



**Universidad
Europea**

Máster en Bioinformática

**Pilot a self-administered Risk Assessment tool led
by the EVITA Platform for Triage and
Identification of eligible Citizens for genetic
counselling, testing and psychological support
under MammoScreen**

Autor: Jose Antonio Sánchez Diaz

Tutor: Vanessa Ferreira

Curso 2023-24

Firmas:

Tutor: _____

Alumno: *Jose Antonio Sánchez Diaz*

Fecha: 06 / 11 / 2024

Agradecimientos

A Pepe, por su valiosa ayuda, y a Alba, por su apoyo incondicional durante todo el proceso; gracias a ambos por estar siempre ahí.

Al equipo de Humanized Solutions, especialmente a Vanessa, cuya guía y dedicación fueron fundamentales, y a Ana y Marta por su valiosa colaboración y orientación en cada etapa del proyecto.

A todos, gracias por su apoyo, compromiso y contribución en el desarrollo de este trabajo.

Indice general

Indice general

Índice de figuras

Indice de tablas

Resumen

Abstract

Introducción

Hipótesis y Objetivos

Metodología

Resultados

Conclusiones y Discusión

Bibliografía

Anexos

Índice de figuras

Figura 1. Diagrama de flujo metodológico del TFM

Figura 2. Gráfica de distribución de grupos de edad

Figura 3. Curva ROC de nuestro modelo

Figura 4. Curva ROC variable generaciones consecutivas

Figura 5. Curva ROC variable generaciones consecutivas

Figura 6. Diagrama de barras que representa la presencia en redes sociales de las asociaciones en cada país

Indice de tablas

Tabla 1.1 Resultados de las búsquedas bibliográficas

Tabla 1.2. Resultado del filtro creado en R para la búsqueda uno

Tabla 1.3. Resultado del filtro creado en R para la búsqueda dos

Tabla 1.4. Resultado del filtro creado en R para la búsqueda tres

Tabla 2.1: Análisis con Respuestas Originales

Tabla 2.2: Análisis con Respuestas Transformadas

Tabla 3. Tabla de contingencia

Resumen

El cáncer de mama hereditario representa una proporción significativa de los casos de cáncer de mama, y su detección temprana es crucial para mejorar los resultados clínicos (Speiser & Bick, 2023). La prevención y el cribado en individuos con predisposición genética requieren estrategias específicas que atiendan sus riesgos y necesidades (Speiser & Bick, 2023). Además, los avances en inteligencia artificial, como el uso de redes neuronales, muestran potencial en la predicción de cánceres hereditarios, aportando herramientas adicionales para la evaluación del riesgo y la toma de decisiones clínicas (Guan, Parmigiani, Braun, & Trippa, 2022).

El proyecto MammoScreen, financiado por la Unión Europea y UK Research and Innovation, tiene como objetivo optimizar el cribado del cáncer de mama hereditario mediante tecnologías innovadoras. En este contexto, se está validando MammoWave, una alternativa no invasiva a la mamografía tradicional. Más información está disponible en <https://mammoscreenproject.eu/>.

Como parte de MammoScreen, la asociación EVITA - Cancro Hereditario desarrolló un cuestionario para evaluar el riesgo de cáncer de mama hereditario (<https://evitacancro.org/>), que recoge datos sobre antecedentes familiares y tipos de cáncer atípicos.

Este Trabajo de Fin de Máster incluyó un análisis exhaustivo del cuestionario mediante regresión logística para identificar patrones de riesgo significativos, así como un análisis descriptivo de las respuestas por grupo de edad. Adicionalmente, se realizaron búsquedas bibliográficas y, empleando conocimientos de programación, se desarrollaron scripts para automatizar el procesamiento de datos. Finalmente, se diseñó un plan de comunicación en colaboración con asociaciones de pacientes para asegurar que los resultados de esta herramienta lleguen a la comunidad y respalden decisiones informadas en salud preventiva.

Palabras clave: cáncer de mama hereditario, evaluación de riesgo genético, cuestionario de riesgo genético, regresión logística, salud preventiva

Abstract

Hereditary breast cancer accounts for a significant proportion of breast cancer cases, and its early detection is crucial for improving clinical outcomes (Speiser & Bick, 2023). Prevention and screening in individuals with genetic predisposition require specific strategies that address their unique risks and needs (Speiser & Bick, 2023). Furthermore, advances in artificial intelligence, such as neural networks, have shown potential in predicting hereditary cancers, providing additional tools for risk assessment and clinical decision-making (Guan, Parmigiani, Braun, & Trippa, 2022).

The MammoScreen project, funded by the European Union and UK Research and Innovation, aims to optimize screening for hereditary breast cancer through innovative technologies. In this context, MammoWave, a non-invasive alternative to traditional mammography, is being validated. More information is available at <https://mammoscreenproject.eu/>.

As part of MammoScreen, the EVITA - Hereditary Cancer Association developed a questionnaire to assess hereditary breast cancer risk (<https://evitacancro.org/>), which collects data on family history and atypical cancer types.

This Master's Thesis included a thorough analysis of the questionnaire using logistic regression to identify significant risk patterns, as well as a descriptive analysis of responses by age group. Additionally, literature searches were conducted, and programming expertise was applied to develop scripts for automating data processing. Finally, a communication plan was designed in collaboration with patient associations to ensure that the results of this tool reach the community and support informed decision-making in preventive health.

Keywords: hereditary breast cancer, genetic risk assessment, genetic risk questionnaire, logistic regression, preventive health

Introducción

1.1. El proyecto MammoScreen

El cáncer de mama es el cáncer más común en las mujeres a nivel mundial, afectando a 1 de cada 8 mujeres. La Organización Mundial de la Salud estima que en 2020, los casos de cáncer de mama en todo el mundo representarán el 24% de todos los cánceres en mujeres, con 2.26 millones de mujeres diagnosticadas con cáncer y 685,000 muertes a nivel global. La detección precoz del cáncer de mama reduce las tasas de mortalidad y los costes sanitarios. (MammoScreen Project. (2024))

Aunque la mamografía es la tecnología estándar para la detección del cáncer de mama tiene limitaciones e inconvenientes, como el uso de radiaciones, la compresión mamaria, el uso limitado en el caso de mujeres con mamas densas y en algunos tipos de cánceres ocultos.

MammoScreen es un proyecto cofinanciado por la Unión Europea y la agencia UK Research and Innovation. Está diseñado con el objetivo de generar evidencia clínica sobre el uso de **MammoWave**, un nuevo mamógrafo basado en frecuencias de microondas, para el cribado del cáncer de mama. El proyecto involucra a siete socios internacionales de Italia, Portugal, España y el Reino Unido.

MammoScreen tiene como objetivo proporcionar una solución innovadora, segura y no invasiva para la detección del cáncer de mama. Para obtener más información, visite <https://mammoscreenproject.eu/> .

La intención de **MammoScreen** es generar pruebas sólidas para utilizar **MammoWave**, un nuevo dispositivo de imagen por microondas, como técnica de cribado del cáncer de mama en programas sanitarios nacionales o regionales. Para reunir evidencia, en el ámbito de **MammoScreen**, se está realizando un ensayo clínico multicéntrico y prospectivo donde participarán 10000 mil de mujeres en al menos 10 centros de reclutamiento europeos, acompañado de una app centrada en el usuario final. Los impactos previstos son significativos:

- Las mujeres se beneficiarán de un método de cribado más seguro, inclusivo (sin limitaciones de edad), cómodo (sin compresión mamaria), accesible y preciso.
- Además, los responsables de las políticas sanitarias obtendrán pruebas cruciales para revisar y, potencialmente, ampliar los actuales programas de cribado del cáncer de mama.

1.2. La Participación de la asociación EVITA- Cancro Hereditario en MammoScreen

La participación del paciente en proyectos europeos es fundamental para garantizar que las soluciones de salud desarrolladas estén verdaderamente centradas en el usuario. Involucrar a los pacientes en cada etapa del proyecto, desde la investigación hasta el desarrollo y la implementación, permite abordar mejor sus necesidades y expectativas, mejorando así los resultados y la calidad de vida. Esta colaboración asegura que las innovaciones en salud sean más efectivas, inclusivas y sostenibles.

En los últimos años, se ha reconocido cada vez más el valor e importancia de involucrar a las asociaciones y a los ciudadanos en la investigación y desarrollo de soluciones innovadoras en salud, especialmente en la tecnología médica (MedTech). Más información sobre la participación de pacientes en MedTech en la referencia primera y referencia segunda. Siguiendo la Guía de MedTech Europe sobre las interacciones entre la industria de tecnología médica y las organizaciones de pacientes, este proyecto ha sido diseñado para incluir la voz de los pacientes y ciudadanos. El enfoque colaborativo de **MammoScreen** empodera a los individuos para participar activamente en la configuración del futuro de la atención médica, impulsando avances y mejorando los resultados para los pacientes. Para obtener más información, visite <https://mammoscreenproject.eu/citizens/>

Descripción de EVITA

Un socio relevante de **Mammoscreen** es la asociación de pacientes **EVITA - Cancro Hereditario**. EVITA ha desarrollado la **Plataforma EVITA**, diseñada para beneficiar a cualquier ciudadano preocupado por su riesgo individual de cáncer, portadores de mutaciones genéticas de alto riesgo y pacientes oncológicos. La plataforma ofrece un enfoque centrado en el ciudadano y también está disponible para profesionales de salud e investigadores. Garantiza la máxima protección de datos personales y orienta sobre la necesidad de asesoramiento genético. Próximamente, ofrecerá consultas virtuales de asesoramiento genético y psicológico. Más información en <https://evitacancro.org/> .

La Plataforma EVITA (EP, EVITA Platform)

EVITA ha desarrollado la Plataforma **EVITA** (Más información en <https://evitacancro.org/evita-platform/>), diseñada para beneficiar a cualquier ciudadano preocupado por su riesgo individual de cáncer, portadores de mutaciones genéticas de alto riesgo y pacientes oncológicos. La plataforma ofrece un enfoque centrado en el ciudadano, también accesible para profesionales de salud e investigadores.

Además, garantizando la máxima protección de datos personales, la Plataforma EVITA orienta a las personas sobre la necesidad de asesoramiento genético mediante un cuestionario basado en su historia personal y familiar de cáncer. Permite a los usuarios almacenar su historial clínico en su área personal y próximamente ofrecerá consultas virtuales de asesoramiento genético y psicológico.

1.3. Participación e Implicación de Pacientes y Público (PPIE) en proyectos europeos

Hoy en día, la atención de las farmacéuticas y los proyectos europeos está en invertir cada vez más recursos en el **PPIE** (Patient and Public Involvement and Engagement) involucrando así la voz de los pacientes en la investigación y desarrollo asegurándose de que su opinión y sus necesidades se tienen en cuenta, lo cual es crucial para la aceptación pública y el éxito a largo plazo de estas intervenciones.

Humanized Solutions (HS) es una empresa especializada en el uso de métodos de Ciencias Sociales y Humanidades (**SSH**) para liderar actividades de Participación e Implicación de Pacientes y Público (**PPIE**). Se dedican a mejorar la participación activa de los pacientes y el público en proyectos de salud, asegurando que sus necesidades y experiencias se reflejan en el desarrollo de soluciones innovadoras y efectivas. A través de la colaboración en consorcios de investigación y la prestación de servicios personalizados, **Humanized Solutions** transforma la forma en que se implementan y evalúan las iniciativas de salud. **HS** está trabajando conjuntamente con **EVITA** y otros colaboradores del proyecto, para involucrar la voz de los pacientes en los Work Packages (WPs) 5 y 7 que se están llevando a cabo en **MammoScreen**.

Con el fin de mejorar el proceso de evaluación de riesgo en pacientes y aumentar la precisión del cribado de cáncer de mama hereditario, este proyecto se centra en analizar un cuestionario específico que permita identificar mejor a las personas con alto riesgo genético. Este cuestionario tiene como objetivo facilitar el cribado y seguimiento de pacientes, optimizando el proceso de detección temprana y brindando una herramienta que apoye tanto a profesionales de la salud como a pacientes en la toma de decisiones informadas.

La participación e implicación de pacientes y público (**PPIE**) es fundamental para el desarrollo de esta herramienta, ya que garantiza que el cuestionario esté alineado con las necesidades y expectativas de los usuarios finales. La colaboración con mujeres (tanto pacientes con cáncer de mama como mujeres sanas que participan en programas de cribado) permite obtener una visión integral sobre cómo el cuestionario puede adaptarse mejor a sus experiencias y necesidades. **Además, esta participación contribuye a identificar posibles áreas de mejora en el cuestionario, lo cual es esencial para que la herramienta sea comprensible, accesible y eficaz en diferentes contextos.**

El cuestionario fue desarrollado como una herramienta preliminar para explorar las necesidades y experiencias de un grupo diverso de personas en relación con el cribado y la predisposición genética al cáncer de mama hereditario. Su estructura busca capturar una variedad de factores relevantes, como antecedentes familiares de cáncer en distintas generaciones, diagnósticos en edades tempranas y tipos de cáncer menos comunes. Al indagar sobre estos aspectos, el cuestionario permite obtener una visión inicial de los perfiles de riesgo.

El enfoque de este cuestionario se extiende más allá de la identificación de predisposición genética. La amplitud de los datos obtenidos a través de este cuestionario permitirá refinar la atención en el cribado de cáncer de mama, promoviendo intervenciones más ajustadas a las características demográficas y familiares de los pacientes.

1.4. Cuestionario sobre las Necesidades de los Ciudadanos en Relación con el Cáncer de Mama

El Cuestionario sobre las Necesidades de los Ciudadanos en Relación con el Cáncer de Mama se divide en dos secciones principales: la primera se enfoca en la evaluación del riesgo genético de cáncer de mama hereditario y la segunda en las necesidades informativas y tecnológicas de los pacientes durante el proceso de cribado.

La sección de **risk assessment** (evaluación de riesgo) recopila información detallada sobre antecedentes familiares, factores de riesgo específicos y características hereditarias del cáncer, con el objetivo de identificar patrones relevantes en la predisposición genética. **Este Trabajo de Fin de Máster se centra exclusivamente en el análisis de esta primera sección, aplicando un enfoque estadístico para identificar patrones de riesgo significativos.**

La segunda parte del cuestionario tiene como propósito ofrecer una visión preliminar de las necesidades y experiencias de un amplio grupo de participantes, antes de profundizar en subgrupos específicos mediante talleres. Para esta sección, el cuestionario se distribuyó antes de un seminario web y abarcó participantes de entre 24 y 80 años, proporcionando una muestra representativa tanto para el cribado tradicional de mamografía como para el uso potencial del dispositivo **MammoWave** en poblaciones más jóvenes.

Para capturar de forma integral las perspectivas de los participantes, el equipo de **HS, ELAROS** y **EVITA** acordó un enfoque de investigación mixta, que combina métodos cualitativos y cuantitativos, permitiendo obtener datos detallados y conclusiones generalizables. El cuestionario fue así diseñado como una fase inicial de un proceso de co-creación, seguida por talleres para profundizar en los hallazgos y adaptar mejor los recursos y tecnologías de apoyo en el contexto del cáncer de mama hereditario.

1.4.1. Evaluación del Riesgo Genético de Cáncer Hereditario

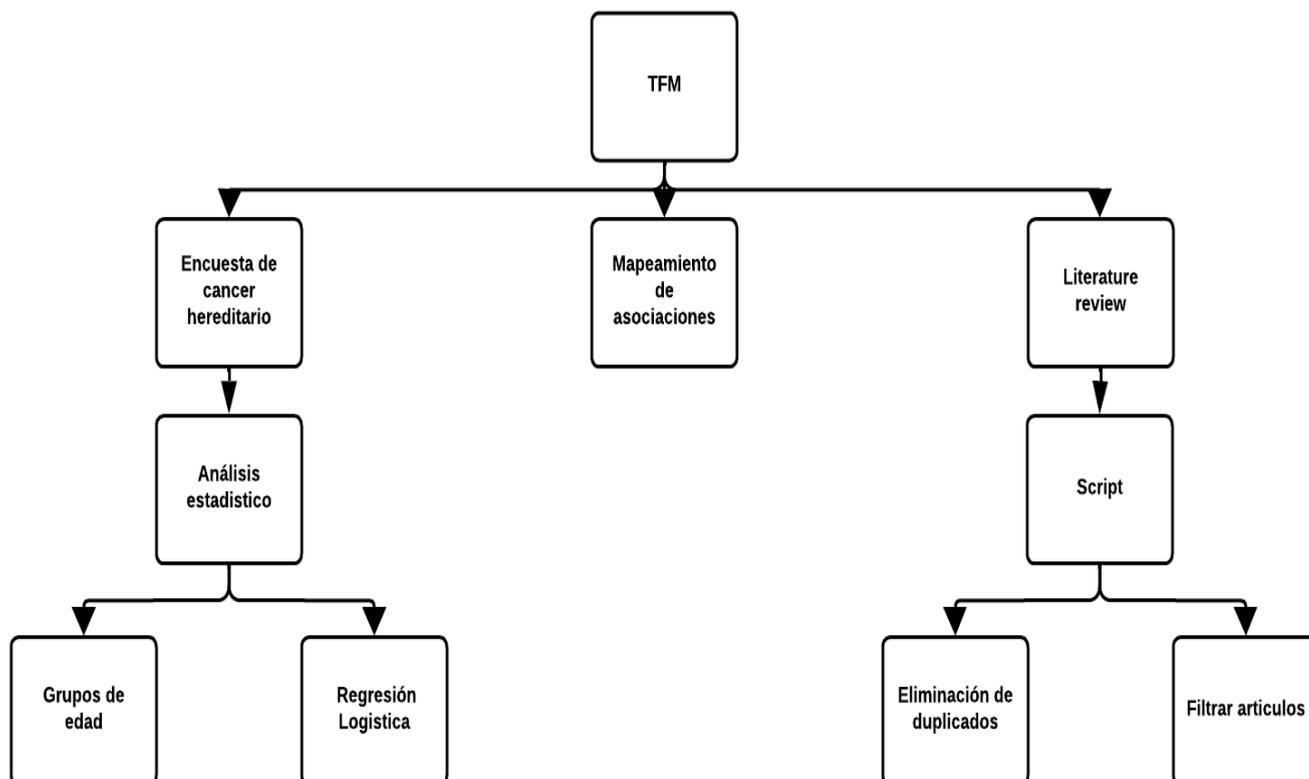
La evaluación del riesgo genético de cáncer hereditario es una herramienta crucial en la medicina moderna, permitiendo identificar a individuos con predisposición genética a desarrollar ciertos tipos de cáncer. Este proceso implica el uso de cuestionarios específicos diseñados para recoger información detallada sobre antecedentes familiares y personales de cáncer, así como la realización de pruebas genéticas cuando es necesario (Speiser & Bick, 2023). El objetivo principal de esta evaluación es ofrecer un enfoque personalizado en la prevención y detección temprana del cáncer, permitiendo a los pacientes y a sus familias tomar decisiones informadas sobre su salud. La efectividad y precisión de los cuestionarios utilizados son fundamentales para garantizar evaluaciones confiables y útiles en la práctica clínica (Guan, Parmigiani, Braun, & Trippa, 2022).

En el presente estudio, el cuestionario fue administrado en inglés y dirigido a mujeres adultas, incluyendo tanto a pacientes con cáncer de mama como a mujeres sanas, con y sin predisposición genética. El periodo de recolección de datos abarcó desde el 22 de agosto de 2023 hasta el 11 de noviembre de 2023, y se obtuvieron 433 respuestas válidas tras la exclusión de dos participantes que no otorgaron su consentimiento. Este enfoque riguroso en el diseño y aplicación del cuestionario respalda el análisis estadístico y la relevancia de los hallazgos en la práctica clínica, contribuyendo al desarrollo de herramientas precisas para la evaluación del riesgo y la toma de decisiones clínicas informadas.

En este Trabajo de Fin de Máster se evaluarán las siguientes cuestiones:

1. Análisis de los resultados del **Cuestionario de Evaluación del Riesgo Genético Hereditario de Cáncer** mediante regresión logística, evaluando su eficacia estadística para identificar factores de riesgo significativos.
2. **Análisis del porcentaje de respuestas según el grupo de edad** para identificar posibles patrones de riesgo en diferentes etapas de la vida.
3. Realización de unas **búsquedas bibliográficas exhaustivas** para identificar artículos relevantes que permitan comparar el funcionamiento de los scripts desarrollados en este TFM, cuyos resultados serán presentados y discutidos en secciones posteriores.
4. **Creación de un plan de comunicación y divulgación** que asegure que los resultados y beneficios de la herramienta de autoevaluación del riesgo genético de cáncer hereditario lleguen de manera efectiva a la comunidad.
5. **Mapeo de asociaciones de pacientes** para contactarlas y hacerlas conocedoras del plan de comunicación, así como de los resultados obtenidos en este TFM.

Figura 1. Diagrama de flujo metodológico del TFM: Estructura metodológica del Trabajo de Fin de Máster.



1.5 Comité Ético y Gobernanza de Datos

Los documentos asociados al cuestionario del **Proyecto MammoScreen** fueron revisados y aprobados por el Comité de Ética y Gobernanza de Datos del proyecto, asegurando que la recolección y uso de datos cumplen con las normativas éticas estipuladas por la Comisión Europea. **Humanized Solutions** solicitó y obtuvo acceso a estos datos con fines de análisis, respetando los lineamientos éticos y legales del proyecto europeo.

El Comité de Ética y Gobernanza de Datos de **MammoScreen** tiene como objetivo garantizar que el manejo de los datos recolectados se realice con transparencia, salvaguardando la confidencialidad de los participantes y aplicando altos estándares éticos en la protección de la privacidad y el consentimiento informado. Este comité también proporciona asesoramiento en torno a las buenas prácticas en la gestión de datos, identificando y mitigando posibles sesgos o conflictos éticos en los métodos de investigación. La supervisión del comité abarca todo el ciclo de vida del proyecto, desde la recolección de los datos hasta su análisis y publicación, para asegurar la integridad y la ética en cada etapa.

Para facilitar la transparencia, se ha incluido el enlace directo a la página oficial del proyecto en <https://mammoscreenproject.eu/governance/>, donde se detallan los principios éticos y normativas de gobernanza aplicadas. De este modo, se garantiza que la documentación cumple con los requisitos de calidad y transparencia demandados por los organismos de regulación y seguimiento.

Hipótesis y Objetivos

2.1 Pregunta de investigación principal:

¿En qué medida es útil el cuestionario de Evaluación del Riesgo Genético de Cáncer Hereditario para identificar factores de riesgo relevantes, y cómo se compara su eficacia en la identificación de estos factores dentro del modelo estadístico empleado?

Objetivo principal:

Evaluar la eficacia del cuestionario mediante pruebas estadísticas, identificando las preguntas que muestran una correlación significativa con el diagnóstico de cáncer y seleccionando aquellas que mejor predicen factores de riesgo hereditario.

2.2 Preguntas secundarias de investigación:

1. **Pregunta secundaria:** ¿Cómo varía la distribución de respuestas en el cuestionario según el grupo de edad de los encuestados?
Objetivo secundario: Calcular y analizar el porcentaje de respuestas para cada grupo de edad, con el fin de identificar patrones de riesgo y de respuesta específicos en distintos rangos etarios.
2. **Pregunta secundaria:** ¿Se puede automatizar el proceso de eliminación y filtrado de artículos obtenidos en un formato Excel obtenido de Publish or Perish en base a palabras clave del abstract y así agilizar el proceso de las búsquedas de literatura?
Objetivo secundario: Implementar y evaluar scripts que permitan la automatización de la limpieza y selección de artículos de acuerdo con criterios de relevancia identificados en el abstract.
3. **Pregunta secundaria:** ¿Cuáles son las organizaciones y asociaciones que se dedican a la investigación del cáncer de mama y el cáncer hereditario?
Objetivo secundario: Desarrollar un plan de comunicación y divulgación para informar a asociaciones de pacientes sobre los resultados del TFM y fomentar su participación en la difusión de estos hallazgos.

Metodología

3. Metodología

3.1. Estrategias de Búsqueda Bibliográfica

Para realizar una revisión exhaustiva de la literatura sobre el cáncer hereditario, centrada en la evaluación de riesgos, la concienciación pública y el asesoramiento genético, se desarrollaron varias estrategias de búsqueda. El propósito fue analizar sistemáticamente la literatura existente para entender mejor los factores clave y barreras que afectan el conocimiento y percepción del cáncer hereditario, así como los enfoques actuales en comunicación y educación en esta área. Las estrategias de búsqueda para cada una de ellas se pueden encontrar en **Anexo I, Anexo II y Anexo III**.

Búsqueda Inicial y Ajuste de Términos de Búsqueda:

La búsqueda inicial se realizó en **PubMed** y se dividió en tres basándose en subtemas, utilizando términos específicos y ajustando parámetros como el tipo de documento y el período de publicación.

En respuesta a la baja cantidad de resultados para ciertos temas específicos, se amplió el alcance de la búsqueda hacia **Google Scholar** para asegurar una mayor cobertura bibliográfica.

Uso de Herramientas para Optimización de Búsqueda:

Publish or Perish:

Utilizada para ajustar parámetros de búsqueda, identificar publicaciones clave y evaluar el impacto de los artículos mediante métricas bibliométricas como el índice h y el índice g (Harzing, 2007).

Automatización con Scripts en R:

Se desarrollaron dos scripts que automatizan la organización y filtrado de resultados:

1. Un script para la **eliminación de duplicados**, asegurando una base de datos depurada.
2. Otro script que aplica **criterios de inclusión y exclusión** específicos, categorizando artículos según palabras clave.

Los scripts se pueden consultar en el **Anexo VIII**.

Procedimientos de Búsqueda y Selección de Artículos

1. **Definición de Palabras Clave y Términos de Búsqueda:**
 - Se elaboró una lista de palabras clave para cada enfoque de revisión, detallada en los anexos.
2. **Selección de Bases de Datos y Ejecución de Búsquedas:**
 - Las búsquedas se realizaron inicialmente en **PubMed**. Cuando fue necesario (búsqueda tres), se expandieron a **Google Scholar** para temas con baja representación en PubMed.
3. **Filtros de Inclusión y Exclusión:**
 - Se aplicaron criterios de inclusión y exclusión con base en el título y el resumen de los artículos. Esto permitió concentrarse en estudios relevantes para los objetivos de la revisión, eliminando artículos tangenciales.
4. **Extracción y Clasificación de Datos:**
 - **Extracción de Datos:** Se recopiló información relevante de cada estudio mediante formularios estandarizados.
 - **Clasificación Temática:** Los estudios se categorizaron en áreas temáticas según su relevancia y enfoque, dividiendo los resultados en tres principales categorías: Evaluación de Riesgos en Cáncer Hereditario, Conciencia y Actitudes hacia el Asesoramiento Genético, y Campañas Educativas sobre el Cáncer Hereditario.

Análisis de Datos

1. **Agrupación y Comparación de Resultados:**
 - Los artículos se organizaron por categorías para facilitar su análisis comparativo.
2. **Síntesis de Resultados:**
 - La información fue sintetizada para identificar patrones y diferencias en la percepción y conocimiento del cáncer hereditario, asesoramiento genético y campañas educativas.
3. **Automatización del Análisis Bibliométrico:**
 - Los scripts en R permitieron la generación de informes bibliométricos (disponible en el anexo), que facilitan la identificación de tendencias en la literatura y el impacto de publicaciones en el campo.

Validación y Revisión

1. Consulta con Tutores y Revisión por Pares:

- Los hallazgos fueron discutidos con tutores y compañeros para obtener retroalimentación adicional.

2. Revisión Final y Edición:

- La metodología y resultados se revisaron de manera integral, incorporando sugerencias y realizando ajustes finales para asegurar la claridad y precisión del análisis.

3.2 Estrategia del análisis del Cuestionario

Codificación y Tratamiento de Variables

Las respuestas del cuestionario se codificaron en varios niveles para reflejar fielmente la variedad de respuestas. "Si" se codificó como **1**, "No" como **0**, y "No lo sé" se trató como una categoría separada "**-1**", permitiendo conservar las respuestas de incertidumbre y analizar patrones que incluyan estas variaciones. Aunque se exploró profundamente como influía el "No lo sé" y se analizó también su conjunción con el "No".

Análisis Estadístico

Distribución por Grupos de Edad y Análisis de Porcentajes:

Las participantes fueron clasificadas en seis grupos de edad específicos, para evaluar las variaciones en las respuestas y detectar patrones de riesgo asociados a cada grupo etario. La división de grupos de edad sigue la metodología empleada por Zhang et al. (2012), lo que facilita la comparación con estudios previos sobre factores de riesgo en diferentes cohortes de edad. Los porcentajes de respuesta para cada grupo etario se presentan detalladamente en el **Anexo VII**.

Prueba Exacta de Fisher:

Dado que la prueba chi-cuadrado resultó inadecuada debido a frecuencias bajas en ciertas celdas, se empleó la prueba exacta de Fisher para evaluar la asociación entre las respuestas relacionadas con el riesgo genético y la variable dependiente de diagnóstico de cáncer de mama . **Anexo V**.

Regresión Logística:

Se aplicó un modelo de regresión logística, tomando como **variable dependiente** el diagnóstico de cáncer de mama (**Si/No**), e incluyendo como variables predictivas aquellos factores que presentaron asociaciones significativas, tales como el historial de cáncer en generaciones consecutivas y diagnósticos en familiares varones **Anexo V**.

Presentación de Resultados y Validación

Reporte de Resultados:

Los hallazgos se presentan en tablas y gráficos, donde se reflejan tanto los resultados de los análisis de regresión como las distribuciones porcentuales, facilitando una interpretación visual clara de los patrones de riesgo observados.

Validación:

Se realizaron revisiones cruzadas y un reanálisis de variables clave para garantizar la consistencia y robustez de los resultados obtenidos, asegurando la fiabilidad en las conclusiones del análisis.

Para una lectura más detallada sobre el proceso de análisis, los porcentajes específicos y los resultados obtenidos, se puede consultar el **Anexo VII** y el **Anexo V**. Esta estrategia, que incluye un tratamiento detallado de las respuestas, análisis estadístico y validación rigurosa, asegura una evaluación completa y precisa del riesgo genético de cáncer de mama, proporcionando una base sólida para la toma de decisiones en salud preventiva.

3.4. Estrategia inicial de comunicación y divulgación

Una estrategia de comunicación efectiva es clave para maximizar el impacto del proyecto de evaluación de riesgo de cáncer de mama de la plataforma **EVITA**, permitiendo sensibilizar sobre la detección temprana y el manejo del riesgo. A continuación, los aspectos centrales de esta estrategia:

1. Objetivos de Comunicación

- **Difusión de Objetivos:** Comunicar los objetivos y logros del proyecto a públicos clave.
- **Colaboraciones:** Crear redes con asociaciones y organizaciones en cáncer de mama.
- **Educación y concienciación:** Promover la importancia de la predisposición genética en el cáncer.

2. Público Objetivo

- **Científicos y Profesionales:** Investigadores, empresas biotecnológicas y hospitales.
- **Instituciones Financieras:** Agencias de financiamiento y ONG.
- **Público General:** Personas interesadas en la ciencia y la salud.

3. Mensajes Clave

- **Impacto en Ciencia y Sociedad:** Resaltar cómo los hallazgos pueden influir en la concienciación y prevención del cáncer.
- **Innovación y Aplicación Práctica:** Explicar las herramientas innovadoras (como **EVITA** y **MammoScreen**) y su aplicación en la práctica clínica.
- **Colaboración y Financiación:** Incentivar la participación de nuevos socios en el proyecto.

4. Canales de Comunicación

- **Publicaciones Científicas y Conferencias:** Compartir resultados en revistas y eventos especializados.
- **Redes Sociales:** Publicaciones en LinkedIn y blogs profesionales.
- **Webinars y Sitio Web del Proyecto:** Webinars para profundizar en los hallazgos y un sitio web de actualización continua.

5. Plan de Acción y Medición de Impacto

- **Creación de Contenidos:** Desarrollo de artículos, infografías y videos.
- **Medición de Resultados:** Monitorear impacto a través de métricas como menciones, citas y cobertura mediática.

6. Evaluación y Retroalimentación

- **Encuestas y Análisis de Métricas:** Obtener opiniones y ajustar la estrategia según los resultados de alcance y participación.

Esta estrategia busca una comunicación integral y efectiva, maximizando el alcance de los resultados de **EVITA** y fortaleciendo la concienciación sobre el riesgo genético en cáncer de mama.

3.3.1. Mapeo de asociaciones

Aunque no se ha llevado a cabo un análisis exhaustivo de asociaciones de pacientes, en la presente sección se incluye una breve descripción de la metodología y los pasos a seguir en proyectos futuros. Este enfoque preliminar permite establecer una base de trabajo y una demostración de cómo se pretende realizar evaluaciones de asociaciones de pacientes para maximizar el impacto y la colaboración en futuros proyectos.

Metodología Propuesta

Paso 1: Identificación de Asociaciones Clave

En primer lugar, se realizó una consulta inicial para identificar algunas de las asociaciones más relevantes de cáncer de mama en países específicos (Francia, Alemania, Portugal, Italia, España y Reino Unido). Esta búsqueda preliminar proporciona una base de datos inicial sobre las organizaciones que podrían ser fundamentales en la divulgación y sensibilización en cada región.

Paso 2: Búsqueda y Recolección de Información

Para proyectos futuros, se propone una búsqueda exhaustiva que abarque las siguientes herramientas:

- **Google:** Para obtener información sobre la misión, visión y actividades principales de cada asociación.
- **LinkedIn:** Para identificar redes profesionales y explorar la estructura organizativa de las asociaciones.
- **Otros motores de búsqueda:** Como complemento, para garantizar una cobertura más completa de cada organización.

Paso 3: Recolección de Características de las Asociaciones

La fase de recopilación de información se diseñará para incluir aspectos relevantes como:

- **Presencia en Redes Sociales:** Documentación de las plataformas activas (Facebook, Twitter, Instagram, etc.).
- **Idiomas Disponibles:** Identificación de los idiomas en los que cada asociación comunica su información.
- **Comité Ético o Científico:** Verificación de comités que aporten credibilidad a las actividades de la asociación.
- **Foco en Tipos de Cáncer:** Determinación de si la asociación se dedica exclusivamente al cáncer de mama o aborda otros tipos.
- **Otras Características:** Registro de programas de apoyo, actividades educativas y campañas.

Paso 4: Organización en Tabla Excel

Se estructurará la información en una tabla Excel, facilitando así su análisis y permitiendo un enfoque sistemático y comparable en proyectos posteriores.

Paso 5: Análisis Descriptivo de Datos en R

Para futuros análisis, los datos serán cargados en R, donde se realizarán:

- **Estadísticas Descriptivas:** Para comprender la distribución de características entre las asociaciones.
- **Visualización de Datos:** Representación gráfica de las diferencias y patrones.
- **Análisis Comparativo:** Comparación de asociaciones por país para identificar enfoques y actividades relevantes.

Este enfoque sistemático y detallado permite no solo mapear las asociaciones de cáncer de mama de manera efectiva, sino también analizar y comprender mejor sus características y su impacto a nivel nacional e internacional información importante de cara a conocer mejor lo que se está llevando a cabo en el área del cáncer de mama y las campañas que se llevan a cabo a través de las asociaciones.

Resultados

4. Resultados:

4.1. Revisiones de literatura

Tabla 1.1 Resultados de las búsquedas bibliográficas

Revisiones final	Resultados iniciales	Duplicados	Resultados finales
Risk assessment	154	42	112
Awareness	349	165	184
Campañas educativas	122	0	122

Tabla 1.2. Resultado del filtro creado en R para la búsqueda uno:

	Genetic risk assessment	Questionnaire	Breast cancer	Awareness , attitudes or perceptions	Review
Risk assessment	3	10	86	31	19

Tabla 1.3. Resultado del filtro creado en R para la búsqueda dos:

	Genetic screening or counseling	Familial or hereditary cancers	Breast cancer	Awareness, attitudes or perceptions	Review
Awareness	68	43	132	62	182

Tabla 1.4. Resultado del filtro creado en R para la búsqueda tres:

	Breast cancer	Cancer	Review
Campañas educativas	23	109	4

1. Resultados Generales de las Búsquedas Bibliográficas (Tabla 1.1)

La Tabla 1.1 resume los resultados iniciales de cada búsqueda, mostrando el número de artículos relevantes después de eliminar los duplicados:

- **PPIE en cáncer:** De los 154 artículos iniciales, 42 eran duplicados, lo que resultó en 112 estudios únicos. Esto demuestra un interés creciente en la participación pública y de pacientes en el contexto del cáncer.
- **Awareness (Concienciación):** Con un total de 349 resultados iniciales y 184 artículos únicos, la concienciación es el tema con mayor cantidad de estudios, subrayando su relevancia en términos de percepción pública y prevención del cáncer.
- **Campañas educativas:** Se identificaron 122 estudios únicos, todos provenientes de una búsqueda realizada en Google Scholar, ya que en PubMed se encontró una disponibilidad limitada de artículos relevantes en esta área. Es importante señalar que, debido a esta restricción en las fuentes, se optó por una búsqueda más general en Google Scholar para obtener una base inicial de información.

2. Análisis Detallado de "PPIE en cáncer" (Tabla 1.2)

Para la búsqueda sobre participación pública y de pacientes (PPIE) en cáncer, se observan las siguientes distribuciones temáticas:

- **Cáncer de mama** es el tema más común, con 86 artículos, lo que refleja un enfoque específico en este tipo de cáncer dentro de los estudios de participación pública.
- **Conciencia, actitudes o percepciones** también destaca con 31 artículos, indicando interés en comprender cómo las actitudes del público influyen en la participación en iniciativas de investigación en cáncer.
- **Revisiones:** Se encontraron 19 artículos de revisión, lo cual sugiere que existe una base consolidada en el tema de PPIE en cáncer.
- **Evaluación de riesgo genético y cuestionarios** aparecen en menor medida, con solo 3 y 10 artículos, respectivamente, lo que apunta a posibles áreas de investigación emergente en el campo de la participación en salud oncológica.

3. Análisis Detallado de "Awareness" (Tabla 1.3)

La búsqueda sobre concienciación mostró una variedad de temas relevantes:

- **Cáncer de mama** aparece como el tema predominante, con 132 estudios, consolidando su relevancia en la literatura sobre concienciación.
- **Revisiones:** Se encontraron 182 artículos de revisión, indicando un esfuerzo sustancial por sintetizar el conocimiento sobre concienciación en cáncer, lo cual puede facilitar el desarrollo de intervenciones basadas en evidencia.
- **Conciencia, actitudes o percepciones** está bien representado, con 62 estudios, reflejando el interés en entender las percepciones y actitudes de las personas hacia el cáncer.
- Otros temas como **consejería o tamizaje genético y cáncer hereditario o familiar** también son significativos, con 68 y 43 artículos, respectivamente, subrayando la importancia de estos factores en la concienciación.

4. Análisis Detallado de "Campañas Educativas" (Tabla 1.4)

La búsqueda sobre campañas educativas se realizó exclusivamente en Google Scholar debido a la falta de información suficiente en PubMed. Esta decisión afectó el enfoque y los resultados de la búsqueda, resultando en una aproximación más amplia pero menos específica:

- **Cáncer en general y cáncer de mama** fueron los temas principales, con 109 y 23 artículos, respectivamente. Esto refleja una orientación general hacia el papel de las campañas educativas en la prevención del cáncer, aunque es probable que existan vacíos específicos debido al enfoque general de la búsqueda.
- **Revisiones:** Solo se encontraron 4 artículos de revisión, lo cual sugiere que existe poca consolidación en la literatura sobre el impacto de las campañas educativas en cáncer. Esto podría ser una consecuencia de la limitación de fuentes y de la necesidad de búsquedas más detalladas en bases de datos especializadas en futuras investigaciones.

Interpretación General

Los resultados reflejan que la literatura sobre participación pública, concienciación y campañas educativas en cáncer se centra fuertemente en el cáncer de mama, lo cual podría reflejar la alta prevalencia y el interés por la prevención y detección temprana de esta enfermedad. La abundancia de revisiones, especialmente en el área de concienciación, indica un esfuerzo por consolidar y sintetizar el conocimiento existente, lo cual puede orientar intervenciones futuras.

En el caso de **campañas educativas**, la búsqueda más general en Google Scholar permitió obtener una base inicial de información, pero es necesario realizar búsquedas más específicas en bases de datos adicionales para un análisis más completo. La elección de Google Scholar fue una solución práctica debido a la limitada disponibilidad de literatura en PubMed sobre este tema, lo que sugiere que este es un campo menos explorado en las bases de datos biomédicas tradicionales.

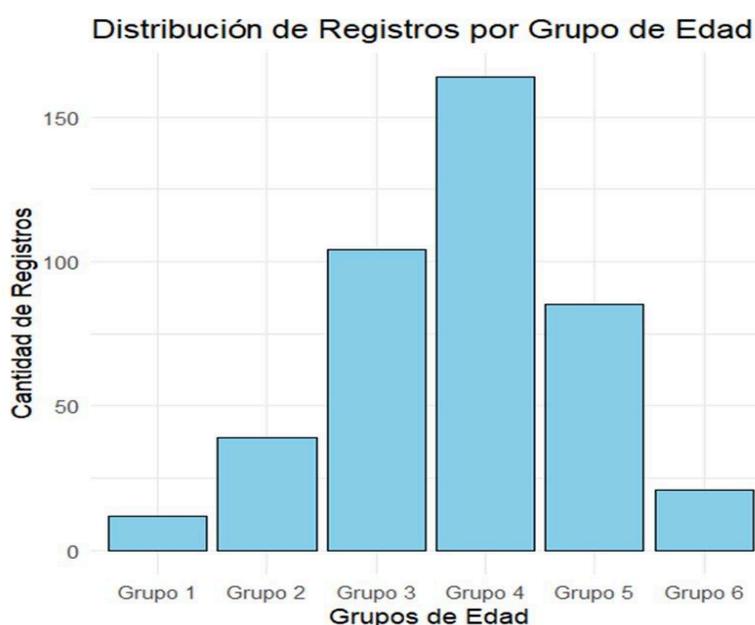
En conclusión, este análisis preliminar proporciona una base de datos y una metodología inicial que pueden guiar futuras investigaciones, especialmente en la identificación de áreas menos exploradas, como el impacto de las campañas educativas y la integración de herramientas como cuestionarios y evaluaciones de riesgo genético en la participación pública y de pacientes con cáncer. Para una explicación más detallada **Anexo IV**.

4.2 Análisis del cuestionario

4.2.1 Análisis de los porcentajes

Como se ha demostrado en estudios previos que emplearon metodologías similares, categorizamos a los participantes en seis grupos de edad distintos.

Figura 2. Gráfica de distribución de grupos de edad



Categorizamos a los participantes en seis grupos de edad distintos de la siguiente manera:

- **Grupo 1 (18–29 años)**
- **Grupo 2 (30–39 años)**
- **Grupo 3 (40–49 años)**
- **Grupo 4 (50–59 años)**
- **Grupo 5 (60–69 años)**
- **Grupo 6 (70 años o más)**

Para cada grupo, calculamos la distribución porcentual de sus respuestas a cada pregunta de la encuesta. Los resultados de este análisis se presentan a continuación, proporcionando un desglose detallado de cómo respondió cada grupo de edad a las distintas preguntas del cuestionario.

- **Grupo 1:** 12 participantes
- **Grupo 2:** 39 participantes
- **Grupo 3:** 104 participantes
- **Grupo 4:** 164 participantes
- **Grupo 5:** 85 participantes
- **Grupo 6:** 21 participantes

A continuación, se presentará únicamente el análisis de los porcentajes correspondientes a las preguntas relevantes para los análisis posteriores. Para consultar todos los resultados detallados con su respectiva interpretación, se puede referir al **Anexo VII**.

Dos o más familiares paternos diagnosticados con el mismo tipo de cáncer:

La mayoría de los grupos respondió "**No**" (rango entre 24% y 58%), seguido de "**No aplica**" y "**No lo sé**" en menor medida. El porcentaje de "**Sí**" varía entre un 15% y 21%, siendo más frecuente en grupos de edad media.

Cáncer en dos o más generaciones consecutivas de la familia:

En general, el 33% al 58% respondió "**No**", mientras que "**Sí**" tuvo mayor representación en los grupos de edad media y mayor, alcanzando hasta el 43%. Las respuestas de "**No lo sé**" se mantuvieron en niveles bajos (8%-14%).

Cáncer bilateral en la familia:

Predominan las respuestas "**No**" en todos los grupos (42%-62%), seguidas por "**No aplica**" y un menor porcentaje de "**Sí**" (entre 8% y 14%).

Familiares varones diagnosticados con cáncer de mama:

La respuesta mayoritaria fue "**No**" en todos los grupos (52%-75%), con "**Sí**" en porcentajes mínimos (1%-8%). La opción "**No aplica**" y "**No lo sé**" presentan frecuencias menores, sin variaciones significativas entre grupos.

4.2.2 Construcción de un modelo de regresión logística

4.2.2.1 Prueba exacta de Fisher

Dado el tipo de respuestas en el cuestionario, la mayoría eran categóricas y nominales, sin un orden específico (por ejemplo, "Sí", "No", "No lo sé"). Algunas preguntas también presentaban una estructura dicotómica, limitándose a respuestas de "Sí" o "No". Las respuestas faltantes, representadas como "NA", debían manejarse adecuadamente para asegurar la validez del análisis de datos. Para el análisis estadístico, estas respuestas categóricas se codificaron en valores numéricos, asignando, por ejemplo, "Sí" = 1, "No" = 0, y tratando "No lo sé" como un valor intermedio o nulo, según los requerimientos del análisis.

Inicialmente, la prueba de chi-cuadrado se consideró el método estadístico apropiado para analizar la relación entre variables categóricas. No obstante, al aplicarla, R mostró la advertencia "**Chi-squared approximation may be incorrect**", común cuando las frecuencias esperadas en algunas celdas son bajas, invalidando la aproximación chi-cuadrado. Esto pudo deberse a un tamaño de muestra pequeño o a la presencia de categorías infrecuentes que generaban celdas vacías o escasamente pobladas en las tablas de contingencia.

Como solución, se utilizó la **prueba exacta de Fisher**, más adecuada para tablas de contingencia pequeñas o con frecuencias bajas, ya que proporciona resultados exactos sin depender de la aproximación chi-cuadrado. Esta prueba permitió obtener resultados más fiables para evaluar la asociación entre las variables del cuestionario y el diagnóstico de cáncer de mama (Kim, 2017).

Para presentar los resultados de la prueba exacta de Fisher, se emplearán dos tablas de análisis. En la **Tabla 2.1**, se muestran los resultados obtenidos manteniendo las respuestas en sus categorías originales: "Sí", "No" y "No lo sé". Este enfoque permite capturar los matices de las respuestas, incluyendo la incertidumbre reflejada en "No lo sé". En la **Tabla 2.2**, se presentan los resultados tras la transformación de los datos, donde "No" y "No lo sé" se agrupan en una misma categoría binaria. Este ajuste simplifica las categorías y elimina la ambigüedad, permitiendo un análisis más directo.

Los resultados más significativos de cada análisis se destacan a continuación. Para obtener una vista detallada de todos los resultados, consulte el **Anexo V**.

Tabla 2.1: Análisis con Respuestas Originales

Pregunta	p valor	Interpretación
¿Hay dos o más familiares paternos diagnosticados con el mismo tipo de cáncer?	0.0152	Significativa: Indica que un historial de cáncer paterno puede ser un predictor relevante para el riesgo de cáncer de mama.
¿Hay cáncer en dos o más generaciones consecutivas?	0.0056	Significativa: Señala que el cáncer hereditario en generaciones consecutivas es un importante factor de riesgo.
¿Hay cáncer de mama en varones de la familia?	0.0770	No significativa, pero con tendencia: Sugiere una posible relación que puede investigarse más a fondo.

En esta segunda tabla, los resultados se obtuvieron tras simplificar las respuestas al agrupar "No" y "No lo sé" en una sola categoría binaria. Este análisis elimina la incertidumbre, permitiendo un enfoque más claro y directo en la relación entre antecedentes familiares y el riesgo de cáncer de mama. Al simplificar las categorías, algunos factores muestran asociaciones más claras y estadísticamente significativas, facilitando una interpretación más robusta de los datos.

Tabla 2.2: Análisis con Respuestas Transformadas

Pregunta	p valor	Interpretación
¿Algún familiar ha tenido cáncer bilateral?	0.0559	Marginalmente significativa: La transformación incrementa su relevancia, sugiriendo que podría considerarse como predictor marginal.
¿Hay cáncer en dos o más generaciones consecutivas?	0.0007	Significativa: Refuerza la hipótesis de que los patrones hereditarios de cáncer en múltiples generaciones son un predictor clave.
¿Hay cáncer de mama en varones de la familia?	0.0453	Significativa: En el análisis transformado, este factor muestra relevancia en el riesgo de cáncer de mama.

Estos resultados destacan algunos factores familiares que parecen tener una relevancia considerable en el riesgo de cáncer de mama, especialmente cuando analizamos los datos en su forma original y luego transformada.

La presencia de cáncer en generaciones consecutivas sobresale en ambos análisis como un predictor fuerte, lo cual refuerza la idea de un componente hereditario sólido en el riesgo de cáncer.

Además, aunque en el análisis original el cáncer de mama en varones familiares sólo sugería una posible relación, al simplificar las categorías se vuelve estadísticamente significativo, subrayando la relevancia de considerar estos casos poco comunes en el contexto familiar.

Finalmente, el cáncer bilateral también se muestra como un posible factor marginal que merece un análisis más profundo. En conjunto, estos hallazgos apuntan a ciertos patrones familiares que, con más investigación, podrían ser clave en la predicción y evaluación de riesgo para cáncer de mama.

4.2.2 Matriz de Confusión y Estadísticas Clave

1. Intercepto

El intercepto en un modelo de regresión representa el valor predicho cuando todas las variables independientes son cero. En este caso, su valor es de -0.08441 con un p-valor de 0.58906, lo que indica que no es significativo y no contribuye sustancialmente a predecir el riesgo de cáncer de mama. Este resultado sugiere que el intercepto no mejora el ajuste del modelo.

2. Exactitud (Accuracy)

La exactitud mide la proporción de predicciones correctas sobre el total de observaciones, permitiendo evaluar la capacidad global del modelo para clasificar correctamente. La exactitud del modelo es del 60.62%, indicando un rendimiento moderado en la clasificación de los individuos. Este nivel de exactitud es adecuado pero no sobresaliente, lo que sugiere la necesidad de mejorar el modelo o añadir más variables predictoras (Holmes y Huber, 2019).

3. Índice Kappa

El índice Kappa cuantifica el grado de concordancia entre las predicciones y los valores reales, ajustando por el acuerdo esperado por azar. Con un valor de 0.216, este índice indica un acuerdo bajo, aunque estadísticamente significativo, sugiriendo que el modelo tiene una capacidad limitada para mejorar sobre el azar. Este índice es útil en modelos donde el acuerdo más allá del azar es crucial para validar la precisión de las predicciones (Holmes y Huber, 2019).

4. Sensibilidad y Especificidad

La sensibilidad mide la proporción de verdaderos positivos (casos correctamente identificados de cáncer de mama), mientras que la especificidad refleja los verdaderos negativos (individuos sanos clasificados correctamente). En este modelo, la sensibilidad es del 64.29%, mientras que la especificidad es del 57.84%. Estos valores indican que el modelo tiene un rendimiento moderado en la detección de ambos grupos, con una leve ventaja en la identificación de casos positivos (Holmes y Huber, 2019).

5. Valores Predictivos Positivo (VPP) y Negativo (VPN)

El valor predictivo positivo (VPP) representa la precisión de las predicciones positivas, mientras que el valor predictivo negativo (VPN) indica la precisión de las predicciones negativas. En este modelo, el VPP es de 53.57%, y el VPN es de 68.15%, lo que significa que el modelo es más fiable al clasificar correctamente los individuos sin cáncer de mama. Estas métricas son fundamentales para interpretar la utilidad del modelo en el contexto de predicciones clínicas (Holmes y Huber, 2019).

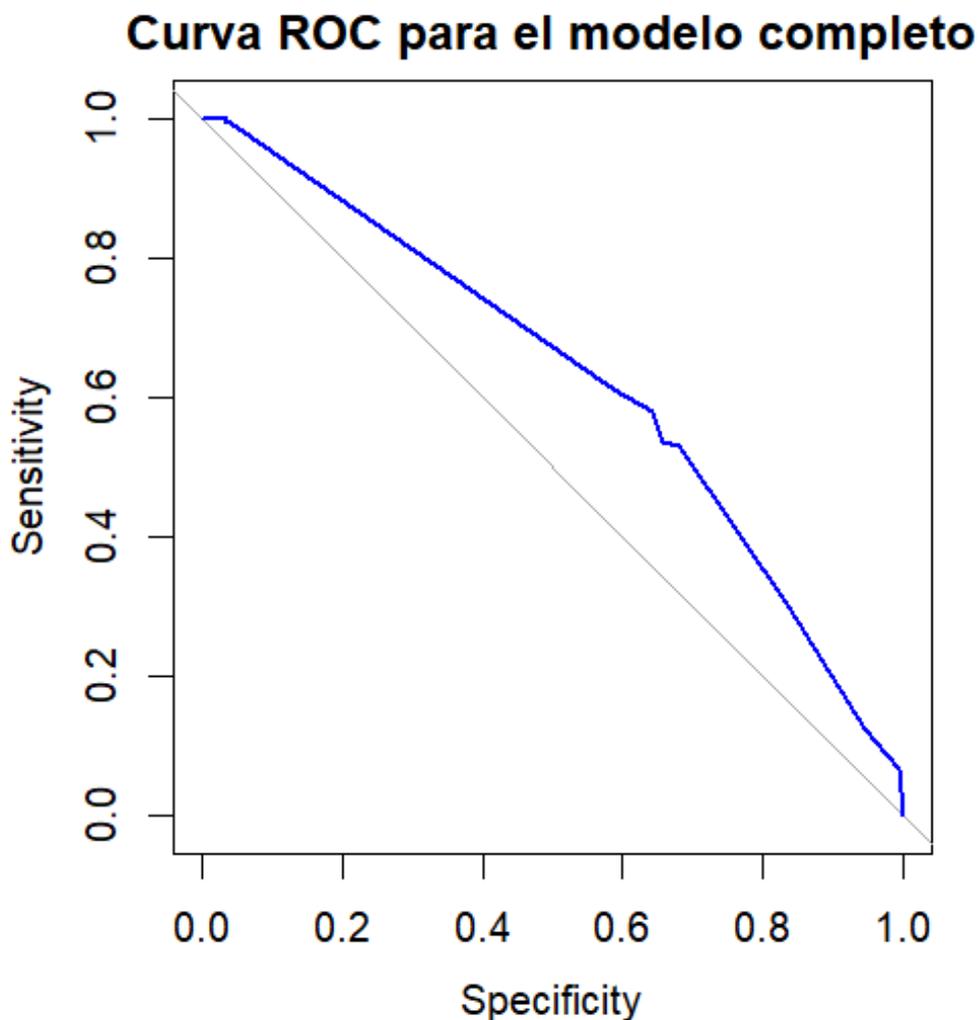
6. Test de McNemar

El Test de McNemar evalúa las diferencias entre errores de clasificación, especialmente entre falsos positivos y falsos negativos. En este caso, el p-valor de 0.01701 indica una diferencia significativa, lo que sugiere que el modelo puede tener un sesgo en sus predicciones, potencialmente clasificando incorrectamente más casos en un sentido que en otro (Holmes y Huber, 2019).

4.2.3 Curva ROC

La **curva ROC (Receiver Operating Characteristic)** es una herramienta fundamental para evaluar la capacidad de un modelo de clasificación en la distinción entre dos clases, como en este caso, la presencia o ausencia de cáncer de mama. Esta gráfica muestra la relación entre la tasa de verdaderos positivos (sensibilidad) y la tasa de falsos positivos, mientras que el área bajo la curva (AUC) mide su rendimiento: un AUC cercano a 1 indica un buen desempeño, mientras que un AUC cercano a 0.5 sugiere que el modelo no es mejor que el azar. Como explican Holmes y Huber en *Modern Statistics for Modern Biology*, tanto la curva ROC como el AUC son métricas clave para interpretar y comparar el poder discriminante de variables predictoras en modelos de clasificación (Holmes y Huber, 2019).

Figura 3. Curva ROC de nuestro modelo



Análisis de la curva ROC y rendimiento del modelo de regresión logística

El análisis de la curva ROC y otras métricas del modelo de regresión logística muestra que este modelo proporciona un rendimiento adecuado en la clasificación del riesgo de cáncer de mama. Aunque la curva ROC no es ideal y se aproxima a la diagonal, el modelo mantiene una capacidad razonable para distinguir entre individuos con y sin cáncer de mama.

Variables Significativas: Dos variables clave han demostrado ser estadísticamente significativas:

- **Cáncer en generaciones consecutivas:** Con un odds ratio de 2.13 y un p-valor significativo de 0.00312, esta variable indica que un historial familiar de cáncer en generaciones consecutivas está fuertemente asociado con un mayor riesgo de cáncer de mama.
- **Cáncer de mama en familiares varones:** Con un odds ratio de 0.09 y un p-valor de 0.03060, esta variable muestra una relación inversa. Aunque el cáncer de mama en hombres es menos común, su presencia puede influir en el riesgo general.

Rendimiento General del Modelo: El rendimiento del modelo es satisfactorio, con un AIC de 433.45 que indica un buen ajuste para este conjunto de datos. Aunque el AUC cercano a 0.5 sugiere que la curva ROC no es óptima, el modelo muestra una capacidad predictiva que puede complementarse con la incorporación de variables adicionales o el análisis de interacciones. La presencia de variables no significativas, como el historial de cáncer en la familia paterna y el cáncer bilateral, resalta que no todos los antecedentes familiares aportan el mismo nivel de predictividad en este contexto.

Para una explicación completa y detallada de estos resultados y métricas del modelo, consulte el **Anexo V**.

4.2.4 Tablas de contingencia

Las **tablas de contingencia** se utilizaron para evaluar la relación entre ciertos factores de riesgo y la probabilidad de desarrollar cáncer de mama. Estas tablas permiten observar la asociación entre variables categóricas como el historial familiar de cáncer en generaciones consecutivas, la presencia de cáncer en varones familiares y los casos de cáncer bilateral, en comparación con el diagnóstico de cáncer de mama.

La estructura de la tabla incluye las predicciones del modelo (sano o enfermo) y los diagnósticos reales, permitiendo evaluar la precisión de las clasificaciones. Además, las gráficas ROC para cada variable permiten analizar su capacidad predictiva individual en la identificación de casos de cáncer.

Tabla 3. Tabla de contingencia

	Predicted = 0	Predicted = 1	Total
Real = 0 (Sano)	90	50	140
Real = 1 (Enfermo)	78	107	185
Total	168	157	325

La **Tabla 3 de Contingencia** muestra cómo el modelo clasifica los casos de cáncer de mama:

- **Predicciones Correctas:** 107 personas fueron correctamente predichas como enfermas y 90 como sanas.
- **Errores:** El modelo clasificó incorrectamente a 50 personas como enfermas cuando estaban sanas (falsos positivos) y a 78 personas como sanas cuando en realidad estaban enfermas (falsos negativos).

La tabla permite evaluar la **exactitud** general del modelo y sus capacidades de **sensibilidad** y **especificidad**, proporcionando una visión clara de su rendimiento en la identificación de casos de cáncer.

4.2.5 Curvas ROC para las variables individuales

Las curvas ROC de cada variable por separado muestran un bajo poder predictivo, lo que indica que ninguna de ellas, individualmente, puede discriminar bien entre casos de cáncer de mama y no casos. Esto sugiere que el poder predictivo del modelo depende de la combinación de todas las variables. En particular:

- **Historial de cáncer paterno:** Significativa en el análisis, pero con bajo poder predictivo individual.
- **Cáncer bilateral:** Marginalmente significativa, pero con poca capacidad predictiva individual.
- **Cáncer en generaciones consecutivas:** Asociada significativamente, pero con capacidad predictiva limitada por sí sola.
- **Cáncer de mama en varones biológicos:** Muestra un poder predictivo muy bajo.

Figura 4. Curva ROC variable generaciones consecutivas

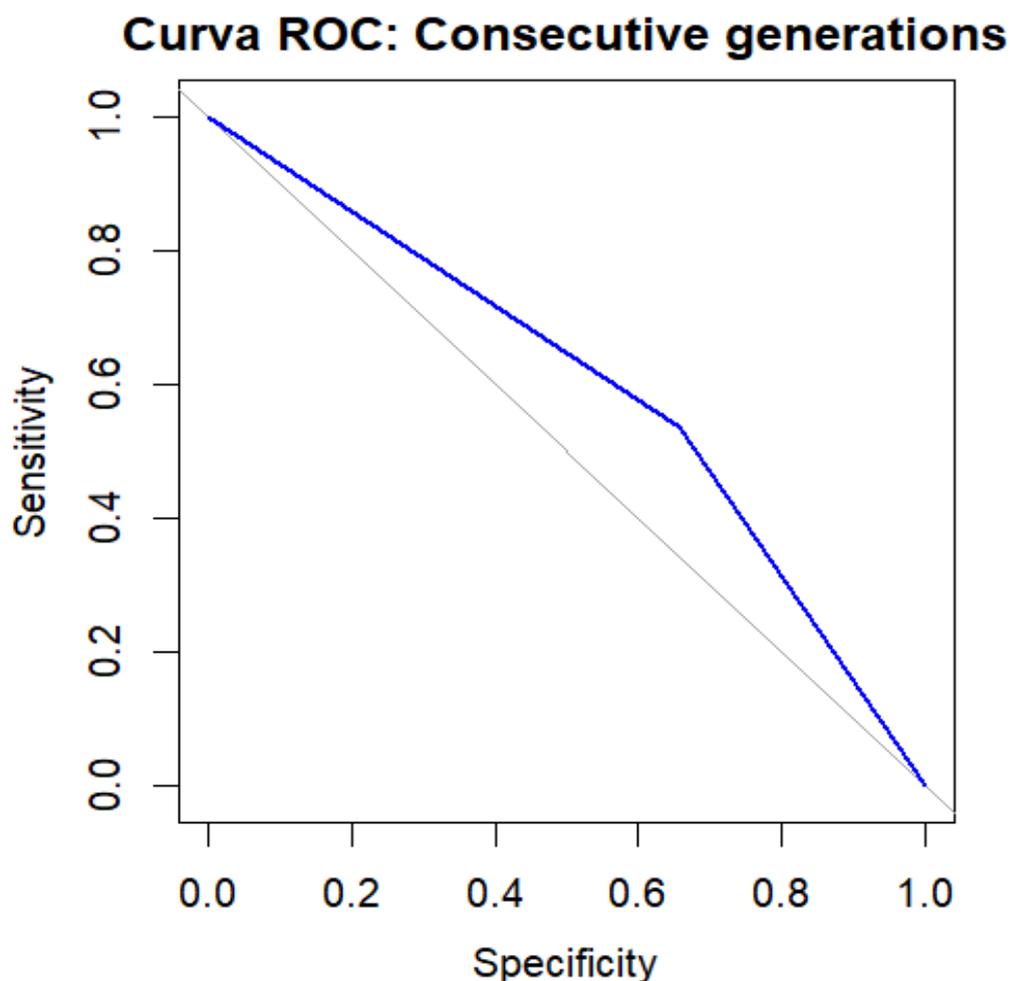
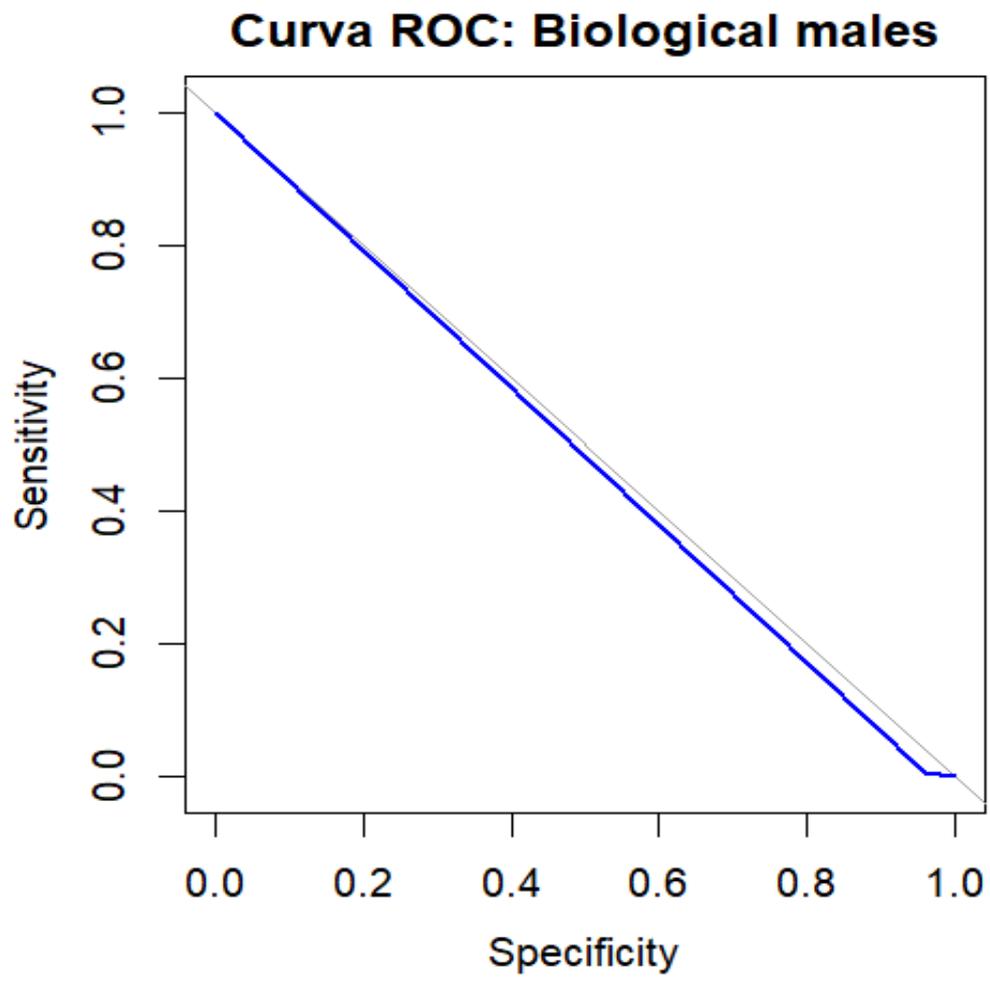


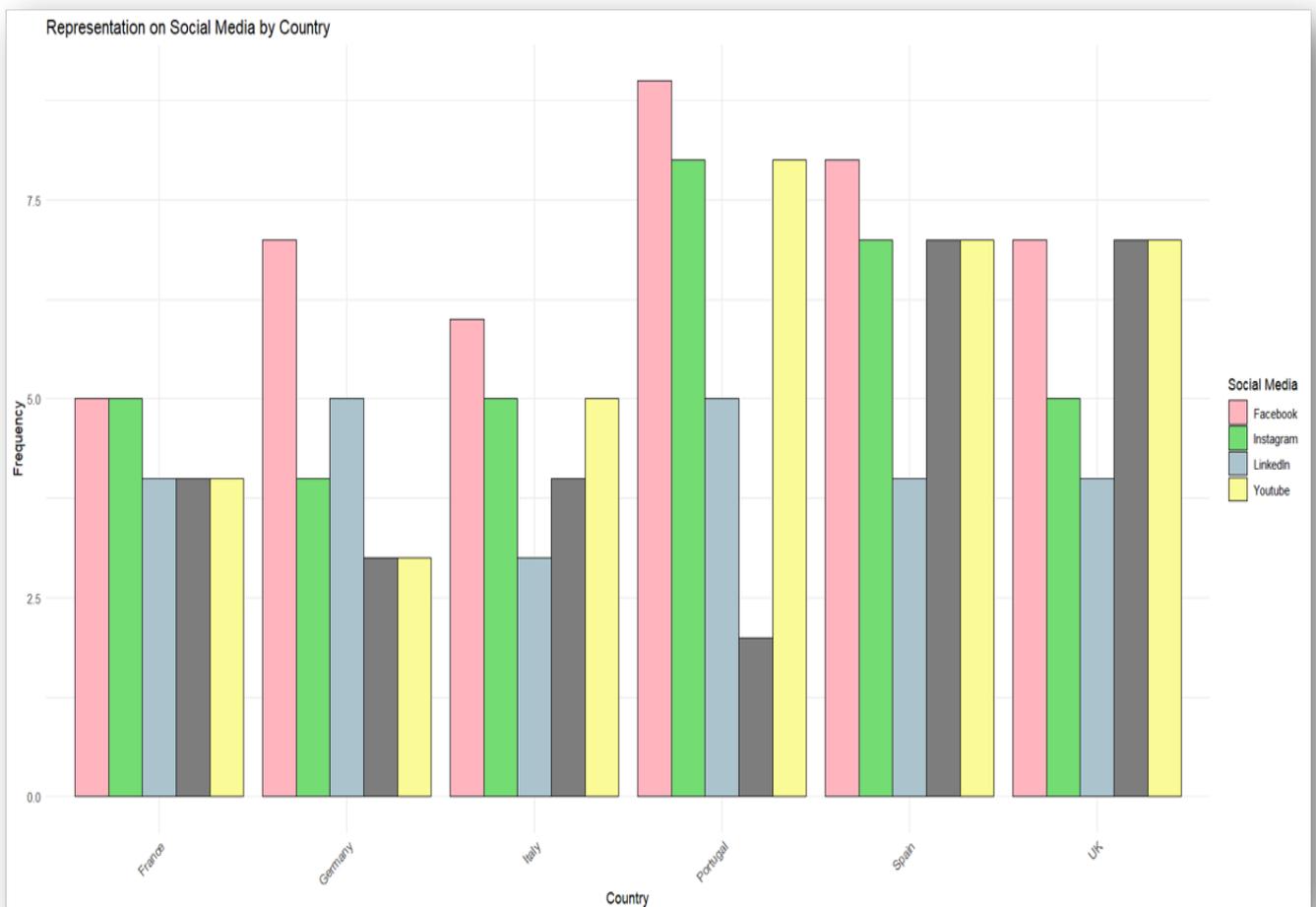
Figura 5. Curva ROC variable generaciones consecutivas



4.3 Análisis de asociaciones

Esta gráfica preliminar ilustra el tipo de análisis que se planea realizar para identificar las tendencias de comunicación en redes sociales de las asociaciones de cáncer de mama en cada país estudiado. Este enfoque permite comparar cómo las organizaciones utilizan diferentes plataformas para sus campañas de sensibilización y apoyo.

Figura 6. Diagrama de barras que representa la presencia en redes sociales de las asociaciones en cada país



Se realizarán más análisis futuros de este tipo para obtener más conclusiones de nuestras búsquedas.

Los datos se encuentran en el Anexo VI.

Conclusiones y discusión

5.1. Conclusión de las búsquedas bibliográficas

Para la primera parte de las conclusiones, centrémonos en analizar los resultados obtenidos de las búsquedas bibliográficas.

Para una comprensión detallada de los resultados obtenidos en cada una de las estrategias de búsqueda y un análisis completo de los estudios revisados, se puede consultar el **Anexo IV**, donde se exponen en profundidad los hallazgos, metodologías y discusiones de cada tema. A continuación, se presentan los resúmenes de las tres búsquedas temáticas realizadas:

Búsqueda 1: Evaluación de Riesgos en Cáncer Hereditario

En resumen, los resultados de esta primera búsqueda reflejan una alta representación del cáncer de mama en el ámbito de la evaluación de riesgos en cáncer hereditario, mientras que otros tipos de cáncer hereditario reciben menos atención. La aplicación de filtros específicos evidenció una gran cantidad de artículos, especialmente revisiones y estudios sobre cáncer de mama, lo que confirma el interés académico en esta área y la necesidad de desarrollar herramientas de evaluación más amplias y diversificadas para cubrir otros tipos de cáncer hereditario. Esta conclusión inicial establece una base para explorar áreas específicas menos estudiadas y enfatiza la importancia de herramientas que no sólo evalúen el riesgo, sino que también consideren la percepción y aceptación del público hacia estas herramientas.

Búsqueda 2: Conciencia y Actitudes hacia el Asesoramiento Genético

En conclusión, la revisión sobre la conciencia y las actitudes hacia el asesoramiento genético destaca el creciente interés en comprender las percepciones y barreras que afectan el acceso a estos servicios. Si bien el cáncer de mama sigue siendo el tipo de cáncer más estudiado, otros cánceres hereditarios también se abordan, aunque en menor medida. La clasificación de los estudios mediante los filtros reveló temas clave como el cribado, la conciencia pública y las barreras emocionales y sociales, que son determinantes en la aceptación del asesoramiento genético. Estos hallazgos resaltan la necesidad de campañas educativas y estrategias de acceso inclusivas, que ayuden a sensibilizar a la población sobre los beneficios del asesoramiento genético, al tiempo que se promueven intervenciones adaptadas a diversas poblaciones para mejorar su disposición y acceso a estos servicios.

Búsqueda 3: Campañas Educativas sobre el Cáncer Hereditario

En resumen, la tercera búsqueda destacó la escasez relativa de estudios sobre campañas educativas específicas para el cáncer hereditario y mostró la importancia de diversificar las fuentes para captar una visión más completa del tema. A pesar de que las campañas educativas tienen el potencial de aumentar el conocimiento público sobre los riesgos genéticos, existen limitaciones en su alcance y efectividad, especialmente en poblaciones subrepresentadas. Los resultados obtenidos indican que mejorar la accesibilidad y precisión en la comunicación sobre el cáncer hereditario, especialmente en cuanto a riesgos genéticos, es fundamental. La creación de campañas educativas más inclusivas y bien fundamentadas científicamente podría cerrar la brecha de conocimiento y aumentar la concienciación pública, facilitando así una mayor participación en la evaluación y el asesoramiento genético en poblaciones de alto riesgo.

En este trabajo, nos centramos específicamente en los resultados de los dos primeros filtros de la búsqueda 1, con el fin de comparar las preguntas de nuestro cuestionario con las de otros cuestionarios desarrollados en el contexto de la evaluación de riesgos genéticos en cáncer hereditario.

En particular, revisamos estudios como:

- *"Genetic Risk Assessments in Individuals at High Risk for Inherited Breast Cancer in the Breast Oncology Care Setting"* (Pal & Vadaparampil, 2012)
- *"Facilitating decision-making in women undergoing genetic testing for hereditary breast cancer: BRECONDA randomized controlled trial results"* (Sherman et al., 2017)
- *"Recruitment, genetic counseling, and BRCA testing for underserved women at a public hospital"* (Lee et al., 2005)
- *"Assessing Controversial Direct-To-Consumer Advertising for Hereditary Breast Cancer Testing: Reactions from Women and Their Physicians in a Managed Care Organization"* (Mouchawar et al., 2005)
- *"Communication of BRCA1 and BRCA2 genetic test results to health care providers following genetic testing at a tertiary care center"* (Ready et al., 2011).

Estos artículos ofrecen una base sólida para contrastar y evaluar las metodologías y enfoques utilizados en otros estudios, lo que enriquece el desarrollo y la validación de nuestro propio cuestionario en el marco de la evaluación de riesgos genéticos. Para una interpretación más detallada de cada artículo, se puede consultar el **Anexo IV**.

En las conclusiones de este trabajo, se subraya la eficacia del cuestionario diseñado para la evaluación de riesgos genéticos en cáncer hereditario, especialmente al compararlo con los cuestionarios de estudios revisados en la literatura. A diferencia de otros cuestionarios que se centran principalmente en la historia familiar básica y en la percepción del riesgo, el cuestionario aquí desarrollado aborda de forma exhaustiva y detallada aspectos complejos, como el **historial familiar en múltiples generaciones, diagnósticos de cáncer en edades tempranas, casos de cáncer bilaterales o raros, y presencia de mutaciones conocidas en la familia.**

Además, a diferencia de herramientas como la de **BRECONDA**, que se orientan a reducir el conflicto emocional en la toma de decisiones, este cuestionario se enfoca en proporcionar una evaluación inicial precisa, sin requerir una intervención extensa de consejería genética, lo cual lo hace accesible y aplicable en distintos contextos poblacionales. Frente a otros estudios que usan cuestionarios simplificados, como los desarrollados para poblaciones desatendidas, este cuestionario ofrece un enfoque innovador al incluir preguntas que permiten captar patrones hereditarios complejos, facilitando así una evaluación más completa.

En resumen, el cuestionario combina precisión y accesibilidad en su diseño, permitiendo no solo identificar factores de predisposición genética de manera integral, sino también ofreciendo a los usuarios un recurso de autoevaluación detallado que, comparado con los cuestionarios revisados, cubre un espectro más amplio de factores de riesgo. Para un análisis en mayor profundidad sobre esta comparación y los estudios revisados, se puede consultar el **Anexo IV**.

5.2. Conclusión porcentajes y grupos de edad

En este apartado de discusión, y teniendo en cuenta las limitaciones de la muestra en cuanto a heterogeneidad y significancia estadística en algunos grupos, podemos extraer algunas conclusiones clave:

1. **Variabilidad entre Grupos de Edad:** Aunque la segmentación en seis grupos etarios permite observar ciertas tendencias sobre la predisposición genética y la historia familiar de cáncer, algunos grupos, especialmente los de mayor edad, presentan una cantidad de datos limitada. Esto afecta la robustez estadística de ciertos análisis, pero permite identificar patrones preliminares en la percepción de riesgo a lo largo de distintas etapas de la vida.
2. **Significancia Estadística Limitada:** La falta de un tamaño de muestra homogéneo y suficiente en todos los grupos restringe la capacidad de generalizar los hallazgos. Sin embargo, estas observaciones iniciales pueden guiar futuras investigaciones, donde una muestra ampliada permitiría comparaciones más sólidas y válidas entre grupos de edad.
3. **Patrones de Respuesta y Percepción del Riesgo:** Los grupos intermedios (30-59 años) parecen tener una mayor disposición a reconocer antecedentes familiares de cáncer en comparación con los grupos más jóvenes o mayores. Esta tendencia sugiere una posible mayor conciencia del riesgo en estas edades, aunque debe interpretarse con precaución debido a las limitaciones estadísticas de algunos segmentos.
4. **Eficacia del Cuestionario para Captar Antecedentes Complejos:** Pese a las limitaciones, el cuestionario es efectivo para captar antecedentes genéticos complejos, como la presencia de cánceres raros o diagnósticos en familiares varones. Estos factores aportan una perspectiva integral del riesgo y demuestran el valor del cuestionario en la detección inicial de predisposición genética en diversas poblaciones.

En conjunto, aunque los resultados deben considerarse exploratorios, aportan una base para mejorar y ampliar el uso del cuestionario en muestras futuras. Para un análisis detallado de cada grupo de edad y sus respuestas, se puede consultar el **Anexo VII**.

5.3 Conclusión modelo de regresión y pruebas estadísticas

En el análisis realizado, algunas preguntas del cuestionario mostraron tener un **carácter predictivo** claro en relación con el riesgo de desarrollar cáncer de mama, lo que significa que su presencia o ausencia se asocia estadísticamente con una mayor o menor probabilidad de diagnóstico de esta enfermedad. Identificar estas preguntas con carácter predictivo es valioso, ya que permite focalizar futuros estudios en los factores familiares y personales más relevantes y desarrollar herramientas de cribado más precisas.

Preguntas con Carácter Predictivo

Entre las preguntas del cuestionario, aquellas que destacaron por su capacidad para predecir un posible riesgo incluyen:

1. **Historial de cáncer en generaciones consecutivas:** La presencia de cáncer en dos o más generaciones consecutivas mostró una asociación significativa con el diagnóstico de cáncer de mama, lo que refuerza la importancia de los patrones hereditarios.
2. **Cáncer de mama en familiares varones:** Este factor resultó significativo y, aunque es un caso menos frecuente, su presencia puede estar asociada con mutaciones genéticas hereditarias específicas, como BRCA.
3. **Cáncer en múltiples familiares paternos:** Esta variable también presentó un carácter predictivo, sugiriendo que el historial de cáncer en el lado paterno podría ser relevante en la evaluación del riesgo.

Implicaciones para Futuros Estudios

Las preguntas que han mostrado carácter predictivo en este análisis pueden orientar investigaciones futuras hacia un desarrollo más detallado de cuestionarios y herramientas de evaluación de riesgos genéticos, con especial énfasis en estos factores. Además, incluir una muestra más diversa y ampliar el análisis a otros tipos de cáncer hereditario podría mejorar el alcance y precisión del cuestionario, optimizando su utilidad en el cribado temprano y en la toma de decisiones informadas sobre pruebas genéticas.

Para una interpretación completa de los resultados estadísticos y de las asociaciones específicas observadas, se puede consultar el **Anexo V**.

5.4 Conclusión búsqueda de asociaciones

Las asociaciones de cada país muestran una clara inclinación hacia ciertas redes sociales, lo que refleja tanto sus objetivos de comunicación como las particularidades de sus audiencias. En general, Facebook e Instagram emergen como las plataformas predominantes, probablemente debido a su alcance y popularidad para conectar con un público amplio y diverso. Sin embargo, también se observan variaciones interesantes en el uso de otras plataformas que sugieren estrategias ajustadas a las necesidades y perfiles de cada país:

- **Francia:** Facebook e Instagram tienen frecuencias similares, mientras que YouTube muestra un nivel ligeramente más bajo. LinkedIn tiene la menor representación.
- **Alemania:** Instagram tiene la mayor frecuencia, seguido por Facebook y LinkedIn con una frecuencia similar. YouTube es la plataforma con menor representación en este país.
- **Italia:** Instagram tiene una frecuencia alta, mientras que LinkedIn y Facebook están por debajo de Instagram. YouTube tiene una frecuencia similar a LinkedIn y Facebook.
- **Portugal:** Instagram muestra la mayor frecuencia, superando notablemente a las otras redes. YouTube y LinkedIn tienen frecuencias más bajas, y Facebook es el menos utilizado entre estas redes.
- **España:** Las frecuencias de Facebook, Instagram y YouTube son bastante similares, mientras que LinkedIn es ligeramente más baja.
- **Reino Unido (UK):** Las redes sociales están equilibradas en frecuencia, con YouTube ligeramente por encima de las demás. Facebook, Instagram y LinkedIn tienen frecuencias similares.

Este análisis general sugiere que, mientras que Facebook e Instagram son ampliamente utilizadas para alcanzar a un público general, LinkedIn y YouTube cumplen un papel más específico en función de los objetivos de cada país. Estas observaciones aportan una base para futuros análisis en profundidad sobre cómo se adaptan las estrategias digitales en cada contexto.

6. Declaración obligatoria del uso de herramientas de IA.

Se declara que se ha utilizado y se utilizará una herramienta de inteligencia artificial como asistente en la redacción de este TFM. Esta herramienta se emplea con el objetivo de mejorar la estructura y la claridad del texto. Los contenidos generados por la IA fueron y serán revisados minuciosamente para asegurar su precisión y coherencia con el objetivo del estudio. Además se utilizará **Rstudio** para el análisis estadístico de nuestros datos.

Bibliografía

Couch Health. (n.d.). *The power of patient engagement in MedTech*. Recuperado de <https://couchhealth.agency/the-power-of-patient-engagement-in-medtech/>

Evita. (n.d.). *Evita Platform*. Recuperado de <https://evitacancro.org/evita-platform/>

Guan, B. Z., Parmigiani, G., Braun, D., & Trippa, L. (2022). Prediction of hereditary cancers using neural networks. *Annals of Applied Statistics*, 16(1), 495-520. <https://doi.org/10.1214/21-aos1510>

Holmes, S., & Huber, W. (2019). *Modern Statistics for Modern Biology*. Cambridge University Press. Disponible en: <https://www.cambridge.org/core/books/modern-statistics-for-modern-biology/>

Harzing, A. W. (2007). *Publish or Perish*. Recuperado de <https://harzing.com/resources/publish-or-perish>

Kim, H. Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative Dentistry & Endodontics*, 42(2), 152-155. <https://doi.org/10.5395/rde.2017.42.2.152>

Lee, R., Beattie, M., Crawford, B., Mak, J., Stewart, N., Komaromy, M., Esserman, L., Shaw, L., McLennan, J., Strachowski, L., Luce, J., & Ziegler, J. (2005). Recruitment, genetic counseling, and BRCA testing for underserved women at a public hospital. *Genetic Testing*, 9(4), 306-312. <https://doi.org/10.1089/gte.2005.9.306>

MammoScreen Project. (s.f.). *Governance*. Recuperado de <https://mammoscreenproject.eu/governance/>

Mouchawar, J., Laurion, S., Ritzwoller, D. P., Ellis, J., Kulchak-Rahm, A., & Hensley-Alford, S. (2005). Assessing controversial direct-to-consumer advertising for hereditary breast cancer testing: reactions from women and their physicians in a managed care organization. *American Journal of Managed Care*, 11(10), 601-608.

Pal, T., & Vadaparampil, S. T. (2012). Genetic Risk Assessments in Individuals at High Risk for Inherited Breast Cancer in the Breast Oncology Care Setting. *Cancer Control*, 19(4), 255-266. <https://doi.org/10.1177/107327481201900402>

Patient Focused Medicine Development. (n.d.). *Patient engagement in MedTech*. Recuperado de <https://patientfocusedmedicine.org/patient-engagement-in-medtech/>

Ready, K., Arun, B. K., Schmeler, K. M., Uyei, A., Litton, J. K., Lu, K. H., Sun, C. C., & Peterson, S. K. (2011). Communication of BRCA1 and BRCA2 genetic test results to health care providers following genetic testing at a tertiary care center. *Familial Cancer, 10*(4), 673-679.

<https://doi.org/10.1007/s10689-011-9460-z>

Sherman, K. A., Kilby, C. J., Shaw, L. K., Winch, C., Kirk, J., Tucker, K., & Elder, E. (2017). Facilitating decision-making in women undergoing genetic testing for hereditary breast cancer: BRECONDA randomized controlled trial results. *Breast, 36*, 79-85.

<https://doi.org/10.1016/j.breast.2017.10.001>

Speiser, D., & Bick, U. (2023). Primary prevention and early detection of hereditary breast cancer. *Breast Care (Basel), 18*(6), 448-454.

<https://doi.org/10.1159/000533391>

Zhang, Q., Liu, L., Wang, F., Mu, K., & Yu, Z. (2012). The changes in female physical and childbearing characteristics in China and potential association with risk of breast cancer. *BMC Public Health, 12*, 368.

<https://doi.org/10.1186/1471-2458-12-368>

Anexos

Anexo I: Estrategia de Búsqueda - Documento Uno
[Estrategia de Busqueda 1.pdf](#)

Anexo II: Estrategia de Búsqueda - Documento Dos
[Estrategia de Busqueda 2.pdf](#)

Anexo III: Estrategia de Búsqueda - Documento Tres
[Estrategia de Busqueda 3.pdf](#)

Anexo IV: Resumen de la Búsqueda Final de Literatura
[Final Literature Search .pdf](#)

Anexo V: Informe del Análisis Estadístico
[Informe del Análisis Estadístico .pdf](#)

Anexo VI: Lista de Asociaciones para el Cáncer de Mama
[List of associations for Breast Cancer J&A \(1\).xlsx](#)

Anexo VII: Porcentajes de Grupos Etarios
[Porcentajes grupos etarios .pdf](#)

Anexo VIII: Scripts para Búsquedas Bibliográficas
[Scripts búsquedas bibliográficas.pdf](#)

Anexo IX: Plan de Comunicación y Diseminación
[Communication and Dissemination Plan.pdf](#)