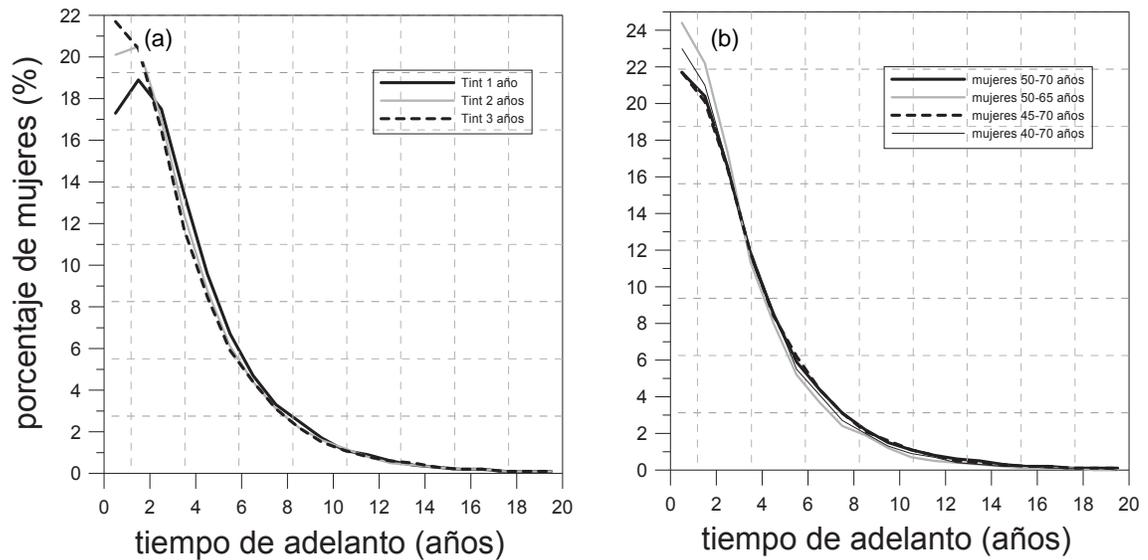


Figura 2.11: Valores de la curva de adelanto, *lead time*, obtenidos por simulación (a) para el rango de edades 50-70 años y tiempos de intervalo 1, 2 y 3 años y (b) comparación de diferentes rangos de edades con tiempo de intervalo 3 años. En todos los casos se estiman medias entre 2 y 3 años.

La contrastación de estas distribuciones resulta extremadamente complicada pues no son directamente medibles. No es posible conocer cuándo sería detectado clínicamente un tumor que aparece en cribado; a lo sumo sería posible determinar el valor medio de este adelanto comparando tiempos de detección del grupo de cribado con los del grupo de control. Actualmente se acepta un valor medio de consenso de 3.33 años para el *lead time* [Duf13, Etz14], suponiendo para la distribución funciones exponenciales o de Weibull con factor de forma 0.5 o 2. Todas estas tienen un aspecto similar a las mostradas en la Figura 2.11 obtenidas de nuestro programa.

Tabla 2.8: Valores medios del tiempo de adelanto resultado de un ajuste log-normal de las curvas obtenidas en nuestras simulaciones. Se muestran los valores obtenidos para las dos configuraciones típicas 40-70 y 50-70 años con tiempos de intervalo de 1, 2 y 3 años. En la última columna se pondera por la sensibilidad global.

Configuración	Tiempo de intervalo	Lead time medio	Lead time ponderado
40-70 años	1 año	4.14 años	3.48 años
	2 años	3.97 años	2.98 años
	3 años	4.02 años	2.65 años
50-70 años	1 año	5.04 años	4.38 años
	2 años	4.67 años	3.69 años
	3 años	4.67 años	3.36 años

Debe tenerse en cuenta que este valor medio de 3.33 años incluye aquellos casos en los que no hay ningún adelanto en la detección o, dicho de otro modo, incluye los casos

que no se detectan en el cribado. Si realizamos una media ponderada del *lead time* con el valor de la sensibilidad global para estimar esta corrección, resultan los valores mostrados en la cuarta columna de la Tabla 2.8. Puede apreciarse como los valores de las diferentes configuraciones se aproximan a este valor de consenso.

La implementación práctica de este formalismo basado en la curva de adelanto y la sensibilidad global, se realiza muestreando el tiempo de adelanto según el histograma con ayuda de la distribución uniforme. Para ello, primero se muestrea si el tumor es detectado en cribado mediante la sensibilidad global y, si lo es, se sortea el tiempo de adelanto, que debe cumplir dos condiciones:

1. La edad de adelanto debe quedar dentro del intervalo de edades que comprende el cribado, es decir entre  $E_{\text{MIN}}$  y  $E_{\text{MAX}}$ .
2. El tamaño resultante del adelanto en la detección debe superar el umbral de detectabilidad por mamografía prefijado en 2 mm.

Si el valor obtenido en el sorteo no cumple ambas condiciones, se rechaza volviendo a sortear hasta obtener un valor realista (algoritmo de rechazo).

Puesto que las curvas de adelanto se calcula hasta 20 años, aquellos casos en que la edad de detección clínica se tal que supere  $E_{\text{MAX}} + 20$ , se consideran “tumores de crecimiento lento” y las historias de estas mujeres se tratan como si no tuvieran tumor, calculándose la muerte por causas diferentes al cáncer de mama.

En nuestro segundo modelo, por tanto, la detección en el cribado es introducida de forma “condensada” a través de la sensibilidad global y la curva de adelanto o *lead time*. En estos dos *inputs* está contenida la información referente a la configuración del cribado y las características de las mujeres en ese intervalo de edad, aplicándose luego de manera estocástica en la simulación del cribado.

#### 2.3.4. Supervivencia tras detección y tratamiento

Un aspecto central de esta segunda iteración de nuestro programa de simulación es el cálculo de la supervivencia tras detección y tratamiento loco-regional de la enfermedad. Si bien el tamaño tumoral es ampliamente reconocido como el más potente predictor de la supervivencia [Kop98], no existe el mismo consenso sobre su utilización exclusiva frente a otras candidatas como podrían ser los diferentes subtipos histológicos o los marcadores tumorales como el p53, HER-2, receptores de progesterona o estrógenos, etc.

El trabajo de Michealson *et al.* [Mic02] sobre la población de 1352 mujeres recopilada por Van Nuys, con el 95 % de los tumores detectados entre 1980 y 1990, permite afirmar que la supervivencia de los pacientes de carcinoma de mama invasivo es una función directa del tamaño tumoral, independiente del método de detección, desautorizando, por tanto, la idea de que la detección en cribado podría encontrar tumores invasivos que con el mismo tamaño que los detectados clínicamente, tendrían una menor letalidad. Esta independencia se revela en dos aspectos:

1. Poblaciones de mujeres con el mismo tamaño de detección tienen la misma supervivencia, independientemente del método de detección.
2. Los valores de supervivencia calculados analíticamente son igualmente precisos para tumores detectados clínicamente o mamográficamente, incluyendo aquellos procedentes del cribado.

La correlación matemática entre tamaño tumoral y supervivencia es consistente con un mecanismo biológico en el que la metástasis letal a distancia ocurre por eventos discretos de diseminación que, para tumores de carcinoma de mama invasivo en el momento de la cirugía, tiene una probabilidad de ocurrencia de aproximadamente  $10^{-9}$  por célula. Estos eventos, que ocurrirían una vez por célula, están asociados a mutaciones en el momento de la división de la célula [Kuu97], con una probabilidad de ocurrencia por célula,  $p$ , de forma que la probabilidad de que no se de un evento de este tipo para  $N$  células sería:

$$F = (1 - p)^N, \quad (2.31)$$

Para valores pequeños de  $p$  y grandes de  $N$  puede ser aproximada a la distribución de Poisson:

$$F = (1 - p)^N \cong e^{-Np}, \quad (2.32)$$

El número de células,  $N$ , es función del diámetro,  $D$ , y de la densidad de células,  $s$ :

$$N = s \frac{4}{3} \pi \left( \frac{D}{2} \right)^3, \quad (2.33)$$

Por otra parte, la probabilidad  $p = -\ln(F)/N$ , se puede ajustar fácilmente a partir de los datos de supervivencia de Tabar *et al.* [Tab92, Tab00, Tab00b], Tubiana y Koscielny [Tub90, Tub91, Tub99] y Koscielny y Tubiana [Tub84], asumiendo forma esférica y densidad  $s$  de  $10^8$  células por cc. Con los valores calculados de la probabilidad,  $p$ , y los tamaños tumorales, se puede realizar un ajuste del tipo:

$$p = a N^b, \quad (2.34)$$

que muestra una declinación con el tiempo de la probabilidad de estos eventos ya que el valor de  $b$  resulta negativo. Por tanto, la fracción de supervivencia se podrá expresar:

$$F = e^{-Np} = e^{-a(s\frac{\pi}{6})^{b+1} D^{3b+3}}, \quad (2.35)$$

De igual forma, es posible parametrizar el ajuste de las curvas de supervivencia Kaplan-Meier del grupo de mujeres de Van Nuys con la siguiente expresión teórica:

$$F = e^{-QD^Z}, \quad (2.36)$$

donde  $F$  es la fracción de supervivencia,  $D$  tamaño tumoral,  $Q$  y  $Z$  son constantes del ajuste ( $Z$  sin dimensiones y  $Q$  [longitud] $^{-Z}$ ).

De forma similar, si introducimos el tiempo y definimos como  $C$  el número de células presente por el número de días que el tumor ha acumulado desde un punto temporal y, asumiendo un crecimiento exponencial en los tumores clínicos del tipo:

$$N = N_0 e^{rt}, \quad (2.37)$$

con  $N_0 = 1$ , número de células en  $t = 0$ , y  $r = \ln(2)/t_D$ , con  $t_D$  tiempo de duplicación del tumor, se puede escribir:

$$C = \int_0^t e^{rt} dt \simeq \frac{N}{r}, \quad (2.38)$$

Sustituyendo  $C$  por  $N$  es posible realizar un desarrollo similar al anterior con  $p' = rp$ . Este nuevo mecanismo reflejaría procesos biológicos que ocurren a lo largo del tiempo, como serían desprendimiento o re inserción celular, angiogénesis, etc. En cualquier caso, es posible ajustar la supervivencia a una ecuación de la forma (2.35), que refleje la fracción de supervivencia a un tiempo prefijado en función, únicamente, del tamaño tumoral. En nuestro caso, hemos llevado a cabo este ajuste usando los datos de Van Nuys para supervivencias entre 2.5 y 15 años, obtenemos así los valores de  $Q$  y  $Z$  que se indican en la Tabla 2.9. Estos ajustes se muestran en la Figura 2.12.

Tabla 2.9: Resultado del ajuste de los datos al modelo exponencial dado por la ecuación (2.35) para 2.5, 5, 7.5, 10, 12.5 y 15 años de supervivencia. Los datos corresponden a Van Nuys [Sil00] obtenidos de 1352 pacientes con carcinomas de mama invasivos detectados entre 1966 y 1990.

Sup.(años)	$Q$	$Z$
2.5	0.00009±0.00003	1.894±0.087
5	0.00033±0.00004	1.836±0.029
7.5	0.0017±0.0002	1.500±0.023
10	0.0030±0.0004	1.375±0.038
12.5	0.0041±0.0003	1.333±0.022
15	0.0057±0.0005	1.334±0.025

Este ajuste podría presentar algunas objeciones desde el punto de vista teórico. Así, puesto que los datos no están aleatorizados por edad, esta podría ser una variable de confusión, reflejando el ajuste únicamente un adelanto sin más en el momento de la detección, que tendría como consecuencia una mayor supervivencia artificial, puesto que no retrasaría el momento de la muerte (lo que se conoce como *lead time*). Además, el ajuste de la población de Van Nuys se refiere únicamente a tumores invasivos. Son excluidos del análisis los CDis que, como sabemos, son detectados en mamografía con una mayor frecuencia y suelen estar asociados a un mejor pronóstico. Por otra parte, el efecto de la mamografía de cribado estaría incluido en el análisis, puesto que el 95% de los casos incluidos en este estudio son diagnosticados entre los años 1980 y 1990, con programas de cribado en funcionamiento.

El modo en el que se introduce el modelo de supervivencia para tumores invasivos en nuestro modelo de simulación tratará de solventar las objeciones anteriores, de forma que si el adelanto en la detección que supone el cribado no se refleja en la reducción de mortalidad, habrá que suponer que se trata de un adelanto “artificial”, únicamente

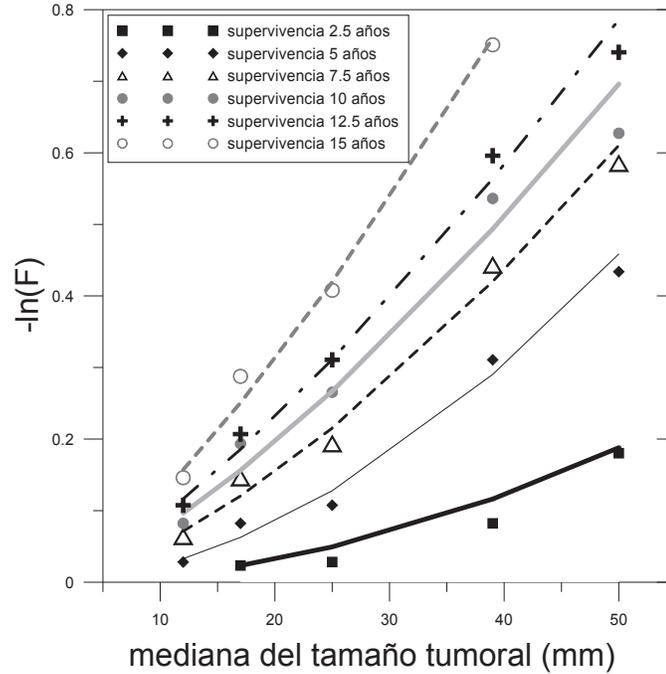


Figura 2.12: Representación de los ajustes para diferentes supervivencias en función de la mediana del tamaño tumoral. Los datos pertenecen al estudio de Van Nuys [Sil00].

asociado al *lead time*, de forma que la ganancia de tiempo en el tratamiento no retrasa la edad de muerte por cáncer de mama. Por otra parte, la justificación teórica realizada por Michaelson *et al.* [Mic02] para los ajuste es un argumento potente a favor de una ganancia biológica asociada al adelanto en la detección.

Para calcular la supervivencia en nuestro modelo hacemos uso de las probabilidades condicionadas, de forma que la probabilidad de que una mujer con tumor detectado de tamaño  $D$  sobreviva  $a + 1$  años desde la detección, cuando ha sobrevivido hasta  $a$  se puede escribir como:

$$F(a \rightarrow a + 1) = \frac{F(a + 1)}{F(a)}, \quad (2.39)$$

Las dos probabilidades de supervivencia del segundo miembro se pueden calcular interpolando en los ajustes realizados. Mediante la expresión anterior se calcula la supervivencia anual en los 15 primeros años tras la detección y tratamiento loco-regional [Kar99], muestreando esta supervivencia anual mediante la distribución uniforme.

### 2.3.5. Evaluación de la mortalidad por cáncer de mama. Incertidumbre

Para la evaluación de la mortalidad se fija el año 100 a partir del momento en que comienza la simulación del año de nacimiento. Las mujeres ingresan en el programa cuando alcanzan la edad mínima fijada en el mismo y aceptan someterse al cribado de acuerdo con un porcentaje de aceptación también prefijado en la simulación. Como indicamos anteriormente, la distribución de año de nacimiento se modifica de forma que se tenga

una distribución de edades coherente en este año de evaluación. De esta manera, la propia estructura de la simulación elimina los efectos de borde asociados al comienzo de un programa (rondas de prevalencia, menores incidencias en diferentes grupos de edad, etc.), pues el momento de comienzo del mismo no está fijado en el tiempo. En las simulaciones se acumula la mortalidad por cáncer de mama en este año de evaluación, contabilizándose las mujeres vivas en el mismo. De esta forma es posible estimar el riesgo de mortalidad por cáncer de mama anual (RR) mediante la siguiente expresión:

$$\text{RR} = \frac{M_0 - M_C}{M_0}, \quad (2.40)$$

$M_0$  es la mortalidad sin programa de cribado y  $M_C$  la mortalidad con el programa de cribado instaurado. A la hora de estimar la reducción de mortalidad que supuestamente proporcionan los programas de cribado mamográfico, debe calcularse la misma sin programa de cribado. Esta reducción de mortalidad sin programa es debida únicamente a la detección clínica o sistemática, forzándose en el programa mediante una sensibilidad global nula. De esta forma se realizan 100 simulaciones de 5 millones de historias cada una y se evalúa la incertidumbre en la mortalidad con un factor de cobertura  $k=3$ .

## 2.4. MODIFICACIONES PARA EL ESTUDIO Y LA SIMULACIÓN DE LOS ENSAYOS ALEATORIOS CONTROLADOS

Los ensayos aleatorios controlados son, hasta el momento, la mejor herramienta para abordar la justificación de las diferentes configuraciones de los programas de cribado mamográfico [UKP12, Got04], si bien se centran fundamentalmente en estimar la reducción de mortalidad por cáncer de mama, siendo mucho más difícil extraer de ellos conclusiones sobre el sobrediagnóstico. Estos ensayos no están exentos de sesgos que ponen en duda tanto su validez interna (metodológica) como externa (aplicación a otras poblaciones). De entre estos problemas podemos destacar aquellos asociados al tamaño de la muestra [Kop93], que, si bien puede estimarse estadísticamente, supone un problema logístico a la hora de reunir cohortes del tamaño requerido. Otros problemas que comprometen la validez interna son la asignación más o menos aleatoria de los miembros del grupo de cribado y control [Got09], las dificultades y sesgos asociados a la asignación de causa de muerte y el tiempo de seguimiento [Got04]. Todos estos sesgos pueden ser eliminados mediante la simulación Monte Carlo, que permite investigar cuáles son aquellos ensayos que aportan más evidencia desde el punto de vista de la validez interna. Además, puesto que en las simulaciones no se han introducido diferencias poblacionales (distribución de edades, incidencia, prevalencia, distribuciones de densidad de mama, proporción de tumores invasivos/*in situ* subyacente, etc.), estas simulaciones permitirán conocer cuáles de entre ellos son los de mayor validez externa.

Hemos pretendido simular los principales ensayos aleatorios controlados de cribado mamográfico llevados a cabo hasta la fecha, recopilados en Gøtzsche *et al.* [Got04]. Así, partiendo del diseño desarrollado en los apartados previos, hemos realizado modificaciones adicionales para simular estos ensayos. Estas modificaciones suponen:

1. Adecuar la manera de generar tanto la distribución de adelanto como la sensibilidad global para reproducir ensayos aleatorios.
2. Fijar un momento en el tiempo para el comienzo del ensayo, en el cual, las mujeres que tengan edades comprendidas  $[E_{\text{MIN}}, E_{\text{MAX}}]$  serán asignadas tanto al grupo de cribado como al de control.
3. Seguir la evolución de estos grupos cerrados, acumulando el número de muertes por cáncer de mama en el periodo de seguimiento establecido (13 años).
4. Estimar el riesgo relativo de mortalidad por cáncer de mama entre los dos grupos y estimar la variabilidad asociada al tamaño de la muestra.

Respecto al punto 1, hay que hacer notar que a diferencia de la simulación de los programas reales  $[E_{\text{MIN}}, E_{\text{MAX}}]$ , no representan sólo las edades de comienzo y final del programa, sino el intervalo de edades de las mujeres que pueden someterse al ensayo al comienzo del mismo. Para no alterar la distribución de edades al comienzo del ensayo, se muestrean las edades de las mujeres y solo se consideran aquellas que tienen edades comprendidas en el intervalo definido por el ensayo real. Más aún, tanto las distribuciones de adelanto como la sensibilidad global se calculan para un periodo entre  $a_0$  y  $a_0 + n^{\circ}\text{rondas} \times T_{\text{int}}$ , siendo  $a_0$  (edad de comienzo en el ensayo) la que estará comprendida entre  $E_{\text{MIN}}$  y  $E_{\text{MAX}}$ ,  $n^{\circ}\text{rondas}$  el número de rondas del que consta el ensayo y  $T_{\text{int}}$  el tiempo de intervalo entre exámenes.

En lo que respecta a los puntos 2 y 3, en la versión previa que simula los programas poblacionales de cribado, se evaluaba la mortalidad anual en un momento del tiempo elegido, de forma que la distribución de edades de las mujeres fuera acorde con una distribución realista [INE07]. El momento de inicio del programa no estaba fijado respecto al anterior: las mujeres entraban en el programa si alcanzaban la edad  $E_{\text{MIN}}$  libres de enfermedad y “aceptaban” (de acuerdo con el porcentaje de participación) someterse al cribado. En este modelo previo, está implícita la idea de que el programa corre desde un tiempo indefinido, evitando así “efectos de borde” asociados al comienzo del programa. Sin embargo, a la hora de simular los ensayos aleatorios, necesitamos un punto temporal del comienzo del ensayo, por lo que para no modificar la estructura anterior tomamos  $y_{\text{eval}}$  como el comienzo del ensayo. Este punto sigue cumpliendo la propiedad que se le exigió, a saber, tener una distribución de edades correcta de la población y, además, es el comienzo de la acumulación de mortalidad en un periodo  $y_{\text{eval}} + 13$  años. Para conseguir los grupos cerrados característicos de los ensayos con una modificación mínima de la estructura de la simulación, las mujeres de las que se simula su historia y que llegan a  $y_{\text{eval}}$  con una edad fuera del intervalo son descartadas (ver Figura 2.13).

Se simulan tantas historias como mujeres hay en los grupos de cribado y control. El grupo de control se simula colocando una sensibilidad global de detección por cribado de 0, de esta forma solo se produce detección clínica en este grupo. Para estimar la variabilidad asociada al tamaño de los ensayos aleatorios reales se realizarán  $10^3$  simulaciones del tamaño del grupo de cribado y control de forma independiente, estimándose el riesgo relativo RR como:

$$\text{RR} = \frac{M_c/N_c}{M_0/N_0}, \quad (2.41)$$

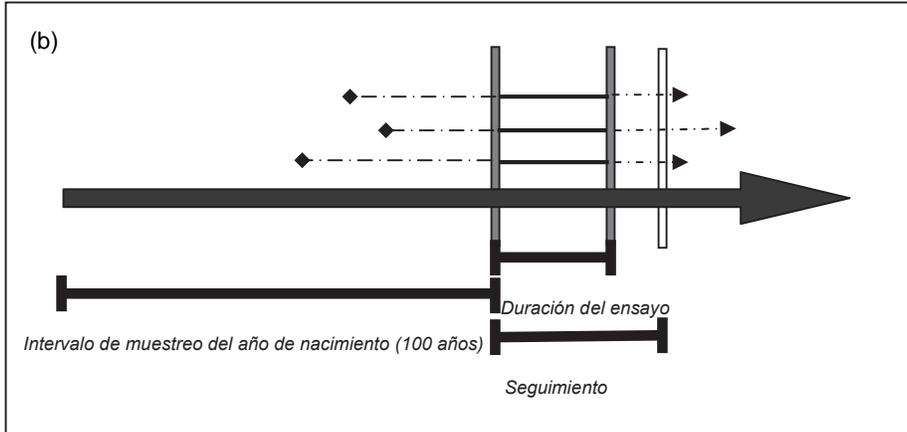
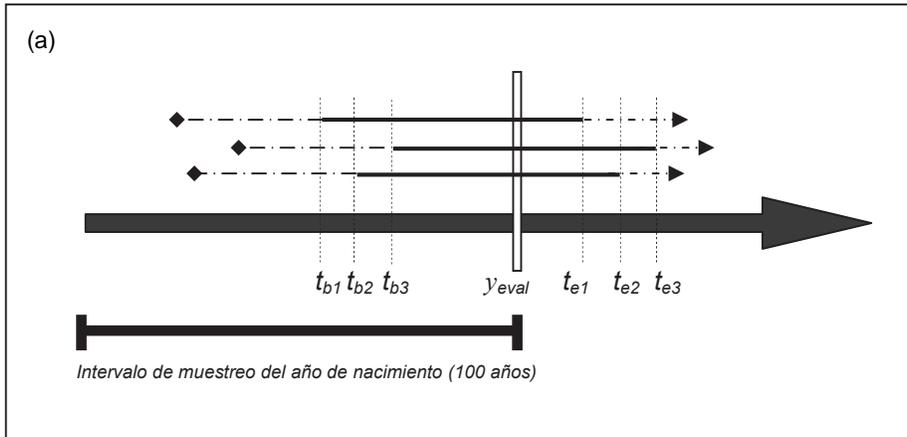


Figura 2.13: (a) Esquema de la estructura de simulación para el estudio de la reducción de mortalidad en los programas de cribado reales.  $t_{b1}$ ,  $t_{b2}$  y  $t_{b3}$  son los momentos en los que las diferentes mujeres alcanzan  $E_{MIN}$  y se someten al programa.  $t_{e1}$ ,  $t_{e2}$  y  $t_{e3}$  son los momentos en que alcanzan  $E_{MAX}$  y abandonan el mismo. (b) Esquema de la estructura de simulación para el estudio de la reducción de mortalidad en los ensayos aleatorios controlados. Aparecen indicados la duración del ensayo y el límite del seguimiento del mismo.

donde numerador y denominador se obtienen de la media de las  $10^3$  muestras de tamaño  $N_C$  o  $N_0$  simuladas. La evaluación de los ensayos aleatorios requiere además estimar la variabilidad del valor de RR asociada al tamaño de los grupos de cribado y control. Para eso, comparamos cada una de las  $10^3$  muestras del grupo de cribado con otra del grupo de control, evitando cualquier tipo de correlación entre ambas. Obtenemos así una distribución muestral de tamaño  $10^3$  para el RR. Con esta distribución construimos el intervalo de confianza del 95 %, determinando el límite superior por el valor del RR por encima del cual quedan un 2.5 % de los casos y el inferior como aquel que deja por debajo un 2.5 % de los casos. Previamente se han ajustado las distribuciones a una log-normal. En este punto hay que señalar que el riesgo relativo se refiere a las mujeres que son invitadas al cribado respecto al grupo de control, independientemente de que se sometan o no a todas las rondas del cribado.

En todas las simulaciones hemos utilizado distribución de edades, incidencia, prevalencia, proporción de tumores subyacente invasivos/*in situ* y tamaño medio de detección clínica idénticas a las empleadas en las simulaciones previas. Los parámetros particulares de cada ensayo incluidos en las simulaciones se detallan en la Tabla 2.10.

Tabla 2.10: Configuración de los ensayos aleatorios usada en las simulaciones. (1) 65 %, 58 %, 50 % y 40 % en cada ronda.(2) Tint efectivo entre 1.5 y 2 ensayo real. (3) Global 70 %. (4) Primera ronda. PP por ronda 100 %, 89 %, 88 %,87 % y 86 % en las 5 primeras rondas. (5) Primera ronda. PP por ronda 100 %, 90 %, 89 %,88 % y 87 % en las 5 primeras rondas. (6) Entre 24 y 28 meses. (7) Global 80 %. (8) 85 %, 78 %, 79 %, 77 % y 75 % en 5 primeras rondas. Corresponden al grupo de edad 39-49 años. (9) 85 %, 78 %, 79 %, 77 % y 75 % en 5 primeras rondas. Corresponden al grupo de edad 39-49 años. (10) Porcentaje de mujeres que se someten al menos a un examen, no necesariamente el primero. Tomamos la aproximación 68 %, 64 %, 60, 56 % y 50 % a partir de la 5ª ronda. (11) 2 años mujeres entre 40-49 y 3 mayores de 50. (12) 91-94 % menores de 60, 50-80 % mayores de 60 años. (13) Entre 2 y 2.5 años. (14) 90 % primera ronda 80 % segunda.

Ensayo	Grupo de edad	Tiempo intervalo (años)	Tamaño grupo cribado	Tamaño grupo de control	Porcentaje participación (1ª ronda)	Número de rondas
New York HIP	40-64	1.0	31000	31000	65 % <sup>(1)</sup>	4
Malmö I	45-69	1.75 <sup>(2)</sup> eff.	21088	21195	74 % <sup>(3)</sup>	8
Canada I	40-49	1.0	25214	25216	100 % <sup>(4)</sup>	5
Canada II	50-59	1.0	19711	19694	100 % <sup>(5)</sup>	5
Estocolmo	39-65	2.2 <sup>(6)</sup> eff.	40000	20000	82 % <sup>(7)</sup>	4
Goteborg	39-59	1.5	26000	26000	84 % <sup>(8)</sup>	5
UK Age Trial	39-41	1.0	53000	106000	81 % <sup>(9)</sup>	10
Kopparberg	40-75	2.5 <sup>(10)</sup> eff.	47389	22658	85 % <sup>(11)</sup>	4
Ostergötland	40-75	2.25 <sup>(12)</sup> eff.	47001	45933	85 % <sup>(13)</sup>	3 <sup>(14)</sup>

Para cada ensayo aleatorio simulado se calculan la distribución de adelanto y la sensibilidad global con simulaciones de 5 millones de historias donde se configura la edad de las mujeres, el tiempo de intervalo entre pruebas y el número de rondas del ensayo. Se ha realizado una modificación de este programa para introducir el efecto de diferentes participaciones por ronda. Esta modificación consiste en introducir un muestreo uniforme

en función de la participación en cada ronda para la detección del tumor, previo a la probabilidad de detección, de esta forma se generan curvas de adelanto y, fundamentalmente, sensibilidades globales más reducidas que tienen en cuenta las diferentes participaciones por ronda. Con estos *inputs* junto con el intervalo de edades, tamaño de los grupos de cribado y control y porcentaje global de participación, se realizan las simulaciones de los ensayos aleatorios reales.

## Capítulo 3

# APLICACIONES DEL MODELO

En este capítulo aplicaremos el modelo desarrollado en el capítulo anterior a tres aspectos fundamentales relacionados con el cribado mamográfico:

1. Reproducción de los programas de cribado mamográfico reales en términos de resultados asociados a la detección, este es el objeto de nuestra primera publicación sobre la materia [For10]. Seguidamente se llevará a cabo un análisis de la sensibilidad del modelo, analizando la influencia de las principales entradas del programa y su criticidad. Este apartado concluye con un estudio del sobrediagnóstico, puesto que este fenómeno está asociado a diferencias en la detección y aparición clínica de cánceres en una población con y sin programa de cribado instaurado [For19].
2. Estudio de la reducción de mortalidad que se atribuye a estos programas mediante la introducción de la supervivencia tras tratamiento, aspecto fundamental de los programas de cribado, en el que se utilizarán las curvas de adelanto y sensibilidad global descritas en el capítulo anterior, junto con la supervivencia tras detección y tratamiento [Zam12]. Para completar este apartado, se lleva a cabo un análisis coste-utilidad en función de la configuración del programa de cribado y del porcentaje de participación en el mismo.
3. Simulación de los principales EAC llevados a cabo hasta la fecha como principal fuente de evidencia sobre el cribado en términos de la reducción de mortalidad que proporcionan y su intervalo de confianza. De esta forma, podremos obtener conclusiones sobre la calidad de los mismos y de su capacidad para ser aplicados a otras poblaciones [Zam16].

### 3.1. RESULTADOS RELACIONADOS CON LA DETECCIÓN. SIMULACIÓN DE PROGRAMAS REALES, ANÁLISIS DE LA SENSIBILIDAD Y DEL SOBREDIAGNÓSTICO

En esta sección se profundiza en el modelo desarrollado, abordando los aspectos relacionados exclusivamente con la detección. Pretendemos comparar las salidas del modelo con los resultados de los programas reales en términos de tasas de detección anuales, proporción de tumores invasivos/*in situ*, proporción de cánceres de intervalo y otros parámetros asociados a la detección. Previamente se realizará un estudio de la consistencia del modelo (análisis de sensibilidad) para encontrar aquellos *inputs* de mayor influencia en las salidas del mismo. Finalmente, se abordará el problema del sobrediagnóstico en términos puramente numéricos, asociado al exceso de detección que supondría el cribado frente a la detección clínica.

De cara a poder verificar la capacidad de nuestro modelo para reproducir los resultados de los programas reales, haremos uso de la información proporcionada por el programa de cribado inglés, NHSBSP (*National Health System Breast Screening Programme*) [UKP12], que lleva realizándose desde el año 1988 de manera ininterrumpida y proporciona resultados consistentes y bien estructurados. Actualmente, el NHS invita a cribado a mujeres entre 50 y 70 años cada 3 años.

Respecto a la estimación del sobrediagnóstico, es difícil obtener información consistente para comparar con nuestros resultados. Una de las principales fuentes de información son los EAC, de entre los cuales, los más adecuados para estimar este fenómeno son Canada I, Canada II [Mil02, Mil00] y Malmö [And88]. Frente a estos, los estudios observacionales presentan numerosos sesgos que dificultan una correcta estimación del sobrediagnóstico, fundamentalmente asociados con la distribución de edad de los grupos a comparar, que puede afectar al resultado. Existen dos perspectivas fundamentales en los EAC a la hora de estimar el sobrediagnóstico:

1. Perspectiva poblacional. Proporción de cánceres diagnosticados durante el periodo de cribado y resto de la vida en mujeres invitadas al mismo que son sobrediagnóstico.
2. Perspectiva individual o de riesgo. Probabilidad de que un cáncer diagnosticado durante el cribado sea un sobrediagnóstico.

Más allá de los resultados particulares de cada EAC, se realizan otras estimaciones del sobrediagnóstico a partir de varios de ellos, como son las de Moss *et al.* [Mos05], empleando 8 EAC y las de Gøtzsche [Got04], que utiliza los resultados de 6 ensayos, junto a revisiones bibliográficas, como las de Biesheuvel *et al.* [Bie07] y Puliti *et al.* [Pul11].

### 3.1.1. Simulación de los resultados de detección en programas reales. Estudio de las diferentes configuraciones

Con los *inputs* descritos en el punto 2.2, calculamos la tasa de detección anual por cribado mamográfico en un programa poblacional con mujeres entre 50 y 70 años y tiempos de intervalo de 1, 2 y 3 años. Posteriormente adelantamos la edad de inicio en el programa a 40 años y, finalmente, situamos la edad de finalización en 65 años para estudiar el controvertido asunto de la edad de comienzo y finalización del cribado desde el punto de vista de las magnitudes asociadas a la detección.

En las tres configuraciones de edades estudiadas se puede observar el elevado resultado de la tasa de detección anual en la primera ronda (ronda de prevalencia) respecto a las siguientes (rondas de incidencia). Cabe destacar la oscilación de la misma para tiempos de intervalo de 2 y 3 años, que es debida a la diferente incidencia del cáncer de mama en los distintos grupos de edad efectivos (ver Figura 3.1). Así, para un tiempo de intervalo de 3 años, al comenzar el programa, entrarán todas las mujeres con 50 o más años, mientras que en la segunda ronda del programa sólo lo harán aquellas que tengan 50 años, pues las de mayor edad entraron en la ronda anterior, de forma similar sucede en la tercera ronda. En el cuarto año de observación volverán a someterse al cribado las mujeres que lo hicieron en el primer año más las nuevas que cumplan 50 años. Las diferencias en las tasas de detección son debidas, por tanto, a la diferente incidencia del primer grupo de mujeres, que cuentan con edades mayores de 50 años. Cuando pasen 20 años (70-50), estas mujeres saldrán del programa, con lo que la distribución de edades se igualará en los tres grupos y, como consecuencia, la tasa de detección se estabilizará. Los resultados globales de las tres configuraciones (50-70 años, 40-70 años y 50-65 años) con tiempos de intervalo 1, 2 y 3 años se muestran en la Tabla 3.1 (incertidumbres entre paréntesis expresadas con una confianza del 99 %,  $k=3$ ).

Para llevar a cabo estas simulaciones se ha utilizado la curva de incidencia canaria [SCS96] junto con una distribución subyacente de tipos histológicos en el *onset* de 37.2 % (ductal invasivo), 2.8 % (lobular invasivo), 51.0 % (CDi) y 9.0 % (lobular *in situ*), que es la distribución original a la que se ajustó el modelo y genera una proporción de tumores invasivos/*in situ* en la detección de cribado de alrededor del 70/30 (tiempo de intervalo de 3 años). Los resultados finales se obtienen del ajuste lineal tras la estabilización del programa (20, 30 y 15 años respectivamente). En aras a verificar la consistencia del modelo, hemos simulado los resultados de la configuración del programa de cribado inglés NHSBSP [Ber06b], consistente en mujeres entre 50 y 70 años y tiempo de intervalo 3 años, cuyos resultados se reportan de forma clara y sistemática. Para ello, introducimos la curva de incidencia inglesa y actuamos sobre la proporción de tumores invasivos/*in situ* subyacente para lograr proporciones de detección en cribado en torno a 80/20. La proporción empleada es 51.15 %, 3.85 %, 38.25 % y 6.75 % para tumores ductales invasivos, lobulillares invasivos, CDi y lobulillar *in situ* respectivamente. Los resultados obtenidos para este programa se muestran en la última fila de la Tabla 3.1.

En el programa NHS se informa de una tasa de detección de 6.6 por  $10^3$ , ligeramente superior a la obtenida en nuestras simulaciones y una proporción de tumores invasivos/*in situ* de 78/22, comparable a la obtenida en las simulaciones, que es de 79/21. Debe tenerse en cuenta que la tasa de detección de 6.6 por  $10^3$  corresponde al año 2000/2001, aumentando hasta 8.2 en el ejercicio 2007/2008. Este aumento puede ser debido a cambios en la

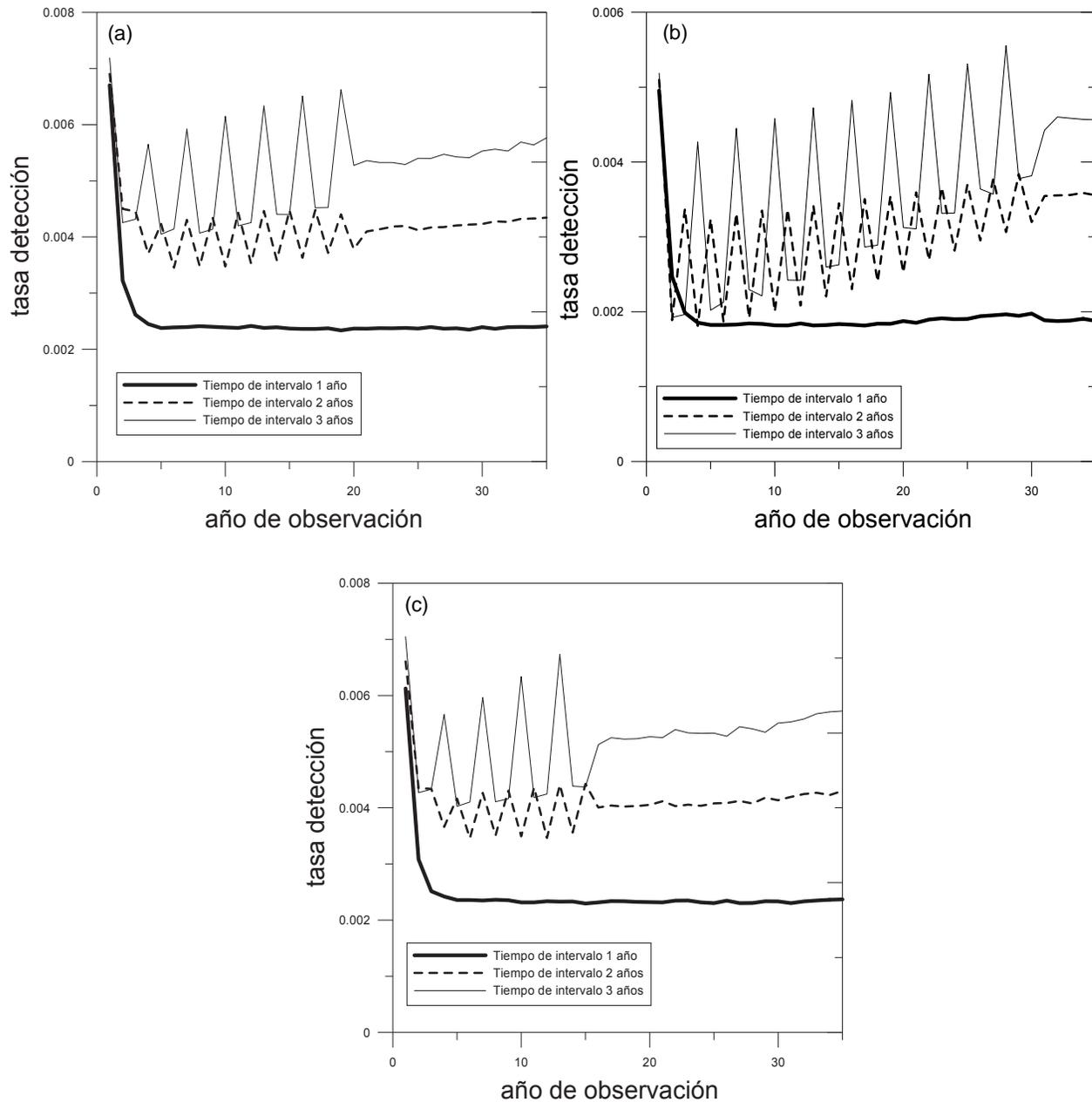


Figura 3.1: Tasa de detección anual del programa de cribado (tumores detectados/mujeres sometidas al programa cada año) para tiempos de intervalo de 1, 2 y 3 años con mujeres entre (a) 50 y 70 años, (b) 40 y 70 años y (c) 50 y 65 años. Se aprecia el efecto de las incidencias asociadas a los diferentes grupos de edad que se someten al cribado anual.

incidencia en los últimos años junto con la introducción de una mejora en la sensibilidad por la introducción de dos vistas en la mamografía [Ber06b].

Respecto a la tasa de cánceres de intervalo por cada  $10^3$  mujeres sometidas al cribado, hemos obtenido resultados para tiempos de intervalo 1, 2 y 3 años (incidencia y distribución subyacente de tumores invasivos/*in situ* inglesa) de 0.209(6), 0.80(4) y 1.80(5), donde las tasas de detección para tiempos de intervalo 2 y 3 se han multiplicado respectivamente por 2 y 3 puesto que estaban dadas en tasa anual. Las estimaciones del NHS son de 0.45 para tiempos de intervalo menores de 1 año, 0.65 para tiempos de intervalo menores de 2 años y entre 1.2 y 1.3 para tiempos de intervalo entre 2 y 3 años. En cualquier caso, estos valores cuentan con una elevada incertidumbre por su carácter estimativo, así como por la dificultad de asignar estos cánceres a una ronda concreta, especialmente para tiempos de intervalo de 1 año.

Finalmente, en lo que se refiere a la proporción de tumores invasivos/*in situ* para estos cánceres de intervalo, se reproduce el valor de 0.96/0.04 del NHS consecuencia de la mayor proporción de cánceres *in situ* en estadios tempranos. Respecto al tamaño medio, obtenemos un valor de 21.3(2) mm frente a los 19.5 mm reportados como mediana en el informe inglés.

### 3.1.2. Análisis de sensibilidad

Una vez comprobada la capacidad del modelo para reproducir un programa real como el NHSBSP, hemos llevado a cabo un análisis de sensibilidad con objeto de encontrar los *inputs* que tienen una mayor relevancia en las salidas del programa, actuando sobre los mismos de manera realista, mediante modificaciones razonables que no supongan llevar al modelo a parametrizaciones absurdas, sin significado real. Los resultados se han presentado en forma de número total de tumores acumulados (en cribado e intervalo), resultado de someter a 5 millones de mujeres de entre 50 y 70 años al programa de cribado con tiempo de intervalo 3 años y resultados por ronda en el año de observación 21, tras la estabilización del programa. Las simulaciones se repiten 10 veces para estimar las incertidumbres ( $k=3$ ) y se comparan con los resultados del programa inglés simulado previamente.

Los *inputs* sobre los que se actúan se describen la Tabla 3.2 y los resultados de estas modificaciones se muestran en las Tablas 3.3 y 3.4 y en la Figura 3.2.

De los resultados mostrados se desprende que el modelo es relativamente insensible a la variación de las entradas descritas, con la excepción del número total de tumores detectados en cribado, tasa de detección por ronda y proporción de tumores invasivos/*in situ* de los cánceres de intervalo, que son muy sensibles a la variación de los dos parámetros que controlan la detección clínica, esto es: tamaño medio de detección clínica y umbral de detección clínica. Fundamentalmente al umbral de detección, que al descender hasta 5 mm produce un déficit en los tumores de detectados en cribado y, por tanto, en la tasa de detección, descendiendo también la proporción de tumores invasivos/*in situ* de intervalo, detectándose más tumores *in situ* al hacerse más probable la detección en fases más tempranas, donde predominan estos tipos. La proporción de tumores invasivos/*in situ* en cribado, sin embargo, prácticamente permanece inalterada.

T.int (años)	cribado					intervalo						
	Tasa det.	Diam. medio inv.	Diam. medio <i>in situ</i>	Prop. inv.	Edad media det.	Tasa cánceres	Diam. medio inv.	Diam. medio <i>in situ</i>	Prop. inv.	Edad media	VPP	
	50-70	1	0.00238(1)	8.49(4)	6.48(2)	0.59(1)	59.6(2)	0.000176(5)	25.3(7)	17.3(7)	0.88(1)	58.7(2)
	2	0.00412(2)	10.10(9)	7.2(1)	0.66(0)	60.1(7)	0.00037(1)	22.2(3)	17.4(9)	0.94(0)	60.0(3)	10.7(1)
	3	0.00533(8)	11.05(3)	7.53(6)	0.70(1)	59.3(2)	0.00058(1)	21.5(2)	17.9(7)	0.96(1)	59.7(2)	12.7(1)
40-70	1	0.00190(6)	7.97(2)	6.36(4)	0.58(1)	58.1(3)	0.000134(5)	20.4(2)	16.2(4)	0.87(1)	56.7(2)	5.49(7)
	2	0.0035(2)	9.7(1)	7.13(5)	0.65(1)	59(1)	0.000319(5)	19.6(2)	18(1)	0.94(1)	58.4(6)	9.1(1)
	3	0.0045(5)	10.9(1)	7.50(6)	0.70(1)	60(1)	0.00057(2)	20.1(2)	17.6(3)	0.96(0)	59.0(7)	11.1(4)
50-65	1	0.00232(2)	8.8(1)	6.52(4)	0.60(1)	56.9(1)	0.000184(3)	27.9(7)	17(1)	0.89(1)	56.4(2)	6.37(4)
	2	0.0041(3)	10.26(6)	7.26(4)	0.67(1)	56.9(4)	0.00037(2)	23.1(4)	17.8(8)	0.94(1)	57.1(1)	9.9(1)
	3	0.0050(7)	11.12(4)	7.53(7)	0.70(1)	57.5(8)	0.00057(2)	21.8(3)	18.4(9)	0.96(1)	58.1(3)	11.9(3)
NHSBP	3	0.00589(9)	10.39(7)	7.48(9)	0.79(1)	59.3(2)	0.00057(2)	21.3(2)	19.1(1)	0.96(1)	59.6(9)	13.9(2)

Tabla 3.1: Resumen de los resultados de las simulaciones para las tres configuraciones de edad estudiadas (50-70, 40-70 y 50-65 años) con tiempos de intervalo 1, 2 y 3 años. Las columnas son: tasa de detección (tumores detectados/mujeres sometidas al cribado), diámetro medio de los tumores invasivos (mm), diámetro medio de los tumores *in situ* (mm), proporción de tumores invasivos, edad media de la detección en cribado, tasa de cánceres de intervalo (cánceres detectados clínicamente/mujeres sometidas al cribado), diámetro medio de los tumores invasivos de intervalo (mm), diámetro medio de los tumores *in situ* de intervalo (mm), proporción de tumores invasivos de intervalo, edad media aparición de los tumores de intervalo y valor predictivo positivo repectivamente. Si nos fijamos en una configuración típica como la de 50-70 y tiempo de intervalo 3 años, comprobamos que los resultados asociados a la detección son muy realistas: tasa de detección en cribado de alrededor del 5 por 1000, diámetro de los tumores invasivos detectados en cribado de 11 mm, proporción tumores invasivos en cribado del 70% y valor predictivo positivo del cribado (VPP) del 12.7%. En la última fila se incluyen los resultados de las simulaciones asociados al programa inglés con intervalos de edades entre 50 y 70 años y tiempo de intervalo de 3 años. Los resultados de tasa de detección (5.9), proporción de tumores invasivos/*in situ* en cribado (79/21) y tamaño medio de tumores invasivos detectados en cribado (21.3 mm), son similares a los del programa real [Ber06b]. Las incertidumbres se indican entre paréntesis con una confianza del 99%.

Tabla 3.2: Descripción de los diferentes *items* sobre los que se ha actuado para verificar la estabilidad del programa (análisis de sensibilidad).

<i>item</i>	Modelo original	Modelo modificado
1.	Modelo de radioinducción	No se incluye
2.	Incidencia cáncer de mama inglesa	Se utiliza la incidencia canaria
3.	Umbral de detección en mamografía 3 mm	Se utiliza 5 mm
4.	Umbral de detección clínica 15 mm	Se utiliza 5 mm
5.	Tamaño medio de detección clínica 20 mm y umbral 15 mm	Se utilizan 15 mm y 5 mm resp.
6.	Tamaño medio de detección clínica 20 mm y umbral 15 mm	Se utilizan 25 mm y 5 mm resp.
7.	Incidencia radioinducidos	Se utiliza 10 veces la original
8.	Sensibilidad mamografía original	Incremento 10 % original
9.	Sensibilidad mamografía original	Incremento 20 % original
10.	Sensibilidad mamografía original	Incremento 30 % original
11.	Prevalencia del 12 %	Se utiliza 13 %
12.	Proporción tumores invasivos/ <i>in situ</i> inicial 55/45	Se utiliza esta proporción constante sin paso <i>in situ</i> → invasivo
13.	Proporción tumores invasivos/ <i>in situ</i> inicial 55/45	Se utiliza una evolución lineal en la proporción de tumores <i>in situ</i> desde el 45 % inicial a el 1 %
14.	Distribución densidad de la mama original	Densidad constante e igual a la de las mujeres entre 50 y 59 años
15.	Distribución densidad de la mama original	Densidad constante e igual a la de las mujeres más jóvenes
16.	Distribución densidad de la mama original	Densidad constante e igual a la de las mujeres de más edad
17.	Distribución de edades original	Todas las mujeres con 50 años al comienzo del programa
18.	Mortalidad por causas diferentes al cáncer de mama original	Se utiliza la mortalidad en el Reino Unido en 2005

Finalmente, la proporción de tumores subyacentes invasivos/*in situ* influye tanto en la proporción detectada en cribado como en los tumores de intervalo. Cuando esta proporción subyacente se mantiene constante sin permitir el paso de tumores *in situ* a invasivos, se obtienen valores absurdos de la proporción de cánceres de intervalo (0.36) y poco realista en cribado (0.65). Basta, sin embargo, con introducir una evolución simple, por ejemplo lineal, entre el 45 % inicial y el 1 % final de la proporción de tumores *in situ*, para que el sistema se comporte de forma realista.

A la vista de los resultados del análisis de sensibilidad, se concluye que es importante establecer un umbral creíble y realista de detección clínica. En este sentido, el valor escogido de 15 mm produce resultados consistentes en la proporción de tumores invasivos/*in situ* de los cánceres de intervalo (alrededor de 95/5). Finalmente, la proporción de tumores invasivos/*in situ* subyacente permite ajustar la proporción detectada en cribado, teniendo poco efecto en la tasa de detección, que, a su vez, puede llegar a alcanzar valores de cerca de 0.0007, aumentando la probabilidad de detección (hasta un 30 %). Esto último puede dar cuenta del aumento de la tasa de detección en el programa inglés en los últimos años, junto con un posible aumento de la incidencia basal.

Tabla 3.3: Primera parte del análisis de sensibilidad del modelo. Se representa el valor de cada una de las salidas indicadas en la primera columna cuando: (1) no se incluye la radioinducción de tumores; (2) se introduce la curva de incidencia canaria; (3) utilizamos un umbral de detección por mamografía de 5 mm; (4) bajamos el umbral de detección clínica a 5 mm; (5) aumentamos el tamaño de detección clínica a 15 mm; (6) aumentamos el tamaño de detección clínica a 25 mm; (7) modelo de radioinducción con una incidencia 10 veces la original; (8) sensibilidad de detección por mamografía 10 veces mayor que la original; (9) sensibilidad de detección por mamografía 20 veces mayor que la original. Tum.= tumores, inv.= tumores invasivos, int.=tumores de intervalo, Prop.=Proporción, tot.=totales, sc.=cribado, det.=detección.

	original	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nº tum. inv. tot. sc.	109193 (309)	109303 (268)	102868 (396)	106594 (433)	78525 (277)	67187 (211)	86856 (229)	109406 (264)	116191 (311)	119898 (269)
Diámetro tum. inv. sc.	10.43 (0.01)	10.4 (0.0)	10.5 (0.0)	11.3 (0.0)	10.5 (0.0)	9.9 (0.0)	10.9 (0.0)	10.4 (0.0)	10.2 (0.0)	10.1 (0.0)
Nº tum. <i>in situ</i> tot. sc.	28608 (149)	28477 (197)	26898 (121)	23817 (148)	21383 (107)	19485 (122)	22703 (129)	28474 (194)	31151 (167)	33623 (154)
Diámetro tum. <i>in situ</i> sc.	7.48 (0.02)	7.49 (0.02)	7.49 (0.02)	8.62 (0.02)	7.03 (0.02)	6.81 (0.02)	7.18 (0.02)	7.49 (0.02)	7.45 (0.02)	7.41 (0.02)
Nº tum. totales sc.	137801 (343)	137780 (333)	129765 (414)	130411 (458)	99908 (297)	86672 (244)	109558 (263)	137880 (328)	147341 (353)	153520 (296)
Prop. tum. inv./ <i>in situ</i> sc.	0.79 (0)	0.79 (0.00)	0.79 (0.0)	0.82 (0.00)	0.79 (0.00)	0.78 (0.00)	0.79 (0.00)	0.79 (0.00)	0.79 (0.00)	0.78 (0.00)
Nº tum. inv. tot. int.	37093 (164)	37067 (193)	35006 (219)	39823 (215)	56598 (258)	63144 (191)	51971 (324)	37370 (192)	30434 (183)	26317 (120)
Diámetro tum. inv. int.(mm)	21.6 (0)	21.63 (0.10)	21.81 (0.08)	21.38 (0.11)	13.23 (0.05)	10.40 (0.03)	16.41 (0.12)	21.65 (0.10)	21.59 (0.05)	21.71 (0.05)
Nº tum. <i>in situ</i> tot. int.	1524 (19)	1535 (42)	1431 (25)	1619 (34)	15206 (159)	18813 (178)	12763 (118)	1534 (42)	1418 (20)	1318 (24)
Diámetro tum. <i>in situ</i> int.(mm)	18.3 (0.3)	18.3 (0.3)	18.3 (0.2)	18.2 (0.2)	6.4 (0.0)	6.1 (0.0)	6.7 (0.0)	18.3 (0.3)	18.2 (0.3)	18.1 (0.3)
Prop. tum. inv./ <i>in situ</i> int.	0.961 (0.001)	0.960 (0.001)	0.961 (0.01)	0.961 (0.001)	0.788 (0.001)	0.770 (0.001)	0.803 (0.001)	0.961 (0.001)	0.955 (0.001)	0.952 (0.001)
Tum. inv. totales	146285 (321)	146370 (346)	137874 (373)	146417 (441)	135124 (234)	130331 (251)	138827 (241)	146776 (353)	146624 (304)	146215 (268)
Tum. <i>in situ</i> totales	30132 (162)	30012 (190)	28329 (116)	25436 (160)	36589 (185)	38298 (209)	35466 (169)	30008 (188)	32569 (169)	34941 (159)
Tum. totales	176417 (359)	176382 (395)	166202 (391)	171853 (469)	171712 (298)	168629 (326)	174292 (295)	176784 (400)	179193 (348)	181156 (312)
Tasa detección	0.0059 (0.0001)	0.0058 (0.0001)	0.0056 (0.001)	0.0055 (0.0001)	0.0043 (0.0001)	0.0037 (0.0001)	0.0047 (0.0001)	0.0058 (0.0001)	0.0063 (0.0001)	0.0065 (0.0001)
Diámetro tum. inv.(mm)	10.4 (0.1)	10.4 (0.1)	10.4 (0.1)	11.3 (0.1)	10.5 (0.1)	9.8 (0.1)	10.8 (0.2)	10.4 (0.1)	10.2 (0.1)	10.0 (0.1)
Diámetro tum. <i>in situ</i> (mm)	7.5 (0.1)	7.5 (0.1)	7.5 (0.1)	8.6 (0.1)	7.0 (0.1)	6.8 (0.1)	7.2 (0.2)	7.5 (0.1)	7.4 (0.1)	7.3 (0.1)
Prop. tum. inv./ <i>in situ</i>	0.79 (0.01)	0.80 (0.01)	0.79 (0.01)	0.82 (0.01)	0.78 (0.01)	0.77 (0.01)	0.79 (0.01)	0.80 (0.01)	0.78 (0.01)	0.78 (0.01)
Edad media det.	59.2 (0.1)	59.1 (0.1)	59.2 (0.2)	59.3 (0.1)	59.2 (0.1)	59.2 (0.1)	59.1 (0.1)	59.1 (0.1)	59.2 (0.1)	59.1 (0.1)

Tabla 3.4: Segunda parte del análisis de sensibilidad del modelo. Se representa el valor de cada una de las salidas indicadas en la primera columna cuando: (10) sensibilidad de detección por mamografía 30 veces mayor que la original; (11) prevalencia del 13%; (12) proporción tumores invasivos/*in situ* subyacente 55/45 y constante; (13) proporción tumores invasivos/*in situ* subyacente 55/45 evolucionando de forma lineal hasta un valor de 1 para los tumores *in situ* en el último tramo de edades; (14) distribución de densidad de mama constante e igual al valor entre 50 y 59 años; (15) distribución de densidad de mama constante e igual al valor correspondiente a las mujeres más jóvenes; (16) distribución de densidad de mama constante e igual al valor correspondiente a las mujeres mayores; (17) distribución de edades constante e igual a 50 años al comienzo del cribado; (18) mortalidad por cáncer del Reino Unido en 2005.

	original	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Nº tum. inv. tot. sc.	109193 (309)	121282 (269)	117682 (483)	82876 (264)	104469 (305)	108932 (304)	102163 (225)	112476 (315)	117434 (223)	105378 (324)
Diámetro tum. inv. sc.	10.43 (0.01)	9.9 (0.0)	10.4 (0.0)	9.4 (0.0)	10.3 (0.0)	10.4 (0.0)	10.5 (0.0)	10.4 (0.0)	10.4 (0.0)	10.4 (0.0)
Nº tum. <i>in situ</i> tot. sc.	28608 (149)	36037 (124)	30896 (174)	45478 (227)	31541 (178)	28447 (158)	25549 (132)	31124 (138)	30934 (166)	27474 (129)
Diámetro tum. <i>in situ</i> sc.	7.48 (0.02)	7.38 (0.02)	7.49 (0.02)	9.87 (0.02)	8.11 (0.02)	7.48 (0.02)	7.51 (0.02)	7.44 (0.02)	7.48 (0.02)	7.48 (0.02)
Nº tum. totales sc.	137801 (343)	157319 (296)	148578 (513)	128354 (349)	136010 (353)	137379 (343)	127713 (261)	143600 (344)	148368 (278)	132852 (349)
Prop. tum. inv./ <i>in situ</i> sc.	0.79 (0)	0.77 (0.00)	0.79 (0.00)	0.65 (0.00)	0.77 (0.00)	0.79 (0.00)	0.80 (0.00)	0.78 (0.00)	0.79 (0.00)	0.79 (0.00)
Nº tum. inv. tot. int.	37093 (164)	24019 (124)	39976 (1799)	16500 (93)	37100 (170)	37250 (165)	43918 (155)	33014 (162)	41182 (161)	35609 (111)
Diámetro tum. inv. int.(mm)	21.6 (0)	21.93 (0.05)	21.70 (0.06)	22.00 (0.15)	21.93 (0.04)	21.64 (0.05)	21.79 (0.06)	21.59 (0.05)	21.21 (0.07)	21.68 (0.05)
Nº tum. <i>in situ</i> tot. int.	1524 (19)	1231 (26)	1646 (38)	29168 (191)	2858 (40)	1529 (18)	1660 (24)	1420 (19)	1689 (49)	1459 (44)
Diámetro tum. <i>in situ</i> int.(mm)	18.3 (0.3)	18.1 (0.4)	18.2 (0.3)	22.4 (0.1)	16.7 (0.2)	18.3 (0.3)	18.4 (0.2)	18.2 (0.4)	18.3 (0.2)	18.2 (0.3)
Prop. tum. inv./ <i>in situ</i> int.	0.961 (0.001)	0.951 (0.001)	0.960 (0.001)	0.361 (0.02)	0.928 (0.001)	0.961 (0.01)	0.964 (0.001)	0.959 (0.001)	0.961 (0.001)	0.961 (0.001)
Tum. inv. totales	146285 (321)	145300 (310)	157657 (426)	99376 (284)	141569 (278)	146182 (326)	146081 (259)	145490 (341)	158616 (341)	140987 (340)
Tum. <i>in situ</i> totales	30132 (162)	37268 (127)	32542 (183)	74646 (343)	34400 (206)	29976 (170)	27209 (146)	32544 (141)	32623 (174)	28934 (160)
Tum. totales	176417 (359)	182568 (335)	190199 (464)	174022 (446)	175968 (346)	176158 (368)	173290 (297)	178034 (369)	191239 (382)	169921 (375)
Tasa detección	0.0059 (0.0001)	0.0067 (0.0001)	0.0064 (0.0001)	0.0055 (0.0001)	0.0058 (0.0001)	0.0059 (0.0001)	0.0055 (0.0001)	0.0062 (0.0001)	0.0059 (0.0001)	0.0059 (0.0001)
Diámetro tum. inv.(mm)	10.4 (0.1)	9.8 (0.1)	10.4 (0.1)	9.4 (0.1)	10.3 (0.1)	10.4 (0.1)	10.5 (0.1)	10.3 (0.1)	10.4 (0.1)	10.4 (0.1)
Diámetro tum. <i>in situ</i> (mm)	7.5 (0.1)	7.3 (0.1)	7.4 (0.1)	9.8 (0.2)	8.1 (0.1)	7.5 (0.1)	7.5 (0.1)	7.4 (0.1)	7.5 (0.2)	7.5 (0.1)
Prop. tum. inv./ <i>in situ</i>	0.79 (0.01)	0.77 (0.1)	0.79 (0.01)	0.64 (0.1)	0.76 (0.01)	0.79 (0.01)	0.80 (0.01)	0.78 (0.01)	0.79 (0.01)	0.79 (0.01)
Edad media det.	59.2 (0.1)	59.1 (0.1)	59.1 (0.1)	59.2 (0.1)	59.2 (0.1)	59.2 (0.1)	59.2 (0.2)	59.1 (0.1)	59.2 (0.1)	59.1 (0.1)

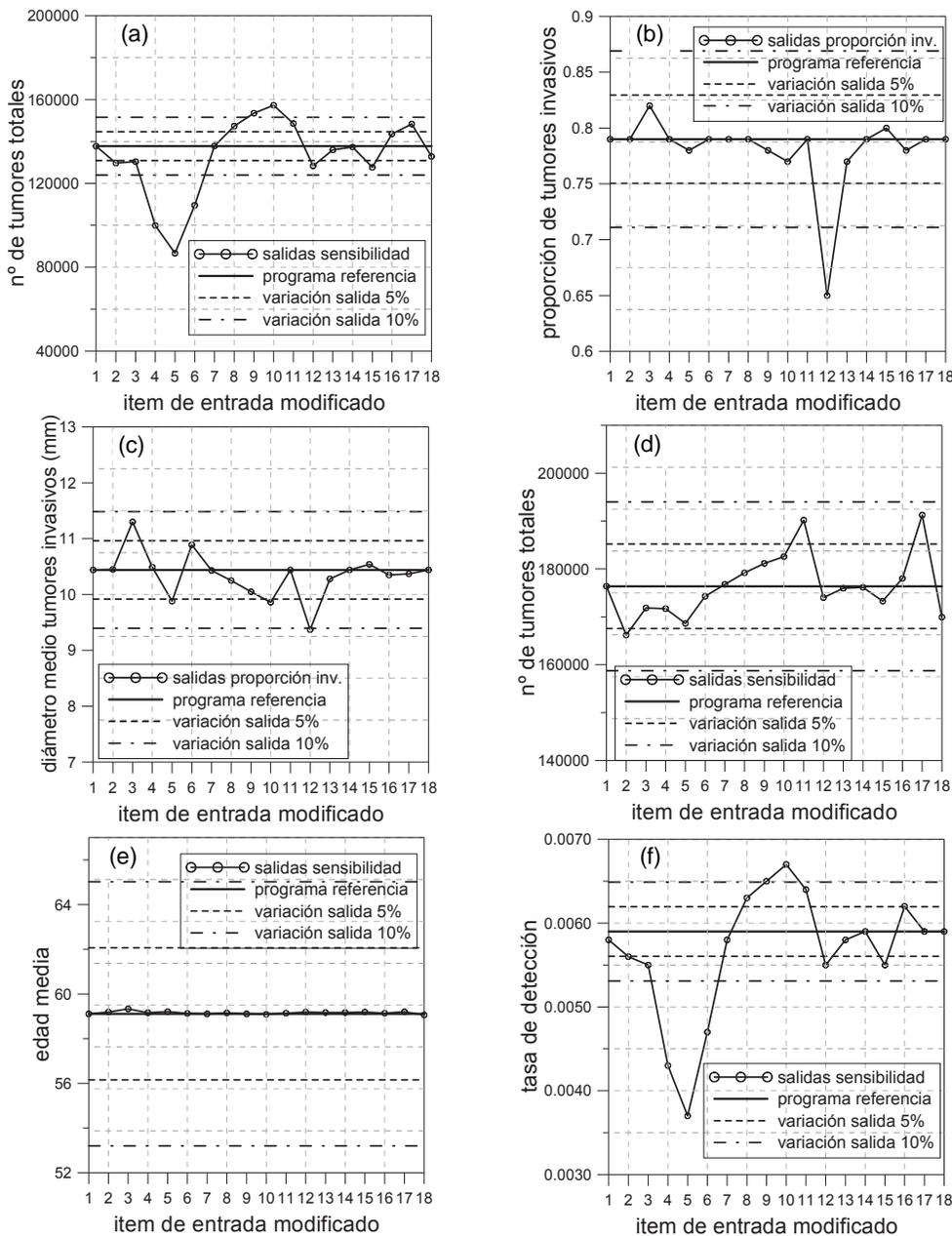


Figura 3.2: Resultados del análisis de sensibilidad. (a) Número total de tumores detectados en cribado en la configuración 50-70 años y tiempo de intervalo de 3 años (5 millones de mujeres sometidas al cribado) en función de los *items* descritos anteriormente. Se aprecia la influencia del tamaño medio de detección clínica cuando se baja a 15 mm. En menor medida, el aumento de la sensibilidad de detección de la mamografía un 30% llega a producir una salida un 5% superior a la original. (b) Proporción de tumores invasivos detectados en cribado. Resulta muy estable salvo cuando se mantiene una proporción subyacente 55/45 sin evolución. (c) Diámetro medio de los invasivos detectados en cribado. Esta salida se mantiene estable  $\pm 5\%$  para cualquiera de los valores de entrada. (d) Número de tumores totales que aparecen el programa de cribado, incluyendo los cánceres de intervalo. (e) Edad media de las mujeres a las que se les detecta un tumor mediante cribado. (f) Tasa de detección, muy afectada por la detección clínica y, en menor medida, por la sensibilidad de la mamografía.

### 3.1.3. Estudio del sobrediagnóstico

Como indicamos en la introducción, el sobrediagnóstico es un aspecto central de la justificación (análisis coste-utilidad) del cribado mamográfico de cáncer de mama. Puesto que se entiende por sobrediagnóstico la detección de aquellos tumores que aparecen en el cribado y que no lo harían clínicamente a lo largo de la vida de la mujer, bien porque son de crecimiento muy lento o bien porque la mujer fallecería por causas diferentes al cáncer de mama antes de su detección [UKP12, Got09], es posible obtener conclusiones en esta materia con la versión de nuestro programa que trata únicamente con aspectos relacionados con la detección.

Para abordar el problema, hemos realizado una modificación en el código que permite introducir un programa de cribado durante un cierto periodo de tiempo en la población de mujeres, cesando luego. Contabilizamos el número total de tumores detectados en cribado en las mujeres sometidas al programa anualmente y acumulados desde el comienzo del mismo, así como aquellos detectados clínicamente, siendo denominados de intervalo si ocurren dentro del programa. Tras comenzar el mismo, se incorporan anualmente las mujeres que alcanzan  $E_{\text{MINIMA}}$  (50 años) hasta que alcanzan  $E_{\text{MAXIMA}}$  (70 años) o terminan el programa. En el segundo caso, se continúa con detección clínica, terminando la historia en el primero de los supuestos. El número de tumores totales contabilizados se compara con los detectados clínicamente en el mismo grupo de edad sin programa de cribado introducido. Estas simulaciones se han llevado a cabo para tiempos de intervalo de 1, 2 y 3 años. De la misma forma se simula el intervalo de edades 40 a 70 años con un programa de cribado que dura 20 años.

El sobrediagnóstico se estima mediante la diferencia entre el número total de tumores (cribado e intervalo), en el caso de que hubiera un programa de cribado, y el número total de tumores detectados clínicamente sin cribado dividido por el número de tumores detectados en cribado e intervalo una vez finalizado el programa [UKP12]. El momento de evaluación de esta magnitud debe ser tal que la diferencia estimada en el numerador se haya estabilizado. Esta definición permite una estimación consistente del sobrediagnóstico, haciéndola independiente de la duración del programa y del punto elegido para el cálculo, siempre que esté suficientemente alejado del final del programa. En efecto, si no hubiera sobrediagnóstico, sino simplemente adelanto de la edad de detección, las curvas de tumores acumulados de la población cuando se somete al cribado y cuando no transcurrirían paralelas, tras un transitorio inicial, pues el número de tumores en ambos casos aumentaría al mismo ritmo. Como se aprecia en la Figura 3.3, estos tramos no son paralelos. Este exceso de tumores se puede escribir, por tanto, de manera lineal (sin ordenada en el origen, pues el exceso de tumores es 0 al inicio del programa):

$$\text{exceso} = N_{\text{mujeres}} \times p_{\text{sobr.}} \times t_{\text{prog.}} , \quad (3.1)$$

donde  $N_{\text{mujeres}}$  es el número total de mujeres sometidas al cribado,  $p_{\text{sobr.}}$  es la probabilidad de sobrediagnóstico y  $t_{\text{prog.}}$  la duración del programa. En otras palabras, para una población sometida al cribado, el exceso de tumores que produce el cribado es proporcional a la duración del programa. Puesto que, como se aprecia en la Figura 3.3, la curva acumulada de tumores de la población sometida al cribado es también lineal mientras dura el mismo, el número de tumores al final del mismo se podrá escribir como:

$$N_{\text{tumores crib.}} = m_{\text{crib.}} \times t_{\text{prog.}} , \quad (3.2)$$

$N_{\text{tumores crib.}}$  número total de tumores acumulados en el programa de cribado para una duración del programa de cribado  $t_{\text{prog.}}$  y  $m_{\text{crib.}}$  es la constante de proporcionalidad.

Por lo tanto, el cociente entre estas dos magnitudes es constante e independiente de la duración del programa y del punto de evaluación del exceso de tumores. Una vez cesado el programa y transcurrido el *lead time*, las curvas transcurren paralelas asintóticamente:

$$\frac{\text{exceso}}{N_{\text{tumores crib.}}} = \frac{N_{\text{mujeres}} \times p_{\text{sobr.}}}{m_{\text{crib.}}}, \quad (3.3)$$

Tras el cese del programa de cribado aparece un déficit en la detección clínica de tumores durante un periodo similar al *lead time*, consecuencia del adelanto en la detección que supone el cribado. Esto introduce cierta asimetría en la simulación de las mujeres fuera del cribado, pues deberían detectarse durante este periodo transitorio menos tumores clínicamente en las mujeres procedentes del programa que en aquellas que nunca se sometieron al mismo. Nuestro modelo no es capaz de simular más allá de los 70 años de edad (ver Figura 2.4 del Capítulo 2), por lo que habrá una pequeña sobreestimación del sobrediagnóstico asociada al desequilibrio introducido por no contabilizar estas contribuciones que, como hemos señalado, no son iguales.

En la Figura 3.3 se muestran los gráficos correspondientes al número de tumores detectados acumulados y por año de observación desde el comienzo del programa de cribado en las configuraciones 50-70 años, y duración del programa de 10 años y 40-70 años y duración 20 años. No se han incluido las barras de incertidumbre para no dificultar la visualización de los resultados. Estas incertidumbres están alrededor del 0.3 % en el caso acumulado y el 2 % en la anual. Se aprecia la magnitud del sobrediagnóstico en tiempos de intervalo de 1, 2 y 3 años, pues la curva que representa el número de tumores acumulados en mujeres sometidas al cribado permanece por encima de la correspondiente a la de ausencia de cribado asintóticamente tras finalizar el mismo. Para estos mismos grupos, se observa en la Figura 3.4 que la tasa de detección de tumores anual se iguala a la correspondiente a la de ausencia del programa de cribado, tras 10 años desde la finalización del programa.

Los valores del sobrediagnóstico, tal como lo definimos al inicio de esta sección, obtenidos de la zona asintótica (a partir del año de observación 20) van desde el 20 % al 11.5 % para la configuración 50-70 años, y tiempos de intervalos de 1 a 3 años, y desde el 16.8 % al 10 % en el caso 40-70 años. Se ha incluido una estimación de los resultados para un porcentaje de participación más realista del 80 %. En esta estimación se ha supuesto que el número de tumores detectado tras el cese del programa permanece aproximadamente igual que en la participación completa (ver Tabla 3.5). En cualquier caso, estos resultados de entre el 11 % y el 20 %, que traducidos a programas reales (tiempos de intervalo de 2 o 3 años y participaciones en el entorno del 80 %) rondarían entre el 7 % y el 10 %, son consistentes con los valores indicados en la bibliografía, pese a la enorme variabilidad e incertidumbre de estas estimaciones.

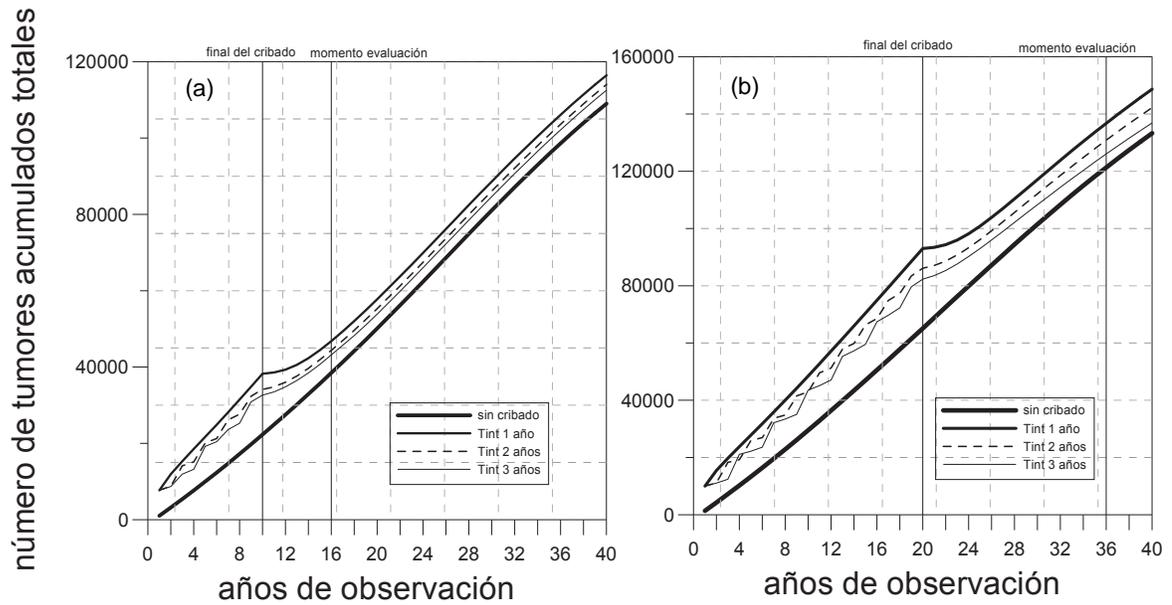


Figura 3.3: Número total de tumores acumulados (cribado más intervalo) en un programa de cribado de mujeres para tiempos de intervalo de 1, 2 y 3 años para (a) mujeres entre 50 y 70 años y (b) mujeres entre 40 y 70 años. La línea negra representa los tumores acumulados sin programa de cribado. Se aprecia el cese del programa a los 10 y 20 años respectivamente, seguido por un transitorio de más de 2 años tras el que las curvas siguen paralelas asintóticamente a la de ausencia de cribado.

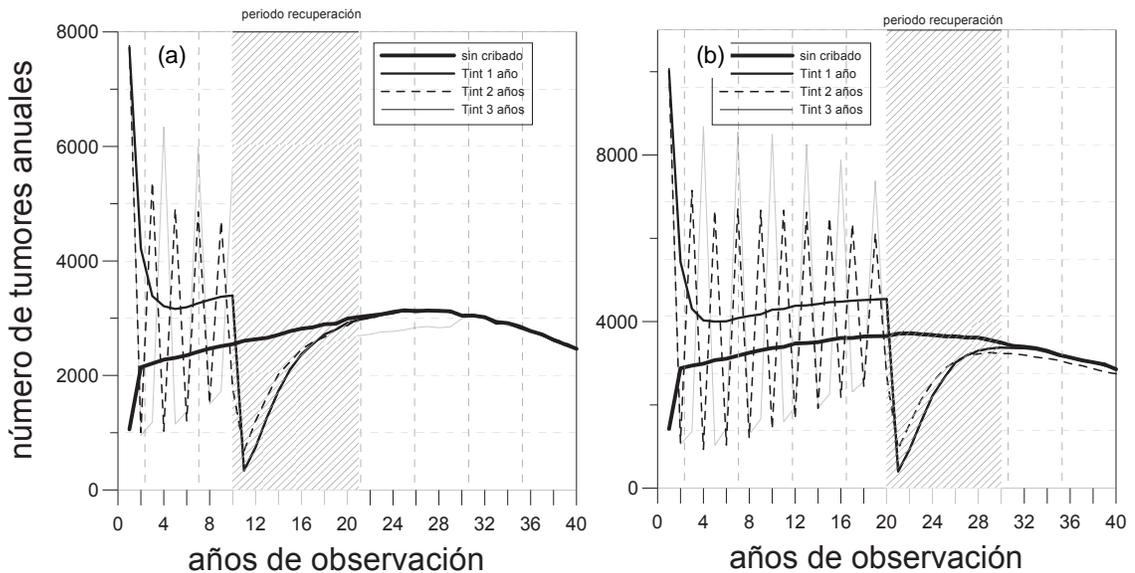


Figura 3.4: Tasas anuales (cribado+intervalo) para tiempos de intervalo 1, 2 y 3 años del programa (a) de 50-70 años y (b) de 40-70 años frente a la ausencia de cribado. Se aprecia más claramente el periodo transitorio tras el cese del programa.

El informe Cochrane [Got09] establece un riesgo absoluto de sobrediagnóstico/sobretatamiento de 0.5 % basado en los EAC; esto significaría que 10 de cada  $2 \times 10^3$  mujeres invitadas a cribado recibirían sobretatamiento. Los resultados del ensayo de Malmö [And88] para mujeres entre 55 y 69 años y los de Canada I y II [Mil02, Mil00] arrojan valores del 11 % desde el punto de vista poblacional (porcentaje de todos los tumores detectados que son sobrediagnóstico) y un riesgo del 19 %.

Existen algunos estudios observacionales basados en los resultados proporcionados por programas implantados durante años. Es el caso del programa británico NHSBSP [UKP12], que proporciona datos fiables, aunque limitados a su población y configuración. Según este, el cribado mamográfico supone una muerte evitada por cada 250 mujeres invitadas al programa durante 20 años (mujeres mayores de 50 años). Así, si partimos de una tasa de aceptación/participación usual en el entorno del 80 % y una tasa de detección en las rondas de incidencia (posteriores a la primera ronda o ronda de prevalencia) de 4 cánceres detectados por cada  $10^3$  mujeres y año, obtendríamos 640 tumores detectados por cada  $10^4$  mujeres durante 20 años, es decir, 40 muertes evitadas por cada  $10^4$  mujeres durante 20 años. A su vez, este mismo informe, basándose en los 3 EAC que no realizan mamografía al grupo de control al final del seguimiento, sitúa el sobrediagnóstico en un 19 %, entendido como proporción de los tumores detectados en cribado que son sobrediagnosticados, por tanto, se tendrían 122 sobrediagnósticos por cada  $10^4$  mujeres durante 20 años. En otras palabras, 3 sobrediagnósticos/sobretatamientos por cada muerte evitada o un 1 % de las mujeres invitadas a cribado.

Recientemente, Bleyer y Welch [Ble12] han publicado un estudio poblacional basado en el análisis de los resultados de detección de los programas de cribado llevados a cabo en EE.UU. durante los últimos 30 años en mujeres mayores de 40 años usando datos SEER del periodo 1976 al 2008. En este estudio se encuentra que la introducción del cribado mamográfico ha supuesto un aumento del 100 % en los casos detectados en estadios tempranos (CDiS y tumores invasivos locales), pasando de 112 a 234 por  $10^5$  mujeres, mientras que la reducción de los casos en fases tardías es sólo del 8 % (de 102 a 94 casos por cada  $10^5$  mujeres). En una primera aproximación, sólo 8 casos de los 112 tumores detectados en estadios tempranos se esperaría que progresaran a estadios más avanzados. Para la estimación del sobrediagnóstico se evalúan varios aspectos que pueden afectar a los resultados, como son los cambios en la incidencia basal obtenidos de la población menor de 40 años (no sometida al cribado) y el efecto de la terapia de reemplazo hormonal, que se supone descontado a partir del año 2006. Con estas estimaciones, se presenta un sobrediagnóstico de entre el 22 % y el 31 % de todos los casos detectados. Este estudio contradice los resultados ofrecidos por los EAC, que consideran muy sesgados para este tipo de estimaciones.

Cabe señalar en este punto que la detección del carcinoma ductal *in situ* (CDiS) no es equivalente a sobrediagnóstico. Aproximadamente un 10 % de las mujeres operadas diagnosticadas de CDiS desarrollan un tumor invasivo en la misma zona [Kop98]. Existe, por tanto, una evolución del tipo CDiS hacia tipos invasivos, al menos en un porcentaje no despreciable de casos.

Los resultados que obtenemos en nuestras simulaciones referentes únicamente a los tumores invasivos muestran que, para la configuración de 50 a 70 años y tiempos de intervalo 1 y 2 años, el número de tumores invasivos es igual en la población sometida al cribado y en la sometida a detección clínica; sin embargo, para tiempos de intervalo de

3 años, aparece un déficit de tumores invasivos en el grupo de cribado del 13% (referido nuevamente al número total de tumores detectado al final del cribado, incluidos los tumores *in situ*). Hay, por tanto, un 13% de tumores *in situ* que son detectados en cribado y que de otra forma evolucionarían a invasivos, interrumpiendo el cribado la historia natural de estos tumores. En esta configuración se tendrá un exceso de tumores *in situ* en el grupo de cribado del 24% pero, por contra, habrá un menor número de tumores invasivos. Para la configuración de 40 a 70 años se tiene, en los tres tiempos de intervalo, un déficit de tumores invasivos de alrededor del 7% en la población de mujeres sometida al cribado. Los resultados globales se muestran en la Tabla 3.5.

Tabla 3.5: Valores de sobrediagnóstico tal y como se define en el texto para las dos configuraciones estudiadas y los tiempos de intervalo considerados en cada caso. Los valores de participación del 100% se obtienen de la simulación. Los resultados para 80% de participación son una estimación. Las incertidumbres se dan con un factor de cobertura de  $k=3$ .

$t_{int}$	sobrediagnóstico			
	[50,70] años; (t=10 años)		[40,70] años; (t=20 años)	
	100 % part.	80 % part.	100 % part.	80 % part.
<b>1 año</b>	0.200±0.009	0.14	0.168±0.005	0.13
<b>2 años</b>	0.150±0.010	0.10	0.123±0.005	0.07
<b>3 años</b>	0.115±0.007	0.06	0.100±0.005	0.07

Este análisis de los tumores invasivos muestra que el sobrediagnóstico se minimiza para tiempos de intervalo de 3 años, mostrando un déficit de tumores invasivos en el grupo sometido al cribado en el rango de 50-70 años y tiempo de intervalo 3 años, siendo uniforme, aunque de menor magnitud, en el rango 40 a 70 años. Este déficit de tumores invasivos es de especial importancia, pues indica que el cribado junto con los tratamientos asociados, más allá de adelantar el diagnóstico es capaz, en algunas configuraciones, de interrumpir la evolución de los tumores *in situ* a invasivos. En la Figura 3.5 se ha representado este efecto para tiempos de intervalo de 1 año.

### 3.2. RESULTADOS RELATIVOS A LA REDUCCIÓN DE MORTALIDAD EN LOS PROGRAMAS DE CRIBADO

El informe inglés NHS [UKP12] realiza una estimación de la reducción del riesgo absoluto (RRA) de mortalidad por cáncer de mama que supone el programa de cribado, entendido como la diferencia en los riesgos de mortalidad con y sin cribado. Para esta estimación es necesario utilizar datos de incidencia, descontando el efecto del cribado que lleva años realizándose en esta población. Por tanto, si la mortalidad en mujeres entre 55 y 79 años es del 1.7% y suponemos una reducción de mortalidad debida al cribado del 20%, tendremos una reducción del riesgo absoluto de:

$$\text{RRA} = 1.70/0.80 - 1.70 = 0.43\%, \quad (3.4)$$

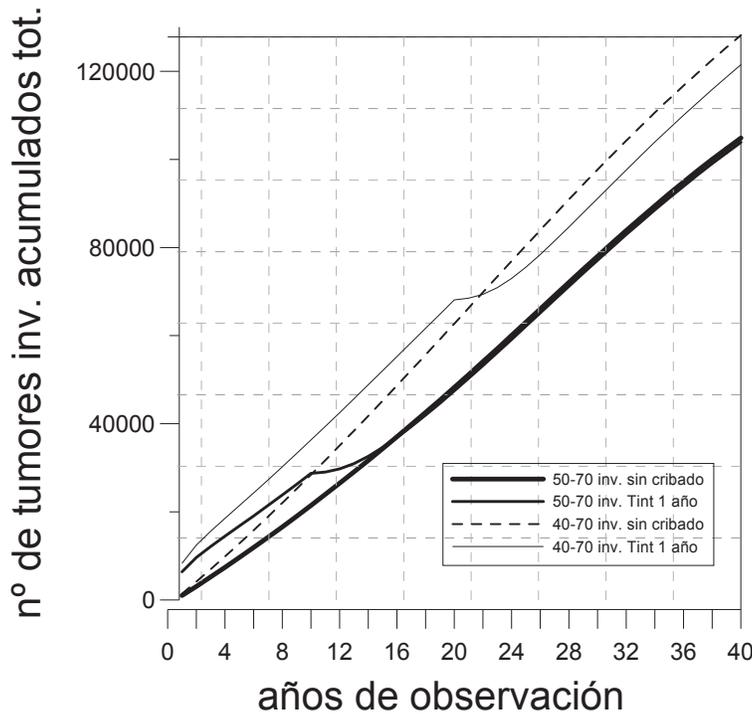


Figura 3.5: Tumores acumulados invasivos (cribado+intervalo) para el grupo 50-70 años y 40-70 con tiempos de intervalo 1 año. En el primer caso se aprecia como las curvas del cribado y la de ausencia del mismo se igualan asintóticamente tras el transitorio, de forma que no hay exceso de tumores invasivos detectados asociados al cribado. Todo el exceso sería atribuible a los tumores *in situ*. En la segunda configuración se aprecia el déficit de invasivos en el grupo sometido al cribado.

y, por consiguiente, el número de mujeres invitadas al programa durante 20 años (mayores de 50) necesarias para prevenir una muerte *Number Needed to Treat* (NNT), será:

$$\text{NNT} = 1/0.43 = 233, \quad (3.5)$$

La reducción de riesgo absoluto lleva implícita, a diferencia del riesgo relativo, una mayor dependencia con la población a la que se aplica puesto que, como hemos visto, es necesario el dato de la incidencia en la población sin cribado, dando esto lugar a una gran variabilidad entre las diferentes estimaciones. La revisión Cochrane 2009 [Got09] realiza la misma estimación pero basándose en 3 EAC (Canada I y II, Malmö y *UK Age trial*). Obtiene del grupo de control del último de los EAC un riesgo de mortalidad del 0.33 %, con una reducción de mortalidad relativa proporcionada por estos ensayos del 15 %. De estos resultados estima:

$$\text{RRA} = 0.33 - 0.33/0.85 = 0.05, \quad (3.6)$$

$$\text{NNT} = 1/0.05 = 2000, \quad (3.7)$$

Resulta, por tanto, una estimación hasta un orden de magnitud mayor. La clave está en el riesgo basal o incidencia sin cribado. Posiblemente el valor de 0.33 % sea excesivamente bajo, consecuencia de un grupo de mujeres más jóvenes (menor incidencia) y un seguimiento insuficiente (13 años). Otras estimaciones muestran un amplio rango de resultados [Ind12]:

- Duffy *et al.* [Duf10] basándose en un seguimiento del EAC *Swedish Two-county* de 22 años obtiene un valor de NNT de 113.
- *US Task Force* [Nel09] presenta un valor de 1339 para mujeres entre 50-59 años y 377 para el grupo 60-69 años.
- *Canadian Task Force* [CTF11] informa un NNT de 720 en el grupo 50-69 años.
- Beral [Ber11] propone un valor de 400 mujeres durante 10 años en el grupo etario 50-70.

Finalmente, los resultados en términos de riesgo relativo de los diferentes EAC son compilados en diferentes meta-análisis. De entre todos estos destacan [Ind12] (los intervalos de confianza están dados al 95 %):

- Cochrane [Got09] (13 años de seguimiento) que reporta un resultado de reducción de mortalidad relativa de 0.81 (0.74-0.87). Si se excluye el grupo de mujeres menores de 50 años el resultado es 0.77 (0.69-0.86). Finalmente, considerando sólo los EAC de mayor evidencia (Canada I y II, Malmö y *UK Age trial*) resultaría un valor de 0.90 (0.79-1.02).
- *US Task Force* [Nel09] aporta un resultado de 0.86 (0.75-0.99) en el grupo de edad 50-59 años y 0.68 (0.54-0.87) para el de 60-69 años.
- *Canadian Task Force* [CTF11] incluye mujeres entre 50 y 69 años con un valor de 0.79 (0.68-0.90).
- Duffy [Duf12] proporciona un valor de 0.79 (0.73-0.86).

### 3.2.1. Validez de la sensibilidad global y curva de adelanto

Los resultados de la sensibilidad global para las diferentes configuraciones de los programas de cribado considerados en esta segunda parte se resumen en la Tabla 3.6 (ver Capítulo 2).

No es sencillo verificar estos valores con estimaciones reales, pues los programas de cribado no suelen proporcionar parámetros comparables, sin embargo, algunos de los ensayos controlados más importantes sí ofrecen datos que permiten valorar la consistencia de los resultados generados por nuestro programa. Así, la recopilación hecha por Humphrey *et al.* [Hum02] muestra los resultados de las revisiones sistemáticas de los principales ensayos aleatorios, calculando la sensibilidad global estimada como el número de cánceres detectados en el cribado dividido por el número total de cánceres que aparecen en el periodo de seguimiento, lo que permite llevar a cabo una comparación con nuestros resultados. Esta comparación se detalla en la Tabla 3.7.

De la comparación de nuestros resultados con las revisiones se puede concluir:

Tabla 3.6: Resultados de la estimación de la sensibilidad global para diferentes configuraciones de los programas de cribado. Se aprecia claramente la menor sensibilidad global de la configuración de edades 40-70 años, que incluye mujeres más jóvenes y, por tanto, con mamas más densas y peor detectabilidad. Se observa también una disminución con el tiempo de intervalo.

Rango de edades	Tiempo de intervalo	Sensibilidad global(%)
50-70 años	1 año	87
	2 años	79
	3 años	72
40-70 años	1 año	84
	2 años	75
	3 años	66
50-65 años	1 año	87
	2 años	79
	3 años	72

1. El valor estimado para el ensayo New York HIP es significativamente más bajo que para el resto de los ensayos considerados (39 %). Nuestra estimación con diferentes participaciones por ronda (aunque relativas a un 100 % en la primera) es también baja (48.6 %). Este valor anormalmente bajo podría atribuirse a la baja calidad de la mamografía llevada a cabo en este ensayo, lo que disminuiría su validez externa [Hum02].
2. El resultado ofrecido por el ensayo de Malmö (61 %) es más elevado que el simulado (54.3 %), pero debe señalarse que este último se ha calculado aplicando un factor global de 0.7 para reflejar la participación media en las diferentes rondas. Además, el número de rondas de cribado simulado es de 8 frente a las sólo 2 rondas que incluyen los ensayos. El mayor peso de la ronda de prevalencia respecto a las de incidencia en el valor reportado por los ensayos puede ser responsable de esta discrepancia, al contemplar más rondas de incidencia que incluyen tumores más pequeños y, por tanto, más difíciles de detectar [Mus98].
3. Los valores para *Swedish Two-County* y Estocolmo son estimaciones directas, no proceden de las revisiones sistemáticas, por lo que presentan una mayor incertidumbre que impide sacar conclusiones más precisas.
4. Los valores estimados para Canada I y II son comparables a las simulaciones, pudiendo ser atribuida la diferencia a una mala calidad de la mamografía [Kop93].

En cualquier caso, nuestra sensibilidad global está definida con respecto a los tumores que son “detectables” por mamografía, entendidos como aquellos que alcanzan los 2 mm en algún momento durante el cribado, en el caso de los ensayos entre  $E_{\text{MIN}}$  y  $E_{\text{MIN}} + n^{\circ}$  rondas  $\times T_{\text{int}}$ .

Finalmente, realizamos un último test de nuestras distribuciones de adelanto de detección que tiene que ver con la relación entre el *lead time* y el sobrediagnóstico. En efecto, se entiende que valores del tiempo de adelanto muy elevados están asociados indefectiblemente a sobrediagnóstico desde el momento que se adelanta tanto la fecha del diagnóstico por

Tabla 3.7: Valores de sensibilidad global estimada a partir de revisiones sistemáticas de publicaciones con diferentes seguimientos de los ensayos de mayor validez externa. Los valores presentados para *Swedish Two-County* y Estocolmo (2) proceden directamente de revisiones recientes de buena calidad, sin aplicar ningún modelo de ajuste. Los valores de sensibilidad global simulada entre paréntesis se han calculado introduciendo las diferentes participaciones por ronda relativas a un 100 % de participación en la primera, salvo el caso del Malmö (1) donde se ha aplicado un factor 0.7 para reflejar la participación global promedio.

Estudio	Sensibilidad global estimada (%)	Sensibilidad global simulada (p.p. ronda)(%)
New York HIP 40-64 años (4 rondas)	39	78.2 (48.6)
Malmö 45-69 años (2 rondas)	61	77.5 (54.3) <sup>(1)</sup>
<i>Swedish Two-County</i> 40-74 años	77 <sup>(2)</sup>	72.4 Kopparberg; 72.9 Ostergötland (64.9)
Estocolmo 40-64 años	68 <sup>(2)</sup>	73.5
Canada I 40-49 años (4 rondas)	61	79.7 (68.4)
Canada II 50-59 años (4 rondas)	66	80.9 (70.5)

criado a la detección clínica que, probablemente, la mujer moriría por causas diferentes al cáncer de mama antes del hipotético diagnóstico clínico. Utilizando la estimación del sobrediagnóstico dada por [Ble12] del 31 % para mujeres mayores de 40 años, calculamos en nuestra distribución de adelanto el valor medio del mismo para la cola, a partir de este porcentaje (ver Figura 3.6). Encontramos para el caso de 40-70 y tiempo de intervalo 1 años un valor de 9 años, valor que coincide con tiempo medio de adelanto para los casos de sobrediagnóstico calculado por [Etz14]. A la luz de estas evidencias indirectas, nuestras curvas de adelanto son razonables desde el punto de vista teórico.

### 3.2.2. Estudio de las diferentes configuraciones y de la participación en los programas de cribado en términos de reducción de mortalidad

En este apartado se ha pretendido dar un paso más en la simulación, introduciendo en la misma la supervivencia tras detección y tratamiento loco-regional. De esta forma, es posible estudiar la mortalidad con y sin un programa de cribado instaurado. Con esta modificación hemos simulado 3 configuraciones con tiempos de intervalo de 1, 2 y 3 años para estudiar:

1. Valor de referencia de reducción de mortalidad para una configuración típica (mujeres de entre 50 y 70 años).
2. Efecto del adelanto en la edad de inicio (mujeres de entre 40 y 70 años).
3. Efecto de adelanto en la edad de finalización (mujeres de entre 50 y 65 años).
4. Efecto de los diferentes tiempos de intervalo (1, 2 y 3 años).
5. Efecto del porcentaje de participación.

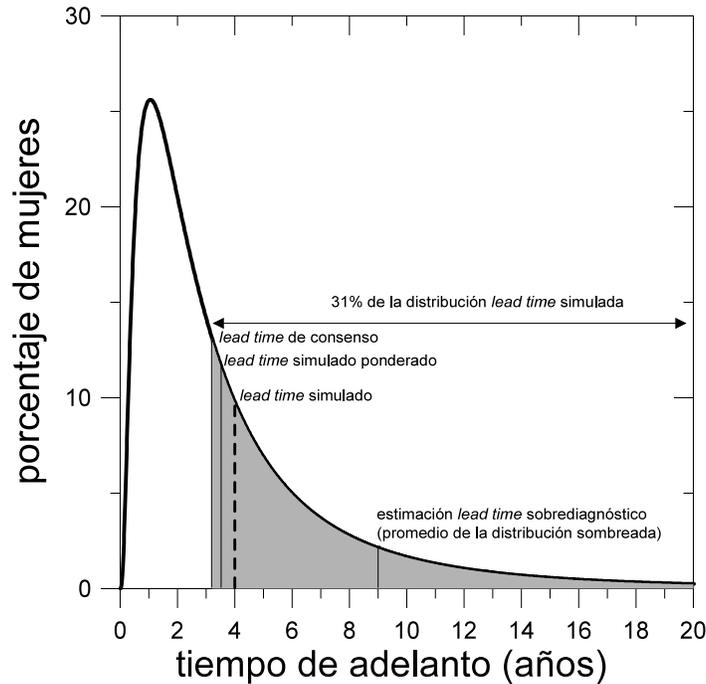


Figura 3.6: Distribución de adelanto de detección simulada para el caso de rango de edades 40-70 años y tiempo de intervalo 1 año. El área sombreada representa el área de la curva con el 31 % de los casos de mayor *lead time*. El promedio de esta área supone un valor de 9 años.

Se define la reducción de mortalidad o riesgo relativo de cáncer de mama (RR) como:

$$RR = \frac{M_0 - M_C}{M_0}, \quad (3.8)$$

El valor obtenido para la mortalidad por cáncer de mama sin programa de cribado es de  $21.0 \pm 0.2$  muertes por  $10^5$  mujeres y año para la población general. Si bien la mortalidad por cáncer de mama es fuertemente dependiente de la población de mujeres concreta que se estudia, podríamos esperar que la reducción de mortalidad supuestamente producida por un programa de cribado tenga un carácter más general e independiente de los detalles poblacionales. Aún así, el valor encontrado por el *Australian Institute of Health and Welfare* (AIHW) es de 22.1 muertes por  $10^5$  mujeres y año [Aus09] (intervalo de confianza 95 % 21.3-23.0), que está claramente en consonancia con el valor obtenido en la simulación. Recientemente Autier *et al.* [Aut11] han obtenido valores similares para las poblaciones de Suecia y Noruega, alcanzándose valores de hasta 30 para otros vecinos europeos.

Si el resultado se expresa en términos de mortalidad estandarizada por edad (ASR) de forma que, en lugar de referirse a mujeres vivas en el año de cálculo, se estima respecto a las mujeres con edades por encima de los 25 años y, teniendo en cuenta que en España este valor es aproximadamente del 75 % de los casos, se obtendría una mortalidad de alrededor de 28 mujeres por  $10^5$  y año. Estos valores están influidos por el cribado en los estudios poblacionales por estar ampliamente implantados los programas de cribado mamográfico en los principales países industrializados.

La simulación explícita de los programas se ha realizado para las configuraciones críticas de 50-70 años, 40-70 años y 50-65 años con tiempos de intervalo realistas de 1, 2 y 3 años. Los grupos de edades elegidos responden a las configuraciones típicas, adelanto de la edad de comienzo a los 40 años y adelanto de la de finalización a los 65. Por su parte, los tiempos de intervalo deben responder al criterio de ser del orden del *lead time* para producir un aumento significativo en la edad de detección, pero no tan cortos que sean inefectivos. Los resultados del RR por cáncer de mama en las configuraciones estudiadas se muestran en la Tabla 3.8.

Tabla 3.8: Resultados de la reducción de mortalidad por cáncer de mama para las tres configuraciones estudiadas con tiempos de intervalo de 1, 2 y 3 años. Se han calculado para diferentes porcentajes de participación en el cribado.

	Tiempo intervalo (años)	Porcentaje Participación 10 %	Porcentaje Participación 30 %	Porcentaje Participación 50 %	Porcentaje Participación 70 %	Porcentaje Participación 100 %
<b>50-70</b>	<b>1</b>	4	10	18	23	33
	<b>2</b>	3	9	14	20	29
	<b>3</b>	3	8	13	18	26
<b>40-70</b>	<b>1</b>	4	12	20	28	40
	<b>2</b>	3	10	18	24	35
	<b>3</b>	3	9	15	21	30
<b>50-65</b>	<b>1</b>	3	9	14	20	29
	<b>2</b>	3	8	13	18	25
	<b>3</b>	2	7	11	16	23

Se comprueba que tiempos de intervalo más cortos (1 años) suponen una mayor reducción de la mortalidad, sin embargo, estos tiempos de intervalo entre mamografías suponen un elevado esfuerzo a nivel logístico, junto a costes emocionales intangibles para las mujeres. Para estudiar cuál es la configuración que permite alargar más los tiempos de intervalo con un menor deterioro del RR, referimos los mismos para las diferentes configuraciones al valor correspondiente del porcentaje de participación más realista del 70 % y tiempo de intervalo 1 año. Se comprueba que la configuración 50-65 es la que mantiene una proporción más alta para tiempos de intervalo 3 años respecto a 1 año (Figura 3.7). Sería, por tanto, este rango de edades el menos sensible al aumento del tiempo de intervalo, perdiéndose tan sólo 4 puntos de reducción de mortalidad. Este resultado está asociado a la eliminación del cribado de las mujeres entre 65 y 70 años, para las que hay una elevada incidencia del cáncer de mama, lo que hace algo más insensible la configuración al tiempo de intervalo.

Para estudiar el efecto del rango de edades, representamos ahora el RR relativo a la configuración de 40-70 años para el mismo tiempo de intervalo, encontrando un resultado mejor de la configuración 50-70 años, especialmente para tiempos de intervalo 3 años, en los que se llega hasta un 86 % del valor correspondiente en la configuración 40-70 años. Una vez más podemos comprobar el efecto importante que tiene eliminar a las mujeres de entre 65 y 70 años del programa a efectos de RR de mortalidad por cáncer de mama.

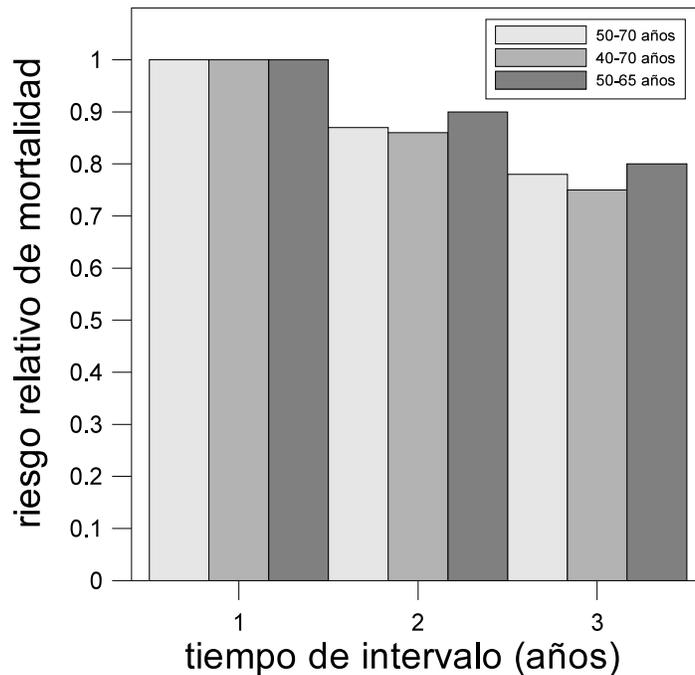


Figura 3.7: Efecto del tiempo de intervalo en la RR por cáncer de mama. Se aprecia una menor reducción del mismo cuando aumenta el tiempo de intervalo para la configuración de edades de 50 a 65 años.

En términos generales, podemos decir que la inclusión de mujeres desde 40 a 49 años supone una ganancia en RR de entre el 3% y el 5%, mientras que el efecto de la inclusión de mujeres entre 65 y 70 años supone entre el 2% y 3%. La importancia de este tramo de solo 5 años frente a los 10 años entre 40 y 50 es debida a la diferente incidencia del cáncer de mama en estos grupos de edad, siendo muy elevada en el grupo 65 a 70 años, hasta un 35% más para la incidencia correspondiente a la población inglesa, a pesar de incluir la mitad de años [Ber06b].

### 3.2.3. Análisis Coste-Utilidad

En este punto realizamos un breve y simplificado estudio coste-utilidad de los programas de cribado en función de la reducción de mortalidad que producen en la población teniendo en cuenta el detrimento que supone el sobrediagnóstico/sobretretamiento asociado a estos programas en términos de reducción de años de vida “efectivos”. Para ello, utilizamos el concepto de “años de vida ajustados por calidad” (*QALY*) [Tor86], con un factor de detrimento de 0.83 para el caso de cáncer de mama propuesto por Plevritis *et al.* [Ple06]. Este factor aplica a los años vividos tras la detección y tratamiento del cáncer de mama, reduciendo la “calidad” de los mismos respecto a los vividos por mujeres sanas o asintomáticas.

Supondremos para los cálculos una esperanza de vida de 86 años, similar a la de las mujeres españolas en 2014 [INE14] y una edad media de mortalidad por cáncer de mama de 66 años [AEC]. La edad media de detección por cribado la fijaremos en el entorno de los 60 años (ver resultados simulaciones) y aceptamos que hay 3 sobrediagnósticos por

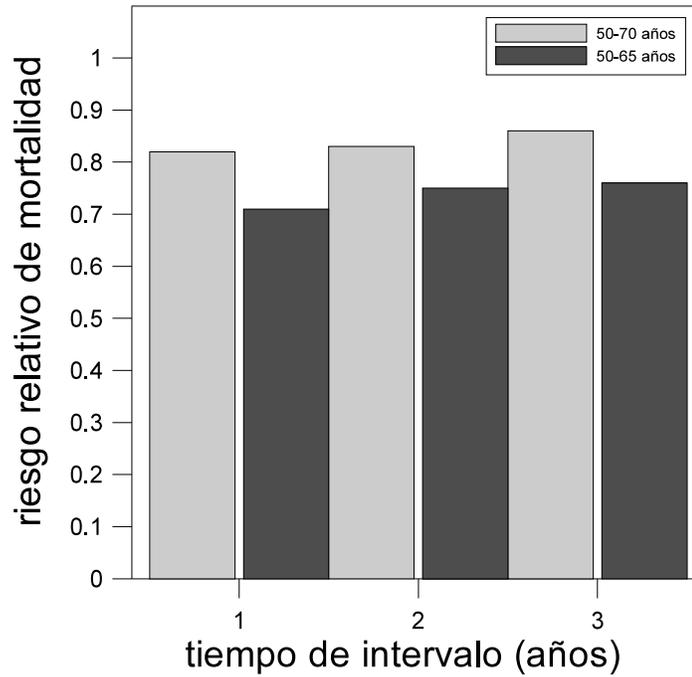


Figura 3.8: Comparación de RR respecto al de la configuración 40-70 para los mismos tiempos de intervalo. Se aprecia un mejor rendimiento de la configuración 50-70 años frente a la de 50-65 años.

cada muerte evitada [Got09]. Finalmente, utilizaremos una tasa de mortalidad anual por cáncer de mama de 20 mujeres por cada  $10^5$  (ver resultados anteriores).

En este ambiente, por cada 100 muertes por cáncer de mama, el grupo de mujeres que se sometiera a cribado tendría una ganancia en  $QALY$  de:

$$QALY_c = (86 - 66) \times RR \times 0.83 = RR \times 16.6 \text{ años}, \quad (3.9)$$

En cuanto al sobrediagnóstico/sobret ratamiento, el grupo de cribado tiene una pérdida de  $QALY$  respecto al no sometido al cribado por cada 100 muertes de:

$$QALY_c = RR \times 3 \times (86 - 60) \times (0.83 - 1) = -RR \times 13.3 \text{ años}, \quad (3.10)$$

Tendríamos por tanto un efecto neto global de:

$$QALY_c = RR \times 3.3 \text{ años}, \quad (3.11)$$

Si trasladamos estos años “efectivos” a días ( $QALD$ ) tendríamos que por cada  $5 \times 10^5$  mujeres sometidas a cribado anualmente (100 muertes por año es 5 veces la mortalidad anual por cáncer de mama, luego será la mortalidad correspondiente a  $5 \times 10^5$  mujeres y año):

$$QALD = RR \times 3.3 \times 365 = RR \times 1205 \text{ días}, \quad (3.12)$$

Estos días “efectivos” ganados deben multiplicarse por 20 para las mujeres que se someten al cribado en la configuración 50-70 años (pues se someterán en promedio 20 años al cribado), 30 para 40-70 años y 15 para 50-65 años. Por lo que el RR en cada una de las 3 configuraciones necesario para alcanzar al menos 1 día por cada mujer invitada al cribado sería:

$$RR_{50-70} = 500000/24100 = 21, \quad (3.13)$$

$$RR_{40-70} = 500000/36150 = 14, \quad (3.14)$$

$$RR_{50-65} = 500000/18075 = 28, \quad (3.15)$$

Estos serían los valores de RR mínimos deseables para alcanzar un día “efectivo” de ganancia del cribado. Este valor es consistente con el resultado del informe WHO 2002, [IAC02] que, si bien no supone un umbral, puede servir como referencia para decidir que RR y, por tanto, qué tasas de participación mínimas debería alcanzar un programa de cribado. Si representamos la ganancia en días “efectivos” en función del RR para las diferentes configuraciones tenemos los resultados mostrados en la Figura 3.9. Con este criterio y, a la luz de nuestros resultados, las participaciones mínimas deberían ser las indicadas en la Tabla 3.9.

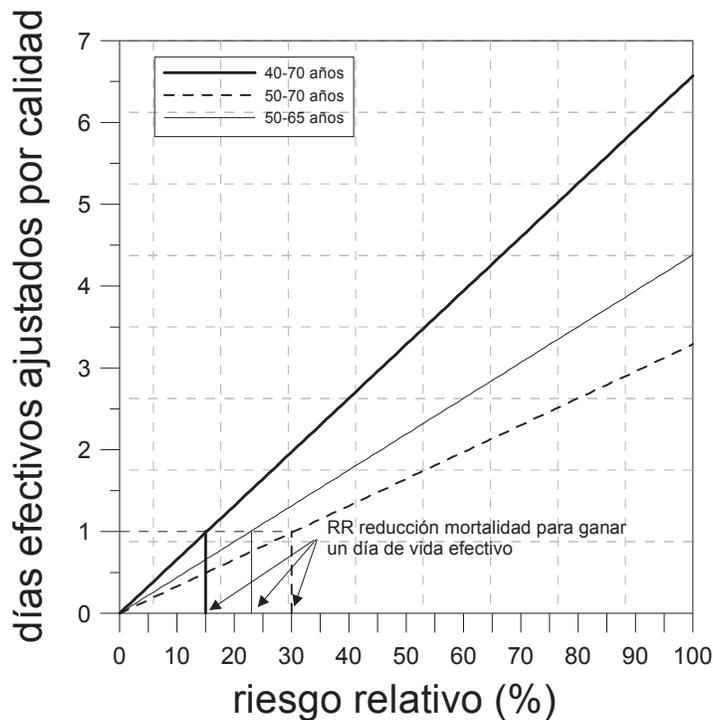


Figura 3.9: Días efectivos ganados por mujer invitada a cribado mamográfico en función del RR para las diferentes configuraciones estudiadas. Se comprueba que para los valores de RR supuestos en el entorno del 20 % la ganancia en días efectivos ronda el valor de 1 para todas las configuraciones.

De los anteriores resultados podemos concluir:

1. En la configuración 50-70 años son necesarios porcentajes de participación mayores que en la 40-70 años para una misma ganancia, siendo deseables participaciones de

Tabla 3.9: Porcentaje de participación mínimo para cada configuración simulada que proporciona una ganancia de 1 día efectivo por mujer invitada al cribado. Este valor es el estimado en el informe WHO [IAC02].

Configuración	Tiempo intervalo (años)	Porcentaje Participación para ganancia 1 día “efectivo”
50-70	1	70 %
	2	80 %
	3	89 %
40-70	1	38 %
	2	43 %
	3	50 %
50-65	1	100 %
	2	–
	3	–

no menos del 70 % para tener una justificación superior al día “efectivo” por mujer invitada al cribado.

2. La configuración de 40-70 años con tiempo de intervalo 3 años requiere de una participación mínima del 50 % con el mismo criterio anterior.
3. Las configuraciones de 50-65 años serían descartables pues no pueden alcanzar siquiera el umbral mínimo establecido de 1 día efectivo.

Esta estimación debería interpretarse de la siguiente forma: si por cada  $10^5$  mujeres hay 20 muertes anuales por cáncer de mama en la población sin un programa de cribado instaurado, si un porcentaje (participación) del grupo de edad considerado se somete al programa de cribado, la mortalidad total por cáncer de mama se reduce un valor igual al RR. Los años o días “efectivos” ganados deben afectar a toda la población (las  $10^5$  mujeres), pues se asume que, sea cual sea la edad de la mujer, al llegar a la edad de participación lo hará en un porcentaje fijado.

Estos resultados están basados en un detrimento por cáncer de mama de 0.83 asociado al cálculo de *QALY*. Este concepto, si bien controvertido, permite introducir el detrimento causado por una enfermedad (y su tratamiento) en los análisis coste-utilidad. Para su estimación se utilizan diferentes métodos basados en encuestas que intentan evaluar cómo se traduce este estado de salud en años de vida “efectivos”. La justificación del cribado es crítica, especialmente con aquellos parámetros de mayor incertidumbre como son el efecto del sobrediagnóstico, establecido en 3 mujeres por cada muerte evitada y en el factor, sin duda más artificial, del detrimento asociado al cáncer de mama antes referenciado. Los valores límite para la justificación del cribado en un análisis como el que hemos realizado se sitúan en 0.80 para el detrimento causado por cáncer de mama y de 3.8 sobrediagnósticos/sobretamientos por muerte evitada. Valores inferiores del primero o

superiores del segundo no permitirían alcanzar el criterio establecido de 1 día “efectivo” por mujer invitada al cribado.

En este análisis sencillo de justificación coste-utilidad se revela la magnitud que la introducción del sobrediagnóstico en el estudio puede tener, a pesar de que no se han tenido en cuenta otros efectos adversos como los falsos positivos que, sin duda, pueden suponer un detrimento aún mayor que el estimado.

### 3.3. SIMULACIÓN DE LOS PRINCIPALES ENSAYOS ALEATORIOS CONTROLADOS DE CRIBADO MAMOGRÁFICO

Los EAC que se consideran de mayor evidencia son: NY HIP [Sha85], *Swedish Two-county* [Tab81], Canada I [Mil02] y II [Mil00], Estocolmo [Fri86], Goteborg [Bju97], *UK Age trial* [Mos06] y, en menor medida, el de Edimburgo [Ale99]. En la Tabla 2.10 del Capítulo 2 se resumen las características de estos EAC.

Estos ensayos aleatorios controlados presentan algunas deficiencias o sesgos que pueden poner en cuestión tanto su validez interna (metodológica), como externa (aplicación a otras poblaciones), entre todos ellos destacamos los siguientes:

1. **Tamaño de la muestra y *end-point*.** En el diseño de los ensayos controlados debe decidirse el tamaño de los grupos de control e intervención (cribado) de acuerdo con la magnitud del efecto que se espera medir. Debido a la reducida magnitud que tiene la mortalidad por cáncer de mama en la mortalidad general (aproximadamente 21 muertes por cada  $10^5$  mujeres y año), el tamaño necesario para evaluar cambios en esta resultaría en grupos de control y de intervención absolutamente inviables. Se acepta, por tanto, un objetivo menos ambicioso, intentando estimar reducción de mortalidad por cáncer de mama como *end-point*, realizándose las estimaciones con un efecto supuesto de entre el 20 % y el 40 %.
2. **Aleatorización.** Para que los resultados del EAC sean correctos desde el punto de vista de la estadística inferencial, la asignación de la población a los grupos de control y cribado debe ser perfectamente aleatoria. Las asignaciones basadas en censos (fecha de nacimiento), poblaciones o zonas geográficas (*clusters*) pueden estar afectadas por sesgos asociados a parámetros socio-económicos o diferentes riesgos basales. Sin duda la aleatorización ideal es la individual. En este sentido, el ensayo de Edimburgo presenta una aleatorización sub-óptima, pues resultó que el 26 % de las mujeres del grupo de control pertenecían al estrato socio-económico más alto frente a un 53 % en el grupo de cribado [Ale94].
3. **Sesgo de selección.** Introduce diferencias asociadas al carácter voluntario de la participación en un ensayo. Para limitar este efecto se decide evaluar la mortalidad por cáncer de mama en mujeres “invitadas” al cribado, pues se presupone cierto sesgo en aquellas mujeres que aceptan la invitación [Zac04, Mos06].

4. **Adjudicación de la causa de muerte.** Este es otro sesgo importante para el que debe establecerse un “doble ciego” (donde ni los individuos participantes ni los investigadores saben quién pertenece al grupo de control y quién al de intervención), pues existe la tendencia a adjudicar la causa de muerte al cáncer de mama en mujeres sometidas al mismo en caso de duda. Este “doble ciego” puede establecerse mediante un comité independiente que certifique la causa de la muerte o la utilización de certificados nacionales de mortalidad. El ensayo *Swedish Two-county* está afectado de este sesgo [Got09]. Hay, además, cierta propensión a diagnosticar el cáncer de mama en el grupo de cribado frente al de control.
5. **Tiempo de seguimiento.** Es conocido que los efectos del cribado tendrían una latencia de al menos 5 años [Kop98]. Por eso tiempos de seguimiento muy limitados no darían una idea completa de la magnitud del efecto. Se recomiendan tiempos de seguimiento entre 10 y 15 años.
6. **Edad.** Los resultados de los EAC dependen del grupo de edad considerado en el mismo. Generalmente se asume un riesgo constante entre 50-70 años [Nys02] y menor para mujeres menores de 50 años. Los resultados de los EAC se suelen dar segregados en estos dos grupos. Como discutiremos más adelante, esta división no deja de ser ciertamente artificial.

Otros sesgos se refieren a la exclusión de mujeres diagnosticadas previamente con ganglios afectos o a la introducción de la autoexploración en la valoración de los resultados. No debe olvidarse además que estos ensayos, en su mayoría, se desarrollaron hace alrededor de 30 años, por lo que se añade la incertidumbre sobre extrapolación a un escenario como el actual en el que los tratamientos adyuvantes son más efectivos y la técnica de diagnóstico mucho más eficiente. Una manera de obtener la máxima información de estos EAC es utilizar los meta-análisis, que introducen diferentes pesos para compilar los resultados de los ensayos, de forma que tienen más importancia aquellos de mayor corrección metodológica.

### 3.3.1. Simulación de la reducción de mortalidad e intervalos de confianza de los EAC

A partir de las modificaciones introducidas en el programa principal buscamos simular los principales ensayos aleatorios llevados a cabo hasta la fecha, de forma que podamos testear su validez interna y externa. Los resultados de las simulaciones de los ensayos aleatorios se muestran en términos de RR de mortalidad por cáncer de mama en mujeres invitadas a cribado (valor medio de las  $10^3$  simulaciones de tamaño igual al grupo de cribado,  $N_C$  y de control,  $N_0$ ), junto con la estimación del intervalo de confianza del 95 %, obtenido de la desviación estándar muestral.

No hemos pretendido reproducir los valores absolutos de mortalidad reportados en los ensayos reales, para ello deberíamos actuar fundamentalmente sobre la supervivencia tras tratamiento loco-regional y, quizás, sobre las curvas de incidencia y prevalencia. En este sentido, como se observa en la Figura 3.10 [Got04], la tasa de mortalidad estandarizada en Canada por año y  $10^5$  mujeres era de alrededor de 23 en los años 80, cuando se realizan la mayor parte de los ensayos aleatorios, bajando hasta 16 en el año 2002 (año de publicación de los resultados de supervivencia tras tratamiento utilizados en las simulaciones).

Tabla 3.10: Resultados de las simulaciones de los principales EAC. La primera columna muestra los resultados reales reportados. La segunda son los resultados de las simulaciones usando la participación de la primera ronda. La tercera muestra el efecto de los diferentes porcentajes de participación por ronda (hasta 5 rondas). La cuarta columna introduce una distribución de detección sintomática suavizada con un umbral de 3 mm. Finalmente la última columna se calcula con valores anormalmente bajos de la sensibilidad global. Entre paréntesis se muestran los intervalos de confianza al 95 % de confianza.

EAC	Resultados		Sim. MC		Sim. MC		Sim. MC		Sim. MC	
	reales		p. 1 <sup>a</sup> ronda		p. rondas		suavizadas		baja Sens.	
	RR	IC 95 %	RR	IC 95 %	RR	IC 95 %	RR	IC 95 %	RR	IC 95 %
<b>NY(HIP)</b>	0.83	(0.70;1.00)	0.79	(0.57;1.05)	0.87	(0.63;1.16)	-	-	-	-
<b>Malmö</b>	0.81	(0.61;1.07)	0.81	(0.56;1.13)	0.82	(0.57;1.14)	-	-	-	-
<b>Kopp.</b>	0.58	(0.45;0.76)	0.79	(0.58;1.06)	-	-	-	-	-	-
<b>Ostergot.</b>	0.76	(0.61;0.95)	0.76	(0.57;1.00)	0.77	(0.59;1.00)	-	-	-	-
<b>Can. I</b>	0.97	(0.74;1.27)	0.64	(0.42;0.94)	0.69	(0.47;1.00)	0.71	(0.45;1.10)	0.97	(0.65;1.40)
<b>Can. II</b>	1.02	(0.78;1.33)	0.67	(0.45;0.97)	0.71	(0.49;1.02)	0.73	(0.49;1.05)	0.98	(0.68;1.38)
<b>Estocolmo</b>	0.73	(0.50;1.06)	0.77	(0.55;1.05)	0.78	(0.55;1.05)	-	-	-	-
<b>Goteborg</b>	0.75	(0.58;0.98)	0.73	(0.51;1.01)	0.76	(0.53;1.06)	-	-	-	-
<b>UK Age</b>	0.83	(0.66;1.04)	0.71	(0.52;0.95)	0.82	(0.62;1.07)	-	-	-	-

Nuestros resultados absolutos de mortalidad por cáncer de mama acumulados en 13 años en el grupo de control para el ensayo Canada I son de 66 mujeres y de 60 para Canada II. Si aplicamos una corrección simple basada en el cociente de tasas de mortalidad indicadas en el párrafo anterior y utilizamos la desviación estándar resultante de las  $10^3$  simulaciones de tamaño  $N_0$  para estimar el intervalo de resultados probables (95 % CI), resultan valores de 95 (71-118) para Canada I y 86 (63-109) para Canada II. Estos resultados son compatibles con los 108 informado en [Mil02] y 105 [Mil00], de forma que gran parte de la diferencia en los valores de mortalidad absoluta podría explicarse atendiendo a los cambios en la mortalidad basal.

Si hablamos en términos de reducción de mortalidad entre los grupo de cribado y control, los valores absolutos de mortalidad prácticamente no tendrán efecto, pudiendo estudiar de forma más nítida el efecto de las diferentes configuraciones del cribado, más allá de detalles poblacionales o coyunturales. Los resultados de RR de mortalidad por cáncer de mama en mujeres invitadas al cribado obtenidos en las simulaciones frente a los reales se muestran en la Figura 3.11 y Tabla 3.10. En esta última, se muestran los resultados con diferentes grados de aproximación: en la primera columna aparecen los resultados utilizando el valor de participación para la primera ronda como participación global. Los resultados bajo esta primera aproximación se reproducen muy bien para todos los ensayos aleatorios salvo para el casos de los ensayos canadienses, que son los que muestran una menor o nula reducción de mortalidad y el ensayo de Kopparberg, perteneciente al global *Swedish Two County*.

En una segunda aproximación (segunda columna de resultados), hemos introducido en el modelo de generación de las curvas de adelanto y sensibilidad la diferente participación por rondas (hasta la 5<sup>a</sup> ronda), añadiendo un muestreo adicional uniforme en cada ronda con una probabilidad igual a la participación en cada una de ellas. Con esta modificación

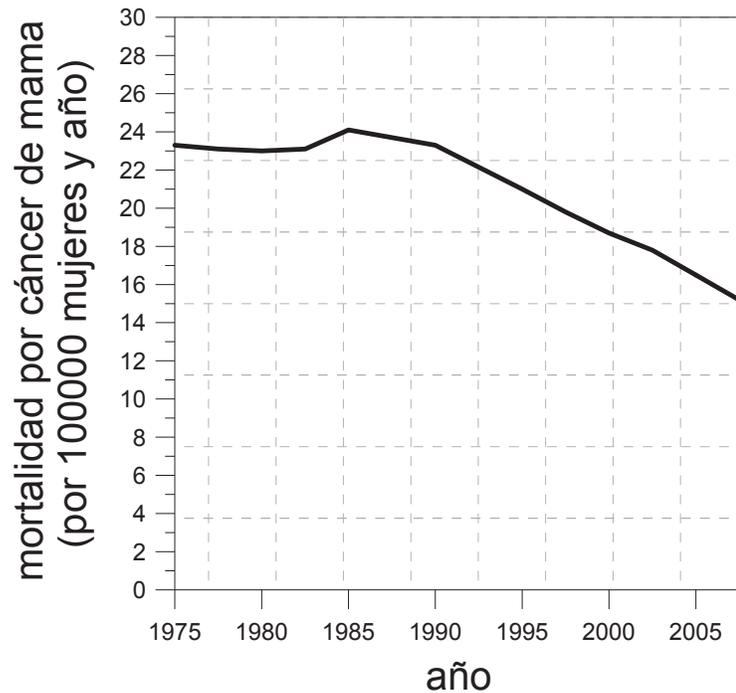


Figura 3.10: Tendencia de mortalidad anual por cáncer de mama en Canadá estandarizada por edad y  $10^5$  mujeres.

se obtiene valores de sensibilidad global más bajos, que reflejan la menor participación por ronda. Estos nuevos *inputs* se introducen en los programas de simulación (no se dispone de este dato para el ensayo de Kopparberg), resultando una pequeña mejora en el acuerdo con los ensayos reales, pero estando aún lejos de explicar las diferencias encontradas en los ensayos canadienses.

Debe tenerse en cuenta que en los ensayos canadienses no se compara mamografía frente a detección clínica, sino mamografía, exploración clínica y mamografía inicial frente a autoexploración más mamografía inicial. Las mujeres sometidas al estudio son adiestradas previamente en la autoexploración. Basándonos en esto, hemos pretendido comprobar la hipótesis de que esta autoexploración y mamografía iniciales se traducirían en una mejor detectabilidad clínica. Para esto hemos actuado sobre el modelo de detección sintomática de nuestras simulaciones, que sigue una distribución logarítmica [Ale94, Zac04] sobre el que se establece un umbral de detectabilidad clínica de 15 mm (todos los tumores que resulten detectados clínicamente con un valor inferior a estos 15 mm son llevados a este umbral).

Para intentar hacer compatibles nuestra distribución de tumores invasivos detectados sintomáticamente con la reportada en el ensayo Canada I para el grupo de control [Mil00], hemos rebajado el umbral de detectabilidad a 3 mm, aplicando un factor de 0.09 a los tumores detectados clínicamente con diámetro igual o inferior a 3 mm para corregir el artefacto de acumulación generado en este valor de detección y suavizar así la distribución. De esta manera obtenemos una distribución de tumores invasivos en el grupo de control que comparamos con la reportada en Canada I, en la que hemos introducido la incertidumbre correspondiente a la estimación de proporciones [And88]. No se pretende

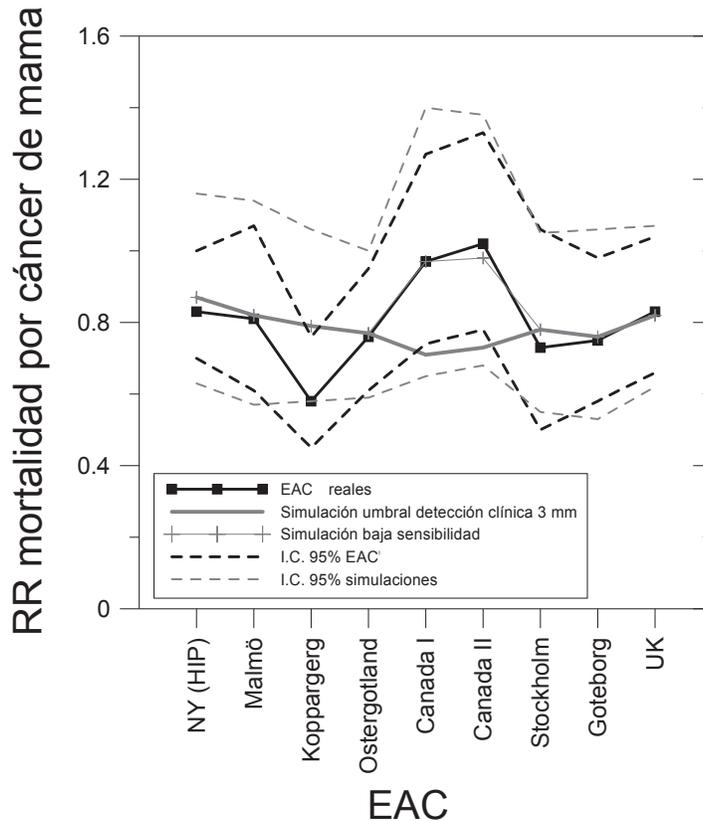


Figura 3.11: Comparación de los resultados de los ensayos reales frente a las simulaciones MC. Se muestran además los resultados para los ensayos canadienses usando valores bajos de la sensibilidad global para simular el efecto de la baja calidad de la mamografía.

reproducir la curva original, sino llevar nuestro modelo de detección clínica (autoexploración en este caso) a valores compatibles con los informados en los grupos de control de los ensayos canadienses que, como se puede comprobar en la Figura 3.12, tiene una elevada incertidumbre.

Introduciendo este modelo de detección clínica en nuestras simulaciones de los ensayos canadienses no se observa gran mejora en la adecuación de resultados a los valores de los ensayos reales (2 puntos porcentuales). Si aplicamos el modelo únicamente a la detección clínica en el grupo de control, manteniendo el umbral de 15 mm para el grupo de cribado, los valores de RR pasan a ser de 0.77 (0.47;1.07) para Canada I y 0.76 (0.49;1.03) para Canada II. Esta última aproximación supone considerar que la autoexploración no añade mejora a la mamografía en el grupo de cribado.

Tras realizar un contraste de hipótesis para estos resultados, suponiendo distribuciones normales y doble cola, encontramos que la confianza en obtener un valor como el informado para el ensayo Canada I, debido a la variabilidad por el tamaño muestral sería, con la aproximación anterior, del 6% (18% si no se aplica a grupo de cribado) y del 2% (5% en el segundo caso) para Canada II. A la vista de lo anterior, parece difícil atribuir el desacuerdo a un efecto de la autoexploración.

Buscamos, por tanto, contrastar la hipótesis de que los desacuerdos con estos ensayos son debidos a bajos valores de sensibilidad, asociados tanto a mala calidad de la mamografía

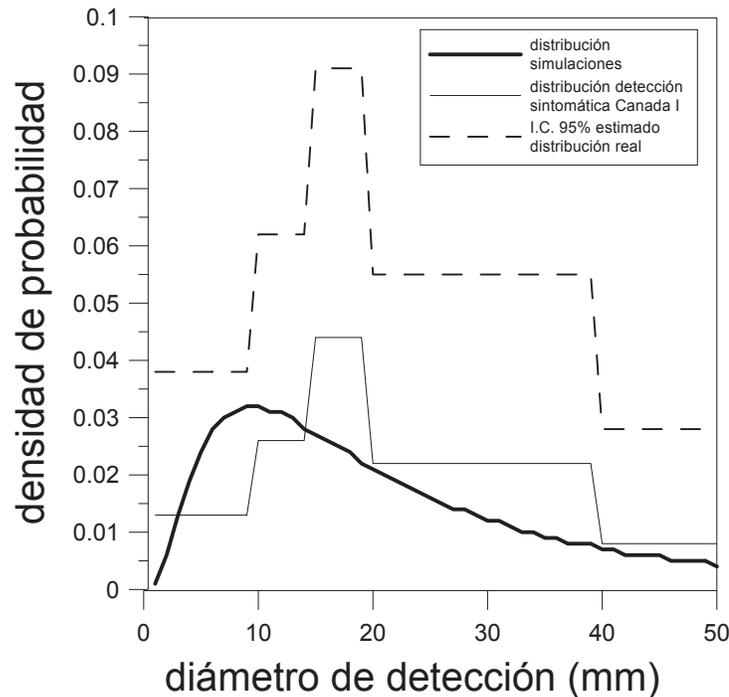


Figura 3.12: Distribución de tumores invasivos detectados en el grupo de control del ensayo real Canada I sometido a cuidados usuales [Mil00] (línea gris) y distribución suavizada de las simulaciones MC con umbral 3 mm (línea negra continua). Los intervalos de confianza (línea negra discontinua) se calculan usando propiedades estadísticas de las proporciones [And88].

grafía como a una deficiente interpretación de las mismas [ICRP91, IAC02, Duf10, Ber11, Ble12]. Atendiendo a esta hipótesis, realizamos simulaciones con valores mucho más bajos de sensibilidad global, parámetro que controla en nuestras simulaciones la eficiencia en la detección a lo largo del cribado. Si situamos este valor en un 10 % (los valores de las simulaciones originales corregidos por la diferente participación por ronda son de 68.4 % y 70.5 % para Canada I y II resp.) obtenemos los valores mostrados en la última columna de la Tabla 3.10.

Respecto al desacuerdo de las simulaciones con el ensayo de Kopparberg (*Swedish Two-County*), podría ser atribuido a defectos metodológicos. Así, se reporta un valor de  $RR=0.58$  (0.45; 0.76) que supondría una reducción de mortalidad del 42 %, sin duda muy elevada, más aún cuando se muestra muy diferente del resultado de Östergötland para edades similares  $RR=0.76$  (0.61; 0.95). Las simulaciones realizadas para el ensayo Kopparberg mostrarían un valor de 0.78 (0.54; 1.02), que parece más realista. En este punto debe hacerse notar que el procedimiento de asignación al azar en este ensayo dio lugar a grupos no comparables [Got09]: la mortalidad por cáncer de mama en el grupo de control fue casi 2 veces más alta en este ensayo que en el paralelo de Östergötland (0.0021 frente a 0.0012,  $P=0.02$ ), lo que introduce un sesgo a favor del cribado. Si introducimos una mortalidad similar en las simulaciones, el resultado sería del orden de  $RR=0.44$ , lo que explicaría en gran medida el resultado del ensayo Kopparberg.

Finalmente, cabría comentar el resultado un tanto paradójico del ensayo NY HIP. Las simulaciones muestran un valor similar al reportado por el ensayo real: 0.86 (0.62;

1.11) frente a 0.83 (0.70; 1.00). Sin embargo, la sensibilidad reportada para este ensayo revela una mala calidad en la mamografía [IAC02], con valores mucho menores que los informados en los demás ensayos. Puesto que en las simulaciones de los diferentes ensayos no se introducen diferencias *a priori* en el modelo de detección mamográfica, más allá de las debidas a la propia configuración del mismo (grupo de edades, tiempo de intervalo y número de rondas), el valor obtenido en las simulaciones debería ser mucho más favorable al cribado que el real. Esto podría haberse compensado por varios sesgos de aleatorización:

1. Casos de cánceres previos en grupo de control reportados 20 años después que debería haber sido excluidos del estudio [Asc04, Sha85].
2. Sesgo en la asignación de causa de muerte, que parecería haber favorecido al cribado [Asc04].

A la vista de los resultados expuestos anteriormente, se puede concluir que las simulaciones de los ensayos aleatorios Malmö, Ostergötland, Estocolmo, Goteborg y *UK Age trial* muestran resultados de RR de mortalidad por cáncer de mama muy similares a los informados. Estos ensayos presentan una buena calidad metodológica. Los resultados son, además, extrapolables a otras poblaciones en configuraciones de programas análogas, puesto que no se han introducido diferencias poblacionales, tasas de incidencia, prevalencia o proporción de tumores invasivos/*in situ* específicos. Mención especial merece el ensayo *UK Age trial* por lo limitado del grupo de edad investigado (39 a 41 años) que, sin embargo, muestra resultados de reducción de mortalidad similares a otros estudios menos específicos.

Respecto a los ensayos canadienses I y II, los resultados informados presentan valores en claro desacuerdo con ensayos similares. Hemos investigado la causa de este desacuerdo, puesto que no parece que pueda ser atribuido a defectos metodológicos [Asc04, Mil00, Fri86, Duf10], al menos en toda su magnitud. El efecto de la autoexploración y mamografía inicial tampoco parece, a la vista de las simulaciones, de suficiente entidad como para explicar la anomalía, como también se indica en [Mil00]. La hipótesis más plausible, a la luz de nuestros resultados, parecería ser una mala calidad en la mamografía, tanto a nivel técnico como interpretativo. Este extremo estaría avalado por diferentes autores [Duf10], algunos de los cuales informan valores de sensibilidad de la mamografía en estos ensayos muy inferiores a los obtenidos en estudios similares.

Los ensayos NY HIP y Kopparberg (correspondiente a *Swedish Two-County*) presentan defectos metodológicos importantes, ligados fundamentalmente a diferencias en la aleatorización entre grupos de control y cribado, que se traducen en grupos no comparables. Deberían, por tanto, ser descartados o, al menos, tener un peso menor en meta-análisis posteriores.

En cuanto a los intervalos de confianza, los límites superiores que determinan la significación son muy similares entre las simulaciones y los proporcionados por los estudios reales, excepción hecha del ensayo Kopparberg, consecuencia de un valor central de RR anómalo. Los intervalos inferiores que determinan el efecto mayor del cribado tienden a ser inferiores (mayor efecto del cribado) en los resultados de las simulaciones que en los ensayos reales. Esto es consecuencia de la asimetría que introduce el cálculo teórico del intervalo de confianza del RR basado en una distribución logarítmica [IAC02], a diferencia del método empleado en nuestras simulaciones, que proporciona un intervalo simétrico

alrededor del valor central del RR suponiendo distribuciones normales. Si aplicamos un modelo logarítmico para el cálculo de nuestras incertidumbres, obtenemos una reproducción muy precisa de los límites inferiores a costa de unos límites superiores algo mayores que los predichos en los ensayos reales.

Así pues, y a la vista de los resultados, una reducción de mortalidad por cáncer de mama en el entorno del 20 % sería razonable agrupando todos los ensayos que consideramos metodológicamente superiores. Además, la validez externa de estos estudios se demostraría elevada, pues no se han introducido diferencias poblacionales en las simulaciones para reproducir los resultados reales, y resultan bastante estables incluso en diferentes grupos etarios.



## Capítulo 4

# CONCLUSIONES Y PERSPECTIVAS

### 4.1. CONSIDERACIONES FINALES

En este trabajo se ha desarrollado un modelo de simulación Monte Carlo de los programas y ensayos aleatorios de cribado de cáncer de mama capaz de reproducir los resultados asociados a la detección de los programas implantados en poblaciones reales y se ha realizado una estimación del sobrediagnóstico consistente con los resultados conocidos. En una segunda fase se ha introducido la supervivencia tras detección y tratamiento loco-regional, lo que ha permitido estimar la reducción de mortalidad que suponen los programas de cribado y explorar las diferentes configuraciones posibles en aras a obtener la más adecuada. Finalmente, se han reproducido con éxito los diferentes ensayos aleatorios llevados a cabo hasta la fecha, con lo que ha podido explicarse el origen de las discrepancias entre ellos de una forma plausible.

Los aspectos diferenciales que presenta nuestro modelo en relación a otros desarrollados previamente se pueden enumerar como sigue:

1. Introducir el menor número posible de parametrizaciones compatible con unos resultados exitosos. En este punto, es conocido que la duración de la fase preclínica y la sensibilidad de la mamografía están fuertemente correlacionadas [Pac95], especialmente a edades tempranas (alrededor de los 40 años). Esta correlación está implícita en la propia definición de la duración de la fase preclínica: mayor sensibilidad implica una detección más temprana y una menor duración, por tanto, de la fase preclínica, dificultando una parametrización independiente de las dos magnitudes, especialmente en función de la edad de las mujeres. En este sentido, Cong *et al.* [Con05] encuentran que la parametrización con la edad de ambas magnitudes no mejora los resultados significativamente. Nuestro modelo de detección por mamografía tiene un carácter más fundamental, pues parametriza esta en función del tamaño de la mama, tipo histológico del tumor y densidad de la mama. Es a través de esta última donde aparece la dependencia con la edad, que en otros modelos se introduce de forma directa. Se trataría, por tanto, de una variable de confusión (correlacionada

con la densidad de la mama). No se incluye, por consiguiente, una dependencia de la duración de la fase preclínica con la edad. Este punto es especialmente relevante cuando los datos proceden de EAC, pues es conocido el sesgo de duración, por el que se detectan preferentemente tumores de crecimiento más lento [Kop98]. En cualquier caso, existen estudios que avalan la diferente duración de la fase preclínica con la edad, permitiendo nuestro modelo introducir esta dependencia mediante la aplicación de diferentes valores medios en función de la edad a la que surge el tumor.

2. El modelo de duración de la fase preclínica es sencillo, modelable con únicamente dos parámetros con significado físico y escasa incertidumbre: el tamaño medio de detección clínica y el umbral de detección clínica. Con estos parámetros se supone una dependencia exponencial limitada para valores pequeños del tamaño de detección. La justificación teórica del modelo supone una visión realista, sin acudir a modelos artificiales y excesivamente parametrizados.
3. Nuestro modelo no utiliza la hipótesis de que los tumores que surgen a edades más tempranas son más agresivos. Los peores resultados del cribado a estas edades se deberían, en esta aproximación, a la menor incidencia y peor sensibilidad de la mamografía, asociada a mayor densidad de las mamas. Así, nuestro modelo genera un tiempo medio de adelanto para mujeres entre 39 y 41 años con tiempo de intervalo 1 año de 3.7 años, y de 4.2 años para mujeres entre 50 y 59, lo que supone medio año de diferencia para mujeres en el entorno de los 40 años. El tiempo medio de adelanto está íntimamente correlacionado con la duración de la fase preclínica (*sojourn time*). Este resultado es consecuencia únicamente de la menor probabilidad de detección en mamas más densas.
4. No se introducen factores *ad hoc*, como la mejora en la supervivencia en función de la modalidad de detección. Puesto que los ajustes de supervivencia en función del tamaño tumoral en la detección están relacionados con la probabilidad de metástasis [Mic99], no parece razonable introducir variables adicionales de confusión.
5. La historia natural de los tumores se parametriza con un modelo de crecimiento basado en un único parámetro libre, distribuido según una log-normal, y una distribución de tumores invasivos/*in situ* inicial con probabilidades de transición *in situ* a invasivos. El ajuste de esta proporción se realiza en virtud de los resultados en la detección de las diferentes rondas de cribado. Este resultado es consistente y no hace hipótesis adicionales sobre la evolución de los tumores. En este punto, si bien los modelos de Markov son útiles, se ha demostrado que sus resultados son fuertemente dependientes de los estados definidos, resultando en algunos casos artificiales [Uhr10]. En nuestro modelo se introduce los CDis y su evolución a invasivos, con la hipótesis de que no producen la muerte si se mantienen como *in situ*. La evolución del tumor es mucho más “natural”, encontrándose los factores pronósticos incluidos de manera implícita en las curvas de supervivencia en función del diámetro de detección, a diferencia del código MISCAN que incluye una evolución entre estados un tanto artificial.
6. Se introduce el efecto de la radioinducción de tumores.

## 4.2. CONCLUSIONES

1. El modelo reproduce de manera consistente los resultados de los programas reales asociados a la detección, y es relativamente insensible a la distribución de tumores subyacentes. Las salidas se muestran especialmente dependientes de los dos parámetros que caracterizan la detección clínica o sintomática, a saber: tamaño medio de detección clínica y umbral de detección. Estos son dos parámetros fácilmente estimables en los programas reales a diferencia de la proporción de tipos subyacente.
2. Con este primer modelo, la estimación del sobrediagnóstico está entre el 10 % y el 20 % de todos los tumores detectados en cribado, siendo algo menor en el rango de edad entre 40 y 70 años, debido a la mayor esperanza de vida de este grupo de edad que supone más tiempo para la aparición de los tumores que se detectan de manera temprana. Se comprueba también el incremento del sobrediagnóstico a medida que disminuye el tiempo de intervalo. Estos valores quedan reducidos hasta el 7 % a 10 % para configuraciones realistas y porcentajes de participación del 80 %. Si se analizan únicamente los tumores invasivos, se encuentra que para mujeres entre 50 y 70 años, los tiempos de intervalo entre 1 y 2 años suponen que todo el exceso de tumores detectados en el grupo de cribado se debe a tumores *in situ*, habiendo un déficit de tumores invasivos para tiempos de intervalo de 3 años de alrededor del 11 % de todos los tumores detectados al final del cribado. Para mujeres entre 40 y 70 años este déficit es menor (alrededor del 7 %) pero, por contra, se daría en las tres configuraciones estudiadas. Este análisis muestra especialmente beneficiosa la configuración de 50-70 años y tiempos de intervalo 3 años, pues disminuye de forma significativa el número de tumores invasivos en el grupo sometido al cribado, a pesar de presentar un exceso de tumores *in situ* del 24 %.
3. En cuanto a la reducción de mortalidad por cáncer de mama atribuible al cribado, nuestras simulaciones arrojan resultados entre el 23 % y el 33 %, según rango de edades y configuración del programa. Para porcentajes de participación realistas, estos valores se reducirían a una horquilla de entre 16 %-23 %. En virtud del análisis coste-utilidad realizado, se comprueba que son necesarios porcentajes de participación mayores en el grupo de edad 50-70 años para obtener la misma ganancia en años de vida efectivos que para el grupo 40-70 años, fijándose un umbral de participación del 70 % para este último grupo de edad de cara a justificar el programa. De este mismo análisis se desprende que el grupo de edad de entre 50 y 65 años debería ser descartado por no obtenerse en ningún caso una ganancia mínima de 1 día efectivo por mujer invitada al cribado.
4. Finalmente, nuestras simulaciones de los principales EAC muestran que los de Malmö, Östergötland, Estocolmo, Goteborg y *UK Age trial* presentan una buena calidad metodológica y sus resultados son extrapolables a otras poblaciones. Los ensayos canadienses podrían explicar sus resultados de escasa o nula reducción de mortalidad debido a la baja calidad de la mamografía, presentando los ensayos NY HIP y Kopparberg defectos metodológicos importantes.

Se ha generado por tanto una herramienta poderosa capaz de aportar luz en el controvertido mundo del cribado mamográfico, mostrándose eficaz a la hora de analizar realidades

complejas que implican aspectos radiobiológicos asociados al crecimiento tumoral, clínicos asociados a la supervivencia tras detección y tratamiento, diagnósticos en cuanto al modelado de la detección mamográfica y clínica, así como demográficos y epidemiológicos e incluso económico/administrativos. Si bien, como dijimos anteriormente, no se puede utilizar como única fuente de conocimiento, si que puede ayudar a discernir aspectos de justificación, configuraciones óptimas y evaluación de resultados, siendo capaz además de encontrar defectos metodológicos que pudiera haberse cometido en otro tipos de estudios experimentales.

### **4.3. PERSPECTIVAS**

En la primera serie de trabajos que han dado lugar a esta tesis se ha pretendido generar un modelo de simulación Monte Carlo estable y capaz de reproducir los principales resultados conocidos de los programas y ensayos aleatorios controlados de cribado mamográfico. Este modelo utiliza distribuciones poblacionales genéricas asociadas a la edad, incidencia de cáncer de mama, prevalencia, mortalidad general, etc., así como parametrizaciones de la sensibilidad de la mamografía y modelo de detección clínica/sintomática genéricas. El siguiente paso sería, por tanto, la introducción en el modelo de distribuciones concretas asociadas a poblaciones conocidas de las que se pretenda saber, con mayor precisión, cuál es la configuración del programa más eficiente en términos puramente cuantitativos asociados a la detección, así como en cuanto a reducción de mortalidad y el sobrediagnóstico. Más allá de esto, permitiría llevar a cabo estudios coste-utilidad mucho más precisos, permitiendo definir y estudiar configuraciones de los programas de cribado, comparanándolos en términos de coste de oportunidad con otras alternativas, y definiendo exactamente los recursos a emplear en virtud del coste marginal que se estimara.

La aplicación a poblaciones de riesgo elevado de cáncer de mama, asociado a entornos socio-económicos o de tipo biológico, como las mutaciones genéticas tipo BRCA1 y 2, podría proporcionar una valiosa información sobre los beneficios del cribado, si es que los hubiera, y sus riesgos. Este es sin duda un ámbito de aplicación en el que este trabajo podría generar conocimiento donde otras herramientas científicas no son útiles, por el limitado tamaño de las poblaciones a estudiar. Bastaría para esto con introducir distribuciones de incidencia conocidas para estos grupos. De igual forma se pueden modelar fácilmente otras terapias mediante la introducción de las oportunas curvas de supervivencia tras tratamiento, así como cualquier modalidad de detección, conocida su sensibilidad y dependencia con las características histológicas, tamaño del tumor y densidad de la mama.

De forma natural, este modelo y el código que lo implementa, podría ser extendido para estudiar cualquier tipo de programa de cribado en materia de salud pública. Como hemos indicado en repetidas ocasiones, el carácter masivo que suponen estos programas, aplicados a poblaciones sanas o asintomáticas, junto con el delicado balance riesgo-beneficio de los mismos, hacen necesario un análisis profundo para una correcta justificación y optimización de los programas. Sería esta, por tanto, una herramienta de gran ayuda para responder a las principales preguntas que todavía hoy quedan abiertas en relación a los programas de cribado.

# Bibliografía

- [ACC] Asociación española contra el cáncer. Disponible en: <https://www.aecc.es/es/todo-sobre-cancer/cancer-mama/evolucion-cancer-mama>. (Acceso Diciembre de 2010).
- [ACR] American College of Radiology (2003). Breast imaging reporting and data system, breast imaging atlas. 4th Ed. Reston VA, E.E.U.U.
- [AEC] Asociación española contra el cáncer. Disponible en: <https://www.aecc.es/sobreelcancer/cancerporlocalizacion/cancermama/paginas/pronostico.sppa>. (Acceso Diciembre de 2010).
- [Ake02] Burnham KP, Anderson DR. (2002). Model selection and multimodel inference a practical information-theoretic approach. 2nd Ed. Nueva York, E.E.U.U.: Springer.
- [Ale99] Alexander FE, Anderson TJ, Brown HK, *et al.* 14-years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999;353:1903-1908.
- [Ale94] Alexander FE, Anderson TJ, Brown HK, *et al.* The Edimburgh randomised trial of breast cancer screening: results after 10 years of follow-up. *Br J Cancer* 1994;70:542-548.
- [And88] Andersson I, Aspegren K, Janzon L, *et al.* Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. *BMJ* 1988;297:943-948.
- [Ane92] Anerlöv C, Medin SO, Lundgren B, *et al.* Breast carcinoma growth rate described by mammographic doubling time and S-phase fraction. Correlations to clinical and histopathologic factors in a screened population. *Cancer* 1992;70:1928-1934.
- [Arr15] Arrospeide A, Rue M, van Ravesteyn NT, *et al.* Evaluation of health benefits and harms of the breast cancer screening programme in the Basque Country using discrete event simulation. *BMC Cancer* 2015;15:671.
- [Asc04] Ascunce N, Barcos A, Ederra M, Redozain N, Murillo A, Mellado M. Programa de detección precoz de cáncer de mama. Resultados de proceso e indicadores de impacto (1990-2002). *An Sist Sanit Navar* 2004;27:319-333.
- [Aus09] Australian Institute of Health and Welfare National Breast and Ovarian Cancer Centre AIHW. (2009). Breast cancer in Australia: an overview. Cancer Series No. 50 Cat. No CAN 46. AIHW. Canberra, Australia.

- [Aut11] Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ* 2011;343-441.
- [Bag10] Bagdonavicius V, Levulienė R, Nikulin M. Goodness-of-fit Criteria for the Cox model from Left Truncated and Right Censored Data. *J Math Sci* 2010;167:436-443.
- [Bai10] Bailey SL, Sigal BM, Plevritis SK. A simulation model investigating the impact of tumor volume doubling time and mammographic tumor detectability on screening outcomes in women aged 40-49 years. *J Natl Cancer Inst* 2010;102:1263-1271.
- [Bar96] Barclay J, Kerlikowske K, Grady D, *et al.* Effect of age, breast density and family history on the sensitivity of first screening mammography. *JAMA* 1996;276:33-38.
- [Beg04] Berg WA, Gutierrez L, Nessaiyer MS, *et al.* Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in Preoperative assessment of breast cancer. *Radiology* 2004;233:830-849.
- [Ber06] Berry DA, Inoue L, Shen Y, *et al.* Modelling the impact of treatment and screening on U.S. breast cancer mortality: a Bayesian approach. *J Natl Cancer Inst Monogr* 2006;36:30-36.
- [Ber06b] Beral V. (2006). Screening for breast cancer in England: past and future. Advisory Committee on Breast Cancer Screening. Sheffield, Reino Unido: NHSBSP Publication No. 61.
- [Ber11] Beral V, Alexander M, Duffy S, *et al.* The number of women who would need to be screened regularly by mammography to prevent one death from breast cancer. *J Med Screen* 2011;18:210-212.
- [Haz01] Hazewinkel M. (2001). Beta-function. *Encyclopaedia of Mathematics*. Springer:ISBN 978-1556080104.
- [Bie07] Biesheuvel C, Barratt A, Howard K, Houssami N, Irwing L. Effects of study methods and biases on estimates of breast cancer overdiagnosis with mammography screening: a systematic review. *Lancet Oncol* 2007;8:1129-1138.
- [Bju97] Bjurstam N, Björnelid L, Duffy SW, *et al.* The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization. *Cancer* 1997;80:2091-2099.
- [Ble12] Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med* 2012;367:1998-2005.
- [Boe98] Boer R, de Koning H, Threlfall A, *et al.* Cost effectiveness of shortening screening interval or extending age range of NHS breast screening programme: computer simulation study. *BMJ* 1998;317:376-379.
- [Box58] Box GEP, Müller ME. A Note on the Generation of Random Normal Deviates. *Ann Math Statist* 1958;29:610-611.
- [Byr95] Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. *SIAM J Sci Comput* 1995;16:1190-1208.

- [Coe90] Coebergh JWW, Crommelin MA, Kluck HM, *et al.* Borstkanker in zuidoost Noord-Brabant en in Noord-Limburg: beloop van incidentie en vervroeging van de diagnose in een niet-gescreende vrouwelijke bevolking 1975-1986. *Ned Tijdschr Geneesk* 1990;134:760-765.
- [Col84] Collette HJA, Day NE, Rambach JJ, Waard F de. Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case-control study. *Lancet* 1984;I:1224-1226.
- [Con05] Cong XJ, Shen Y, Miller AB. Estimation of age-specific sensitivity and sojourn time in breast cancer screening studies. *Stat Med* 2005;24:3123-3138.
- [Cor06] Cortesi L, Turchetti D, Marchi I, *et al.* Breast cancer screening in women at increased risk according to different family histories: an update of the Modena Study Group experience. *BMC Cancer* 2006;6:210.
- [CTF11] Canadian Task Force on Preventive Health Care. Recommendations on screening for breast cancer in average-risk women aged 40-74 years. *Can Med Association J* 2011;183:1991-2001.
- [Dij97] Van Dijck JA, Verbeek AL, Beex LV, *et al.* Breast-cancer mortality in non-randomized trial on mammography screening in women over age 65. *Int J Cancer* 1997;70:164-168.
- [Duf91] Duffy SW, Tabar L, Fagerberg G, *et al.* Breast screening prognostic factors and survival-results from the Swedish two-county study. *Br J Cancer* 1991;64:1133-1138.
- [Duf95] Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. *Stat Med* 1995;14:1531-1543.
- [Duf97] Duffy SW, Day NE, Tabár L, Chen HH, Smith TC. Markov models of breast of breast tumor progression: some age-specific results. *J Natl Cancer Inst Monogr* 1997;22:93-97.
- [Duf10] Duffy SW, Tabár L, Olsen AH, *et al.* Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010;17:25-30.
- [Duf12] Duffy SW, Yen Ming-Fang A, Chen TH, *et al.* Long-term benefits of breast screening. *Breast Cancer Management* 2012;1:31-38.
- [Duf13] Duffy SW, Parmar D. Overdiagnosis in breast cancer screening: the importance of length of observation period and lead time. *Breast Cancer Res* 2013;15:41.
- [Eco07] Last JM. (2007). *A dictionary of public health*. Oxford, Reino Unido: Oxford University Press.
- [Etz14] Etzioni R, Xia J, Hubbard R, Weiss NS, Gulati R. A reality check for overdiagnosis estimates associated with breast cancer screening. *J Natl Cancer Inst* 2014;106.
- [Fei94] Feig SA. Determination of mammographic screening intervals with surrogate measures for women 20-49 years. *Radiology* 1994;193:311-314.

- [Fet01] Fett MJ. Computer modelling of the Swedish two county trial of mammographic screening and trade offs between participation and screening interval. *J Med Screen* 2001;8:39-45.
- [For10] Forastero C, Zamora LI, Guirado D, Lallena AM. A MonteCarlo tool to simulate breast cancer screening programmes. *Phys Med Biol* 2010;55:5213-5229.
- [For19] Forastero C, Zamora LI, Guirado D, Lallena AM. Evaluation of the overdiagnosis in breast screening programmes using a Monte Carlo simulation tool: a study of the influence of the parameters defining the programme configuration. *BMJ Open* 2019;9.
- [Fri86] Frisell J, Glas U, Hellström L, Somell A. Randomized mammographic screening for breast cancer in Stockholm. Design, first round results and comparisons. *Breast Cancer Res Treat* 1986;8:45-54.
- [Fri87] Frisell J, Eklund G, Hellstrom L, Somell A. Analysis of interval breast carcinomas in a randomized screening trial in Stockholm. *Breast Cancer Res Treat* 1987;9:219-225.
- [Fry06] Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittman V, Remington PL. The Wisconsin breast cancer epidemiology simulation model. *J Natl Cancer Inst Monogr* 2006;36:37-47.
- [Gen81] Gentili C, Sanfilippo O, Silvestrini R. Cell proliferation in relation to clinical features and relapse in breast cancer. *Cancer* 1981;48:974-979.
- [Goc15] Gocgun Y, Banjevic D, Taghipour S, *et al.* Cost-effectiveness of breast cancer screening policies using simulation. *Breast* 2015;24:440-448.
- [Got04] Gøtzsche PC. On the benefits and harms of screening for breast cancer. *Int J Epidemiol* 2004;33:56-64; discussion 69-73.
- [Got09] Gøtzsche PC, Nielsen M. (2009). Screening for breast cancer with mammography 2009 (Review). The Cochrane Collaboration. Copenhagen, Dinamarca: John Wiley and Sons.
- [Gre10] Greuter MJ, Jansen-van der Weide MC, Jacobi CE, *et al.* The validation of a simulation model incorporating radiation risk for mammography breast cancer screening in women with a hereditary-increased breast cancer risk. *Eur J Cancer* 2010;46:495-504.
- [Gun12] Gunsoy NB, Garcia-Closas M, Moss SM. Modelling the overdiagnosis of breast cancer due to mammography screening in women aged 40 to 49 in the United Kingdom. *Breast Cancer Res* 2012;14.
- [Hab85] Habbema JD, Van Oortmarsen GJ, Lubbe JT, Van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. *Comput Methods Programs Biomed* 1985;20:79-93.
- [Hak95] Hakkama M, Holli K, Isola J, Olli-Pekka K, Karkkainen A, Visakorpi T, *et al.* Aggressiveness of screen-detected breast cancers. *Lancet* 1995;345:221-224.
- [Hol06] Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006;36:19-25.

- [Hol94] Holland R, Peterse JL, Millis RR, *et al.* Ductal carcinoma in situ: a proposal for a new classification. *Semin Diagn Pathol* 1994;11:167-180.
- [Hsu18] Hsu CY, Yen MF, Auvinen A, Chiu YH, Chen HH. Bayesian negative-binomial-family-based multistate Markov model for the evaluation of periodic population-based cancer screening considering incomplete information and measurement errors. *Stat Methods Med Res* 2018;27:2519-2539.
- [Hua18] Huang X, Li Y, Song J, Berry DA. A bayesian simulation model for breast cancer screening, incidence, treatment and mortality. *Med Decis Making* 2018;38:78S-88S.
- [Hum02] Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;137:347-360.
- [IAC02] International Agency for Research on Cancer. (2002). IARC Handbooks of Cancer Prevention. Breast Cancer Screening. Lyon, Francia: IARC 7.
- [IAR12] International Agency for Research on Cancer IARC. Disponible en: <http://www-dep.iarc.fr>. (Acceso Diciembre de 2012).
- [ICRP91] ICRP 1991. (1990). Recommendations of the International Commission on Radiological Protection. ICRP Publication 60. *Ann. ICRP* 21(1-3).
- [Ind12] Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: and independent review. *Lancet* 2012;380:1778-1786.
- [INE07] INE Instituto Nacional de Estadística. Disponible en: [www.ine.es](http://www.ine.es). (Acceso Diciembre de 2007).
- [INE14] INE Instituto Nacional de Estadística. Disponible en: [www.ine.es](http://www.ine.es). (Acceso Diciembre de 2014).
- [Ish19] Ishende G, Humphreys K. Modelling breast cancer tumor growth for stable disease population. *Stat Methods Med Res* 2019;28:681-702.
- [Jac06] Jacobi CE, Nagelkerke NJ, van Houwelingen JH, de Bock GH. Breast cancer screening, outside the population-screening program, of women from breast cancer families without proven BRCA1/BRCA2 mutations: a simulation study. *Cancer Epidemiol Biomarkers Prev* 2006;15:429-436.
- [Jan95] Jansen JT, Zoetelief J. MBS: a model for risk benefit analysis of breast cancer screening. *Br J Radiol* 1995;68:141-149.
- [Jan97] Jansen JT, Zoetelief J. Optimisation of mammographic breast cancer screening using a computer simulation model. *Eur J Radiol* 1997;24:137-144.
- [Jon03] Jonker MA, Jacobi CE, Hoogendoorn WE, Nagelkerke NJD, de Bock GH, Van Houwelingen JC. Modeling familial clustering breast cancer using published data. *Cancer Epidemiol Biomarkers Prev* 2003;12:1479-1485.
- [Kal85] Kalbfleisch JD, Lawless JF. The analysis of panel data under a Markov assumption. *J Am Stat Assoc* 1985;80:863-871.

- [Kar99] Karrison TG, Ferguson DJ, Meier P. Dormancy of mammary carcinoma after mastectomy. *J Natl Cancer Inst* 1999;91:80-85.
- [Ker96] Kerlikowske KM, Grady DG, Barclay J, Sickles EA and Ernster V. Effect of age, breast density and family history on the sensitivity of first screening mammography. *J Am Med Assoc* 1996;276:33-38.
- [Kon91] De Koning HJ, van Ineveld BM, van Oortmarssen GJ, *et al.* Breast-cancer screening and cost-effectiveness, policy alternatives, quality of life considerations and the possible impact of uncertain factors. *Int J Cancer* 1991;49:531-537.
- [Kop93] Kopans DB, Feig SA. The Canadian National Breast Screening Study: a critical review. *AJR Am J Roentgenol* 1993;161:755-760.
- [Kop98] Kopans DB.(1998). *Breast Imaging*. Philadelphia, E.E.U.U.: Lippincott-raven.
- [Kos84] Koscielny S, Tubiana M, Le MG, *et al.* Breast cancer: relationship between the size of primary tumor and the probability of metastatic dissemination. *Br J Cancer* 1984;49:709-715.
- [Kri06] Kriege M, Brekelmans CT, Boetes C, *et al.* Dutch MRI Screening (MRISC) Study Group. Differences between first and subsequent rounds of the MRISC breast cancer screening program for women with a familial or genetic predisposition. *Cancer* 2006;106:2318-2326.
- [Kuu97] Kuukasjarvi T, Karhu R, Tanner M, *et al.* Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res* 1997;57:1597-1604.
- [Lea05] Leach MO, Boggis CR, Dixon AK, *et al.* MARIBS study group. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: a prospective multicentre cohort study (MARIBS). *Lancet* 2005;365:1769-1778.
- [Lee06] Lee S, Zelen M. A stochastic model for predicting the mortality of breast cancer. *J Natl Cancer Inst Monogr* 2006;36:79-86.
- [Lin95] Lindfors KK, Rosenquist C. The Cost-effectiveness of mammographic screening strategies. *JAMA* 1995;274:881-884.
- [Man06] Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. *J Natl Cancer Inst Monogr* 2006;36:47-55.
- [Mic99] Michaelson JS, Halpern E, Kopans DB. Breast cancer: computer simulation method for estimating optimal intervals for screening. *Radiology* 1999;212:551-560.
- [Mic02] Michaelson J, Silverstein M, Wyatt J. Predicting the survival of patients with breast carcinoma using tumor size. *Cancer* 2002;95:713-723.
- [Mil89] Miller A, Howe GR, Sherman GJ, *et al.* Mortality from breast cancer after irradiation during fluoroscopic examinations in patients being treated for tuberculosis. *N Engl J Med* 1989;321:1285-1289.

- [Mil00] Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-Year Results of a Randomized Trial in Women Aged 50-59 Years. *J Natl Cancer Inst* 2000;92:1490-1499.
- [Mil02] Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002;137:305-312.
- [Mos86] Moskowitz M. Breast cancer: Age-specific growth rates and screening strategies. *Radiology* 1986;161:37-41.
- [Mos05] Moss SM. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled trials of breast cancer screening. *Breast Cancer Res* 2005;7:230-234.
- [Mos06] Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Trial Management Group. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006;368:2053-2060.
- [Muk92] Mukerjee R. Comparison between the conditional likelihood and the usual likelihood ratio tests. *J Royal Stat Soc, Series B* 1992;54:189-194.
- [Mus98] Mushlin AI, Kouides RW, Shapiro DE. Estimating the accuracy of screening mammography: a Meta-analysis. *Am J Prev Med* 1998;14.
- [NCR06] NCR. (2006). Health risks from exposures to low levels of ionizing radiation: BEIR VII Phase 2. National Research Council. Washington DC, E.E.U.U.: National Academy Press.
- [Nel65] Nelder JA, Mead R. A simplex method for function minimization. *Comput J* 1965;7:308-313.
- [Nel09] Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. Preventive Services task Force. *Ann Intern Med* 2009;151:717-737.
- [Nys02] Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909-19.
- [Oor90] Van Oortmarsen GJ, Habbema JD, van der Maas PJ, *et al.* A model for breast cancer screening. *Cancer* 1990;66:1601-1612.
- [Pac95] Paci E, Boer R, Zappa M, *et al.* A model-based prediction of the impact on reduction in mortality by a breast cancer screening programme in the city of Florence, Italy. *Eur J Cancer* 1995;31A:348-353.
- [Pee93] Peer PGM, Van Dijck JAAM, Hendriks JHCL, *et al.* Age-dependent growth rate of primary breast cancer. *Cancer* 1993;71:3547-3551.
- [Pee96] Peer PG, Verbeek AL, Straatman H, Hendriks JH, Holland R. Age-specific sensitivities of mammographic screening for breast cancer. *Breast Cancer Res Treat* 1996;38:153-160.

- [Ple06] Plevritis SK, Kurian AW, Sigal BM, *et al.* Cost-effectiveness of screening BRCA1/2 mutation carriers with breast magnetic resonance imaging. *JAMA* 2006;295:2374-2384.
- [Plev06b] Plevritis SK, Sigal BM, Salzman P, Rosenberg J, Glynn P. A stochastic simulation model of US breast cancer mortality trends from 1975 to 2000. *J Natl Cancer Inst Monogr* 2006;36:86-95.
- [Pre02] Preston AL, Mattsson A, Holmberg E, Shore R, Hildreth NG, Boice JD. Radiation effects on breast cancer risk: a pooled analysis of eight cohorts. *Radiat Res* 2002;158:220-235.
- [Pres02] Preston DL, Mattsson A, Holmberg E, Shore R, Hildreth NG, Boice Jr JD. Radiation effects on breast cancer risk: a pooled analysis of eight cohorts. *Radiat Res* 2002;158:220-235.
- [Pul11] Puliti D, Miccinesi G, Paci E. Overdiagnosis in breast cancer: design and methods of estimation in observacional studies. *Prev Med* 2011;53:131-133.
- [Rav11] Van Ravesteyn NT, Heijnsdijk EA, Draisma G, de Koning HJ. Prediction of higher mortality reduction for the UK Breast Screening Frequency Trial: a model-based approach on screening intervals. *Br J Cancer* 2011;105:1082-1088.
- [Rav15] Van Ravesteyn NT, Stout NK, Schechter CB, *et al.* Benefits and harms of mammography screening after age 74 years: model estimates of overdiagnosis. *J Natl Cancer Inst* 2015;107.
- [Ros06] Rosenberg R, Yankaska B, Abraham L, *et al.* Performance benchmarks for screening mammography. *Radiology* 2006;241:55-66.
- [Sal94] Salked G, Gerard K. Will early detection of breast cancer reduce the costs of treatment?. *Aus J Public Health* 1994;18:388-393.
- [San03] Sant M, Allemani C, Capocaccia R, *et al.* Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. *Int J Cancer* 2003;106:416-422.
- [SCS96] Servicio Canario de Salud. (1996). Dirección general de salud pública. Situación en la comunidad autónoma canaria. Incidencia del cáncer de mama. Registro poblacional del cáncer.
- [Sei16] Seigneurin A, Labarere J, Françoise O, *et al.* Overdiagnosis and overtreatment associated with breast cancer mammography screening: a simulation study with calibration to population-based data. *Breast* 2016;28:60-66.
- [Sha85] Shapiro S, Venet W, Strax P, Venet L, Roeser R. Selection, follow-up, and analysis in the Health Insurance Plan Study: a randomized trial with breast cancer screening. *J Natl Cancer Inst Monogr* 1985;67:65-74.
- [She05] Shen Y, Zelen M. Robust modelling in screening studies: estimation of sensitivity and preclinical sojourn time distribution. *Biostatistics* 2005;6:604-614.
- [Shen05] Shen Y, Parmigiani G. A model-based comparison of breast cancer screening strategies: mammograms and clinical breast examinations. *Cancer Epidemiol Biomarkers Prev* 2005;14:529-532.

- [Shw75] Shwartz M. (1975). Analysis of preventive screening to detect early breast cancer. (Ph D. disertacion). Department of Urban and regional Planning. Ann Arbor (MI), University of Michigan, E.E.U.U.
- [Shw78] Shwartz M. A mathematical model used to analyze breast cancer screening strategies. *Oper Res* 1978;26:937-955.
- [Shw81] Shwartz M. Validation and use of mathematical model to estimate the benefits of screening younger women for breast cancer. *Cancer Detect Prev* 1981;4:595-604.
- [Shw92] Shwartz M. Validation of a model of breast cancer screening: an outlier observation suggests the value of breast self-examinations. *Med Decis Making* 1992;12:222-228.
- [Sig12] Sigal BM, Munoz DF, Kurian AW, Plevritis SK. A simulation model to predict the impact of prophylactic surgery and screening on the life expectancy of BRCA1 and BRCA2 mutation carriers. *Cancer Epidemiol Biomarkers Prev* 2012;21:1066-1077.
- [Sil95] Silverstein MJ, Poller DN, Waisman JR, *et al.* Prognostic classification of breast ductal carcinoma-in-situ. *Lancet* 1995;345:1154-1157.
- [Sil00] Silverstein MJ. The Van Nuys Breast Center: the first free-standing multidisciplinary breast center. *Surg Oncol Clin North Am* 2000;9:159-175.
- [Spr93] Spratt JS, Spratt JA, Fournier D, *et al.* Decelerating growth and human breast cancer. *Cancer* 1993;71:2013-2019.
- [Spr93b] Spratt JA, von Fournier D, Spratt JS, Weber EE. Mammographic assessment of human breast cancer growth and duration. *Cancer* 1993;71:2020-2026.
- [Sur06] Surveillance, Epidemiology and End Results Program SEER\*StatDatabase (incide for 1973-2002). Disponible en: <http://www.seer.cancer.gov>. (Acceso 25 de Abril de 2006).
- [Sze96] Szeto KL, Devlin NJ. The cost-effectiveness of mammography screening: evidence from a microsimulation model for New Zealand. *Health Policy* 1996;38:101-115.
- [Tab81] Tabar L, Gad A. Screening for breast cancer: the Swedish trial. *Radiology* 1981;138:219-222.
- [Tab92] Tabár L, Fagerberg G, Day NE, *et al.* Breast cancer treatment and natural history: new insights from results of screening. *Lancet* 1992;339:412-414.
- [Tab92b] Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update in the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;30:187-210.
- [Tab95] Tabar L, Fagerberg G, Chen HH, *et al.* Efficacy of breast cancer screening by age: new results from the Swedish Two-county trial. *Cancer* 1995;75:2507-2517.
- [Tab95b] Tabar L, Fagerberg G, Chen HS, *et al.* Efficacy of breast cancer screening by age: new results from Swedish two-county trial. *Cancer* 1995;75:2507-2517.

- [Tab96] Tabar L. Breast cancer screening with mammography in women aged 40-49 years. *Int J Cancer* 1996;68:693-699.
- [Tab00] Tabar L, Vitak B, Chen HH, *et al.* The Swedish Two-County Trial 20 years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* 2000;38:625-651.
- [Tab00b] Tabar L, Dean PB, Duffy SW, Chen HH. A new era in the diagnosis of breast cancer. *Surg Oncol Clin North Am* 2000;9:233-277.
- [Tag13] Taghipour S, Banjevic D, Miller AB, *et al.* Parameter estimates for invasive breast cancer progression in the Canadian National Breast screening study. *Br J Cancer* 2013a;108:542-548.
- [Tag17] Taghipour S, Caudrelier LN, Miller AB, Harvey B. Using simulation to model and validate invasive breast cancer progression in women in the study and control groups of the Canadian National Breast Screening Studies I and II. *Med Decis Making* 2017;37:212-223.
- [Tor86] Torrance GE. Measurement of health state utilities for economic appraisal: A review. *Journal of Health Economics* 1986;5:130.
- [Tsu15] Tsunematsu M, Kakehashi M. An analysis of mass screening strategies using a mathematical model: comparison of breast cancer screening in Japan and United States. *J Epidemiol* 2015;25:162-171.
- [Tub84] Koscielny S, Tubiana M, Le MG, *et al.* Breast cancer: relationship between the size of the primary tumor and the probability of metastatic dissemination. *Br J Cancer* 1984;49:709-715.
- [Tub90] Tubiana M, Koscielny S. The natural history of human breast cancer: implications for screening strategy. *Int J Radiat Oncol Biol Phys* 1990;19:1117-1120.
- [Tub91] Tubiana M, Koscielny S. Natural history of human breast cancer: recent data and clinical implications. *Breast Cancer Res Treat* 1991;18:125-140.
- [Tub99] Tubiana M, Koscielny S. The rationale for early diagnosis of cancer-the example of breast cancer. *Acta Oncol* 1999;38:295-303.
- [Uhr10] Uhry Z, Hédelin G, Colonna M, *et al.* Multi-state Markov models in cancer screening evaluation: a brief review and case study. *Stat Methods Med Res* 2010;19:463-86.
- [UKP12] Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;380:1778-1786.
- [Ver84] Veerbek ALM, Hendriks JHCL, Holland R, Mravunac M, Sturmans F, Day NE. Reduction of cancer mortality through mass screening with modern mammography: First results of the Nijmegen project 1987-1981. *Lancet* 1984;I:1222-1224.
- [Vil09] Vilapriño E, Rue M, Marcos-Gragera R, Martínez Alonso M. Estimacion of age- and stage-specific Catalan breast cancer survival function using US and Catalan survival data. *BMC Cancer* 2009;9:99-111.

- [Wee08] Weedon-Fekjaer H, Lindqvist BH, Vatten LJ, Aalen OO, Tretli S. Breast cancer tumor growth estimated through mammography screening data. *Breast Cancer Res* 2008;10.
- [Woo10] Woolf SH. The 2009 Breast Cancer Screening Recommendations of the US Preventive Services task Force 2. *JAMA* 2010;303.
- [Wu18] Wu WY, Nystrom L, Jonsson H. Estimation of overdiagnosis in breast cancer screening using a non-homogeneous multi-state model: A simulation study. *J Med Screen* 2018;25:183-190.
- [Zac04] Zackrisson S, Andersson I, Manjer J, Janzon L. Non-attendance in breast cancer screening is associated with unfavourable socio-economic circumstances and advanced carcinoma. *Int J Cancer* 2004;108:754-760.
- [Zam12] Zamora LI, Forastero C, Guirado D, Martínez-Luna RJ, Lallena AM. A Monte Carlo tool to study the mortality reduction due to breast screening programs. *Med Phys* 2012;39.
- [Zam16] Zamora LI, Forastero C, Guirado D, Lallena AM. A Monte Carlo analysis of breast screening randomized trials. *Phys Med* 2016;32:1609-1614.



# Índice de figuras

2.1. Histograma que representa la distribución de edad de las mujeres españolas en el año 2007. . . . .	40
2.2. Modelo de crecimiento de Spratt ajustado para el seguimiento de 448 tumores con un valor de $b$ promedio de $0.0032 \text{ días}^{-1}$ . . . . .	42
2.3. Comparación de las curvas de tiempo de duplicación (TD) de nuestro modelo con el de Anerlöv. . . . .	43
2.4. Curva de incidencia de cáncer de mama de la población canaria (1996) . . .	44
2.5. Distribución de los tipos de densidad de la mama según la clasificación BI-RADS. . . . .	46
2.6. Curva de detección propuesta y distribuciones de tumores subyacentes y detectados en función del tamaño para reproducir los datos INC. . . . .	50
2.7. Resultado de la comparación de la proporción de tumores detectados en función del tamaño entre nuestras simulaciones con 100 millones de tumores y los resultados proporcionados por INC [Ros06]. . . . .	51
2.8. Representación esquemática la estructura de las simulaciones realizadas para el estudio de la reducción de mortalidad debida al cribado por cáncer de mama. . . . .	56
2.9. Gráfico que representa el proceso de corrección del nacimiento uniforme en el momento de evaluación de la mortalidad. . . . .	57
2.10. Pesos relativos aplicados a la distribución de nacimiento uniforme. . . . .	58
2.11. Valores de la curva de adelanto obtenidos por simulación. . . . .	59
2.12. Representación de los ajustes para diferentes supervivencias en función de la mediana del tamaño tumoral. . . . .	63
2.13. Esquemas de la estructura de simulación para el estudio de la reducción de mortalidad en los programas de cribado reales y los EAC. . . . .	66
3.1. Tasa detección anual (tumores detectados/mujeres sometidas al programa cada año) del programa de cribado para tiempos de intervalo de 1, 2 y 3 años con mujeres entre (a) 50 y 70 años, (b) 40 y 70 años y (c) 50 y 65 años. 72	

3.2. Resultados del análisis de sensibilidad del modelo. . . . .	78
3.3. Número total de tumores acumulados (cribado más intervalo) en un programa de cribado de mujeres para tiempos de intervalo de 1, 2 y 3 años para (a) mujeres entre 50 y 70 años y (b) mujeres entre 40 y 70 años. . . . .	81
3.4. Número total de tumores acumulados (cribado más intervalo) en un programa de cribado de mujeres para tiempos de intervalo de 1, 2 y 3 años para (a) mujeres entre 50 y 70 años y (b) mujeres entre 40 y 70 años. . . . .	81
3.5. Tumores acumulados invasivos (cribado+intervalo) para los grupos 50-70 años y 40-70 años con tiempo de intervalo 1 año. . . . .	84
3.6. Distribución de adelanto de detección simulada para el caso de rango de edades 40-70 y tiempo de intervalo 1 año. . . . .	88
3.7. Efecto del tiempo de intervalo en el RR de mortalidad por cáncer de mama. . . . .	90
3.8. Comparación de RR respecto al de la configuración 40-70 para los mismos tiempo de intervalo. . . . .	91
3.9. Dias efectivos ganados por mujer invitada a cribado mamográfico en función del RR para las diferentes configuraciones estudiadas. . . . .	92
3.10. Tendencia de mortalidad anual por cáncer de mama en Canada. . . . .	97
3.11. Comparación de los resultados de los ensayos reales frente a las simulaciones MC. . . . .	98
3.12. Distribución de tumores invasivos detectados en el grupo de control del ensayo real Canada I. . . . .	99

# Índice de tablas

1.1. Primera parte de la tabla resumen de códigos numéricos de cribado mamográfico. . . . .	35
1.2. Segunda parte de la tabla resumen de códigos numéricos de cribado mamográfico. . . . .	36
2.1. Distribución del parámetro $b$ . . . . .	41
2.2. Valores de mortalidad por causas diferentes al cáncer de mama correspondiente al año 2006 proporcionada por el INE. . . . .	45
2.3. Probabilidad de mantener el tipo de mama según la clasificación BI-RADS al final de cada década. . . . .	47
2.4. Probabilidad de detección en función del tipo histológico del tumor y de la densidad de la mama. . . . .	47
2.5. Tipos histológicos de cáncer de mama y frecuencia de aparición. . . . .	48
2.6. Distribución de tumores suyacente obtenida de las simulaciones para reproducir los datos de tumores detectados en función del tamaño del INC[Ros06].	49
2.7. Probabilidad de mantener el tipo <i>in situ</i> cuando el tamaño del tumor sobrepasa 6, 11, 16 y 20 mm. . . . .	49
2.8. Valores medios del tiempo de adelanto resultado de un ajuste log-normal de las curvas obtenidas en nuestras simulaciones. . . . .	59
2.9. Resultado del ajuste de los datos al modelo exponencial dado por la ecuación (2.37) para 2.5, 5, 7.5, 10, 12.5 y 15 años de supervivencia. . . . .	62
2.10. Configuración de los diferentes ensayos aleatorios usada en las simulaciones.	67
3.1. Resumen de los resultados de las simulaciones para las tres configuraciones de edad estudiadas (50-70, 40-70 y 50-65 años) con tiempos de intervalo 1, 2 y 3 años y el programa inglés NHSBSP. . . . .	74
3.2. <i>Items</i> modificados para el análisis de sensibilidad del modelo. . . . .	75
3.3. Primera parte del análisis de sensibilidad del modelo. . . . .	76

3.4. Segunda parte del análisis de sensibilidad del modelo. . . . .	77
3.5. Valores de sobrediagnóstico tal y como se define en el texto para las dos configuraciones estudiadas y los tiempos de intervalo considerados en cada caso. . . . .	83
3.6. Resultados de la estimación de la sensibilidad global para diferentes configuraciones de los programas de cribado. . . . .	86
3.7. Valores de sensibilidad global estimada a partir de revisiones sistemáticas de publicaciones con diferentes seguimientos de los ensayos de mayor validez externa. . . . .	87
3.8. Resultados de la reducción de mortalidad por cáncer de mama para las tres configuraciones estudiadas con tiempos de intervalo de 1, 2 y 3 años. Se han calculado para diferentes porcentajes de participación en el cribado. . . . .	89
3.9. Porcentaje de participación mínimo para cada configuración simulada que proporciona una ganancia de 1 día efectivo por mujer invitada al cribado. . . . .	93
3.10. Resultados de las simulaciones de los principales ensayos aleatorios controlados. . . . .	96

# Publicaciones, comunicaciones y citas destacadas

En el proceso de elaboración de esta tesis se han realizado publicaciones de algunos de los temas que la componen en revistas de reconocido prestigio internacional en el ámbito de la Física Médica y la Medicina. Además, se han llevado a cabo ponencias y comunicaciones en congresos nacionales e internacionales en forma oral o en póster. Finalmente, señalaremos algunas citas que alguno de los trabajos que componen esta tesis han recibido, bien por la relevancia de la cita o por su nivel de difusión en medios de comunicación o en la propia Internet.

## Publicaciones

- Forastero C, Zamora LI, Guirado D, Lallena AM. A Monte Carlo Tool to simulate breast cancer screening programmes. *Phys Med Biol* 2010;55:5213-5229.
- Zamora LI, Forastero C, Guirado D, Martínez-Luna RJ, Lallena AM. A Monte Carlo tool to study the mortality reduction due to breast screening programs. *Med Phys* 2012;39.
- Zamora LI, Forastero C, Guirado D, Lallena AM. A Monte Carlo analysis of breast screening randomized trials. *Phys Med* 2016;32:1609-1614.
- Forastero C, Zamora LI, Guirado D, Lallena AM. Evaluation of the overdiagnosis in breast screening programmes using a Monte Carlo simulation tool: a study of the influence of the parameter defining the programme configuration. *BMJ Open* 2019.

## Ponencias y Comunicaciones en congresos

- Zamora LI. Riesgos al paciente en Radiodigánóstico. Apartado de programas de cribado mamográfico. Ponencia en el XVII Congreso de la Sociedad Española de Física Médica (SEFM). Alicante, España. 2-5 de junio de 2009.
- Forastero C, Zamora LI, Guirado D, Lallena AM. Optimización de un programa de cribado mamográfico. Comunicación oral en el XVII Congreso de la Sociedad Española de Física Médica (SEFM). Alicante, España. 2-5 de junio de 2009.
- Forastero C, Zamora LI, Guirado D, Lallena AM. Optimization of screening programmes for breast cancer. Póster en el Medical Physics and biomedical engineering World Congress (IFMBE). Munich, Alemania. 7-12 de septiembre de 2009.
- Zamora LI, Forastero C, Guirado D, Lallena AM. Simulación de la reducción de mortalidad por cáncer de mama en programas de cribado mamográfico. Póster en el II Congreso conjunto de la Sociedad Española de Física Médica y la Sociedad Española de Protección Radiológica (SEFM-SEPR). Sevilla, España. 10-13 mayo de 2011.
- Zamora LI, Forastero C, Guirado D, Lallena AM. Estudio de los principales ensayos aleatorios controlados de cribado mamográfico mediante simulación Monte Carlo. Comunicación oral en el IV Congreso conjunto de la Sociedad Española de Física Médica y la Sociedad Española de Protección Radiológica (SEFM-SEPR). Valencia, España. 23-26 de junio de 2015.
- Forastero C, Zamora LI, Guirado D, Lallena AM. Monte Carlo simulation of breast screening programmes. International Conference on Monte Carlo Techniques for Medical Applications (MCMA2017). Napoli, Italia (2017).
- Forastero C, Zamora LI, Guirado D, Lallena AM. Estudio del sobrediagnóstico en cribado mamográfico mediante simulación Monte Carlo. Comunicación oral en el V Congreso conjunto de la Sociedad Española de Física Médica y la Sociedad Española de Protección Radiológica (SEFM-SEPR). Girona, España. 13-16 de junio de 2017.

## Otras citas destacadas

En este apartado señalamos tres citas de relevancia, bien mediática, pues el asunto que nos ocupa es de interés público, o estrictamente científica. En este sentido destacaremos:

1. Charles Day Editor On-line de *Physics Today*, escribió un editorial titulado: *Monte Carlo, Colloids, and Public Health*. Computing in Science Engineering. 2013;15:88. Donde explica: “*My latest brush with Monte Carlo happened last week. Looking for research to write about, I came across a paper by Luis Zamora and his colleagues entitled A Monte Carlo Tool to Study the Mortality Reduction Due to Breast Screening Programs. Screening for breast cancer is difficult and controversial. Its difficult because the principal method, x-ray mammography, cannot by itself determine whether a lesion is malignant. Because of that limitation, follow-up biopsies are essential, but most lesions roughly 4 in 5 turn out to be benign. Controversy surrounds the question of when to start screening. Not only is the disease harder to detect in young women, its also less prevalent. Definitive evidence in favor of screening women aged between 40 and 49 years is lacking. Yet doctors who treat individuals, not populations are reluctant to tell patients under 49 that they dont need a mammogram yet. Why take even a small risk.? The tool that Zamora and his colleagues have built simulates the fate of a population of women who enter a screening program. You can adjust the programs age range and participation rate. Clinically derived parameters, such as the probability of detecting a tumor, are incorporated into the tool. Zamora and his colleagues present their results in graphs and tables, which are hard to summarize in a short column. They predict, for example, that breast cancer mortality can be reduced by 29 percent if 100 percent of women aged 50-70 are screened every two years. But they did discover what appears to be a critical parameter. For a screening program to be effective, its participation rate must be at least 50 percent. In the US, where 16.3 percent of the population lacks health insurance, that target is unfortunately ambitious.*” . <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6478760>.
2. La entrada que tiene por título *Monte Carlo Method* de la WIKIPEDIA en su versión en inglés cita como ejemplo de aplicación del método Monte Carlo nuestro artículo *A Monte Carlo Tool to simulate breast cancer screening programmes* en el apartado de aplicaciones en Biología Computacional, definida como: “*involves the development and application of data-analytical and theoretical methods, mathematical modeling and computational simulation techniques to the study of biological, ecological, behavioral, and social systems [...]*”.
3. La Universidad de Granada en su página WEB oficial publicó la siguiente noticia: “*Científicos españoles diseñan una herramienta informática que permite estudiar la eficiencia de los programas de screening mamográfico.*” Donde se explica: “*El uso de esta herramienta permitiría al sistema sanitario mejorar la eficiencia de los fondos públicos que en la actualidad se invierten en estos programas para el diagnóstico precoz del cáncer de mama En el estudio, publicado en la revista Medical Physics de la American Association of Physicists in Medicine, participa un investigador de la Universidad de Granada[...]*”. El 10 de junio de 2013. [http://secretariageneral.ugr.es/pages/tablon/\\*/noticias-canal-ugr/cientificos-espanoles-disenan-una-herramienta-informatica-que-permite-estudiar-la-eficiencia-de-los-programas-de-screening-mamografico.XC5XdjBKjIV](http://secretariageneral.ugr.es/pages/tablon/*/noticias-canal-ugr/cientificos-espanoles-disenan-una-herramienta-informatica-que-permite-estudiar-la-eficiencia-de-los-programas-de-screening-mamografico.XC5XdjBKjIV).