



Universidad de Oviedo

Programa de Doctorado "Biomedicina y Oncología Molecular"

**Efectividad y eficiencia en los programas poblacionales de cribado
de cáncer de mama**

*Effectiveness and efficiency in population-based breast cancer screening
programs*

TESIS DOCTORAL

Carmen Natal Ramos

Oviedo, 2017



Universidad de Oviedo

Programa de Doctorado "Biomedicina y Oncología Molecular"

**Efectividad y eficiencia en los programas poblacionales de cribado
de cáncer de mama**

*Effectiveness and efficiency in population-based breast cancer screening
programs*

TESIS DOCTORAL

Autora:

Carmen Natal Ramos

Directores

Adonina Tardón García

Xavier Castells Oliveres

Tutora

Carmen Rodríguez Sánchez



RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español/Otro Idioma: Efectividad y eficiencia en los programas poblacionales de cribado de cáncer de mama	Inglés: Effectiveness and efficiency in population-based breast cancer screening programs
2.- Autor	
Nombre: Carmen Natal Ramos	DNI/Pasaporte/NIE:
Programa de Doctorado: Biomedicina y Oncología Molecular	
Órgano responsable: Instituto Universitario de Oncología	

RESUMEN (en español)

Antecedentes y Objetivos

La complejidad creciente de la atención sanitaria, los constantes cambios en la tecnología, la limitación de los recursos y las incertidumbres sobre el valor de las distintas opciones se traducen tanto en dificultades para el abandono de prácticas difundidas que han resultado obsoletas, como para la expansión de tecnologías efectivas, generando en ambos casos problemas de efectividad y de eficiencia.

Nuestro objetivo fue evaluar algunos aspectos relacionados con el uso de las tecnologías sanitarias que podrían determinar la efectividad clínica y la eficiencia en los programas poblacionales de detección precoz de cáncer de mama.

Métodos

Se analizó la incidencia de cáncer de mama en una cohorte de mujeres de la población diana del programa de detección precoz de cáncer de mama de Asturias, en relación con su participación en el mismo. El seguimiento fue de diez años.

Se ha estudiado la incorporación de tecnologías en los programas poblacionales de cribado en una cohorte de mujeres de participantes en programas de cribado poblacional de cáncer de mama de tres comunidades autónomas del Estado español (Asturias, Cantabria y Cataluña) mediante un estudio de tendencias, y un estudio de variabilidad en la práctica clínica mediante el método de análisis de áreas pequeñas.

Resultados

Los riesgos relativos de diagnóstico de cáncer invasivo, in situ y total de las participantes en el programa respecto a las no participantes fueron respectivamente 1,16 (0,94-1,43) , 2,98 (1,16-7,62) y 1,22 (0,99-1,49). El riesgo relativo de participantes frente a no participantes para diagnóstico en estadio I fue de 2,47 (1,55-3,96); para T1 de 2,58 (1,67-3,99) y para afectación ganglionar negativa de 2,11 (1,38-3,23).

La tasa de realización de biopsias cerradas pasó entre los años 2000 y 2008 de 3 por mil a 5,9 por mil. La tasa de realización de cualquier tipo de prueba invasiva pasó entre 1996 y 2011 de 28,1 por mil a 7,9 por mil. Para esta cohorte, la tasa de detección de lesiones benignas se redujo de 23,9 por mil en 1996 a 3,8 por mil en 2011, y la tasa de detección de cáncer invasivo no se vio afectada, estabilizándose en 3,2 por mil.

La variación entre unidades de cribado, medida mediante las razones de

variación entre los percentiles ²⁵⁻⁷⁵ en las tasas de realización de pruebas diagnósticas oscilaron entre 1,68 para la tasa de recitaciones y 3,39 para la realización de citologías y las de las tasas de detección de lesiones benignas, *carcinoma ductal in situ* y cáncer invasivo fueron respectivamente de 2,79, 1,99 y 1,36.

Conclusiones

La participación en el programa de detección precoz de cáncer de mama de Asturias ofrece una mayor probabilidad de diagnóstico en estadio precoz sin suponer un incremento del riesgo de diagnóstico de cáncer de mama invasivo o de todos los diagnósticos de cáncer de mama, aunque incrementa el riesgo de diagnóstico de *carcinoma ductal in situ*.

La introducción de la biopsia con aguja gruesa como prueba de confirmación diagnóstica supone una mejora en la eficiencia de los programas estudiados, ya que reduce las tasas de realización de pruebas invasivas y de detección de lesiones benignas, mientras que las tasas de detección de carcinoma no se modifican.

La variación en las tasas de realización de pruebas invasivas para el diagnóstico de cáncer de mama entre unidades de cribado, en la cohorte estudiada, es alta y no se relaciona con los resultados finales del proceso, que tienen una variabilidad menor. Esto supone que para producir los mismos resultados el consumo de recursos es diferente entre las diferentes unidades de cribado.

Descriptorios

Evaluación en salud; Planes y Programas de Salud; Neoplasias de la mama.

Background and Objectives

The increasing complexity of health care, the constant changes in technology, resource constraints and uncertainties about the value of different options translate both difficulty abandoning widespread practices that have proved obsolete, as for expansion effective technologies, generating in both cases problems of effectiveness and efficiency.

Our aim was to evaluate some aspects in technologies utilization that could determine the clinical effectiveness and efficiency in response to resource consumption in population-based screening breast cancer. programs

Methods

The incidence of breast cancer is analyzed in a cohort of women in the target population of the screening breast cancer program in Asturias Follow-up was ten years.

The incorporation of technologies is analyzed in population screening programs in a cohort of women participating in population screening programs for breast cancer in three Community of Spain, through a study of trends and a study of variability in clinical practice by small areas analysis.

Results

Relative Risks diagnosis of invasive cancer, in situ and total in women that participate in screening program compared to no participants were respectively 1.16 (0.94 to 1.43), 2.98 (1.16 to 7.62) and 1.22 (0.99 to 1.49). The relative risk of participants versus nonparticipants for stage I was 2.47 (1.55 to 3.96); T1 of 2.58 (1.67 to 3.99) and 2.11 negative (1.38 to 3.23) lymph node involvement.

The rate of perform closed biopsies increased between 2000 and 2008 of 3 ‰ to 5.9 ‰. The rate of any invasive test decreased between 1996 and 2011 from 28.1 ‰ to 7.9 ‰., while detection rates of benign lesions decreased from 23.9 ‰ in 1996 to 3.8 ‰ in 2011, and the rate of invasive cancer detection stabilized at 3.2 ‰.

The reasons for variation among the 25-75 percentiles in rates of diagnostic tests for screening units ranged from 1.68 to 3.39 for recitations and fine-needle aspiration cytology and detection rates of benign lesions, *ductal carcinoma in situ* and invasive cancer were respectively 2.79, 1.99 and 1.36.

Conclusions

Participation in the program of early detection of breast Asturias cancer does not increase the risk of diagnosis of invasive breast or all diagnoses of breast cancer, even if it increases the risk of diagnosis of *ductal carcinoma in situ* and it provides a higher probability of diagnosis in early stages.

The introduction of core biopsy as diagnostic confirmation test proof a reduction of testing rates and a decrease in detection rates of benign lesions, whereas carcinoma detection rates are not changed.

The variation in the performance of invasive tests rates for the diagnosis of breast cancer among screening units included in the study is high and is not related to the final results of the process, with less variability. Therefore, the resource utilization between screening units is different, to produce equal results.

Descriptors

Health Evaluation; Health Programs and Plans; Breast Neoplasm.

Il progresso! Sempre tardi arriva.

Cinema Paradiso. Giuseppe Tornatore

*Whenever I am reminded of death - and it happens every day - I think of my own,
and this makes me try to work harder.*

John Berger. A fortunate Man

AGRADECIMIENTOS

A lo largo de estos años de trabajo en los servicios sanitarios he contado con el inestimable apoyo y ayuda de muchas personas a las que deseo expresar mi agradecimiento.

A las personas a través de las cuales tuve mi primer contacto con el cribado de cáncer de mama, en 1993: el programa que se realizaba para la población del Área Sanitaria IV, en el Hospital Monte Naranco, Beatriz Vidal y todo el equipo de Radiodiagnóstico, que continúan siendo motor del programa en Asturias. A Martín Caicoya, que introdujo el cribado de cáncer de mama en Asturias, que despertó mi interés por la epidemiología y gracias a quien encontré una profesión.

A las compañeras y compañeros del Programa de detección precoz de cáncer de mama de Asturias en los Servicios Centrales: Miguel Prieto, Elena Blázquez y María Luisa García y a todos los profesionales de las unidades de cribado; gracias a todos por su aportación y su buen hacer.

A Elvira Muslera, impulsora de la evaluación de tecnologías sanitarias, compañera y amiga, con quien siempre es un placer compartir un nuevo proyecto.

A Adonina Tardón, esta tesis me ha dado la oportunidad de conocerla de cerca y de disfrutar aprendiendo a su lado.

A Ana Fernández-Somoano, su apoyo y aliento han sido fundamentales para llegar aquí.

Una parte de esta tesis es fruto del proyecto compartido del grupo BELE. Quiero agradecer a todo el grupo, especialmente a Xavier Castells, Marta Román, Isabel Torá y su generosidad e iniciativa.

A las mujeres que durante los últimos 26 años han acudido a los diferentes programas de detección precoz de cáncer de mama, ellas son el motivo para mejorar permanentemente nuestro trabajo.

Y finalmente dedico este trabajo a Na, que con paciencia enseñó a leer y animó a querer aprender a las criaturas de mi familia durante tres generaciones.

FINANCIACIÓN

La realización de esta tesis doctoral ha contado con diferentes ayudas.

- Fondo de Investigaciones Sanitarias (FIS-ISCIII) FEDER en la convocatoria del año 2011 Instituto de Salud Carlos III FEDER (PI:11/01296), para financiar el proyecto de investigación “Determinantes, evolución y características biológicas del cáncer de mama detectado en una cohorte de mujeres cribadas”. El referido estudio comparte el registro de mujeres, fuente de información de los estudios realizados para el desarrollo de los objetivos 2, 3, 4 y 5.
- Ayudas para financiación de una publicación concedidas por el Instituto Universitario de Oncología del Principado de Asturias (IUOPA), Universidad de Oviedo, y Consorcio Español para la Investigación en Epidemiología y Salud Pública (CIBERESP).

PRESENTACION

La memoria de tesis doctoral que se presenta está incluida en la línea de investigación Estudio experimental de estrategias para el tratamiento y prevención de los procesos neoplásicos, del Programa de Doctorado Biomedicina y Oncología Molecular de la Universidad de Oviedo. Consta de un resumen y los siguientes capítulos: introducción que finaliza con la justificación del estudio; Hipótesis y objetivos; Material y métodos, Resultados, Discusión y Conclusiones.

Con frecuencia se necesitan muchos años para determinar si una intervención sanitaria (una prueba diagnóstica o un tratamiento) es eficaz; es preciso realizar estudios costosos y en ocasiones de alta complejidad logística. Por ello, parece importante reducir el periodo que transcurre desde que una intervención ha probado su eficacia y ha sido autorizada para su uso hasta que este uso se generaliza en la población diana, atendiendo a criterios de eficiencia y equidad.

En este trabajo se abordan algunos aspectos de interés relacionados con como las tecnologías sanitarias se introducen en la práctica clínica y que impacto tiene este proceso en la efectividad y en la eficiencia aplicados a la Prevención secundaria del cáncer de mama.

El primer objetivo aborda la evaluación de la efectividad del Programa de detección precoz de cáncer de mama de Asturias. en términos de incidencia de cáncer de mama. Los siguientes objetivos se centran en cómo se incorporan las tecnologías sanitarias en los programas poblacionales de cribado de cáncer de mama en España, analizando las tendencias en la incorporación de las biopsias cerradas a los mismos, y los efectos que puede tener en sus resultados finales, así como en la variabilidad entre unidades de cribado en relación con el proceso diagnóstico de cáncer.

Mi actividad profesional ha estado vinculada, desde 1993, al Programa de detección precoz de cáncer de mama en el Principado de Asturias, durante los años 2005 a 2012 he sido la responsable del mismo y en esta función he colaborado con otros

centros nacionales e internacionales, incluida la Red de programas de cribado de cáncer de España. Esta red supone el punto de encuentro de los responsables de los programas de cribado de cáncer de las Comunidades Autónomas desde el año 2000, y su objetivo principal es el intercambio de experiencias entre gestores de Programas Poblacionales de Detección Precoz de Cáncer.

También desde el inicio de mi actividad profesional, en el Servicio de Epidemiología Clínica y Medicina Preventiva del Hospital Monte Naranco de Oviedo, trabajé en Evaluación de Servicios Sanitarios, centrando la actividad en este periodo a los servicios hospitalarios. A partir del año 2002, con mi incorporación a la Unidad de Análisis y Programas del Servicio de Salud del Principado de Asturias, en el ámbito de la Evaluación de Servicios Sanitarios centré mis actividades en el estudio de como las innovaciones tecnológicas se incorporan a la práctica clínica, y el papel que supone la difusión del conocimiento en el proceso de introducción de tecnologías sanitarias.

Estos dos campos profesionales me han permitido colaborar en diversos proyectos de investigación y desarrollar esta tesis doctoral.

INDICE GENERAL

ABREVIATURAS Y SIMBOLOS.....	3
INDICE DE TABLAS Y FIGURAS.....	4
1. INTRODUCCION.....	7
1.1. Objetivos de la atención Sanitaria.....	7
1.2. Las intervenciones de prevención secundaria de enfermedades en los servicios sanitarios.....	11
1.3. El cáncer de mama.....	15
1.3.1. Epidemiología del cáncer de mama.....	15
1.3.2. Los programas de detección precoz de cáncer de mama.....	17
1.4. Justificación del estudio.....	23
2. HIPÓTESIS Y OBJETIVOS.....	25
2.1. Hipótesis.....	25
2.2. Objetivos.....	26
3. MATERIAL Y METODOS.....	27
4. RESULTADOS.....	33
5. DISCUSION.....	49
6. CONCLUSIONES.....	55
7. BIBLIOGRAFIA.....	89

ABREVIATURAS Y SIMBOLOS

ANOVA: Análisis de la varianza

AVALIA-T: Axencia Galega de Avaliación de Tecnoloxías Sanitarias

BIRADS: Breast Imaging Reporting and Data System

CI. Carcinoma invasivo

CDIS: Carcinoma ductal in situ

CSV: Coeficiente sistemático de la variación

ETS: evaluación de tecnologías sanitarias

HER2: Receptor 2 de factor de crecimiento epidérmico humano

IARC: International Association of Cancer Registries

INAHTA: *International Network of Agencies for Health Technology Assessment*

OTA: *Office of Technology Assessment*

PCA: Porcentaje de cambio anual

RV: Razón de variación

VPM: Variaciones en la práctica médica

INDICE DE TABLAS Y FIGURAS

TABLAS

Tabla 1. Criterios de cribado (1968 y 2008).

Tabla 2. Beneficios y riesgos del cribado (Holland, 2006).

Tabla 3. Posibles resultados de una prueba de detección precoz de cáncer.

Tabla 4. Población de estudio. Incidencia de cáncer de mama en Asturias 1999-2010.

Tabla 5. Incidencia de cáncer invasivo e in situ entre mujeres participantes y no participantes ajustada por edad. Asturias 1999-2010.

Tabla 6. Evolución al diagnóstico de los carcinomas invasivos según participación en el Programa de detección precoz de cáncer de mama. Asturias 1999-2010.

Tabla 7. Tendencias en incidencia de cáncer de mama in situ, invasivo y total en relación con la participación en el Programa de detección precoz de cáncer de mama. Asturias 1999-2010.

Tabla 8. Estudio de tendencias. Descripción de los estudios realizados y lesiones detectadas en la población estudiada. Valores absolutos y tasas por 1000 estudios de cribado. (Asturias, Cantabria y Cataluña. 1996-2011).

Tabla 9. Estudio de tendencias. Tasas ajustadas por edad de pruebas diagnósticas y lesiones detectadas por 1000 estudios de cribado (Asturias, Cantabria y Cataluña. 1996-2011).

Tabla 10. Estudio de tendencias. OR de las tasas de recitación y pruebas invasivas por tipo de cribado y cuatrienios. (Asturias, Cantabria y Cataluña. 1996-2011).

Tabla 11. Estudio de tendencias. OR de las tasas de realización de distintas pruebas invasivas por tipo de cribado y cuatrienios. (Asturias, Cantabria y Cataluña. 1996-2011).

Tabla 12. Análisis de variabilidad. Descripción de la población de estudio: mujeres participantes, pruebas realizadas y resultados. (Asturias, Cantabria y Cataluña. 2002-2011).

Tabla 13. Variabilidad entre unidades de cribado poblacional de cáncer de mama en las tasas de realización de pruebas de confirmación diagnóstica (x 1000) basado en la estandarización directa. (Asturias, Cantabria y Cataluña. 2002-2011).

Tabla 14. Variabilidad entre unidades de cribado poblacional de cáncer de mama en las tasas de resultados del episodio de cribado (x 1000) basado en la estandarización directa. (Asturias, Cantabria y Cataluña. 2002-2011).

FIGURAS

Figura 1: Tendencias en la incidencia acumulada de cáncer invasivo e in situ en Asturias en relación con la participación en el Programa de detección precoz de cáncer de mama (1999-2010).

Figura 2: Tendencias en la realización de pruebas invasivas y detección de lesiones por tipo de cribado (Asturias, Cantabria y Cataluña. 1996-2011).

Figura 3. Variaciones en las pruebas de confirmación diagnóstica por unidades de cribado. (Asturias, Cantabria y Cataluña. 2002-2011).

Figura 4: Variaciones en las tasas de detección de lesiones por unidades de cribado. (Asturias, Cantabria y Cataluña. 2002-2011).

Figura 5. Relación entre las tasas de realización de pruebas invasivas y las tasas de detección por unidades de cribado. (Asturias, Cantabria y Cataluña. 2002-2011).

1. INTRODUCCION

1.1. Objetivos de la atención sanitaria

La atención sanitaria tiene como objetivos mejorar la salud de la población, de un modo sostenible y garantizando que los servicios llegan a la población que los necesita. En este sentido se han desarrollado diferentes herramientas de evaluación que se utilizan desde variadas perspectivas: la calidad, el impacto, el seguimiento. Todas ellas comparten una serie de dimensiones como son accesibilidad, aceptabilidad, adecuación, capacidad resolutive, efectividad, eficiencia, equidad, longitudinalidad, orientación al paciente, rapidez o seguridad que es necesario evaluar. La evaluación de las diferentes dimensiones de la atención sanitaria para cada proceso, por más pequeño que parezca, nos indicará que mejoras pueden ser aplicables a la organización, gestión, formación y coordinación, para, en definitiva, acortar distancias entre lo eficaz, lo efectivo y lo eficiente.

Las primeras definiciones de eficacia y eficiencia fueron realizadas por la *Office of Technology Assessment* (OTA) en 1978 (OTA, 1978), aunque las definiciones de mayor difusión, son las que establece la *International Network of Agencies for Health Technology Assessment* (INAHTA), cuya versión española ha sido elaborada por por la Axencia Galega de Avaluación de Tecnoloxias Sanitarias (AVALIA-T) (INAHTA, 2006)

La eficacia se define como el beneficio que supone utilizar una tecnología para un determinado problema en condiciones idóneas, por ejemplo, un entorno de laboratorio, dentro del protocolo de un ensayo controlado aleatorio rigurosamente gestionado, o en un “centro de excelencia”.

La efectividad es el grado en que una determinada intervención, procedimiento, régimen o servicio tiene el efecto previsto en circunstancias normales, en lugar de en condiciones controladas. O, más específicamente, evaluación de la razón riesgo/beneficio de una intervención, en un entorno clínico estándar, utilizando

resultados que midan cuestiones importantes para los pacientes (por ejemplo capacidad para realizar las actividades diarias o prolongación de la vida). La efectividad, por tanto está relacionada con la utilización adecuada de las intervenciones, pruebas y tratamientos que han probado su eficacia.

La eficiencia se define como la obtención del máximo beneficio posible a partir de los recursos disponibles. En este sentido relaciona los beneficios de las intervenciones con el esfuerzo realizado para obtener dichos beneficios. La eficiencia, por tanto está relacionada con la cantidad de recursos empleados en la realización de aquellas intervenciones que se consideran efectivas.

Si bien los estudios relacionados con la eficacia podría considerarse que tienen validez universal, el trabajo para alcanzar la efectividad y su evaluación son necesariamente locales y es necesario establecer metodologías claras, que se expliciten de antemano, que ayuden al acercamiento entre lo que resulta eficaz y lo que verdaderamente hacemos cuando esa intervención se pone en práctica. (Ortún, 2006).

Una aproximación a la medida de la eficiencia es el análisis de las variaciones en la práctica médica (VPM). Se han definido como las variaciones sistemáticas en las tasas estandarizadas de un procedimiento médico o quirúrgico a un determinado nivel de agregación de la población. Generalmente, los estudios de VPM utilizan diseños ecológicos y su principal herramienta metodológica es el análisis de áreas pequeñas; intentan relacionar el número de residentes en cada área geográfica que ha recibido un determinado servicio sanitario (por ejemplo, una prueba diagnóstica, una derivación, un tratamiento, una intervención quirúrgica, etc.) en un período de tiempo definido, con la población total de tales áreas en el mismo período. Las tasas de incidencia acumulada obtenidas, generalmente estandarizadas por edad y sexo, se utilizan para valorar si la variabilidad entre áreas implica una diferente utilización de los servicios estudiados, así como la magnitud y relevancia de estas diferencias.

El concepto de VPM incluye también otros aspectos de la variabilidad en la utilización de servicios sanitarios, que no son poblacionales sino de base individual. Estos

estudios también analizan la variabilidad en el uso de servicios, pero en pacientes en situaciones clínicas similares, no en áreas geográficas. A diferencia de los anteriores, se desarrollan sobre una base individual (no poblacional) y las herramientas metodológicas esenciales son las utilizadas para el ajuste de riesgos (y no el análisis de áreas pequeñas). El objetivo suele ser evaluar la efectividad o la eficiencia de tecnologías o de proveedores de servicios sanitarios (centros o profesionales), o buscar determinantes de variabilidad en función de características de los pacientes (como sexo, grupo étnico, nivel socioeconómico), del médico (especialidad, edad, sexo, formación, experiencia, sistema de pago), del hospital (público o privado, rural o urbano, universitario o no, docencia, tamaño, terciarismo) o del sistema sanitario (financiación, organización, cobertura u otras). Aunque ambos tipos de estudios de "variaciones", que simplificando podemos llamar poblacionales e individuales, tienen relaciones entre sí, y ambos muestran que los médicos actúan de forma diferente ante situaciones similares quebrando la creencia tradicional de que los profesionales sanitarios aplican de manera uniforme un tratamiento inequívocamente adecuado ante cada problema de salud, las implicaciones para la política sanitaria y la práctica clínica, así como los abordajes posibles de estos dos tipos de VPM, son sustancialmente diferentes.

Las decisiones sobre el uso de pruebas diagnósticas es compleja, entre las razones comúnmente citadas que influyen en el uso de una determinada tecnología se encuentran: la experiencia personal (Parker, 2013), el efecto del líder (Wright FC, 2011), o la protección frente a los problemas de mala práctica (Elmore, 2005). Todas ellas está asociadas con la hipótesis de los profesionales entusiastas (Chassin, 1993 y/o con la incertidumbre de los profesionales (Wennberg JE, 1982).

Las propuestas básicas de la hipótesis de la incertidumbre pueden resumirse en tres puntos: Las diferencias en morbilidad y otras variables de la población no explican sustancialmente las VPM entre áreas vecinas con estructuras de población similar; la variabilidad es mínima cuando existe acuerdo entre los clínicos sobre el valor de un procedimiento y finalmente, sólo en presencia de incertidumbre (o por ignorancia) los clínicos desarrollan estilos de práctica diferentes, que serían la principal fuente de VPM, tratamientos inadecuados y gasto sanitario innecesario.

Por otra parte está descrito que el simple desarrollo de recomendaciones no es por sí mismo un motivo suficiente para cambiar las prácticas (Caldow J, 2006). La implementación de buenas prácticas implica identificar las barreras y adoptar medidas que permitan superarlas, en la forma de utilizar los recursos desde el punto de vista organizativo (cambio de la organización) y de conductas individuales (cambio cultural).

La complejidad creciente de la atención sanitaria, los constantes cambios en la tecnología, la limitación de los recursos y las incertidumbres sobre el valor de las distintas opciones se traducen tanto en dificultades para el abandono de prácticas difundidas que han resultado obsoletas, como para la expansión de tecnologías efectivas, generando en ambos casos problemas tanto de efectividad como de eficiencia. Algunos factores que pueden influir en el proceso son la propia tecnología, la investigación en servicios sanitarios, la difusión del conocimiento y el análisis de los estilos de práctica (conocimiento desigual y concepto desigual de la ética médica).

El ritmo de difusión de los cambios tecnológicos tiene importantes implicaciones en términos de equidad. Los cambios tecnológicos determinan cambios en la estructura, en los procesos organizativos y en los estilos de práctica, que a su vez condicionan los cambios tecnológicos.

El proceso de desarrollo, adopción, difusión y abandono de tecnologías es heterogéneo y depende de un entramado de causas sociales, financieras y profesionales, tanto en el abandono de tecnologías ineficaces, como el uso de la radiografía de tórax preoperatoria en pacientes de bajo riesgo o el uso de la litotricia para el tratamiento de los cálculos biliares) como en la lentitud con la que algunas tecnologías de alta efectividad y bajo coste tardan en alcanzar el estado de práctica dominante en la profesión, como el uso de la fibrinólisis temprana en el tratamiento del infarto agudo de miocardio o la prevención de complicaciones en el parto prematuro (Lau, 1992; Chalmers, 1989).

La evaluación de tecnologías sanitarias (ETS) es un instrumento basado en la evidencia científica para la toma de decisiones sobre la inclusión de avances

tecnológicos en los sistemas sanitarios. Ahora bien, hay múltiples condicionantes y elementos, provenientes tanto de la propia ETS como de profesionales, políticos y de la sociedad. Es un proceso de análisis e investigación, dirigido a estimar el valor y contribución relativos de cada tecnología sanitaria a la mejora de la salud individual y colectiva, teniendo además en cuenta su impacto económico y social.

1.2. Las intervenciones de prevención secundaria de enfermedades en los servicios sanitarios

La prevención secundaria se basa en la intervención precoz sobre las enfermedades mediante un diagnóstico y un tratamiento precoz, con el fin de mejorar el pronóstico de las personas afectadas en las fases iniciales de una enfermedad. Se aplican procedimientos de diagnóstico precoz para descubrir a las personas enfermas que, de ordinario, son asintomáticos.

El cribado es una tecnología diagnóstica, de prevención secundaria, que se basa en la utilización de pruebas que permitan clasificar a las personas objeto de la intervención como afectadas o al menos sospechosas de estarlo y no afectadas. Las pruebas utilizadas deben ser válidas, reproducibles y aceptables, además de especialmente seguras.

Si bien, las consecuencias de cualquier intervención sanitaria pueden ser beneficiosas o perjudiciales, ninguna intervención sanitaria es inocua y todas ellas pueden originar efectos adversos, probablemente la prevención secundaria deba considerarse con mayor cautela, ya que al dirigirse a poblaciones sanas a las que se invita a participar, los efectos secundarios que podrían a parecer son a cambio de ningún beneficio para estas personas (Muir Gray, 2008).

Los criterios clásicos que debe reunir una enfermedad para que pueda ser incluida en un programa de cribado fueron desarrollados por Wilson y Jungner en 1968 (Wilson, 1968). Aunque estos principios siguen aceptándose cuarenta años después, en el año 2003, la Organización Mundial de la Salud actualizó estos criterios incluyendo aspectos de evaluación de la viabilidad, la efectividad y la idoneidad de

los programas de cribado (Andermann, 2008). En la Tabla 1 se presentan los criterios de Wilson y Jungner y la última actualización, de Andermann.

Tabla 1. Criterios de cribado (1968 y 2008)

Wilson y Jungner 1968	Andermann, 2008
La condición ha de ser un importante problema de salud	El programa de cribado debe responder a una necesidad de salud reconocida.
Debe existir un tratamiento aceptado para los pacientes con la enfermedad	Los objetivos del programa deben estar inicialmente definidos
Los recursos para el diagnóstico y tratamiento deben estar disponibles	Debe haber una población diana definida
Debe existir una fase latente o estadio preclínico	Debe disponerse de una buena evidencia científica de la efectividad del programa
Debe existir un test adecuado para examinar el órgano	El programa debe integrar aspectos educativos, prueba de cribado, servicios diagnósticos y de gestión.
El test ha de ser aceptado por la población	
La historia natural de la condición, incluyendo el desarrollo desde fase latente hasta enfermedad declarada, debe ser entendida	Debe disponerse de sistema de aseguramiento de la calidad, con mecanismos que minimicen los riesgos potenciales del cribado.
Debe existir una política acordada de a quién tratar como pacientes	El programa debe asegurar las decisiones informadas, confidencialidad y respeto a la autonomía.
El coste de encontrar un caso (incluyendo el diagnóstico y tratamiento de pacientes diagnosticados) debe ser económicamente balanceado con relación al tratamiento total	El programa debe promover la equidad y accesibilidad al programa para toda la población diana
El cribado ha de ser un proceso continuo y no de una sola vuelta.	La evaluación del programa debe estar planificada desde el inicio del mismo Los beneficios del programa deben sobrepasar los riesgos

Sin embargo, los cribados, como cualquier intervención sanitaria, tienen riesgos y beneficios que deben ser considerados. Los beneficios parecen claros: la precocidad en el diagnóstico y la intervención mejorará el pronóstico en algunas personas enfermas. Los inconvenientes son más complejos: se alargará la duración de la enfermedad en las personas diagnosticadas cuando no haya modificaciones del pronóstico; puede haber un sobretreatmento para las formas más leves de la enfermedad; las personas enfermas con resultados negativos en la prueba serán

identificadas falsamente como sanas; las pruebas realizadas tienen riesgos asociados; otro inconveniente destacado es que las personas sanas con resultados positivos en la prueba de cribado pueden experimentar preocupación y ansiedad hasta el diagnóstico definitivo, además de ser sometidas a pruebas que no precisan. Net Calonge (1ª Conferencia de Prevención y Promoción de la Salud en la Práctica Clínica en España. Madrid: Ministerio de Sanidad y Consumo, 2007), señala que al aplicar una prueba de cribado pueden pasar cinco cosas y que cuatro de ellas no son buenas. Y esto sin contar los efectos adversos atribuibles a la prueba misma del cribado que, en general, acostumbra ser muy segura. Puede darse un resultado falso negativo, que genera una tranquilidad injustificada y favorece el retraso del diagnóstico, y un falso positivo, lo que comporta una confirmación innecesaria y potencialmente peligrosa y a menudo provoca el llamado efecto etiqueta (Brian, 1978). También puede ocurrir que la detección no conlleve ningún beneficio relevante y sólo se distraigan recursos necesarios para otras intervenciones más pertinentes. Pero la consecuencia más temible es el sobrediagnóstico: cuando el resultado es un verdadero positivo pero irrelevante clínicamente, de manera que el tratamiento precoz es innecesario. Es decir, se trata de alteraciones que no hubieran provocado problemas al paciente si hubieran seguido su curso natural, pero que al ser reveladas por el cribado exigen una intervención. Una lista de beneficios y riesgos del cribado se presenta en la Tabla 2.

Tabla 2. Beneficios y riesgos del cribado (Holland, 2006)

Beneficios	Riesgos y Desventajas
Mejora del pronóstico de los casos detectados	Costes añadidos
Tratamiento menos radical que cura los casos precoces	Mayor tiempo de morbilidad en casos que no mejoran su pronóstico
Ahorro de recursos	Sobret ratamiento de anomalías de pronóstico incierto
Mayor tranquilidad en casos con resultado negativo	Riesgo de efectos adversos por el proceso de cribado
	Falsa tranquilidad en los casos falsos negativos

Todos estos aspectos ponen de relevancia la dimensión ética de los cribados, sobre ella el Consejo Danés de Ética establece tres cuestiones clave en relación con los riesgos y beneficios de los cribados: La primera de ellas aborda en qué medida las pruebas de cribado suponen *medicalizar* a numerosas personas sanas o, por el contrario, son un buen servicio ofrecido para identificar a pocas personas enfermas. La segunda está en relación con asumir el riesgo de que las pruebas de cribado puedan tener falsos resultados, esto es, clasificar como enfermas a las personas sanas y clasificar como sanas a las personas enfermas. Esta circunstancia, que es común a cualquier prueba de diagnóstico, tiene mayor relevancia en las pruebas de cribado, cuya finalidad es clasificar a las personas que la realizan en dos grupos con diferentes probabilidades de tener la enfermedad buscada. Finalmente, la tercera cuestión está en relación con la equidad de los sistemas sanitarios, ya que considera que como se utilicen los programas de detección precoz indica como los sistemas sanitarios distribuyen y priorizan sus recursos (The Danish Council of Ethics, 2001).

Las intervenciones sobre las que actualmente hay acuerdos de que pueden detectar cánceres precozmente y ayudarnos a incrementar la salud y la calidad de vida son solamente tres: la realización de una mamografía cada dos años para las mujeres que tienen entre 50 y 69 años, para detectar precozmente los cánceres de mama, la realización de una citología del cuello del útero, test de Papanicolaou, cada tres años hasta los 65, para las mujeres que tiene relaciones sexuales, para detectar precozmente el cáncer de cuello de útero y, por último, la realización de detección de sangre oculta en las heces para la población entre 50 y 65 años para la detección precoz de cáncer de colon (EU, 2003).

1.3. El cáncer de mama

1.3.1. Epidemiología del cáncer de mama

El cáncer de mama es la neoplasia más frecuente entre las mujeres en el mundo occidental, representando entre el 20 y el 25 % de los tumores que se diagnostican en mujeres. Se estima que en el año 2015 1.790.861 mujeres han sido diagnosticadas de cáncer de mama, en el mundo (Globocan, 2012).

La incidencia de cáncer de mama en España se incrementó hasta el año 2001, esta tendencia, que se asoció a la difusión de los programas poblacionales de detección precoz, se modificó a partir de este en las mujeres menores de 65 años, este aumento y reducción de la incidencia se asocia al diagnóstico de los casos prevalentes al inicio de los programas de cribado poblacionales (Pollán, 2009).

En 2004 (último año del que se dispone de datos de incidencia publicados) se diagnosticaron en Asturias 592 casos, ninguno de ellos en menores de 20 años, 7,9 % en menores de 40 años, entre 40 y 49 años el 17,22 %, entre 50 y 59 años el 20,94 %, entre 60 u y 69 años el 18,92 %, entre 70 y 79 años el 23,81 % y mayores de 79 años el 26,35 %. La tasa ajustada por edad a la población europea fue para ese año de 75,82 por 100.000 mujeres.

La mortalidad también ha disminuido en los países de Europa occidental, incluido España (Desantis, 2015. Ferlay, 2013), debido a la detección precoz y a las mejoras en los procesos diagnósticos y terapéuticos.

En Asturias en el año 2013 fallecieron 187 mujeres por cáncer de mama. Es la décima causa de mortalidad en mujeres, representando el 2,9 % de todas las defunciones, y la segunda entre los 15 y 64 años. La tasa ajustada por edad a la población europea fue de 17,5 por 100.000 mujeres (Mortalidad en Asturias 2013).

Aunque la etiología del cáncer de mama sea multifactorial, desde el punto de vista de la prevención secundaria tiene interés el estudio de los factores de riesgo, que permitan identificar la población diana del diagnóstico precoz.

Sólo en aproximadamente un 5% de la enfermedad es posible identificar factores hereditarios (Wagner, 1999) y habitualmente en los casos diagnosticados no se identifica otro factor de riesgo conocido además de edad y sexo.

Como sucede con la mayoría de los cánceres, la edad es un factor de riesgo muy importante, incrementándose la incidencia con la edad, el 78% de todos los casos ocurre en mujeres de más de 50 años. La probabilidad de que una mujer desarrolle cáncer de mama entre los 40 y los 50 años es 1 entre 68, 1 entre 44 entre los 50 y los 60 y 1 entre 29 entre los 60 y 70 años (Howlader, 2016)

Las mujeres con cáncer de mama previo tienen un riesgo superior, estimado en un 1% anual, de desarrollar otro cáncer de mama. Las mujeres con lesiones de mama preinfiltrantes y displasias, en la cohorte BELE, tienen entre 1,5 y 2 veces más de riesgo (Castells, 2016).

Se han identificado dos genes cuya mutación se ha relacionado con una mayor probabilidad de desarrollar el cáncer de mama. Estos genes se denominan BRCA1 y BRCA2 y según algunos estudios entre el 50% y el 60% de mujeres que han heredado estos genes mutados pueden desarrollar el cáncer antes de los 70 años. Otros genes también relacionados con el cáncer de mama serían el P53 y BARD1, el BRCA3 y el Noey2. Si el cáncer es hereditario el riesgo aumenta a un 5% por año. Menos del 1% de las mujeres con cáncer de mama están asociados a síndromes genéticos tales como Síndrome de Cowden, Síndrome de Li-Fraumeni, Ataxia Telangiectasia, Síndrome de Bloom, Síndrome de Peutz-Jeghers y Síndrome de Werner. (Lindor, 1998). Se han

propuestas estrategias de prevención primaria para aquellas mujeres con alteraciones genéticas, entre las cuales se incluye el consejo genético.

Las mujeres con menarquia temprana o menopausia tardía tienen un riesgo mayor. Igual sucede con aquellas que no tuvieron hijos o los tuvieron después de los 30 años de edad. La terapia hormonal utilizada para aliviar los síntomas de la menopausia, parece aumentar a largo plazo (más de 10 años) el riesgo de sufrir cáncer de mama. Este riesgo es mayor a medida que aumenta la duración del tratamiento, pudiendo estar su diagnóstico dificultado por el aumento de la densidad mamaria que produce el tratamiento hormonal (Rossouw, 2002).

El abordaje terapéutico actual del cáncer de mama descansa en un análisis sólido de los factores pronósticos que orienten la decisión terapéutica. Los factores pronósticos conocidos son la evolución de la enfermedad al diagnóstico y las características biológicas del tumor.

La evolución de la enfermedad al diagnóstico se determina por el estadio, que incluye tamaño del tumor primario, afectación de los ganglios linfáticos y extensión de la diseminación tumoral. Las características biológicas incluyen el tipo histológico, el grado de diferenciación tumoral y marcadores moleculares.

1.3.2. Los programas de detección precoz de cáncer de mama

La detección precoz de cáncer de mama mediante mamografía es una práctica clínica generalizada en los países desarrollados. En la Unión Europea, en 2014 los 28 estados miembros tienen programas de detección precoz, en 22 casos de carácter poblacional y en 20 de los estados con una cobertura completa (Altobelli, 2014).

En España, en 1990 se inició el primer programa en Navarra. Actualmente la cobertura es del 100%. La población diana más frecuente son las mujeres de 50 a 69 años, aunque algunos programas se inician a los 45 años y la prueba de cribado es la mamografía bienal con doble proyección. La gestión depende de los gobiernos autonómicos, aunque se dispone de mecanismos de coordinación entre los diferentes programas y todos ellos tienen carácter poblacional (Ascunce, 2013).

El beneficio de los programas de detección precoz de cáncer de mama es la reducción de la mortalidad, evaluada en ocho ensayos clínicos y 20 revisiones sistemáticas realizadas en los últimos 20 años. Sus resultados permiten afirmar que el cribado bienal mediante mamografía en mujeres entre 50 y 69 años determina una reducción de la mortalidad entre un 20 y un 30 % según los resultados de un metanálisis publicado en 2004 (Smith, 2004). Revisiones sistemáticas posteriores ratifican estos resultados (Pace, 2014). En cuanto a la evaluación de la efectividad en programas concretos, los últimos resultados de Two Contry Trial y el programa ingles estiman la reducción de la mortalidad en 8,8 y 5.7 muertes evitadas respectivamente por cada 1000 mujeres que realizan cribado durante 20 años iniciándolo a los 50 años (Duffy, 2010).

La Tabla 3 presenta los posibles resultados de una prueba de detección precoz, un resultado corresponde al efecto deseado (diagnostico temprano), el resto son efectos adversos que es preciso controlar.

Los resultados de los estudios de efectos adversos son más controvertidos: los mejor estudiados son los falsos positivos, los falsos negativos y el sobrediagnóstico.

Estos inconvenientes no se deben a errores en la realización o interpretación de las pruebas, sino a que las pruebas disponibles son imperfectas, y en el

caso del sobrediagnóstico a la forma que tienen las enfermedades de evolucionar a lo largo del tiempo.

Tabla 3. Posibles resultados de una prueba de detección precoz de cáncer

	CANCER	NO CANCER
PRUEBA POSITIVA	Sobrediagnostico Diagnostico precoz (resultado buscado)	Falso positivo
PRUEBA NEGATIVA	Falso negativo	Ningún beneficio Riesgos de la prueba

El Cáncer de intervalo

Se define cáncer de intervalo como el tumor primario de mama que es diagnosticado después de haber realizado la mamografía de cribado y las pruebas de valoración correspondiente con resultado negativo para malignidad. Este diagnóstico se realiza antes de la siguiente invitación para cribado o en un periodo de tiempo igual al intervalo de cribado en mujeres que superan la edad límite para cribado (Perry, 2006).

Se trata de un grupo heterogéneo de tumores que tienen diferentes causas y que en relación a la última mamografía de cribado se clasifican radiológicamente como intervalo verdadero, signos mínimos, falso negativo y oculto. En caso de no tener acceso a la mamografía diagnóstica o a la última mamografía de cribado, se clasifica como no evaluable (Almazán, 2009). La incidencia del cáncer de intervalo está determinada por la historia natural de la enfermedad, la frecuencia y la sensibilidad de la prueba de cribado. El cáncer de intervalo es inevitable en un programa de cribado, pero su número debe mantenerse tan bajo como sea posible, ya que una proporción elevada de

cáncer de intervalo puede comprometer la efectividad del cribado y disminuir su potencial reducción de la mortalidad

Los resultados falsos positivos

El resultado falso positivo en una prueba de cribado es aquel en el que no se confirma el diagnóstico de cáncer. Es un efecto indeseable de los programas de cribado que debe ser medido, analizado e informado (Perry, 2006).

Los resultados falsos positivos en los programas de detección precoz de cáncer de mama oscilan entre el 10 y el 23 % en cada serie según los estudios (Ascunce, 1994; Nelson, 2009), siendo superiores siempre en las mujeres más jóvenes (Johns, 2010). Tienen efecto en la tasa de respuesta a las siguientes convocatorias (Román, 2011) y están determinados entre otros factores por las características de los profesionales (Alberdi, 2011).

El sobrediagnóstico

Es probablemente el efecto adverso más controvertido en los programas de cribado poblacional. Se define como el cáncer que se diagnostica en el cribado que no hubiera sido diagnosticado si la mujer no hubiera realizado cribado.

La estimación del sobrediagnóstico, por cada mil mujeres que siguen el cribado durante 10 años oscila entre 4,3 casos para el Two-County Trial y 2,3 casos para el programa inglés (Duffy, 2010). Otros autores, en revisiones secundarias, elevan la cifra de sobrediagnóstico hasta el 30 % (Gozstche, 2011).

Una medida indirecta del sobrediagnóstico es la tasa de detección de *carcinoma ductal in situ* (CDIS). Se trata de una proliferación celular maligna que se desarrolla en el interior de las estructuras de la mama que no afecta al estroma circundante. Aproximadamente el 50 % progresarían a cáncer invasivo

(CI) sin tratamiento y entre un 12 y un 20 % de los CDIS tratados con cirugía conservadora y radioterapia tendrán una recurrencia local en un intervalo de 10 años después del tratamiento (Thompson, 2008). Otros estudios de casos y controles sugieren que el curso natural algunos tumores *in situ* diagnosticados en el cribado podrían haber evolucionado a la regresión espontáneamente, basándose en que la incidencia acumulada de cáncer de mama invasivo siempre es menor en el grupo de mujeres que no participan en programas de cribado (Zahl, 2011).

Actualmente no se conocen bien los factores que determinan la progresión del CDIS a enfermedad invasiva, aunque se conoce que está en relación con la expresión génica del tumor. Identificar estos marcadores que identifiquen un mayor riesgo de progresión permitirá mejorar la selección de pacientes en relación con los tratamientos administrados y evitar sobretratamientos.

La detección de CDIS en los programas de cribado es un indicador de calidad de los mismos que debe ser monitorizado. Asimismo estos diagnósticos nos proporcionan información valiosa sobre la evolución de los mismos y las posibles alternativas de tratamiento y seguimiento.

Con estas evidencias los programas poblacionales de cribado de cáncer de mama basados en la mamografía bienal se recomiendan para mujeres entre 50 y 69 años, en países con suficientes recursos, por la Organización Mundial de la Salud (WHO, 2014), por la Unión Europea (EU, 2003).

Garantizar la efectividad de los programas de cribado poblacional de cáncer de mama es una tarea compleja, que requiere un gran desarrollo organizativo. El sistema sanitario debe disponer de información que le permita convocar a las personas destinatarias y evaluar los resultados, de esta forma se puede saber si hay dificultades para el acceso a las pruebas de carácter económico, social o cultural, para intentar evitarlas. La evaluación también permite conocer si algún

problema de calidad impide obtener los resultados esperados y corregir estos problemas. Esto sólo es posible mediante programas específicos organizados y dirigidos desde los servicios sanitarios. Los programas implantados deben tener un criterio poblacional, estar dotados de recursos necesarios y con indicadores calidad que permitan su evaluación.

El proceso de cribado poblacional de cualquier problema de salud incluye las siguientes etapas: identificación e invitación personalizada de toda la población elegible. Realización de las pruebas de cribado. Diagnóstico de las personas con alteraciones en las pruebas de cribado y tratamiento y seguimiento de las personas con lesiones confirmadas (Anttilla, 2015). En todas estas etapas deben cumplirse los criterios de garantía de calidad.

En los programas de cribado de cáncer de mama la uniformidad de criterios en la fase de cribado es alta, en cuanto a la población diana, prueba de cribado y periodicidad de la misma. Sin embargo en la fase de diagnóstico (realización de estudios intermedios y pruebas de confirmación) los resultados por comunidades autónomas informan de que la realización de pruebas adicionales se multiplica por 7 en la comunidad que más hace respecto a la que menos hace, por 3,5 la realización de pruebas invasivas y por 25 la realización de estudios intermedios (Ascunce, 2013).

1.4. Justificación del estudio

La finalidad del trabajo que aquí se presenta es el análisis orientado a la mejora de la efectividad y de la eficiencia de los programas de cribado de cáncer de mama, más allá de la evaluación sistemática de los mismos.

La disponibilidad de cohortes de mujeres cribadas con un seguimiento a largo plazo, con información individualizada sobre el proceso diagnóstico y sobre los resultados obtenidos en cada prueba realizada, permite una evaluación exhaustiva de los programas de cribado, de su impacto en la incidencia de cáncer de mama tanto invasivo como in situ, y del rendimiento de las diferentes pruebas diagnósticas, así como orientar nuevas estrategias para maximizar el beneficio del cribado, quizá con estrategias de prevención más personalizadas en un futuro.

La toma de decisiones para la incorporación de Tecnología Sanitaria en los Servicios de Salud, es a menudo un proceso complejo. Para conseguir que la incorporación de nuevas tecnologías a los sistemas sanitarios se base cada vez más en datos concernientes a su seguridad, efectividad y eficiencia, los canales de difusión de las recomendaciones formuladas por las guías clínicas deben ser lo más amplios, protocolizados y sistemáticos posible, implicando a los profesionales sanitarios, tanto clínico-asistenciales como a los encargados de la administración, planificación y evaluación.

Por estas razones resulta importante evaluar la práctica clínica en el cribado del cáncer de mama, poniendo en relación los resultados obtenidos con los recursos empleados, y así poder valorar el funcionamiento de los programas y de facilitar su mejora, proponiendo cambios en la organización que produzcan un balance riesgo-beneficio adecuado, una garantía de equidad y con un consumo de recursos que permita su sostenibilidad.

2. HIPOTESIS Y OBJETIVOS

2.2. HIPÓTESIS

1. Los programas de detección precoz de cáncer de mama son seguros desde el punto de vista del sobrediagnóstico, medido como un aumento de la incidencia debido a la participación en los mismos. El incremento de la incidencia que se produce en las primeras etapas, está relacionado con el adelantamiento diagnóstico y no se mantiene en el tiempo.
2. El proceso diagnóstico en los programas poblacionales de cribado de cáncer de mama ha mejorado su efectividad y eficiencia desde la incorporación de la biopsia con aguja gruesa como prueba de confirmación diagnóstica.
3. Hay variaciones, en la utilización de recursos, entre las unidades de cribado que dependen de los estilos de práctica médica, y que no necesariamente se traduce en mejores resultados de los programas.
4. Los modelos organizativos de los programas de detección precoz de cáncer de mama pueden determinar diferencias en la efectividad y eficiencia de los mismos.

2.3. OBJETIVOS

Objetivo general

Evaluar algunos aspectos relacionados con el uso de las tecnologías sanitarias que podrían determinar la efectividad clínica y la eficiencia en los programas poblacionales de detección precoz de cáncer de mama.

Objetivos específicos

1. Comparar la incidencia acumulada de cáncer de mama, su evolución temporal y el estadio al diagnóstico entre mujeres participantes y no participantes en un programa poblacional de detección precoz.
2. Determinar las tendencias en el uso de pruebas para el diagnóstico de las lesiones identificadas en la mamografía del cribado.
3. Determinar la relación entre las tasas de utilización de pruebas diagnósticas y las tasas de detección de lesiones benignas, carcinoma ductal in situ y cáncer invasivo.
4. Analizar las variaciones en el uso de pruebas diagnósticas durante el proceso de confirmación diagnóstica entre las unidades de cribado.
5. Evaluar si la variabilidad en el uso de pruebas se asocia a una variabilidad en los resultados.

3. MATERIAL Y METODOS

Objetivo 1

Se realizó un estudio de cohortes retrospectivo de incidencia de cáncer de mama en relación con la participación en el Programa de detección precoz de cáncer de mama de Asturias (Natal, 2015).

La población diana de este programa, en el periodo del estudio, era la de mujeres residentes en la Comunidad autónoma, con edades comprendidas entre 50 y 64 años, la prueba de cribado mamografía bienal con doble proyección en serie inicial y proyección única en series sucesivas. La realización de doble lectura fue irregular. La invitación, postal con oferta de cambio de cita y recordatorio a las no participantes.

La población de estudio fueron las mujeres incluidas en la población diana, en el año 1999 (año en que el programa alcanzó la cobertura poblacional completa), que corresponde a las nacidas antes del 1 de Enero de 1940. Se excluyó a aquellas que después del 31 de diciembre de 2005 seguían acudiendo al programa, por cualquier motivo. Este periodo se estableció con el fin de controlar el adelantamiento diagnóstico que se produce en los programas de detección precoz (Pacci, 2004; Puliti, 2012). Se definieron como participantes aquellas mujeres invitadas que asistieron al menos una vez, y como no participantes las que recibieron al menos una invitación pero no acudieron nunca o las que no recibieron ninguna invitación en el periodo.

El periodo de análisis comprendió los años 1999 a 2010.

El diagnóstico de cáncer de mama se obtuvo del Registro poblacional de tumores de Asturias, con fecha de diagnóstico anterior al 31 de diciembre de 2010. El Registro poblacional de tumores de Asturias es un registro de

base poblacional, que se inició en 1982, y su fuente de información son todos los hospitales de la Comunidad Autónoma.

Subsidiariamente se utilizó el Registro hospitalario de tumores del Servicio de Salud del Principado de Asturias, para completar la información sobre la invasión al diagnóstico. Este registro recoge los casos diagnosticados y/o tratados en los diferentes hospitales pertenecientes a la Red Hospitalaria de Utilización Pública de Asturias.

Se analizó la incidencia acumulada de cáncer de mama y la invasión al diagnóstico en ambos grupos (participantes y no participantes) a través de las siguientes variables: tipo de tumor (invasivo o in situ), tamaño de tumor, afectación ganglionar y estadio, esta última siguiendo los criterios de clasificación del American Joint Committee on Cancer (American Joint Committee on Cancer, 2002).

Se realizó un análisis de la tendencia de la incidencia acumulada entre participantes y no participantes de acuerdo al tipo de tumor (invasivo o in situ). Para estudiar la tendencia de la incidencia acumulada, y detectar la existencia de puntos de cambio durante el período, se han utilizado modelos segmentados de regresión de Poisson, usando el software específico que proporciona el Surveillance Research Program del Instituto Nacional de Cáncer de EE.UU. Proporciona, además, el porcentaje de cambio anual de las tasas en cada uno de los tramos definidos con un intervalo de confianza del 95%. Este porcentaje de cambio anual describe la proporción de incremento o decremento de las tasas por unidad de tiempo.

Objetivos 2, 3, 4 y 5

Para analizar las tendencias y la variabilidad en el uso de pruebas diagnósticas, y su efecto en los resultados de las mismas se realizaron un estudio retrospectivo de tendencias, para estudiar la dimensión temporal y

un estudio de variabilidad de la práctica clínica, para estudiar la dimensión geográfica (Natal, 2016).

La población de estudio se construyó sobre una cohorte de mujeres participantes en los programas de cribado de cáncer de mama en tres CCAA del Estado español (Asturias, Cantabria y Cataluña).

En cada episodio de cribado se recopila de rutina información relacionada con la interpretación mamográfica, con la recitación para una evaluación adicional, cuya finalidad es descartar o confirmar la malignidad, y con las evaluaciones adicionales llevadas a cabo en cada caso. Con esta información se construyó una base de datos única y anonimizada con datos individualizados para cada mujer que asistió al menos a una visita de cribado. El diagnóstico definitivo de cáncer de mama siempre se confirmó histopatológicamente utilizando los códigos de la Clasificación Internacional de Enfermedades para Oncología 10ª Ed. (CIE-10). Los casos de cáncer de mama fueron clasificados como invasivos o CDIS.

Las medidas utilizadas fueron respecto a las pruebas realizadas las tasas estandarizadas por edad, de recitación, de realización pruebas invasivas y de realización de cada tipo de prueba invasiva: punción-aspiración con aguja fina (PAAF), biopsia con aguja gruesa (BAG) y biopsia quirúrgica (BQ). Respecto a los diagnósticos se utilizaron las tasas de detección de lesiones benignas, CDIS y CI.

La tasa de recitación se definió como el número de citas para exploraciones adicionales después de una prueba de cribado, la tasa de realización de pruebas invasivas como el número de exploraciones adicionales invasivas, las tasas de realización de PAAF, BAG y BQ como el número de cada una de las pruebas realizadas, la tasa de detección de lesiones benignas, CDIS y CI como el número de diagnósticos de lesión benigna, CDIS o CI, todas ellas por cada mil episodios de cribado.

Estudio de Tendencias

Se incluyeron, para este estudio los datos de 542.695 mujeres que se habían sometido a al menos un estudio de cribado en el periodo 1996 a 2011. Estas mujeres realizaron un total de 1.763.810 estudios.

Se calcularon las tasas anuales estandarizadas por edad de pruebas diagnósticas y resultados de las mismas y se realizó un análisis de regresión logística incondicional de la tendencia de las tasas anuales para definir la dirección de la tendencia.

Para realizar el análisis de riesgo de realizarse una prueba invasiva a lo largo del periodo de estudio se agruparon las tasas por períodos de 4 años y se calcularon Odds Ratios de manera independiente según la serie inicial o sucesiva, ajustando por edad, antecedentes de lesiones benignas de mama previas y antecedentes familiares de cáncer. Se tomó como referencia el primer periodo (1996-1999).

Los intervalos de tiempo para el análisis se definieron de acuerdo a años de calendario del estudio de cribado (1996-2011) y se agruparon por periodos de cuatro años (1996-1999, 2000-2003, 2004-2007 y 2008-2011).

Todos los análisis estadísticos se realizaron bajo hipótesis bilaterales, tomando como estadísticamente significativo un valor de $p < 0,05$, excepto para OR de tendencias que se asumió un valor $< 0,001$

Análisis de variabilidad

Para el análisis de la variabilidad en el uso de pruebas de confirmación diagnóstica entre unidades de cribado, se realizó un estudio observacional de las variaciones de las tasas estandarizadas de pruebas diagnósticas y

de detección de lesiones en 34 unidades de cribado. El periodo de estudio fueron los años 2002 a 2011, ambos incluidos.

Las unidades de cribado incluidas representan la totalidad de los programas de las tres comunidades autónomas analizadas (Asturias, Cantabria y Cataluña) y el 25% de las unidades mamográficas fijas del Estado español (Red de Programas de cribado de cáncer, 2008). La unidad de cribado, que cuenta con recursos específicos, es el dispositivo sanitario que realiza las exploraciones propias del programa de detección precoz de cáncer de mama a una población diana definida.

Como medidas de resultado se utilizaron las tasas estandarizadas por edad, de recitación, de realización pruebas invasivas y de realización de cada tipo de prueba invasiva: PAAF, BAG y BQ. Se realizó un análisis para los subgrupos de BI-RADS 0, 3, 4 y 5 de las tasas de recitación y de realización de pruebas invasivas. Se excluyeron los BI-RADS 1 y 2, porque los estudios adicionales en estos tienen un carácter muy excepcional. No se realizó el análisis para los subgrupos de BI-RADS en cada tipo de prueba invasiva porque los grupos resultaron demasiado pequeños al desagregar. También se utilizaron las tasas estandarizadas por edad, de detección de lesiones benignas, detección de CDIS y detección de CI.

La estandarización por edad se realizó mediante el método directo utilizando como población de referencia la suma de las poblaciones de todas las unidades de cribado al inicio del estudio. Los estadísticos utilizados en el análisis de variabilidad fueron las razones de variación (RV) entre los percentiles 5-95 y 25-75, y el Componente Sistemático de la Variación (CSV), (Martínez -Lizaga, 2011).

Para analizar el efecto del programa en la variabilidad de las tasas de realización de las diferentes pruebas y en las tasas de detección de

lesiones se realizó un análisis de la varianza de un factor (ANOVA), considerando el programa la variable independiente.

Para la representación gráfica de la variación de las tasas analizadas se realizaron gráficos de puntos en escala logarítmica de media 0. Cada punto corresponde a una unidad de cribado. En el eje de las X se representó cada una de las pruebas analizadas y en el eje de las Y la desviación del logaritmo de la tasa estandarizada con respecto a la tasa media. Esta representación permite comparar en un mismo gráfico tasas con magnitudes muy diferentes.

Para medir la asociación entre las tasas de realización de pruebas invasivas y las tasas de detección se utilizó el coeficiente de regresión. Para la representación gráfica se elaboraron gráficos de burbujas que permiten incluir, además de la asociación entre las tasas analizadas, el volumen de actividad de cada unidad de cribado.

4. RESULTADOS

Objetivo 1

Se incluyeron en el estudio 14.863 mujeres de las cuales el 59,3% había participado al menos una vez en el Programa de detección precoz de cáncer de mama de Asturias. En el periodo de estudio se diagnosticaron 417 cánceres de mama, 388 de tipo invasivo y 29 CDIS. En el grupo de mujeres participantes se diagnosticaron 260 casos (238 CI y 22 CDIS) y en grupo de mujeres que no participaron se diagnosticaron 157 (150 CI y 7 CDIS). La distribución de la población por grupos de edad y de los diagnósticos de cáncer de mama en función de la asistencia al programa se presenta en la Tabla 4.

Tabla 4. Población de estudio. Incidencia de cáncer de mama en Asturias 1999-2010

	Participantes	No participantes	TOTAL
Mujeres	8812	6051	14863
Edad al inicio			
60-61	4215	2802	7017
62-63	3950	2051	6001
64-65	647	1198	1845
Cáncer diagnosticado			
Invasivo	238	150	382
CDIS	22	7	29
Total	260	157	417

En el grupo de mujeres participantes la incidencia acumulada para todos los tumores en el periodo de estudio fue de 29,5 casos por 1000 mujeres, para CI 27,0 y para CDIS 2,5 casos por 1000 mujeres. En el grupo de mujeres no participantes las incidencias acumuladas para los mismos tipos de cáncer fueron respectivamente 25,9, 24,8 y 1,2 casos por 1000 mujeres.

Los RR de diagnóstico de cáncer, invasivo, *in situ* y total, de las participantes en el programa respecto a las no participantes, ajustados por edad fueron

respectivamente 1,16, 2,98 y 1,22, resultando estadísticamente significativo para CDIS (Tabla 5).

Tabla 5. Incidencia de cáncer invasivo e in situ entre mujeres participantes y no participantes ajustada por edad. Asturias 1999-2010

	Participantes		No participantes		RR	IC 95%
	casos	Tasa x1000	casos	Tasa x 1000		
Invasivo	238	27,0	150	24,8	1,16	0,94-1,43
CDIS	22	2,5	7	1,2	2,98	1,16-7,62
totales	260	29,5	157	25,9	1,22	0,99-1,49

En relación con el grado de evolución de la enfermedad al diagnóstico, (Tabla 6) la probabilidad de ser diagnosticada de un carcinoma clasificado como estadio I para las mujeres participantes fue mayor que para las no participantes, (RR=2; IC 95%: 1,36-2,96), de un T1 (RR=1,9; IC 95%: 1,35-2,70) y de no tener afectación ganglionar (RR=1,7; IC 95%: 1,19-2,70).

El mayor crecimiento de la incidencia acumulada se produjo en el periodo 1999-2001, tanto para el grupo de mujeres participantes como para el de no participantes, siendo el porcentaje de cambio anual (PCA) en las mujeres participantes de 92,3% frente a 72,5% en las no participantes. En el periodo 2001-2006 la incidencia siguió aumentando más en el grupo de mujeres participantes y a partir de ese año el incremento anual en la incidencia es mayor en el grupo de mujeres no participantes (Tabla 7 y Figura 1).

Tabla 6: Evolución al diagnóstico de los carcinomas invasivos según participación en el Programa de detección precoz de cáncer de mama. Asturias 1999-2010

Estadio de los carcinomas invasivos según participación						
	Participantes		No participantes		RR	IC 95%
	casos	tasa	casos	tasa		
I	97	11,0	33	5,5	2,0	1,36-2,96
II	73	8,3	41	6,8	1,2	0,81-1,85
III	25	2,8	19	3,1	0,9	0,29-2,83
IV	3	0,3	5	0,8	0,4	0,05-3,20
Desconocido	40	4,5	52	8,6	0,5	0,34-0,80
Total	238	27,0	150	24,8	1,1	0,88-1,35

Tamaño del tumor de los carcinomas invasivos según participación						
	Participantes		No participantes		RR	IC 95%
	casos	tasa	casos	tasa		
T1	124	14,1	45	7,4	1,9	1,35-2,70
T2	54	6,1	36	5,9	1,0	0,15-7,10
T3	6	0,7	6	1,0	0,7	0,10-4,88
T4	5	0,6	2	0,3	2,0	0,01-306,4
Desconocido	49	5,6	61	10,1	0,6	0,38-0,81
Total	238	27,0	150	24,8	1,1	0,88-1,35

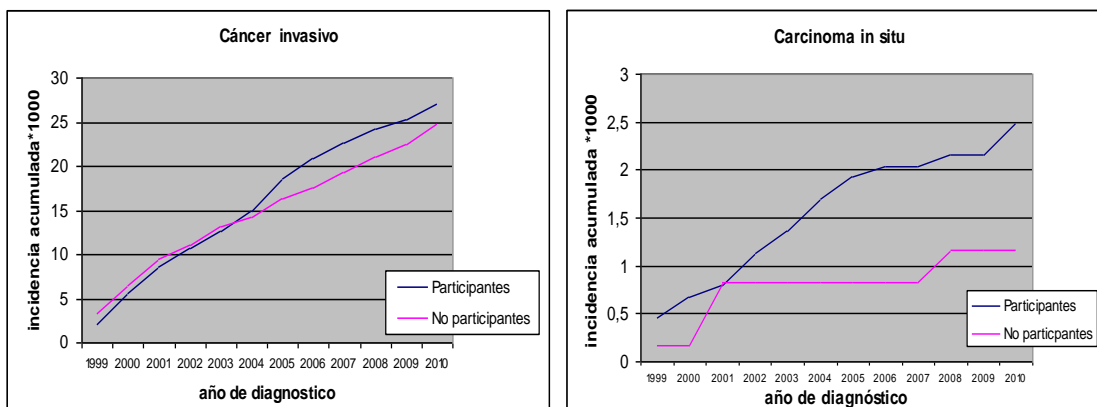
Afectación ganglionar de los carcinomas invasivos según participación						
	Participantes		No participantes		RR	IC 95%
	casos	tasa	casos	tasa		
N(-)	119	13,5	49	8,1	1,7	1,19-2,34
N(+)	69	7,8	41	6,8	1,2	0,75-1,16
Desconocido	50	5,7	60	9,9	0,6	0,40-0,84
Total	238	27,0	150	24,8	1,1	0,88-1,35

Tabla 7. Tendencias en incidencia de cáncer de mama in situ, invasivo y total en relación con la participación en el Programa de detección precoz de cáncer de mama. Asturias 1999-2010

Tipo	Participantes			No participantes		
	Periodo	PCA	IC 95%	Periodo	PCA	IC 95%
CDIS	1999-2004	30,1	24,7-35,6	1999-2001	131,3	64,5-225,4
	2004-2010	4,9	1,7-8,3	2001-2010	6,0	2,7-9,3
Cancer invasivo	1999-2001	101,9	89,2-115,4	1999-2001	69,07	54,5-85,0
	2001-2006	18,1	15,7-20,5	2001-2005	13,26	8,3-18,5
	2006-2010	6,3	4,1-8,5	2005-2010	8,59	6,4-10,8
Total	1999-2001	92,3	81,6-103,5	1991-2001	72,58	58,5-87,9
	2001-2006	18,2	16,1-20,3	2001-2005	12,55	7,9-17,4
	2006-2010	5,9	4,0-7,8	2005-2010	8,57	6,5-10,6

PCA: Porcentaje de cambio anual

Figura 1: Tendencias en la incidencia acumulada de cáncer invasivo e in situ en Asturias en relación con la participación en el Programa de detección precoz de cáncer de mama (1999-2010)



Objetivos 2, 3, 4 y 5

Estudio de tendencias

En el periodo comprendido entre 1996 y 2011 realizaron un nuevo examen 47 de cada 1000 mujeres que hicieron una prueba de cribado (tasa de recitación), y 11,6 de cada 1000 fueron sometidas a alguna prueba invasiva.

Entre las pruebas invasivas la más frecuentemente utilizada fue la PAAF con una tasa de 5,4 por 1000, seguida de BAG, con una tasa de 4,7 por 1000 y finalmente BQ, cuya tasa fue de 1,5 por 1000. Todas las pruebas invasivas se realizaron más frecuentemente en cribado inicial respecto al sucesivo.

Tabla 8. Estudio de tendencias. Descripción de los estudios realizados y lesiones detectadas en la población estudiada. Valores absolutos y tasas por 1000 estudios de cribado. (Asturias, Cantabria y Cataluña. 1996-2011)

	Total	Cribado inicial	Cribado sucesivo
Estudios de cribado	1.763.810	542.695	1.221.115
Mujeres participantes	542.695	542.695	426.826
Recitaciones	83.354 (47,3‰)	42.183 (77,7‰)	41.171 (33,7‰)
Pruebas invasivas	20,255 (11,6‰)	10.356 (19,1‰)	10.166 (8,3‰)
PAAF	9.532 (5,4‰)	5.477 (10,1‰)	4.055 (3,3‰)
BAG	8.268 (4,7‰)	3.451 (6,4‰)	4.817 (3,9‰)
BQ	2.722 (1,5‰)	1.428 (2,6‰)	1.294 (1,1‰)
Lesiones benignas	11.196 (6,3‰)	6.816 (12,6‰)	4.380 (3,6‰)
CDIS	1.197 (0,7‰)	493 (0,9‰)	704 (0,6‰)
Cáncer invasivo	5.726 (3,2‰)	2.122 (3,9‰)	3.605 (2,9‰)

PAAF: Punción-aspiración con aguja fina. BAG: biopsia con aguja gruesa. BQ: biopsia quirúrgica. CDIS: *Carcinoma ductal in situ*

En la Tabla 9 figuran las tasas anuales de recitación, de pruebas invasivas total y por tipos (PAAF, BAG, BQ), tasas de detección de lesiones benignas, CDIS y de CI. La tasa de recitación descendió entre 1996 y 2011 de 60,3‰ a 46,0 ‰, la tasa de realización de pruebas invasivas pasó de 28,1‰ a 7,9‰ en los mismos años y la tasa de realización de BAG, comenzó a incrementarse en el año 2000, a partir de 3,0‰, hasta estabilizarse en 2008 en 5,9‰.

En cuanto a los resultados de las pruebas, la tasa de detección de lesiones benignas pasó de 23,9‰ en 1996 a 3,8‰ in 2011, la de detección de CDIS se estabilizó a partir de 1997 en 0,7‰ y la de cáncer invasivo se fue incrementando hasta 1997 (4,5‰), experimentando una caída a 3,2‰ en el año siguiente y manteniéndose estable hasta el fin del estudio, En la Figura 2 se muestran estas tendencias,

En el análisis de efecto, la recitación tuvo una reducción del 16% para cribado inicial y del 9% para sucesivo, aunque en este último no resultó

estadísticamente significativa. En cuanto al total de pruebas invasivas se redujeron entre un 48 y un 40% respectivamente por tipo de cribado (Tabla 10).

Tabla 9. Estudio de tendencias. Tasas ajustadas por edad de pruebas diagnósticas y lesiones detectadas por 1000 estudios de cribado (Asturias, Cantabria y Cataluña. 1996-2011)

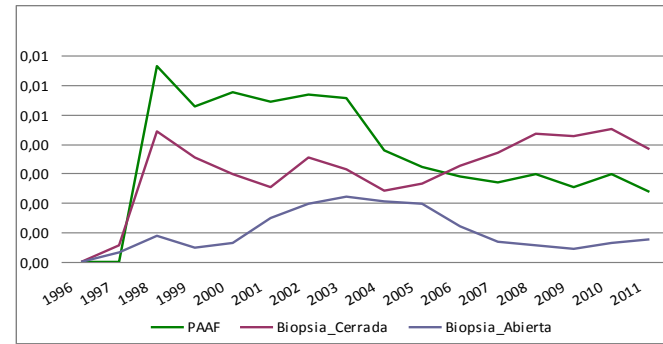
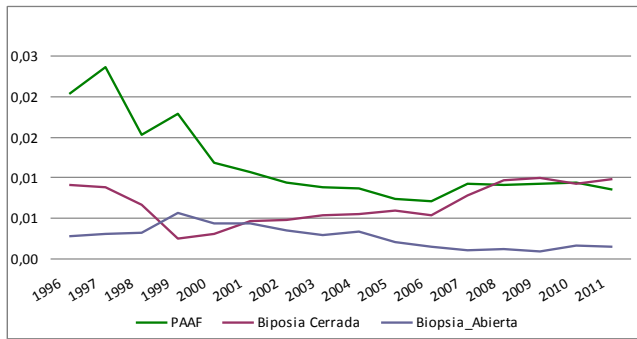
Año	Tasas de pruebas diagnósticas					Tasas de detección		
	Recitación	Pruebas invasivas	PAAF	BAG	BQ	Lesiones benignas	CDIS	Cáncer invasivo
1996	60,3 (55,5-65,1)	28,1 (24,4-31,8)	18,8 (15,8-21,8)	9,5 (7,3-11,7)	2,5 (1,5-3,5)	23,9 (20,6-27,4)	0,4 (0,1-0,7)	4,4 (2,9-5,9)
1997	62,3 (57,8-66,7)	28,4 (25,4-31,3)	20,3 (17,8-22,7)	7,8 (6,3-9,3)	3,1 (1,9-4,2)	25,9 (23,2-28,8)	0,7 (0,3-1,1)	4,5 (3,1-5,9)
1998	39,1 (36,5-41,6)	16,2 (14,5-17,9)	10,4 (9,1-11,8)	5,5 (4,5-6,5)	1,9 (1,3-2,5)	13,1 (11,5-14,6)	0,6 (0,2-0,9)	3,2 (2,4-3,9)
1999	60,5 (58,3-62,7)	17,2 (15,9-18,4)	12,2 (11,2-13,2)	3,1 (2,5-3,7)	3,7 (3,2-4,3)	12,9 (11,9-14,0)	0,6 (0,4-0,8)	3,5 (2,9-4,1)
2000	55,2 (53,4-56,9)	13,3 (12,5-14,2)	9,5 (8,7-10,2)	3,0 (2,6-3,4)	3,4 (2,9-3,9)	9,6 (8,8-10,3)	0,6 (0,4-0,8)	3,2 (2,7-3,6)
2001	46,0 (44,5-47,5)	12,1 (11,3-12,9)	7,8 (7,1-8,4)	3,6 (3,2-4,1)	2,9 (2,7-3,4)	8,3 (7,7-9,0)	0,6 (0,4-0,8)	3,3 (2,8-3,7)
2002	43,3 (42,1-44,6)	11,7 (11,0-12,4)	6,9 (6,4-7,5)	4,1 (3,7-4,5)	2,6 (2,3-2,9)	7,9 (7,3-8,4)	0,6 (0,4-0,8)	3,5 (3,2-3,9)
2003	43,8 (42,6-45,0)	11,2 (10,6-11,8)	6,7 (6,2-7,2)	4,1 (3,7-4,5)	2,5 (2,2-2,8)	7,6 (7,1-8,1)	0,7 (0,6-0,9)	3,1 (2,8-3,5)
2004	36,9 (35,9-38,1)	9,3 (8,8-9,8)	5,2 (4,8-5,7)	3,5 (3,2-3,9)	2,4 (2,2-2,7)	5,7 (5,2-6,1)	0,6 (0,5-0,7)	3,1 (2,8-3,4)
2005	41,8 (40,8-42,8)	8,9 (8,5-9,5)	4,5 (4,2-4,8)	3,7 (3,4-4,1)	2,0 (1,8-2,2)	5,4 (5,0-5,8)	0,7 (0,6-0,8)	2,9 (2,7-3,2)
2006	43,8 (42,8-44,8)	8,9 (8,4-9,3)	4,2 (3,9-4,6)	4,2 (3,9-4,5)	1,3 (1,1-1,5)	4,9 (4,6-5,3)	0,7 (0,6-0,8)	3,2 (2,9-3,5)
2007	45,8 (44,8-46,9)	8,9 (8,5-9,4)	4,1 (3,8-4,5)	4,9 (4,6-5,3)	0,8 (0,7-0,9)	5,2 (4,8-5,5)	0,8 (0,6-0,9)	3,1 (2,9-3,4)
2008	53,2 (52,1-54,2)	10,3 (9,8-10,8)	4,4 (4,1-4,7)	5,9 (5,6-6,3)	0,7 (0,6-0,8)	6,2 (5,8-6,5)	0,8 (0,6-0,9)	3,4 (3,2-3,7)
2009	51,4 (50,3-52,4)	9,1 (8,7-9,6)	3,9 (3,6-4,1)	5,7 (5,4-6,1)	0,6 (0,4-0,7)	5,2 (4,8-5,5)	0,6 (0,5-0,7)	3,3 (3,1-3,6)
2010	52,1 (51,1-53,2)	9,0 (8,6-9,4)	4,2 (3,9-4,5)	5,6 (5,3-5,9)	0,8 (0,7-0,9)	4,8 (4,5-5,1)	0,7 (0,6-0,9)	3,4 (3,2-3,7)
2011	46,0 (45,1-47,0)	7,9 (8,3-3,4)	3,4 (3,2-3,7)	4,9 (4,6-5,2)	0,9 (0,7-1,0)	3,8 (3,5-4,0)	0,6 (0,5-0,8)	3,1 (2,9-3,4)

PAAF: Punción-aspiración con aguja fina. BAG: biopsia con aguja gruesa. BQ: biopsia quirúrgica. CDIS: *Carcinoma ductal in situ*

Figura 2: Tendencias en la realización de pruebas invasivas y detección de lesiones por tipo de cribado (Asturias, Cantabria y Cataluña. 1996-2011)

Pruebas invasivas
Cribado inicial

Cribado sucesivo



Detección de lesiones
Cribado inicial

Cribado sucesivo

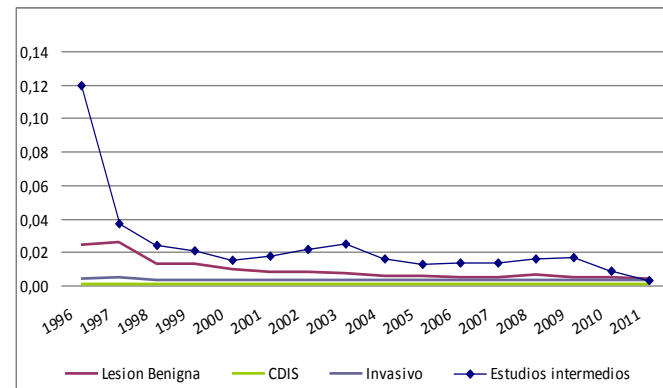
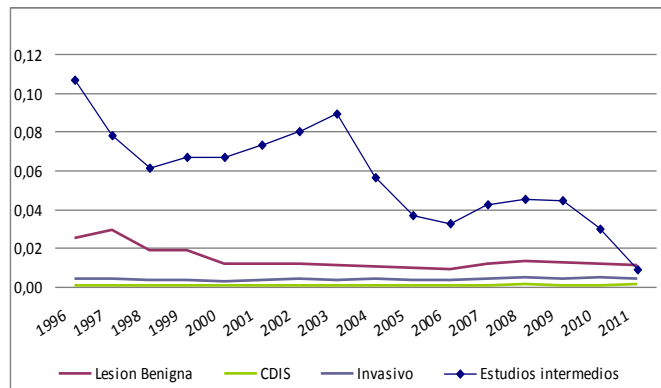


Tabla 10. Estudio de tendencias. OR de las tasas de recitación y pruebas invasivas por tipo de cribado y cuatrienios. (Asturias, Cantabria y Cataluña. 1996-2011)

	Tasa de recitación						Pruebas invasivas					
	Cribado inicial			Cribado sucesivo			Cribado inicial			Cribado sucesivo		
	OR	IC 95%		OR	IC 95%		OR	IC 95%		OR	IC 95%	
1996-1999	1,00			1,00			1,00			1,00		
2000-2003	0,86	0,82	0,89	1,26	1,13	1,40	0,56	0,52	0,60	1,08	0,92	1,27
2004-2007	0,82	0,78	0,86	1,10	0,99	1,23	0,51	0,47	0,55	0,65	0,55	0,76
2008-2011	0,84	0,80	0,88	0,91	0,82	1,01	0,52	0,47	0,56	0,60	0,51	0,70

p-trend<0,001

El uso de BAG se incrementó en un 85% en el periodo 2008-2011 para cribado inicial y en un 94% para cribado sucesivo, aunque su difusión fue lenta: en el cuatrienio 2004-2007 para cribado inicial sólo había aumentado un 26%. Finalmente el abandono del uso generalizado de la PAAF se observa a partir del cuatrienio 2004-2007, con OR de 0,39 para cribado inicial y 0,38 para cribado sucesivo. La reducción de la utilización de la BQ fue más precoz e intensa para cribado inicial, OR de 0,74, 0,41 y 0,14 para los diferentes cuatrienios y solo alcanzó resultados estadísticamente significativos para cribado sucesivo en el último cuatrienio, como puede verse en la tabla 11.

Tabla 11. Estudio de tendencias. OR de las tasas de realización de distintas pruebas invasivas por tipo de cribado y cuatrienios. (Asturias, Cantabria y Cataluña. 1996-2011)

	PAAF						BAG						BQ					
	Cribado inicial			Cribado sucesivo			Cribado inicial			Cribado sucesivo			Cribado inicial		Cribado sucesivo			
	OR	IC 95%		OR	IC 95%		OR	IC 95%		OR	IC 95%		OR	IC95%	OR	IC 95%		
1996-1999	1,00			1,00			1,00			1,00			1,00			1,00		
2000-2003	0,49	0,45	0,53	1,06	0,88	1,26	0,76	0,65	0,89	1,36	0,98	1,89	0,74	0,63	0,88	1,01	0,52	1,96
2004-2007	0,39	0,35	0,43	0,38	0,32	0,46	1,26	1,07	1,48	1,79	1,31	2,45	0,41	0,33	0,51	0,76	0,40	1,44
2008-2011	0,34	0,31	0,38	0,29	0,24	0,35	1,85	1,58	2,16	1,94	1,42	2,65	0,14	0,10	0,20	0,45	0,24	0,87

p-trend<0,001. PAAF: Punción-aspiración con aguja fina. BAG: biopsia con aguja gruesa. BQ: biopsia quirúrgica.

Análisis de variabilidad

Entre los años 2002 y 2011, a las 34 unidades de cribado estudiadas, han asistido 523.480 mujeres, con una participación media de 3 episodios de cribado. Fueron recitadas 46,3 mujeres de cada mil participantes, a 10,6 mujeres de cada mil participantes se les realizó una prueba invasiva. El 43% de las pruebas invasivas fueron PAAF, el 45% BAG y el 12% BQ. A 11,8 mujeres de cada mil participantes se les diagnosticó un CI, un CDIS a 2 de cada mil, una lesiones benignas a 15,7 de cada mil mujeres. Estos datos se presentan en la Tabla 12.

Tabla 12. Análisis de variabilidad. Descripción de la población de estudio: mujeres participantes, pruebas realizadas y resultados. (Asturias, Cantabria y Cataluña. 2002-2011)

UNIDADES DE CRIBADO	34	
MUJERES EXPLORADAS	523.480	
EPISODIOS	1.521.720	
RECITACIONES (por mil mujeres exploradas)	70.383	(46,3)
PRUEBAS INVASIVAS(por mil mujeres exploradas)	16.145	(10,6)
PAAF (%)	6.882	(42,6)
BAG (%)	7.289	(45,1)
BQ (%)	1.974	(12,2)
EPISODIOS CON LESIONES BENIGNAS (Tasa de detección ‰)	8.240	(15,7)
EPISODIOS CON CANCER DE MAMA(Tasa de detección ‰)	6.174	(11,8)
CDIS (Tasa de detección ‰)	1.050	(2,0)
Cáncer invasivo (Tasa de detección ‰)	4.930	(9,4)

PAAF: Punción-aspiración con aguja fina. BAG: biopsia con aguja gruesa. BQ: biopsia quirúrgica. CDIS: *Carcinoma ductal in situ*

La Tabla 13 muestra los estadísticos de variabilidad para tasa de recitaciones, tasa de realización de pruebas invasivas y tasas específicas de realización de PAAF, BAG y BQ. Se encontró una RV_{25-75} de 1,68 para la tasa de recitación y 1,81 para la realización de pruebas invasivas. En cuanto al tipo de prueba invasiva realizada las razones de variación de las unidades de cribado comprendidas entre los percentiles 25-75, en las tasas de realización de pruebas invasivas, oscilaron entre 1,99 para la BAG y 3,39 para la PAAF. Asimismo, la variación sistemática, no esperada por el azar, medida con el

CSV en la tasa de recitación, fue menor que en la tasa de realización de pruebas invasivas (13 y 23% respectivamente).

Tabla 13. Variabilidad entre unidades de cribado poblacional de cáncer de mama en las tasas de realización de pruebas de confirmación diagnóstica (x 1000) basado en la estandarización directa. (Asturias, Cantabria y Cataluña. 2002-2011)

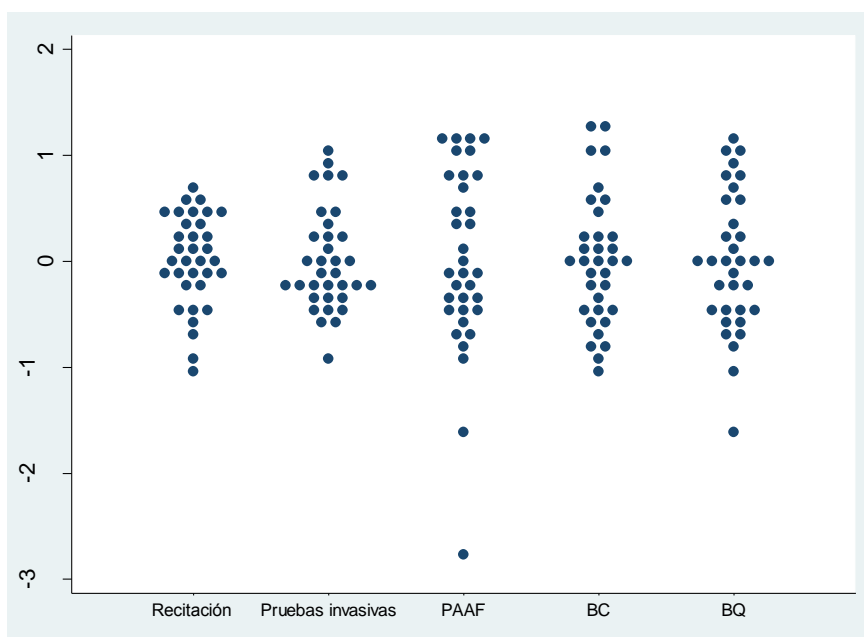
	Recitaciones	Pruebas invasivas	PAAF	BAG	BQ	
Episodios	70.383	14.215	6.882	7.289	1.974	
Tasa cruda	46,25	9,34	4,52	4,79	1,30	
Tasa estandarizada	42,52	8,22	3,39	4,79	1,29	
Tasa estandarizada mínima	13,30	2,89	0,21	3,36	0,25	
Tasa estandarizada máxima	74,15	20,20	10,96	1,21	4,04	
Tasa estandarizada P ₅	18,01	4,08	1,07	11,95	0,53	
Tasa estandarizada P ₂₅	32,40	2,27	2,11	1,40	0,83	
Tasa estandarizada P ₅₀	39,14	6,49	3,00	2,09	1,29	
Tasa estandarizada P ₇₅	54,59	9,55	7,17	3,39	2,19	
Tasa estandarizada P ₉₅	69,29	17,19	10,90	4,15	3,67	
RV ₂₅₋₇₅	1,68	1,81	3,39	1,99	2,6	
RV ₅₋₉₅	3,85	4,21	10,19	7,40	6,94	
CSV	0,13	0,23	0,52	0,35	0,64	
ANOVA (variable independiente: programa de cribado)	R ²	0,44	0,88	0,81	0,91	0,57
	IC 95%	0,27-0,61	0,77-0,99	0,70-0,95	0,82-1,00	0,39-0,73
	p	0,02	<0,0001	<0,0001	<0,0001	0,0012

P₅ P₂₅ P₅₀ P₇₅ P₉₅: Percentil 5, 25, 50, 75 y 95. RV₂₅₋₇₅: Razón de variación entre la tasa estandarizada de la unidad del percentil 75 y la de la unidad del percentil 25. RV₅₋₉₅: Razón de variación entre la tasa estandarizada de la unidad del percentil 95 y la de la unidad del percentil 5. CSV: Coeficiente sistemático de la variación. ANOVA: Análisis de la varianza de un factor. PAAF: Punción-aspiración con aguja fina. BAG: biopsia con aguja gruesa. BQ: biopsia quirúrgica.

el análisis por subgrupos de pruebas invasivas, la menor variación sistemática se encontró en la tasa de realización de BAG, con un CSV del 35%, para la tasa de realización de PAAF fue del 52% y para la tasa de realización de BQ del 64%. En

En la Figuras 3 se observa la distribución de la variación de las tasas analizadas, con una menor dispersión para la tasa de realización de pruebas invasivas y la tasa de recitaciones, que para las tasas de realización de cada una de las pruebas diagnósticas específicas.

Figura 3. Variaciones en las pruebas de confirmación diagnóstica por unidades de cribado. (Asturias, Cantabria y Cataluña. 2002-2011).



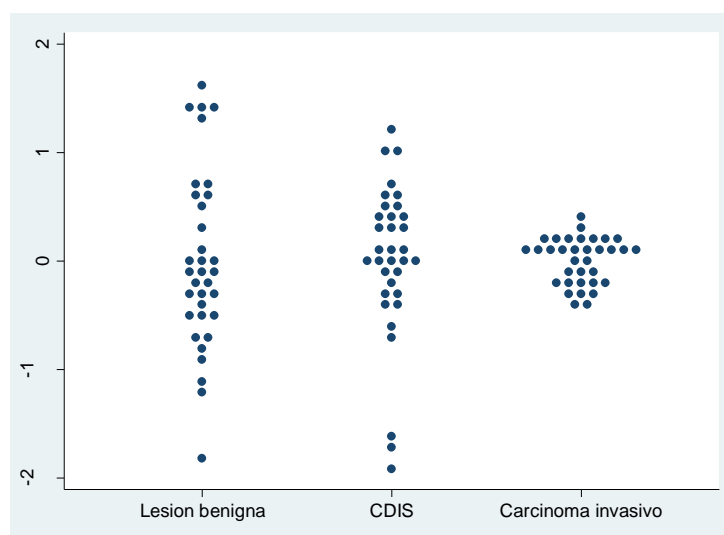
PAAF: Punción-aspiración con aguja fina. BC: biopsia con aguja gruesa. BQ: biopsia quirúrgica.

En el análisis por subgrupos BI-RADS, la tasa de recitaciones presentó una RV_{25-75} de 1,63 para el BI-RADS 0, de 5,94 para el BI-RADS 3, de 1,08 para el BI-RADS 4 y de 2,40 para el BI-RADS 5. En el análisis por subgrupos BI-RADS de la tasa de realización de pruebas invasivas las RV_{25-75} fueron de 2,10 para el BI-RADS 0, 2,20 para el BI-RADS 3, 1,50 para el BI-RADS 4 y 2,38 para el BI-RADS 5 (estos resultados no figuran en las tablas).

Las variaciones en las tasas de detección (lesiones benignas, CDIS y CI) se presentan en la Tabla 14. Las mayores diferencias se encontraron en las tasas de detección de lesiones benignas, con una RV_{25-75} de 2,79, mientras que las RV_{25-75} de las tasas de detección, de CI y CDIS fueron menores de 2. En la Figura 4 se representa esta variación; las tasas de detección de CI de todas las unidades aparecen agrupadas en torno a los valores centrales y no se identificaron valores extremos, mientras que si aparecen en las otras tasas analizadas, mayores para la detección de lesiones benignas. En el ANOVA

realizado con el fin de controlar el efecto del programa al que están asignadas las unidades de cribado, los coeficientes R^2 oscilaron entre 0,91 para la tasa de realización de BAG y 0,44 para la tasa de recitación, todos resultaron estadísticamente significativos. (tablas 13 y 14).

Figura 4: Variaciones en las tasas de detección de lesiones por unidades de cribado. (Asturias, Cantabria y Cataluña. 2002-2011)



CDIS: CArcinoma Ductal in situ

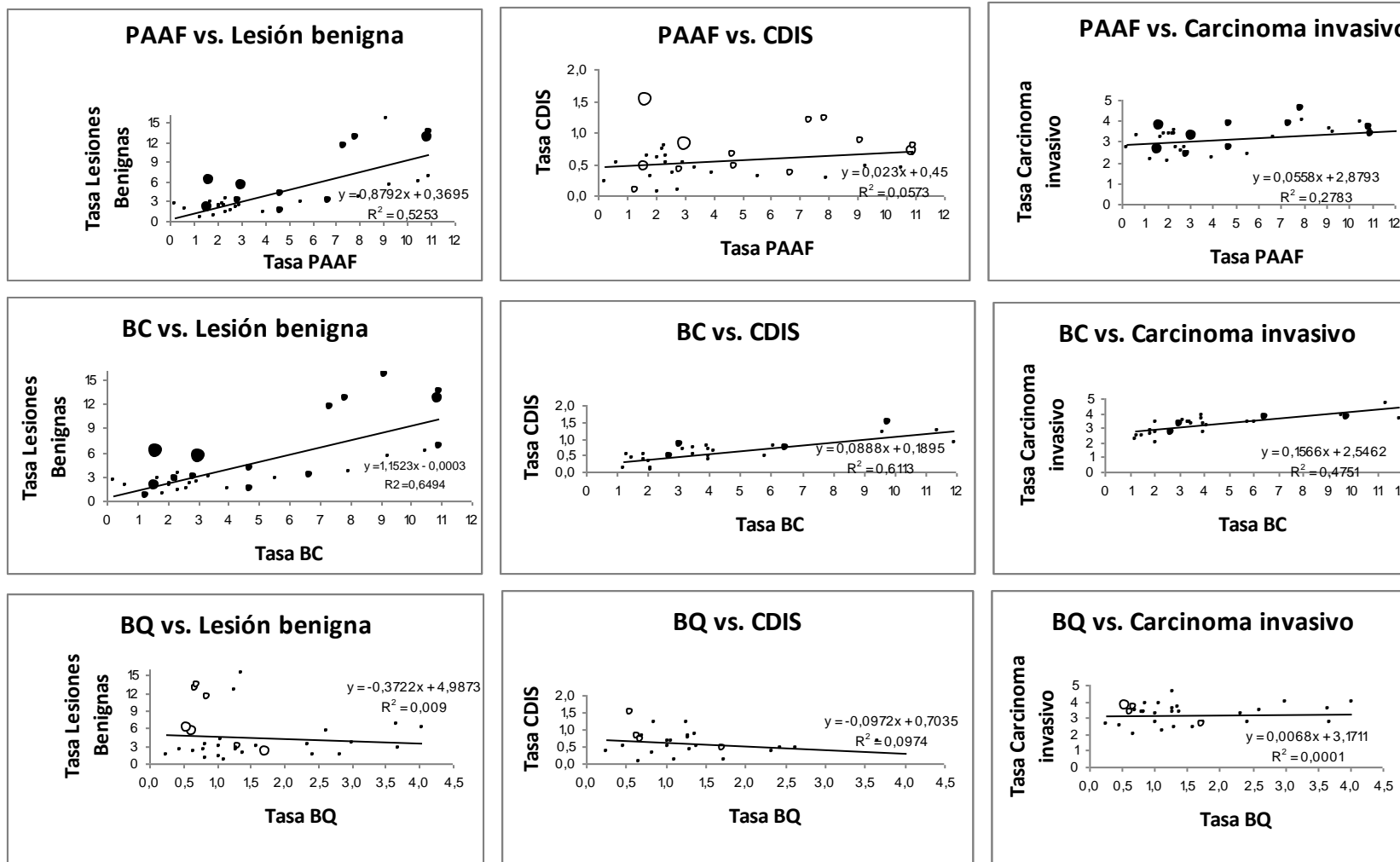
Tabla 14. Variabilidad entre unidades de cribado poblacional de cáncer de mama en las tasas de resultados del episodio de cribado (x 1000) basado en la estandarización directa. (Asturias, Cantabria y Cataluña. 2002-2011)

	Lesiones Benignas	Carcinoma ductal in situ	Cáncer invasivo	
Episodios	8.240	1.050	4.930	
Tasa cruda	5,42	0,69	3,24	
Tasa estandarizada	4,40	0,55	3,18	
Tasa estandarizada mínima	0,54	0,07	2,03	
Tasa estandarizada máxima	15,50	1,50	4,62	
Tasa estandarizada P ₅	0,97	0,09	2,22	
Tasa estandarizada P ₂₅	1,97	0,35	2,65	
Tasa estandarizada P ₅₀	2,82	0,49	3,32	
Tasa estandarizada P ₇₅	5,51	0,70	3,62	
Tasa estandarizada P ₉₅	12,98	1,20	4,09	
RV ₂₅₋₇₅	2,79	1,99	1,36	
RV ₅₋₉₅	13,41	13,73	1,84	
CSV	0,56	0,19	0,03	
ANOVA (variable independiente: programa de cribado)	R ² p IC 95%	0,91 0,82-1,00 <0,0001	0,85 0,73-0,97 <0,0001	0,51 0,33-0,67 0,0052

P₅ P₂₅ P₅₀ P₇₅ P₉₅: Percentil 5, 25, 50, 75 y 95. RV₂₅₋₇₅: Razón de variación entre la tasa estandarizada de la unidad del percentil 75 y la de la unidad del percentil 25. RV₅₋₉₅: Razón de variación entre la tasa estandarizada de la unidad del percentil 95 y la de la unidad del percentil 5. CSV: Coeficiente sistemático de la variación. ANOVA: Análisis de la varianza de un factor

Finalmente, en la Figura 5 se representa la relación entre las diferentes pruebas invasivas realizadas y los resultados del cribado, según el número de episodios en cada unidad. En conjunto las unidades con más actividad tienen tasas de realización de pruebas invasivas menores que las unidades más pequeñas. Se encontraron coeficientes de regresión estadísticamente significativos entre la tasa de realización de PAAF y la de detección de lesiones benignas y de carcinoma invasivo, aunque en esta última fue de 0,28; también resultaron estadísticamente significativas las regresiones entre las tasas de realización de BAG y las tasas de detección de los tres tipos de lesiones estudiadas. Por último, la tasa de realización de BQ no guarda relación con los resultados del proceso de cribado.

Figura 5. Relación entre las tasas de realización de pruebas invasivas y las tasas de detección por unidades de cribado. (Asturias, Cantabria y Cataluña. 2002-2011)



PAAF: Punción Aspiración con aguja fina. BAG: Biopsia con aguja gruesa BQ: Biopsia Quirúrgica. Las burbujas negras representan coeficientes R^2 estadísticamente significativos, y las burbujas blancas coeficientes R^2 estadísticamente no significativos

5. DISCUSIÓN

El análisis de la incidencia acumulada de cáncer de mama y el riesgo de sobrediagnóstico asociado a la participación en el Programa de detección precoz de cáncer de mama de Asturias muestra un incremento del riesgo para el diagnóstico de CDIS, pero no para CI o para la incidencia total de cáncer de mama.

Cinco años después de su último contacto con el programa la diferencia de incidencia sigue siendo más alta en las mujeres participantes, aunque las diferencias se aminoran. Como es esperable, en los primeros años del cribado, cuando se adelanta el diagnóstico, la incidencia acumulada crece a un ritmo más rápido en las mujeres que realizan cribado que entre las mujeres que no participan; asimismo al final del periodo de seguimiento, cuando hay una proporción grande de mujeres que no realizan cribado es más rápido el aumento de la incidencia acumulada en las mujeres no participantes, pero este cambio de tendencia en el grupo de mujeres participantes no alcanza a frenar la tasa tanto como para igualarse a las que no participan, como debería ocurrir si todo el exceso de incidencia se debiera al adelantamiento diagnóstico. Así mismo la tendencia en la detección de CDIS, se frena en las mujeres participantes después de finalizar su participación en el programa.

Los RR encontrados para cualquier tipo de cáncer coinciden con los que presenta, en las mujeres de mayor edad, el seguimiento del ensayo clínico de Malmö (Zackisson, 2006), así como el mayor incremento para cáncer in situ en relación con el invasivo encontrado en el estudio RAFFP (Román, 2013). Asimismo, como ocurre en otros programas (Hofving, 2012; Bleyer, 2012) los diagnósticos se producen en estadios más tempranos entre las mujeres participantes respecto a las no participantes.

La principal fortaleza del análisis de incidencia realizado se encuentra en la y garantía de exhaustividad de la información utilizada, debido a la oportunidad

de disponer de registros adecuados, como las cohortes de mujeres participantes en los programas de detección precoz de cáncer de mama, que han permitido el seguimiento a largo plazo de los mismos. Contar con los registros poblacional y hospitalario de tumores, como ocurre en Asturias, son requisitos imprescindibles para poder medir la efectividad de los programas poblacionales de cribado de cáncer. Están reconocidos los registros poblacionales de tumores como un elemento clave para la evaluación de los Programas de detección precoz de cáncer (Navarro, 2010), con una buena exhaustividad (Villar-Coromina, 2011) y en el caso de Asturias con la ventaja de tener una cobertura de CCAA, que incluye centros privados. Disponer de estos recursos permite aportar a nivel del Estado español los primeros resultados en efecto sobre la incidencia del cáncer de mama en relación con los programas poblacionales de cribado de cáncer de mama (Natal, 2015).

Los programas poblacionales de detección precoz de cáncer de mama tienen ventajas e inconvenientes que ya se han descrito, mientras que las ventajas benefician a las mujeres que tienen un diagnóstico precoz, los inconvenientes afectan a todas las mujeres que se someten a las pruebas de cribado, con independencia de los resultados del mismo. En el programa evaluado las mujeres que participan en el mismo tienen como beneficio una probabilidad mayor, respecto a las no participantes de que se les diagnostique un cáncer de mama en un estadio más temprano, sin que se incremente el riesgo de diagnóstico de cáncer de mama invasivo, aunque como inconveniente tienen un riesgo ligeramente superior de ser diagnosticadas de un cáncer de mama in situ, lo que permitiría esperar, a la luz del conocimiento actualmente disponible, una reducción de la mortalidad como resultado de su participación en el cribado.

Estos hallazgos podrían estar condicionados por las siguientes limitaciones: en primer lugar se trata de una cohorte pequeña, con una distribución por año de nacimiento sesgada con una población pequeña de las mujeres más jóvenes. En conjunto se trata de mujeres que tenían 60 y más años cuando se inició el

programa, que son las cohortes de edad que han tenido suficiente tiempo de seguimiento para ajustar el adelantamiento diagnóstico (Puliti, 2009), pero también sabemos que sólo forman parte de la población elegible aquellas mujeres que no han sido diagnosticadas de cáncer de mama previamente.

Por otra parte el hecho de no participar en un programa de cribado no implica que las mujeres no se realicen mamografías preventivas, ya que se trata de una práctica con una aceptación muy alta entre la población diana, encontrándose en nuestro medio que el 98 % de las mujeres entre 50 y 69 años realizaron una mamografía en los cinco últimos años (Perea, 2012). Por estas razones podría estar infraestimado el incremento de la incidencia del programa.

En cuanto al impacto de la introducción de la BAG en el diagnóstico de lesiones de mama identificadas en el cribado encontramos que supone en el periodo 1996-2011 una reducción del "esfuerzo diagnóstico". En las mujeres con hallazgos de sospecha de cáncer de mama en las mamografías de cribado: se redujeron el número total de recitaciones y el número de pruebas invasivas realizadas. Esta reducción en el uso de pruebas invasivas se corresponde con lo encontrado en otros estudios (Chin-Lenn, 2013).

Esta reducción en el uso de pruebas diagnósticas representa en primer lugar un beneficio para las mujeres, que tienen una oportunidad de diagnóstico menos agresiva y además un beneficio para el sistema sanitario, que consigue los mismos resultados con menos pruebas y por tanto con menor consumo de recursos. Sin embargo, a pesar de sus ventajas, la introducción de la BAG fue lenta y desigual y su uso no se ha generalizado en el contexto de los programas de cribado de cáncer de mama. Otras publicaciones sobre el uso de la BAG en programas de detección precoz de cáncer de mama informan tasas muy variables, entre el 62% y el 90% (Duijim, 2007; Román, 2013).

El análisis geográfico de variabilidad observamos que las unidades de detección precoz de cáncer de mama tienen variaciones en las tasas de realización de pruebas invasivas entre que pueden clasificarse como altas, mientras que las tasas de detección de lesiones tienen una variación baja. Esto puede suponer que, en algunos casos, se realicen pruebas invasivas adicionales evitables. En conjunto se ha encontrado que hacer más recitaciones y/o pruebas invasivas no está asociado con la detección de cáncer de mama y sí con la detección de lesiones benignas. Las unidades que hacen más PAAF tienen tasas de detección de lesiones benignas mayores y las que hacen más BAG tienen mayores tasas para todas las lesiones, pero menores que con la PAAF. Aunque tanto el CDIS, como las lesiones benignas tienen interés por su evolución a carcinoma invasivo (Duffy, 2016; Castells, 2015), es importante tener presente que el objetivo de los programas de detección precoz de cáncer de mama es reducir la mortalidad por cáncer de mama, con los mínimos efectos adversos para las mujeres participantes, y la realización de estudios adicionales en los que finalmente no se diagnostica un cáncer de mama es el efecto adverso más frecuente de los mismos.

En todos los casos es destacable que los CSV son altos, lo que significa que las causas de esta variabilidad son sistemáticas y el efecto atribuible del azar es pequeño.

Clásicamente se definen como factores vinculados a las variaciones en la práctica médica, las deficiencias en la fiabilidad del análisis, los factores asociados a la demanda y los factores asociados a la oferta, considerando tanto la organización como el proveedor directo (Wennberg, 1982), en este caso es el profesional quien decide qué procedimiento diagnóstico utilizar. Las posibles deficiencias en la fiabilidad del análisis, se han descrito en las limitaciones del estudio. En cuanto a los factores asociados a la demanda, se asume que los programas poblacionales eluden su influencia, ya que la oferta es universal y determinada por el sistema sanitario y no por la demanda individual.

La fuente de información para la realización de los análisis de tendencias y variabilidad fue la Cohorte BELE, se trata de una cohorte retrospectiva que incluye a 542.695 mujeres con edades comprendidas entre 50 y 69 años que han participado en al menos un episodio de cribado entre 1996 y 2010 en nueve programas poblacionales de cáncer de mama de tres comunidades autónomas del Estado español (Asturias, Cantabria y Cataluña) aportando un total de 1.763.810 episodios de cribado. Cada mujer se identifica mediante un código único, y se recoge información sobre el tipo de exploraciones adicionales realizadas y sobre el resultado de las mismas en cada episodio. La principal fortaleza de estos estudios es el tamaño de la cohorte analizada, que además como ya se ha descrito anteriormente representa la totalidad de la población que realiza cribado de cáncer de mama en tres comunidades autónomas.

Entre las limitaciones de estos análisis se encuentra en primer lugar la dificultad para generalizar los resultados, aunque las variaciones en el proceso asistencial encontradas, y el bajo efecto del programa en la misma recomiendan que se estudie como se diagnostica y se trata el cáncer de mama, más allá de las fases iniciales de cribado en los diferentes dispositivos asistenciales, contemplando otras unidades de análisis como el centro o área sanitaria. Cada programa cuenta con varias unidades de cribado, en distintos centros sanitarios, con diferentes culturas y modelos de gestión y planificación diferentes. Estas circunstancias podrían repercutir en los procedimientos diagnósticos incluidos en los programas poblacionales de cribado y no necesariamente contemplados en su diseño y evaluación.

Por otra parte, la variabilidad encontrada puede estar afectada, entre otros factores por la variabilidad en la aplicación de la clasificación BI-RADS en la lectura de las mamografías, descrita en otros estudios (Boyer, 2013). Tampoco se han analizado la variabilidad para cada prueba desagregada por BI-RADS, y/o por hallazgo radiológico debido a que los efectivos en cada grupo fueron

muy pequeños. Finalmente, en el análisis de los resultados no fue posible incluir las tasas de cáncer de intervalo (tumores no diagnosticados en el cribado que aparecen antes del cribado siguiente) de las distintas unidades analizadas, que representan una medida de los resultados falsos negativos de los programas; estas tasas podrían ser diferentes entre las diferentes unidades, aunque las informadas para los programas incluidos en este estudio están dentro de las recomendaciones (Prieto-García, 2103; Blanch, 201)

La complejidad en la aplicación de las nuevas tecnologías pueden influir mucho la calidad de los servicios que se prestan a la ciudadanía. La superación de estas limitaciones mediante el apoyo a la implementación de las nuevas tecnologías es importante, por ejemplo dotando a los centros de los recursos diagnósticos y terapéuticos más adecuados y estableciendo medidas que mejoren su aceptabilidad. Estos factores han sido reconocidos como claves desde la introducción de los cribados en la década de 1970 (Wilson, 1968).

6. CONCLUSIONES

1. La participación en el Programa de detección precoz de cáncer de mama de Asturias no incrementa el riesgo de diagnóstico de cáncer de mama total e invasivo aunque si aumenta el riesgo de diagnóstico de *carcinoma ductal in situ*.
2. La detección de *carcinoma ductal in situ* en las mujeres participantes en el Programa de detección precoz de cáncer de mama de Asturias se incrementa durante el período de participación en el cribado, y se frena después de finalizar su participación en el programa.
3. En el Programa de detección precoz de cáncer de mama de Asturias, las mujeres que participan tienen una mayor probabilidad de que se les diagnostique un cáncer en un estadio más temprano respecto a las no participantes.
4. El uso de la biopsia con aguja gruesa para los procesos diagnósticos de los programas de cribado de cáncer de mama españoles estudiados ha sido lenta y no generalizado, con una diseminación mayor en los últimos años.
5. La introducción de la biopsia con aguja gruesa como prueba de confirmación diagnóstica en los programas de cribado de cáncer de mama españoles estudiados supone una reducción de las tasas de pruebas diagnósticas en general, tanto recitaciones, como pruebas invasivas. y una disminución de las tasas de detección de lesiones benignas, mientras que las tasas de detección de carcinoma, tanto invasivo como *in situ* no se modifican.

6. Las unidades de cribado de los programas poblacionales de cribado de cáncer de mama estudiados presentan una variabilidad sistemática muy alta en la realización de pruebas invasivas para el diagnóstico de cáncer de mama. La menor es para las tasas de realización de biopsias cerradas y es independiente del resultado de la clasificación BIRADS de la mamografía de cribado.
7. La variabilidad entre unidades en las tasas de detección de cáncer tanto *carcinoma ductal in situ* como carcinoma invasivo es baja, y moderada en las tasas de diagnóstico de lesiones benignas.
8. Las unidades que más citologías realizan también tienen una tasa de detección más alta de lesiones benignas y carcinoma invasivo y las que realizan más biopsias cerradas tienen mayor tasa diagnóstica de todas las lesiones: lesiones benignas, *carcinoma ductal in situ* y carcinoma invasivo.

IMPLICACIONES PARA LAS POLITICAS DE CRIBADO

Los programas poblacionales de cribado de cáncer de mama han realizado y realizan un gran esfuerzo en la adherencia a las recomendaciones de las guías de práctica clínica en cuanto a la prueba de cribado, población diana y periodicidad de las pruebas, resultando una buena homogeneidad de criterios en todos ellos. La evaluación continua de los programas de cribado y la difusión de los resultados es una herramienta imprescindible para que la población diana de los mismos pueda decidir individualmente sobre su participación.

Entre los criterios de evaluación de los programas poblacionales de cribado de cáncer de mama debería incluirse sistemáticamente la medición de la incidencia poblacional de cáncer de mama invasivo e *in situ*. Estos estudios no son posibles si no se dispone de una buena información sobre el cáncer de mama diagnosticado en el mismo ámbito geográfico y asistencial, los registros

poblacionales y hospitalarios de tumores son un recurso imprescindible para evaluar la efectividad de los programas de detección precoz de cáncer.

Para mantener los beneficios y reducir los riesgos, sería conveniente explorar estrategias de cribado que se adapte en cierta medida al riesgo individual de diagnóstico de cáncer de mama, permitiendo reducir los riesgos poblacionales de los mismos, así como profundizar en el estudio de la biología de los tumores, que permita diseñar tratamientos personalizados, minimizando los riesgos potenciales del sobretratamiento.

Los programas poblacionales de cribado y los hospitales implicados en el proceso diagnóstico deben continuar los esfuerzos para disminuir la variabilidad en el uso de pruebas diagnósticas y difundir el uso de las tecnologías que resultan efectivas. El uso de protocolos comunes a los distintos centros asistenciales, basados en guías de práctica clínica basadas en la evidencia, la dotación adecuada de recursos, la formación de los profesionales y la evaluación del proceso asistencial son las herramientas para garantizar la equidad la eficiencia en la atención sanitaria.

7. BIBLIOGRAFIA

Alberdi RZ, Llanes AB, Ortega RA, Expósito RR, Collado JM, Verdes TQ, Ramos CN, Sanz ME, Trejo DS, Oliveres XC; CFPR (Cumulative False Positive Risk) group (1). Effect of radiologist experience on the risk of false-positive results in breast cancer screening programs. *Eur Radiol.* 2011;21:2083-90.

Almazán R; Alvarez M, Casamitjana M, Fernandez R; Fernández A, Miranda J, Sarriugrate G. Protocolo para la evaluación de los cánceres de intervalo en los programas de detección precoz de cáncer de mama. Red de Programas de detección precoz de cáncer de mama. 2009 (disponible en http://www.cribadocancer.es/images/archivos/mama/documentacion/canceres_intervalo/archivos/pecid.pdf) Accedido el 15/05/2016.

Altobelli E, Lattanzi A. Breast cancer in European Union: an update of screening programmes as of March 2014 (review). *Int J Oncol.* 2014;45(5):1785-92.

American Joint Committee on Cancer: AJCC Cancer Staging Manual. 6th ed. New York, NY. Springer, 2002.

Andermann A, Blancquaert I, Beauchamp Sb, Déry V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bulletin of the World Health Organization.* 2008; 86(4):241-320.

Anttilla A, Lönnberg S, Ponti A, Suonio E, Villain P, Coeberg JW, von Karsa L. Towards better implementation of cancer screening in Europe through improved monitoring and evaluation and greater engagement of cancer registries, *Eur J Cancer.* 2015; 51:241-51.

Ascunce N, del Moral A, Murillo A et al. Early detection programme for breast cancer in Navarra Spain. *Eur J Cancer Prev.* 1994;3 (Suppl 1): 41-8.

Ascunce N, Elfrade J, Salas D, Zubizarreta R, Ederra M, en nombre de Red de Programas de Cribado de Cáncer. Programas de detección precoz de cáncer de mama en España: características y principales resultados. *Med Clin (Barc).* 2013;141:13–23.

Blanch J, Sala M, Ibáñez J, et al. Impact of risk factors on different interval cancer subtypes in a population-based breast cancer screening programme. *PLoS One.* 2014;9(10):e110207.

Bleyer A, Welch G. Effect of three decades of screening mammography on breast cancer incidence. *N Engl J Med.* 2012;367:1998-2005.

Boyer B, Canale S, Arfi-Rouche J et al. Variability and errors when applying the BI-RADS mammography classification. *Eur J Radiol.* 2013;82:388-97.

Brian R, Sackett DL, Wayne D, Gibson E, Johnson AL. Increased absenteeism from work after detection and labeling of hypertensive patients. *N Engl J Med.* 1978; 299: 741-4

Caldow J, Goh T, Kiger A, Reid J, Tucker J, Vale L. An impact assessment of the health technology assessment. NHS Quality Improvement Scotland. 2006.

Castells X, Domingo L, Corominas JM, et al. Breast cancer risk after diagnosis by screening mammography of nonproliferative or proliferative benign breast disease: a study from a population-based screening program. *Breast Cancer Res Treat.* 2015;149:237-44.

Castells X, Torá-Rocamora I, Posso M, Román M, Vernet-Tomas M, Rodríguez-Arana A, Domingo L, Vidal C, Baré M, Ferrer J, Quintana MJ,

Sánchez M, Natal C, Espinàs JA, Saladié F, Sala M; BELE Study Group. Risk of Breast Cancer in Women with False-Positive Results according to Mammographic Features. *Radiology*. 2016;280:379-86.

Council of the European Union. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). *Off J Eur Union*2003 (L327):34-38.

Chalmers I, Enkin M, Keirse M. *Effective care during pregnancy and childbirth*. Oxford. Oxford University Press; 1989.

Chassin MR. Explaining geographic variations. The enthusiasm hypothesis. *Med Care*. 1993;31(5 Suppl):YS37-4.

Chin-Lenn L, Craighead P, Bryant HE, et al, Quality indicators for ductal carcinoma in situ (DCIS) of the breast: development using a multidisciplinary delphi process and its use in monitoring population-based treatment, *J Surg Oncol*, 2013;108:348-51.

DeSantis CE, Bray F, Ferlay J, Lortet-Tieulent J, Anderson BO, Jemal A. International Variation in Female Breast Cancer Incidence and Mortality Rates. *Cancer Epidemiol Biomarkers Prev*. 2015;24:1495-506.

Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, Yen AM, Smith RA. 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.

Duijijm L, Groenewoud J, Roumen R et al, a decade of breast cancer screening in The Netherlands: trends in the preoperative diagnosis of breast cancer, *Breast Cancer Res Treat*. 2007;106:113-9.

Duffy SW, Dibden A, Michalopoulos D, et al. Screen detection of ductal carcinoma in situ and subsequent incidence of invasive interval breast cancers: a retrospective population-based study. *Lancet Oncol.* 2016;17:109-14.

Elmore JG, Taplin SH, Barlow WE, Cutter GR, D'Orsi CJ, Hendrick RE, Abraham LA, Fosse JS, Carney PA. Does litigation influence medical practice? The influence of community radiologists' medical malpractice perceptions and experience on screening mammography. *Radiology.* 2005;236:37-46.

Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer.* 2013;49:1374-403.

Globocan, 2012 (IARC). Estimated cancer incidence, mortality and prevalence worldwide in 2012. Disponible en: http://globocan.iarc.fr/Pages/burden_sel.aspx. Accedido en 09/05/2016.

Gold HT, Dick AW. Variations in treatment for ductal carcinoma in situ in Elderly women. *Med Care.* 2004;42:267-275.

Hanney S, Buxton M, Green C, Coulson D, Raftery J. An assessment of the impact of the NHS Health Technology Assessment Programme. *Health Technol Assess.* 2007;11(53): 1-180.

Holland WW, Stewart S, Masseria C. Policy brief : screening in Europe [en línea]. Geneva: World Health Organization, European Observatory on Health Systems and Policies; 2006.

Hofvind S, Lee CI, Elmore J. Stage-specific breast cancer incidence rates among participants and non participants of a population-based mammographic screening program. *Breast Cancer Res Treat.* 2012;135:291-99.

Howlander N, Noone AM, Krapcho M, Miller D, Bishop K, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2013/, based on November 2015 SEER data submission, posted to the SEER web site, April 2016.

INAHTA: Health Technology Assessment (HTA). Glossary. 2006. Disponible en <http://htaglossary.net/List+All+Terms#E> [acceso 30 de Septiembre de 2015].

Johns L, Moss S. False-positive Results in the randomized controlled trial of mammographic screening from age 40 (Age trial). *Cancer Epidemiol Biomarkers Prev.* 2010;19:2758-64.

Lau J, Antman EM, Jimenez-Silva I, Kupelnick B, Mosteller F, Chalmers TVC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med.* 1992; 327:248-54.

Lindor NM, Greene MH. The Concise Handbook of Family Cancer Syndromes. *JNCI.* 1998;90:1039-71.

Martínez-Lizaga N, Montes Y, Rodrigo I, et al. Metodología del atlas de variaciones en hospitalizaciones potencialmente evitables en el Sistema Nacional de Salud. *Atlas Var Pract Med Sist Nac Salud.* 2011;4:371-8.

Muir Gray JA, Patnick J, Blanks R. Maximising benefit and minimising harm of screening. *BMJ.* 2008;336: 480-3.

Natal C, Caicoya M, Prieto M, Tardón A. [Breast cancer incidence related with a population-based screening program]. *Med Clin (Barc).* 2015;144(4):156-60.

Natal C, Fernández-Somoano A, Torá-Rocamora I, Tardón A, Castells X; en nombre del grupo BELE. [Variations in the diagnostic confirmation process between breast cancer mass screening units]. *Gac Sanit.* 2016;30:265-71.

Navarro C, Martos C, Ardanaz E, Galcerán J, Izarzugaza I, Peris-Bonet R. Population-based cancer registries in Spain and their role in cancer control. *An Oncol.* 2010;21 (Suppl3):iii3-iii13.

Nelson HD. Screening for breast cancer: An update for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2009; 151:727-37.

Pace LE, Keating NL. A systematic assessment of benefits and risk to guide breast cancer screening decisions. *JAMA.* 2014;311:1327-35.

Paci E, Warwick J, Falini P, Duffy SW. Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? *J Med Screen.* 2004;11:23-7.

Perea MD, Castaño-Vinyals G, Altzibar JM, Ascunce N, Moreno V, Tardón A, Pollán M, Sala M en nombre de los investigadores del MCC-Spain. Prácticas de cribado de cáncer y estilos de vida asociados en la población de controles del estudio español multi-caso control (MCC-Spain). *Gac Sanit.* 2012; 26: 301-10.

Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L, et al., editors. *European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis.* European Commission. 4th ed, Luxembourg: Office for Official Publications of the European Communities; 2006.

Prieto García MA, Delgado Sevillano R, Baldó Sierra C et al. Clasificación y características del cáncer de intervalo en el programa de detección precoz de cáncer de mama en el Principado de Asturias. *Radiología.* 2013;55:408-15.

Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge , Zappa M, Paci E and the EUROSCREEN WORKING GROUP. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J Med Screen*. 2012;19 Suppl 1:42-56.

Puliti D, Zappa M, Miccinesi G, Falini P, Crocetti E, Paci E. An estimate of overdiagnosis 15 years after the start of mammographic screening in Florence. *Eur J Cancer*. 2009;45:3166-71.

Ortún Rubio, V. Desempeño y deseabilidad del sistema sanitario: España. *Revista Asturiana de Economía RAE*, 2006;35.

Roman M, Hubbard RA, Sebuodegard S, et al, The cumulative risk of false-positive results in the Norwegian Breast Cancer Screening Program: updated results. *Cancer*, 2013;119:3952-8.

Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. Principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288:321-3.

United States. Congress. Office of Technology Assessment. (1978). *Assessing the efficacy and safety of medical technologies*. Washington: Congress of the United States, Office of Technology Assessment).

Parker A, Shroen AT, Brenin DR. MRI utilization in newly diagnosed breast cancer: a survey of practicing surgeons. *Ann Sug Oncol*. 2013; 20: 2600-6.

Pollán M, Pastor-Barriuso R, Ardanaz E, Argüelles M, Martos C, Galcerán J, et al. Recent changes in breast cancer incidence in Spain, 1980-2004. *J Natl Cancer Inst*. 2009;101:1584-91.

Red de programas de cribado de cáncer.

<http://www.cribadocancer.es/images/archivos/mama/situacion/Mama2008.pdf>.

Accedido el 25/07/2015

Román R, Sala M, De la Vega M, Natal C, Galcerán J, Gonzalez-Román I, Baroja A, Zubizarreta R, Ascunce N, Salas D, Castells X. Effect of false-positives and women's characteristics on long-term adherence to breast cancer screening. *Breast Cancer Res Treat.* 2011; 130:543-52

Román M, Rué M, Sala M, Ascunce N, Baré M, Baroja A, De la Vega M, Galcerán J, Natal C, Salas D, Sánchez-Jacob M, Zubizarreta R, Castells X; Cumulative False Positive Risk Group. Trends in detection of invasive cancer and ductal carcinoma in situ at biennial screening mammography in Spain: a retrospective cohort study. *PLoS One.* 2013;8:e83121.

Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned?. *Radiol Clin North Am.* 2004;42(5):793-806.

Villar-Corromina N, Miró-Queralt J, Cano-Bautista A, Villardell-Gil L, Torres-Barbié P, Marcos-Grajera R. Cáncer cutáneo distinto de melanoma: tendencias de la incidencia poblacional en Girona, 1994-2007. *Med Clin (Barc)* .2011;137:145-51.

Warner E, Heisey RE, Goel V, Carroll JC, McReady DR. Hereditary breast cancer. Risk assessment of patients with a family history of breast cancer. *Can Fam Physician.* 1999;45:104-12.

WHO Position Paper on Mammography Screening. Geneva: World Health Organization; 2014.

Wennberg JE, Barnes BA, Zubkoff M. Professional Uncertainty and the problem of supplier-induced demand. *Soc Sci Med.* 1982;16:811-24.

Wilson JMG, Jungner G. Principles and practice of screening for diseases. WHO (report N° 34); Geneva.1968.

Wright FC, Gagliardi Ar, Frases N, Quan ML. Adoption of surgical innovations: factors influencing use of sentinel lymph node biopsy for breast cancer. *Surg Innov.* 2011; 18:379-86.

Zackrisson S, Andersson I, Janzon L et al. Rate of overdiagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ.* 2006;332:689-92.