



Informes Anticipando

# **PREDICCIÓN DE RIESGO DE ENFERMEDAD EN POBLACIONES EN LA ERA DE LA MEDICINA PERSONALIZADA DE PRECISIÓN**





## Informe Anticipando coordinado por:

### Ángel Carracedo

Director de la Fundación Pública Gallega de Medicina Genómica (Servicio Gallego de Salud). Coordinador del Grupo de Medicina Genómica de la Universidad de Santiago de Compostela.

### Marina Pollán

Profesora de Investigación del Centro Nacional de Epidemiología (CNE) del Instituto de Salud Carlos III. Directora Científica del Área de Epidemiología y Salud Pública del Centro de Investigación Biomédica en Red (CIBERESP).



## Expertos colaboradores:

### María Brión

Investigadora del Servicio Gallego de Salud (SERGAS). Responsable del grupo de Genética Cardiovascular del Instituto de investigación Sanitaria de Santiago (IDIS).

### Carlos Flores

Investigador Principal de la Fundación Canaria Instituto de Investigación Sanitaria de Canarias del Hospital Universitario Nuestra Señora de Candelaria. Director de Área de Genómica del Instituto Tecnológico y de Energías Renovables. Investigador del CIBER de Enfermedades Respiratorias del ISCIII.

### Roberto Pastor

Científico Titular del Centro Nacional de Epidemiología (CNE) del Instituto de Salud Carlos III (ISCIII).

### Agustín Ruiz

Director Científico del Centro de Alzheimer de la Fundación ACE (ACE Alzheimer Center Barcelona).



## Comité Asesor del Observatorio de Tendencias en la Medicina del Futuro:

### Joaquín Arenas

Director del Instituto de Investigación del Hospital Universitario 12 de Octubre (I+12).

### Ángel Carracedo

Director de la Fundación Pública Gallega de Medicina Genómica (Servicio Gallego de Salud). Coordinador del Grupo de Medicina Genómica de la Universidad de Santiago de Compostela.

### Pablo Lapunzina

Jefe de grupo de investigación del Instituto de Genética Médica y Molecular (INGEMM) del IdiPaz. Director científico del CIBERER.

### Fernando Martín-Sánchez

Profesor de Investigación en Salud Digital. Escuela Nacional de Sanidad. Instituto de Salud Carlos III.

Nº de depósito legal: M-27813-2022

©2022 del contenido: Fundación Instituto Roche. Se permite la reproducción parcial, sin fines lucrativos, indicando la fuente y la titularidad de la Fundación Instituto Roche sobre los derechos de la obra.

[www.institutoroche.es](http://www.institutoroche.es)

Con la colaboración de Ascendo Sanidad&Farma.

Cómo citar este informe: Fundación Instituto Roche. Informe Anticipando: Predicción de riesgo de enfermedad en poblaciones en la era de la Medicina Personalizada de Precisión. 2022; ISBN edición online: 978-84-09-45979-7

# Contenidos

<b>RESUMEN EJECUTIVO .....</b>	<b>5</b>
<b>INTRODUCCIÓN.....</b>	<b>9</b>
Modelos de predicción de riesgo de enfermedad.....	11
Aportación de la genómica a los modelos de predicción de riesgo de enfermedad: Estimaciones de Riesgo Poligénico .....	11
Bases de datos poblacionales y el proyecto 1 <i>Million Genomes</i> .....	16
<b>APLICACIONES DE LOS MODELOS DE PREDICCIÓN DE RIESGO CON LA INCORPORACIÓN DE LAS ESTIMACIONES DE RIESGO POLIGÉNICO.....</b>	<b>17</b>
Enfermedades oncológicas.....	18
Cáncer de mama .....	18
Cáncer de colon.....	18
Cáncer genitourinario .....	19
Enfermedades cardiovasculares.....	19
Otras aplicaciones de los PRS en la Medicina del Futuro .....	20
<b>RETOS.....</b>	<b>21</b>
Retos analíticos .....	21
Retos de implementación .....	22
Retos éticos.....	22
<b>CONCLUSIONES Y RECOMENDACIONES .....</b>	<b>25</b>
Recomendaciones .....	25
<b>BIBLIOGRAFÍA .....</b>	<b>27</b>





# PRESENTACIÓN

---

Los Informes Anticipando, elaborados en el marco del Observatorio de Tendencias en la Medicina del Futuro impulsado por la Fundación Instituto Roche, surgen con el objetivo de contribuir a la generación y difusión de los avances en áreas de conocimiento incipiente relacionadas con la Medicina Personalizada de Precisión y que formarán parte de la Medicina del Futuro.

El Observatorio cuenta con un Comité Asesor de expertos formado por el **Dr. Ángel Carracedo**, el **Dr. Joaquín Arenas**, el **Dr. Pablo Lapunzina** y el **Dr. Fernando Martín-Sánchez**. Entre sus funciones, se incluye la selección de las temáticas que abordarán estos informes, la identificación de expertos y la validación de los contenidos.

Este informe que versa sobre la **predicción de riesgo de enfermedad en poblaciones en la era de la Medicina Personalizada de Precisión** está coordinado por el **Dr. Ángel Carracedo** y la **Dra. María Pollán**, y en su elaboración han participado como expertos la **Dra. María Brión**, el **Dr. Carlos Flores**, el **Dr. Agustín Ruiz** y el **Dr. Roberto Pastor**.

El **Dr. Ángel Carracedo** es catedrático de Medicina Legal de la Universidad de Santiago (USC) y fue director del Instituto de Medicina Legal de dicha Universidad (1994 – 2012). Desde 1999 es director de la Fundación Pública Galega de Medicina Xenómica (SERGAS-Xunta de Galicia), director de la Fundación Kaertor para el desarrollo precoz de fármacos, y, miembro del CIBERER, del CIMUS y del IDIS. En la actualidad coordina al Grupo de Medicina Xenómica y sus principales áreas de trabajo incluyen la genética del cáncer, genética de enfermedades raras, la farmacogenómica y las enfermedades psiquiátricas en edad infantil. Es miembro del *board* de Sociedades nacionales e internacionales de Medicina Forense, Genética, Farmacogenómica y del ámbito de la Oncología.

Actualmente, coordina el pilar IMPaCT Genómica de la infraestructura IMPaCT (Infraestructura de Medicina de Precisión asociada a la Ciencia y la Tecnología) y el grupo espejo español de la acción europea 1+M Genomes. También es miembro de diversos organismos reguladores como la EMA, *Forensic DNA Regulator UK*, *DNA ISFG Commission*, Comisión Nacional para el uso forense del ADN (España), Comité de la Cruz Roja Internacional, Consorcio Internacional de Enfermedades Raras (IR-DiRC), entre otros, y forma parte del Comité asesor de más de diez institutos de investigación de ámbito nacional e internacional. Es Doctor *Honoris Causa* por varias universidades de Europa y América.

La **Dra. Marina Pollán** es licenciada en Medicina y Cirugía por la Universidad de Salamanca y doctora en Medicina Preventiva y Salud Pública por la Universidad Autónoma de Madrid. Es profesora de Investigación de Organismos Públicos de Investigación (OPI) en el Centro Nacional de Epidemiología del Instituto de Salud Carlos III y directora científica del CIBER del Área de Epidemiología y Salud Pública (CIBERESP). Como epidemióloga del cáncer, se ha centrado en el cáncer de mama y los factores de riesgo asociados, en particular el estilo de vida y las exposiciones modificables. Durante la pandemia de COVID-19, ha sido la coordinadora científica del estudio seroepidemiológico de base poblacional nacional Estudio Nacional de Sero-Epidemiología de la infección por SARS-CoV-2 (ENE-COVID). Ha colaborado en el programa de diagnóstico precoz de cáncer de mama de Navarra para la evaluación de la utilidad de modelos de predicción de riesgo de enfermedades (en concreto el modelo predictivo de Gail) en la predicción individualizada de este tumor y la necesidad de adaptar dicho modelo a nuestro contexto y de incorporar la densidad mamográfica para lograr mayor discriminación. Actualmente coordina el pilar de Medicina Predictiva de la convocatoria de IMPaCT.

La **Dra. María Brion** es bióloga y doctora en el área de genética forense por la Universidad de Santiago de Compostela. Realizó su postdoctorado en el Instituto Dante Pazzanese de Cardiología y la Universidad de Sao Paulo (Brasil) y desde el año 2008 es coordinadora de la Unidad de Genética Cardiovascular del Instituto de Investigación Sanitaria de Santiago (IDIS). El objetivo principal de su investigación básica y clínica es la identificación de variantes genéticas capaces de predecir el riesgo de enfermedad cardiovascular y muerte súbita cardiaca en niños y adultos jóvenes, permitiendo llevar a cabo una Medicina Personalizada de Precisión para reducir la mortalidad y morbilidad asociada con las cardiopatías, los aneurismas y disecciones de aorta torácica y la cardiotoxicidad asociada a tratamientos antitumorales. Junto a su labor investigadora desarrolla una labor asistencial como genetista responsable de estudios genéticos y asesoramiento genético de la Unidad de Cardiopatías Familiares del Complejo Hospitalario de Santiago de Compostela.

El **Dr. Carlos Flores**, doctor en Biología, es investigador de CIBERES desde 2008. Dirige el grupo de Investigación “Variación genética y enfermedad” del Hospital Universitario Nuestra Señora de Candelaria y es director de la División de Genómica del Instituto Tecnológico y de Energías Renovables. Entre los objetivos generales de su equipo están la identificación de genes de enfermedad en diversas enfermedades raras y complejas, principalmente respiratorias e infecciosas, y el desarrollo de aproximaciones de medicina genómica y su acercamiento a la práctica clínica habitual en el campo de la Medicina Personalizada de Precisión desde el plano del huésped y su microbioma, y desde el plano del patógeno. El equipo es parte de la red de vigilancia genómica para la COVID-19 en Canarias. Son de especial relevancia científica sus contribuciones con mejoras metodológicas en el ámbito de la secuenciación genómica y en la identificación de factores genéticos implicados en la fibrosis pulmonar idiopática, el síndrome de distrés respiratorio agudo post-sepsis, la COVID-19 grave y fenotipos relacionados con el asma, entre otros.

El **Dr. Roberto Pastor** es científico titular en el Centro Nacional de Epidemiología del Instituto de Salud Carlos III desde 2006, y es director académico y profesor de bioestadística en el Máster en Salud Pública de la Escuela Nacional de Sanidad desde 2016. Se licenció en matemáticas por la Universidad del País Vasco, y posee un máster en epidemiología por el Gobierno del País Vasco. Posterior-

mente, se doctoró en Medicina Preventiva y Salud Pública por la Universidad Autónoma de Madrid. Su trabajo se ha centrado principalmente en la epidemiología y la bioestadística, principalmente en la investigación sobre factores de riesgo y modelos de predicción para las enfermedades cardiovasculares y el cáncer. Sus principales líneas de investigación se centran fundamentalmente en el estudio de las relaciones dosis-respuesta, el análisis de datos longitudinales, la estimación de efectos causales a través de modelos estructurales marginales y aleatorización mendeliana, y el desarrollo y validación de modelos de predicción basados en factores de riesgo tradicionales y marcadores genéticos para contribuir a la implementación de estrategias preventivas y de salud pública. Durante la pandemia de COVID-19, ha participado como autor principal en el estudio seroepidemiológico de base poblacional nacional Estudio Nacional de Sero-Epidemiología de la infección por SARS-CoV-2 (ENE-COVID). Adicionalmente, ha colaborado como uno de los investigadores principales en el programa de diagnóstico precoz de cáncer de mama de Navarra para la evaluación de la utilidad del modelo predictivo de Gail en la predicción individualizada del cáncer de mama y la necesidad de recalibrar dicho modelo para su aplicación en distintas poblaciones y de incorporar la densidad mamográfica para lograr mayor discriminación. Actualmente, continúa desarrollando su labor investigadora en el campo de la predicción de riesgo de enfermedades con el objetivo de ampliar conocimientos y generar evidencias que permitan avanzar en la mejora de los modelos predictivos en el marco de la Medicina Personalizada de Precisión.

El **Dr. Agustín Ruiz** es licenciado en Medicina y Cirugía por la Universidad de Sevilla, donde se doctoró en 1999 obteniendo el Premio Extraordinario de doctorado por sus estudios sobre la localización del gen RP25, implicado en la retinopatía neurodegenerativa. Ha dedicado la mayor parte de su carrera profesional a la genética médica y actualmente es director científico de la Fundación ACE (*ACE Alzheimer Center Barcelona*). La Fundación ACE cuenta con un importante biorepositorio en el que están recopilados más de 18.000 de personas con y sin demencia. Esta gran colección ha permitido generar múltiples proyectos de investigación y ha contribuido activamente en el descubrimiento del 80% de los genes conocidos de la Enfermedad de Alzheimer. Además, es importante destacar que ACE contribuyó a la generación del primer



---

estudio de asociación de genoma completo (GWAS) para la enfermedad de Alzheimer en la población española. En 2012, ACE promovió y fundó el consorcio DEGESCO (*Dementia Genetics Spanish Consortium*), del que el Dr. Ruiz es director y co-fundador. DEGESCO cuenta con 23 grupos activos dedicados a la investigación genómica de las enfermedades neurodegenerativas usando una colección genética de más de 35.000 sujetos. Actualmente el consorcio está lanzado una nueva fase de investigación prospectiva con el establecimiento de una plataforma nacional de medicina de precisión (PMP-DEGESCO). Con

el objetivo de lograr un diagnóstico temprano, PMP-DEGESCO está integrando el uso secuencial de marcadores digitales, estimas de riesgo poligénico (PRS) y biomarcadores sanguíneos. Por último, cabe mencionar que el Dr. Ruiz también forma parte del Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), y es coordinador del proyecto internacional *Alzheimer's Disease Apolipoprotein Pathology for Treatment Elucidation and Development* (ADAPTED) de la Iniciativa de Medicamentos Innovadores (IMI) europea.





# RESUMEN EJECUTIVO

---

Los avances en el campo de la Medicina Personalizada de Precisión y la incorporación de las nuevas tecnologías, han permitido profundizar en el conocimiento y la comprensión de las enfermedades y sus causas. Esto supone una gran oportunidad para el desarrollo y mejora en otros campos de la medicina, tales como la Medicina Preventiva y la Salud Pública de Precisión.

Tradicionalmente, estos dos campos, se han servido de modelos de predicción de riesgo, como herramienta principal para realizar determinaciones predictivas sobre la probabilidad de sufrir determinados eventos clínicos a partir de la información disponible sobre factores condicionantes de salud o enfermedad de diferente naturaleza. Sin embargo, el advenimiento de estas nuevas tecnologías en el marco de la Medicina Personalizada de Precisión, ha permitido la incorporación de la información genómica a los modelos de predicción de riesgo, a través de las Estimas de Riesgo Poligénico (PRS, por sus siglas en inglés), abriendo nuevos horizontes para el avance de la Medicina Preventiva y la Salud Pública de Precisión.

Los PRS, que son una medida global del riesgo genético de desarrollar una enfermedad por parte de una persona respecto de la población general, permiten incorporar información sobre variantes en las secuencias de los genes asociadas al riesgo de desarrollar enfermedades. Esto podría contribuir a mejorar de manera sustancial la capacidad predictiva de los modelos de predicción de riesgo desarrollados hasta el momento. Así, se espera que estos PRS constituyan una herramienta clave para la predicción de riesgo tanto a nivel poblacional como de manera personalizada, por ejemplo, guiando la toma de decisiones clínicas, o la selección o priorización de tratamientos en individuos o grupos de pacientes.

Por todo ello, aunque es necesario avanzar en el estudio y validación de los PRS, se prevé que estos tengan un papel relevante en la Medicina del Futuro, si bien, es necesario abordar una serie de retos analíticos, de implementación y éticos por parte de la sociedad, investigadores y profesionales sanitarios, para desarrollar todo el potencial de los PRS en la práctica clínica.





# INTRODUCCIÓN

La mayoría de las enfermedades que afectan al ser humano son complejas y se ven condicionadas por diferentes factores. Estos **factores condicionantes de la salud o enfermedad** pueden ser factores **biológicos** (como factores genéticos o fisiológicos), **ambientales o de exposición** (lo que se conoce como exposoma<sup>a</sup> (para más información, ver [Informe Anticipando sobre Exposoma](#))), o **psicosociales**, como los hábitos de vida, factores psicológicos, socioeconómicos o culturales.<sup>1</sup>

Cuando estos factores se incluyen **en análisis estadísticos** se denominan **variables<sup>b</sup>** y pueden ser empleados para **predecir el riesgo<sup>c</sup> de desarrollar una determinada enfermedad** a través de la creación de **modelos de predicción**. Este tipo de análisis, además de servir para **predecir el riesgo específico de cada persona de padecer una determinada enfermedad**, permite la integración del conocimiento sobre **factores y características individuales** con la información de salud pública para **identificar factores que condicionen el desarrollo de enfermedades a nivel poblacional**.

## MODELOS DE PREDICCIÓN DE RIESGO DE ENFERMEDAD

Los modelos de predicción de riesgo son **estimaciones estadísticas** sobre la **probabilidad de que se produzca un determinado evento clínico** (en este caso,

el desarrollo de una enfermedad) en un tiempo determinado.<sup>2</sup> Para ello, se basan en estudios que **evalúan la asociación y cuantifican el efecto de la presencia de un determinado factor** (denominado **factor de riesgo** cuando se demuestra esta asociación) **sobre el desarrollo de una enfermedad**.

Los **modelos de predicción más utilizados en la práctica clínica utilizan modelos de regresión**, si bien con la aparición de nuevas tecnologías de análisis, se están generando nuevos **modelos** que utilizan métodos sofisticados de **procesamiento de la información**, como algoritmos basados en **inteligencia artificial**, por ejemplo, las **redes neuronales<sup>d</sup>**.

Los modelos de regresión **determinan la relación entre variables** estudiando cómo varía el valor de una variable (a la que se denomina **“variable dependiente”**) a medida que cambia el valor de otra variable (a la que se denomina **“variable independiente”**). Por ejemplo, estos modelos permiten estudiar el efecto del aumento del colesterol en sangre (variable independiente) sobre la tensión arterial (variable dependiente). Cuando las variables independientes son **factores de riesgo de enfermedades**, el **resultado** del modelo de predicción puede ser aplicado a la construcción de una **medida del riesgo global de un individuo de padecer dicha enfermedad**.

Tradicionalmente, la **construcción de los modelos de predicción de riesgo** se lleva a cabo en un **proceso estructurado** en pasos sucesivos (Figura 1):<sup>3</sup>

Figura 1. Flujo de trabajo para la obtención de un modelo predictivo.



El proceso comprende desde la extracción de datos de estudios observacionales para la identificación de factores de riesgo, hasta la validación del modelo, que debe realizarse siempre para poder extender su uso en la práctica clínica. Adaptado de (5)

<sup>a</sup> Conjunto de factores no genéticos a los que un individuo está expuesto a lo largo de toda su vida y que condicionan el estado de salud o enfermedad. <sup>b</sup> Características o aspectos de los individuos que constituyen la muestra sobre la que se realiza un análisis estadístico. Las variables pueden ser cualitativas (como el sexo o el grupo sanguíneo) o cuantitativas (como el número de embarazos o los niveles de colesterol en sangre). <sup>c</sup> El riesgo de una enfermedad es la proporción de nuevos casos en un período de tiempo respecto del total de individuos que no presentan la enfermedad en dicho período de tiempo (ya que los individuos que presentan la enfermedad no son sujetos en riesgo). <sup>d</sup> Modelos computacionales basados en inteligencia artificial que reconocen patrones de todo tipo y realizan categorizaciones en base a dichos patrones.

- **Identificación de factores de riesgo.** Consiste en la caracterización de los determinantes de una enfermedad a través de **estudios observacionales** que pueden ser **de cohortes** o **de casos y controles**. Por un lado, los **estudios de cohortes** seleccionan dentro de la población<sup>e</sup> una muestra<sup>f</sup> a la que se hace seguimiento a lo largo del tiempo para observar la aparición o no de eventos clínicos. Por otro lado, los **estudios de casos y controles** parten de una muestra de personas que presentan el evento clínico en estudio (casos) y otra muestra de personas, proveniente de la misma población, que no lo presenten (controles) y estudian de manera retrospectiva todas las características medibles en ambas muestras (como peso, altura, exposiciones, hábitos, etc.).<sup>2</sup>
  - **Caracterización del riesgo relativo.** El riesgo relativo es una medida del **efecto específico de cada** uno de los **factores de riesgo** asociados a un evento clínico (una enfermedad) en una muestra de sujetos **expuestos frente a los no expuestos** a dicho factor. De esta manera, se consideran factores de riesgo aquellos cuya presencia **determina el incremento relativo del riesgo de los sujetos independientemente de la presencia de otros factores**.<sup>3</sup>
  - **Estimación del riesgo absoluto.** Es el cálculo de la probabilidad o **riesgo global de desarrollar una enfermedad, en un intervalo de tiempo determinado**, como resultado de una combinación **de todos los factores de riesgo** presentes en cada sujeto. Para esta estimación se emplean la incidencia<sup>g</sup> de una enfermedad y la prevalencia<sup>h</sup> de los diferentes factores de riesgo en la población, que son específicas de cada población, por lo que el riesgo absoluto no es una medida generalizable.<sup>3</sup>
  - **Generación del modelo.** Una vez obtenida esta información, se emplea para la **construcción de un modelo estadístico con el objetivo de predecir la probabilidad de que se produzca un determinado evento clínico en los individuos en función del patrón de factores de riesgo** que presentan. Para generar este tipo de modelos, esencialmente se emplean el riesgo relativo y el **riesgo basal**, es decir, el riesgo o probabilidad de enfermar de los sujetos que no están expuestos a ningún factor de riesgo.
  - **Validación del modelo.** El último paso de la generación de un modelo de predicción consiste en verificar su **capacidad predictiva**, es decir, su capacidad de producir resultados reproducibles que expliquen la asociación de los factores de riesgo con la enfermedad. Para ello, se debe **testar el modelo en sujetos diferentes a los que se emplean para su generación**. Idealmente la validación del modelo se debe realizar en una muestra independiente de sujetos de la población o de poblaciones similares (validación externa) a la que se aplicará el modelo. En ausencia de esta muestra, se debe llevar a cabo una validación interna basada en submuestras o muestras repetidas en la misma población (validación interna, cruzada o *Bootstrap*). Generalmente, la capacidad predictiva de los modelos suele evaluarse atendiendo a dos criterios:
    - **Calibración:** capacidad para **estimar** de manera precisa **la incidencia de un determinado evento clínico en un intervalo temporal en una muestra de la población distinta** a la empleada para el desarrollo del modelo. Es decir, un modelo bien calibrado, será capaz de predecir con precisión el número de casos que ocurrirán en la población. La calibración es un criterio poblacional, es decir, que no contempla la capacidad de identificar los sujetos concretos que sufrirán la enfermedad.<sup>3</sup>
    - **Discriminación.** Se define como la capacidad de un modelo para **diferenciar entre los sujetos que desarrollarán la enfermedad o el evento en el futuro de los que no**, de manera que cuanto mayor sea la discriminación mejor será el modelo.
- De cara a la **generación de modelos predictivos eficientes**, es importante disponer de datos a partir de los cuales identificar factores condicionantes de la salud o la enfermedad.<sup>2,3</sup> Una vía para avanzar en la mejora de estos modelos, es el descubrimiento de nuevos factores que permitan realizar asociaciones más precisas con el



riesgo de padecer enfermedades. En los últimos años, gracias al **desarrollo de tecnologías en el marco de la Medicina Personalizada de Precisión, como la genómica, y al avance en los métodos de análisis de grandes cantidades de datos**, ha sido posible **identificar nuevos factores** a partir de distintas fuentes de información que contribuyen a nutrir los modelos de predicción de riesgo.

## **APORTACIÓN DE LA GENÓMICA A LOS MODELOS DE PREDICCIÓN DE RIESGO DE ENFERMEDAD: ESTIMAS DE RIESGO POLIGÉNICO**

El **estudio de la genómica** ha permitido obtener información sobre la **base genética de las enfermedades** e identificar alteraciones en uno o varios genes como **mecanismos subyacentes a numerosas patologías**. En este sentido, la incorporación de la información sobre diferentes factores genéticos, como las **variantes genómicas**<sup>1</sup> asociadas al desarrollo de enfermedades, contribuye a **mejorar la capacidad predictiva de los modelos de predicción de riesgo** y, por lo tanto, a su futura **aplicación en la práctica clínica**.<sup>4,5</sup>

Desde el punto de vista genético, las enfermedades pueden ser **“mendelianas”** (aquellas que están determinadas por una alteración en genes de elevada penetrancia<sup>k</sup>, es decir, alteraciones que confieren una elevada probabilidad de padecer la enfermedad), o **complejas** (cuyo desarrollo está condicionado por la presencia de múltiples variaciones genéticas y/o de otros factores no genéticos).<sup>4</sup> La detección de enfermedades mendelianas se puede realizar en la práctica clínica habitual mediante **pruebas de diagnóstico genético** para la identificación de mutaciones causantes de patologías o que aumentan el riesgo de padecer una enfermedad.

Sin embargo, en la mayoría de las enfermedades comunes el componente “mendeliano” explica solo una parte de la heredabilidad.<sup>3</sup> Estos casos se deben principalmente a variaciones genéticas relativamente frecuentes en la población, de baja o media penetrancia<sup>k</sup> y que confieren un riesgo bajo de manera individual, pero que

en conjunto con el efecto de otros genes puede resultar considerable.

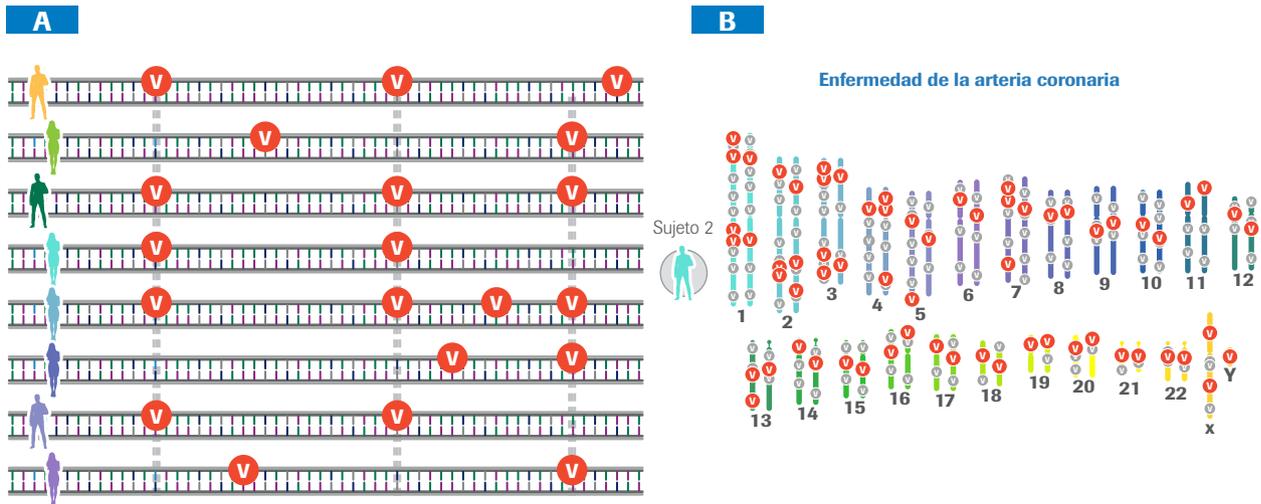
De hecho, la mayoría de las enfermedades comunes presentan un componente mendeliano, con variaciones genéticas de elevada penetrancia, y un componente más complejo con genes de escasa penetrancia que, en su conjunto y por su interacción con el ambiente, confieren un riesgo variable de enfermedad por lo que, para la detección de enfermedades complejas, se deberán tener en cuenta otros factores.

En este sentido, los **estudios genéticos familiares** permiten establecer el grado en el que una enfermedad concreta está determinada genéticamente (es decir, su heredabilidad). Estos estudios han permitido **identificar mutaciones concretas** asociadas a patologías como, por ejemplo, las **mutaciones en los genes BCRA1 y BCRA2 asociados al cáncer de mama y a otros tumores**. Se ha visto que la mutación de estos genes en el genoma de **una persona aumenta considerablemente su predisposición a padecer cáncer de mama**, por lo que constituye un factor de riesgo de desarrollar la enfermedad.<sup>3</sup> De hecho, en la práctica clínica se emplean pruebas genéticas rutinarias para la detección de estas mutaciones en familias para la detección temprana y el abordaje clínico de personas de alto riesgo. Sin embargo, estas mutaciones son muy infrecuentes y de alta penetrancia, y la mayoría de las pacientes con cáncer de mama no presentan este tipo de mutaciones, sino que el desarrollo del cáncer de mama está influido por la combinación de otros factores.<sup>3</sup>

Por otro lado, para conocer variantes comunes de baja penetrancia se emplean los **estudios de asociación de genomas completos o GWAS**, por sus siglas en inglés (*Genome Wide Association Studies*), que permiten **identificar múltiples variantes** como, por ejemplo, los **polimorfismos en un único nucleótido** o **SNPs**<sup>m</sup>, por sus siglas en inglés (*Single Nucleotide Polymorphism*), asociadas al riesgo de desarrollar enfermedades.<sup>3</sup> De hecho, este tipo de estudios en los que se comparan genomas de individuos han permitido **identificar variantes presentes de manera más frecuente en personas con una enfermedad determinada**. Por ejemplo, en el caso de la enfermedad coronaria, se han identificado más de 300 variantes genómicas que están presentes en mayor medida en personas con esta patología.<sup>6</sup>

<sup>1</sup>Pequeñas variaciones en la secuencia del genoma en localizaciones concretas que diferencian el genoma de una persona del resto de individuos. <sup>k</sup>Proporción de individuos que presentan una mutación causante de una patología determinada y muestran síntomas clínicos de esa patología. <sup>l</sup>Estimación de la variación de un rasgo fenotípico en una población que se debe a la variación genética entre individuos en esa población. <sup>m</sup>Se trata de un tipo de polimorfismos puntuales que afectan a un único par de bases y que aparecen de manera frecuente en la población general. Si bien los SNPs se producen de manera natural en un gen, secuencia de ADN, proteína o cromosoma, y no suelen tener efectos adversos sobre el individuo de manera aislada, la presencia de múltiples SNPs puede tener implicaciones en el riesgo de desarrollar de una enfermedad.

Figura 2. El papel de las variantes genómicas en el desarrollo de enfermedades.



2A) Representación de variantes genómicas de diferentes individuos. Las variantes están representadas con la letra V. Esta imagen muestra que en algunos casos distintos individuos pueden presentar las mismas variantes y otras son específicas de cada individuo. 2B) Representación de las variantes de un individuo asociadas al riesgo de enfermedad coronaria. En la imagen, las variantes con la letra "V" y en color naranja, representan aquellas que se asocian con la enfermedad y aquellas en gris, representan otras variantes genómicas que no están asociadas con esta enfermedad en concreto. Adaptado de (7).

Esta información sobre variantes genéticas puede incorporarse al desarrollo de modelos de predicción de riesgo y contribuir a mejorar su eficacia en la estimación del riesgo de enfermedades poligénicas. En este contexto, nacen las **Estimas de Riesgo Poligénico** o **PRS**, por sus siglas en inglés (*Polygenic Risk Scores*).

Los **PRS** son una **medida global del riesgo genético** de desarrollar una enfermedad por parte de una persona respecto de la población general. Estos PRS **se calculan al sumar** de manera ponderada el **efecto individual sobre el riesgo de cada una de las variantes genéticas asociadas a una patología concreta**,<sup>7</sup> así, para cada combinación de variantes genéticas posibles que pueden encontrarse en las personas con una patología se obtiene un valor de riesgo. En este sentido, es importante tener en cuenta que al igual que en el proceso de creación de modelos de predicción, los valores obtenidos **deben ser validados en una población independiente** de la que se ha utilizado para su cálculo. Posteriormente, las estimas de riesgo poligénico pueden ordenarse de menor a mayor para la clasificación de **los PRS en cuantiles**<sup>n</sup>, lo cual

permite **estratificar a la población en diferentes niveles de riesgo** para cualquier enfermedad.<sup>7-9</sup>

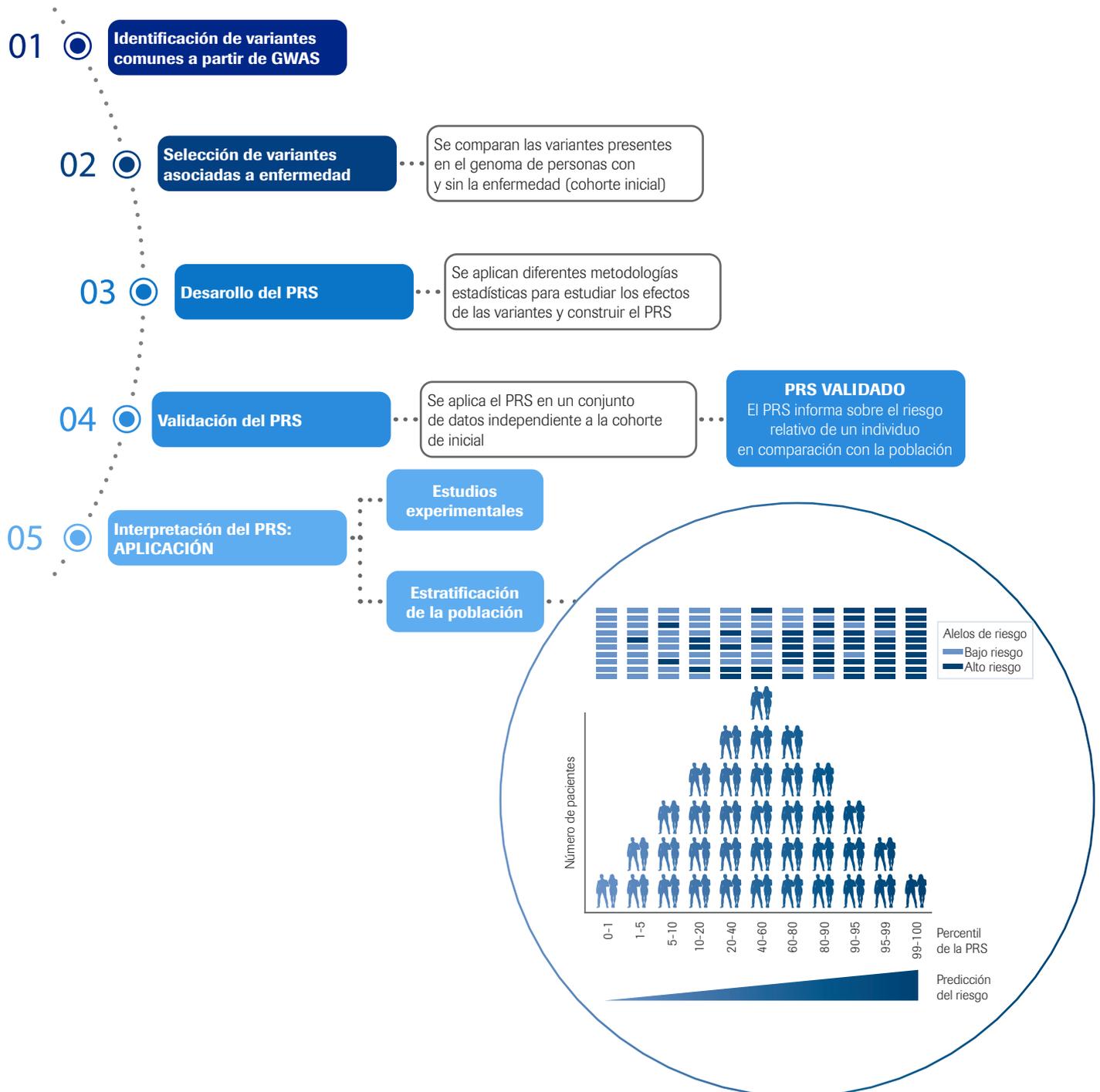
La estimación de riesgo genético de las enfermedades comunes cobró un impulso mayor gracias a la obtención de datos mediante los GWAS, que permitieron calcular riesgos de enfermedad en diferentes patologías comunes.<sup>2</sup>

En este contexto, con el objetivo de contrastar la utilidad de los PRS que se generen y apliquen a la práctica clínica, se están **desarrollando iniciativas que buscan establecer estándares a la hora de aplicar PRS en la práctica clínica**. Recientemente se ha publicado una serie de **recomendaciones para el reporte de PRS de manera estandarizada**, por el *Clinical Genome Resource (ClinGen) Working Group* y el *Polygenic Score Catalog (PGS Catalog)*, conocidas como *Polygenic Risk Score Reporting Standards (PRS-RS)*, que tienen por objetivo definir los criterios mínimos que deben reflejarse en los estudios para una correcta interpretación de los PRS y para que los resultados sean reproducibles a lo largo de todo el proceso de desarrollo.<sup>10,11</sup>

<sup>n</sup>Los cuantiles son una medida de posición que consiste en ordenar los valores de un conjunto de datos (en este caso los valores de los PRS) de menor a mayor y dividirlos en partes de igual tamaño. Hablamos de percentiles cuando la muestra se divide en 100 partes; de deciles cuando se divide en 10 partes; y de cuartiles cuando se divide en 4 partes. En este sentido, para una muestra de 100 personas, el percentil 70 es el valor del PRS que es igual o superior al 70% del total de los valores. De este modo, es posible establecer el nivel de riesgo (alto, medio o bajo) de una persona para una determinada enfermedad en base a su valor de PRS si se compara con el riesgo de la enfermedad en la población general.



Figura 3. Construcción de los PRS.



Esquema del proceso para la obtención de PRS y su interpretación. Adaptado de (7, 8 y 9)

## **BASES DE DATOS POBLACIONALES Y EL PROYECTO 1 MILLION GENOMES**

---

Debido a la variabilidad genética entre individuos, así como a la diversidad y complejidad de las enfermedades conocidas, para el cálculo de los PRS deben emplearse muestras poblacionales de gran tamaño que permitan identificar aquellos SNPs asociados a enfermedades. La capacidad predictiva de los modelos de predicción de riesgo aumentará a medida que se dispongan más datos poblacionales de los que se disponen actualmente. Dado que la mayoría de los estudios se han realizado sobre población europea de origen caucásico y del Reino Unido, la capacidad predictiva es mayor en esta población que en otras europeas y mundiales.<sup>12</sup> Esto dificulta la extrapolación de los PRS a otras poblaciones, por lo que ha surgido la necesidad de **creación de bases de datos poblacionales que permitan representar una imagen general del genoma de todas las poblaciones que confluyen en Europa.**

Un ejemplo es el **proyecto “1 Million Genomes” (1+MG)**, que nace con el objetivo principal de **favorecer y asegurar el acceso a los datos genómicos y clínicos correspondientes a toda la población europea para conseguir mejorar la investigación, avanzar en la Medicina Personalizada de Precisión y mejorar las políticas sanitarias en base a esta información.** Para ello, el proyecto busca secuenciar un millón de genomas para tener una base de datos poblacional accesible y compartida entre diferentes países europeos.<sup>13</sup> A partir de este proyecto, han surgido iniciativas como, el **proyecto “Beyond 1 Million Genomes” (B1MG)**, impulsada por el compromiso de 23 países europeos y dirigida a crear una red de datos genéticos y clínicos europeos compartidos para coordinar y apoyar el 1+MG.<sup>14</sup>

La información generada a partir de estos proyectos será crucial, tanto para determinar cuáles son los elementos del genoma relevantes para que se produzca un evento clínico (factores genéticos condicionantes), como para la obtención de modelos de predicción de riesgo extrapolables a la población.



# APLICACIONES DE LOS MODELOS DE PREDICCIÓN DE RIESGO CON LA INCORPORACIÓN DE LAS ESTIMAS DE RIESGO POLIGÉNICO

Actualmente existen modelos de predicción de riesgo que ya se utilizan en la práctica clínica con diferentes objetivos, si bien, dichos modelos están basados esencialmente en factores de riesgo no genéticos y su capacidad predictiva es limitada. En los últimos años, se ha avanzado hacia una mayor comprensión de las enfermedades y el riesgo de desarrollarlas mediante el estudio y conocimiento del genoma humano y el aumento de información fenotípica y genotípica disponible, que han contribuido a descubrir **la base poligénica de las enfermedades**. Estos avances se están incorporando progresivamente al campo de la predicción de riesgos, contribuyendo al desarrollo y mejora de los modelos, que se espera que en el futuro permitan establecer estrategias de prevención personalizadas y precisas.

Los PRS irrumpen en el escenario de la Medicina del Futuro como una nueva manera de estimar el riesgo genético de desarrollar patologías. El creciente número de evidencias que apoyan la asociación de los PRS con la enfermedad permitirá la generación de modelos de predicción integrados que los combinen con otros factores predictores de riesgo como, por ejemplo, factores demográficos, clínicos, biomarcadores, etc. De esta manera, se espera que estos nuevos modelos cambien el paradigma de la atención sanitaria a nivel poblacional e individual dentro del sistema sanitario.<sup>5</sup> A nivel de salud pública, permiten **estratificar de una manera más precisa a la población según el riesgo genético para desarrollar**

**una enfermedad**, de manera que la información obtenida a partir de los PRS podría contribuir a la **optimización de los recursos sanitarios y dirigir las estrategias de prevención desde la salud pública**. A nivel individual, también permiten obtener una **medida cuantitativa del riesgo individual** de un sujeto para una determinada enfermedad, esto, contribuirá a la **toma de decisiones clínicas personalizadas** para la **implementación de estrategias preventivas** a diferentes niveles en base al riesgo.<sup>5</sup>

Así, **la integración de los PRS en los modelos de predicción de riesgo**, les confiere aplicaciones clínicas que pueden clasificarse de la siguiente manera:<sup>15</sup>

- **Prevención primaria.** La posibilidad de estratificar a la población según el riesgo genético de desarrollar una enfermedad, así como de conocer el riesgo de cada individuo permite, por un lado, **priorizar y dirigir las estrategias preventivas**. Estas estrategias pueden estar dirigidas a **mejorar el conocimiento y la concienciación** sobre determinados factores de riesgo como, por ejemplo, diseñando campañas de prevención del tabaquismo o la obesidad dirigidas a determinados sectores de la población o individuos con un mayor nivel de riesgo. Por otro lado, puede contribuir a **implementar intervenciones terapéuticas de manera personalizada** que puedan frenar la aparición o avance de la enfermedad antes incluso de que se desarrolle.

- **Prevención secundaria.** El cribado de enfermedades se ha basado tradicionalmente en la edad y algunos factores de riesgo y ha demostrado ser una buena estrategia para mejorar el **diagnóstico precoz de determinadas patologías**. Sin embargo, la mejora de los modelos de predicción de riesgo con la incorporación de PRS permitirá una **mejora los criterios de inclusión de los cribados actuales**.

Todo ello pone de manifiesto la relevante contribución de la **Medicina Personalizada de Precisión** a la **Medicina Preventiva** y a la **Salud Pública de Precisión**<sup>o</sup> ya que se podría **identificar a los individuos más susceptibles de beneficiarse de estrategias para abordar o prevenir enfermedades en la población, intervenir precozmente a los pacientes de manera individualizada**, orientarlos hacia estilos de vida saludables, etc., y realizar acciones que contribuyan a **reducir la carga de las enfermedades en los sistemas de salud y un uso óptimo de los recursos sanitarios**.<sup>5,16</sup>

A continuación, se presentan algunos ejemplos de modelos de predicción de riesgo habituales de aplicación en la práctica clínica, fundamentalmente en enfermedades oncológicas y cardiovasculares. Además, se presentan algunos de los PRS que se están desarrollando e incorporando progresivamente a la práctica y que se prevé que mejorarán sustancialmente los modelos de predicción y el abordaje preventivo de diferentes patologías. Asimismo, se destacarán otras aplicaciones derivadas del desarrollo y el estudio de los PRS dejando entrever el importante papel que jugará en el marco de la Medicina Personalizada de Precisión del futuro.

## ENFERMEDADES ONCOLÓGICAS

### CÁNCER DE MAMA

El primer modelo desarrollado en el campo del cáncer de mama fue el **modelo de Gail**. Este modelo, que tenía por objetivo **identificar a las mujeres con mayor riesgo de cáncer de mama se ha utilizado para seleccionar participantes sanas que podrían beneficiarse de la quimioprevención con tamoxifeno<sup>p</sup> en base a 6 factores de riesgo**: número de hijos, edad de la menarquia, edad de inicio de la menopausia, terapia hormonal sustitutiva, biopsias benignas previas e historia familiar.<sup>17,18</sup> Sin embar-

go, la capacidad discriminativa de este modelo es limitada, por ello es de gran utilidad crear modelos que incluyan otro tipo de factores de riesgo, como los genéticos, de cara a su posible aplicación como herramienta en la toma de decisiones sobre la prevención o el cribado.<sup>19</sup>

En este contexto, la **incorporación de las variantes de riesgo genético** ha permitido conocer otros mecanismos de desarrollo del cáncer de mama que, integrado con otros factores condicionantes de la salud que pueden modificarse, se pueda contribuir a la mejora del abordaje preventivo de este tipo de cáncer. Por ejemplo, el *Breast Cancer Association Consortium* (BCAC) ha creado una base de datos de más de 100.000 casos de cáncer de mama y más de 100.000 controles libres de la enfermedad, a partir de la cual se han identificado, hasta ahora, 313 SNPs<sup>20</sup> y existen modelos basados en esta información que ya están aplicándose en la práctica clínica. Un ejemplo de ello es BOADICEA (*Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm*), un modelo que incluye mutaciones genéticas, PRS, la historia familiar y otros factores de riesgo para la predicción de riesgo de desarrollar cáncer de mama o de ovario.<sup>21,22</sup> Este modelo, ha sido aprobado para su implementación en la práctica clínica en Europa a través de la herramienta CanRisk con el objetivo de mejorar el cribado del cáncer de mama y evitar efectos adversos asociados, como el sobrediagnóstico.<sup>22</sup>

Por ello, se están iniciando ensayos clínicos basados en simulaciones preliminares de la aplicación del modelo BOADICEA para el cribado de cáncer de mama de manera personalizada en base a factores de riesgo en Estados Unidos y en Europa. Uno de estos estudios es el denominado *My Personal Breast Screening* (MyPeBS), un estudio clínico internacional en una muestra de 85.000 mujeres voluntarias cuyo objetivo es comparar la estrategia de cribado basada en la aplicación del modelo BOADICEA frente a estrategias de cribado estándar actuales. El estudio busca valorar la efectividad y viabilidad de este enfoque personalizado que tiene en cuenta los factores epidemiológicos, de estilo de vida, exposición hormonal y factores genéticos para evitar los falsos positivos, el sobrediagnóstico, o el sobretratamiento<sup>r</sup> asociados frecuentemente a las mamografías.<sup>23</sup>

### CÁNCER DE COLON

En la actualidad, existen **programas de cribado de cáncer de colon basados en factores de riesgo** como

<sup>o</sup>La Salud Pública de Precisión consiste en la mejora de la capacidad para prevenir enfermedades, promover la salud y reducir las diferencias de salud en la población mediante la aplicación de métodos y tecnologías emergentes que permitan conocer la situación global de las personas (enfermedad, patógenos, exposiciones, comportamientos) y la susceptibilidad en la población, así como el desarrollo de políticas y programas de aplicación específicos con el objetivo de adaptar las intervenciones preventivas a ciertos grupos de riesgo y mejorar la salud general poblacional. Esta disciplina, que abarca la Medicina Personalizada de Precisión, permite desarrollar soluciones personalizadas para mejorar el abordaje de la salud, poniendo el foco en la predicción y prevención de enfermedades, integrando la variabilidad genética interindividual, estilos de vida y factores medioambientales, además de todas aquellas acciones que los ciuda-



la edad, la historia familiar, enfermedad inflamatoria intestinal, etc., en base a los cuales se emiten recomendaciones individuales. Estos programas podrían mejorarse notablemente con la incorporación de PRS, como se ha visto en un estudio que comprobó la utilidad clínica de un PRS basado en 27 SNPs para guiar el cribado. Se ha propuesto que el uso de estos SNPs en combinación con la historia familiar podría contribuir considerablemente a estimar la edad a la que los individuos alcanzan el riesgo absoluto, es decir, aquella edad en la que su riesgo de desarrollar la enfermedad es mayor.<sup>24</sup> En este estudio, se vio que los pacientes en el percentil superior alcanzan el riesgo absoluto 10 años antes (42 años en hombres; 47 años en mujeres) que los que se encuentran en el percentil inferior (52 años en hombres; 58 años en mujeres). Así se demuestra que **la incorporación de PRS para el cáncer colorrectal** permite **estratificar pacientes de forma más precisa** y sería útil incorporarlos a los modelos de predicción actuales. De hecho, podría mejorarse esta capacidad predictiva con el estudio e identificación de un mayor número de SNPs, contribuyendo a una toma de decisiones sobre el cribado más informada. Esto permitirá personalizar a la edad a la que se realiza el cribado de cáncer colorrectal, mejorando la capacidad de diagnóstico precoz de los programas de cribado y reduciendo las potenciales complicaciones derivadas de la colonoscopia.<sup>24</sup>

## CÁNCER GENITOURINARIO

Otro ejemplo de la potencial utilidad de los modelos predictivos es en aquellas patologías para las que no existen programas de cribado, como es el caso del cáncer de vejiga. Se conoce que este tipo de cáncer está asociado a un componente ambiental importante, principalmente el tabaco y las exposiciones laborales, por lo que las decisiones preventivas están dirigidas en este sentido. Recientemente, se ha planteado la posibilidad de que el impacto de estas medidas preventivas sea diferente en función del riesgo genético. Para ello, utilizando la información de 12 variantes genéticas (11 SNPs y 1 delección<sup>5</sup>) se observó que el cese del consumo de tabaco reduce en mayor medida el riesgo absoluto de desarrollar cáncer de vejiga a 30 años en pacientes cuyo valor de PRS se encontrase en el cuartil superior.<sup>25</sup>

## ENFERMEDADES CARDIOVASCULARES

Los modelos de predicción de riesgo cardiovascular son ampliamente conocidos en la práctica clínica, donde se utilizan habitualmente. De hecho, en 1948, se inició un estudio prospectivo por el *National Heart Institute* en el que se empezaron a recoger datos sobre las enfermedades cardiovasculares, características y factores de predisposición, así como los efectos de dichos factores. A partir de la cohorte utilizada en este estudio, se ha elaborado el modelo SCORE2 (*Systematic Coronary Risk Evaluation*), recientemente incorporado en las guías clínicas europeas de la Sociedad Europea de Cardiología.<sup>26</sup> El modelo SCORE2, desarrollado por la Sociedad Europea de Cardiología,<sup>27</sup> estima el riesgo de padecer complicaciones o de mortalidad por causas cardiovasculares en 10 años teniendo en cuenta factores como el sexo, hipertensión, colesterol elevado, tabaco, etc.<sup>26</sup> El SCORE2 ofrece la oportunidad de **crear e implementar intervenciones para la reducción del riesgo** de estas patologías **de manera temprana**. En este sentido, se prevé que la incorporación del riesgo genético a este modelo con la identificación de variantes genéticas que se puedan emplear para el desarrollo de PRS, mejore de manera significativa las estimaciones, permitiendo mejorar y **personalizar el abordaje preventivo y terapéutico**.

Hasta el momento, **los PRS más avanzados en el campo de las enfermedades cardiovasculares** son los desarrollados para la **enfermedad coronaria**. El consorcio *Coronary ARtery Disease Genome-wide Replication and Meta-analysis Consortium* (CARDIoGRAMplusC4D), identificó 153 SNPs asociados de manera significativa con la enfermedad coronaria.<sup>28</sup> Estos hallazgos sugieren que la integración de los PRS en la predicción de riesgo cardiovascular, y en concreto de enfermedad coronaria, puede ser de gran interés al permitir potencialmente asignar nuevos valores a las categorías de riesgo sobre las que se establecen recomendaciones para la reducción del riesgo.<sup>10</sup> Concretamente, hay 38 PRS para el riesgo cardiovascular y otros problemas asociados, ya registrados en el *PGS Catalog*, si bien, de cara a su futura integración en los modelos de predicción de riesgo actuales, se deben seleccionar el número óptimo de SNPs y validar su aplicación.<sup>29</sup>

danos pueden llevar a cabo para la protección de su salud. <sup>5</sup>Se trata de un medicamento que bloquea los efectos del estrógeno, utilizado comúnmente como tratamiento complementario o preventivo para mujeres con un riesgo elevado de padecer cáncer de mama en base al modelo de Gail. <sup>6</sup>Se trata del diagnóstico de cánceres asintomáticos cuya evolución natural nunca producirá daños significativos en el paciente en el transcurso de su vida, debido a un crecimiento lento o nulo. <sup>7</sup>Tratamiento de los cánceres que han sido sobrediagnosticados y que puede suponer riesgos graves asociados a la potencial toxicidad de los medicamentos utilizados. <sup>8</sup>Alteración genética consistente en la pérdida de un segmento de ADN.

En este contexto, de cara a la futura integración de estos PRS en los modelos, se están realizando ensayos clínicos piloto con estos PRS, como el estudio HEART (*Healthcare Evaluation of Absolute Risk Testing*), iniciado recientemente en Inglaterra en 1.000 pacientes. Este estudio busca analizar el efecto de la incorporación de un PRS al análisis QRisk<sup>t</sup> para la estimación del riesgo cardiovascular y así evaluar la utilidad clínica y el potencial de la integración de estos modelos en práctica habitual de atención primaria.<sup>30</sup>

## **OTRAS APLICACIONES DE LOS PRS EN LA MEDICINA DEL FUTURO**

La identificación de las diferentes variantes genéticas asociadas a enfermedades concretas contribuye a conocer mejor los procesos fisiopatológicos, y con ello poder identificar nuevas características biológicas, marcadores de diagnóstico y pronóstico, y nuevas dianas terapéuticas. En este sentido, la incorporación de PRS a los modelos de predicción de riesgo tiene un gran potencial para estratificar a la población de forma más precisa y personalizada, y calcular la probabilidad de beneficiarse de determinados tratamientos de manera que se puedan priorizar las terapias en base a dicho cálculo. De este modo, los PRS en la Medicina del Futuro, además de mejorar la prevención primaria y secundaria, pueden ser relevantes para:<sup>15</sup>

- **Diagnóstico.** Es posible emplear los PRS para clasificar las enfermedades o subtipos de enfermedades que padecen las personas y descartar otros posibles diagnósticos. Esta aplicación puede resultar especialmente beneficiosa en aquellas enfermedades en las que el diagnóstico se asocia a un proceso costoso o prolongado en el tiempo, y por lo tanto que repercuta negativamente en la salud de la persona. Un ejemplo de ello es la reciente publicación de un PRS para la valoración clínica del riesgo de fractura en pacientes con osteoporosis. Se ha visto que la inclusión de este PRS en los programas de cribado de riesgo de fracturas por osteoporosis, aumentaba la capacidad de identificación de individuos con un alto riesgo de fractura. De esta manera, es posible reducir el número de individuos que deben someterse a valoraciones clínicas sin disminuir la sensibilidad y especificidad del diagnóstico y aumentando la eficiencia del proceso.<sup>31</sup>

- **Pronóstico.** Es posible desarrollar PRS para **estimar el riesgo de sufrir eventos adversos asociados a enfermedades**, incluido el propio avance de las mismas. Un ejemplo reciente desarrollado para la COVID-19 demostró la existencia de numerosas variantes comunes en genes que modulan la severidad de la enfermedad causada por SARS-CoV-2.<sup>32</sup> Estas variantes están estrechamente asociadas con un aumento del riesgo de hospitalización y sintomatología grave, especialmente en individuos con otros factores de riesgo como la edad avanzada. Esto sugiere que el potencial desarrollo de PRS para COVID-19 podría ayudar a priorizar la inmunización o tratamiento con anticuerpos monoclonales en aquellos individuos de mayor riesgo cuando el suministro de vacunas es limitado.<sup>32,33</sup>
- **Tratamiento.** Los PRS pueden emplearse en el futuro para la **predicción de la respuesta al tratamiento** de un paciente o población, permitiendo dirigir las estrategias terapéuticas en función de la susceptibilidad de los pacientes para beneficiarse de ellas, así como la **identificación de nuevas dianas terapéuticas a partir de la información genética**. Por ejemplo, en el caso de la enfermedad de Alzheimer, cuya heredabilidad, es decir, el grado en el que la patología está determinada genéticamente, se sitúa entre el 60% y el 80%, el Consorcio Europeo de Alzheimer y Demencias ha reunido, recientemente, un conjunto de datos de GWAS de 20.464 pacientes de Alzheimer y 22.244 controles de 15 países europeos. A partir de esta información, se han identificado 83 SNPs relacionados de manera significativa con la enfermedad de Alzheimer y otras demencias asociadas. A partir de esta información se ha desarrollado un PRS y se está evaluando su aplicación prospectiva en poblaciones y pacientes con deterioro cognitivo leve. Este estudio ha detectado, junto con la edad como factor de riesgo, una asociación significativa del PRS con el desarrollo de Alzheimer independiente del estatus de gen de apolipoproteína E<sup>4</sup>. Este tipo de estudios sugieren que la caracterización del escenario genético en el que se desarrolla la enfermedad de Alzheimer puede contribuir a mejorar el conocimiento sobre la fisiopatología de la enfermedad, abriendo nuevos horizontes a la identificación de dianas terapéuticas y desarrollo de nuevas terapias basadas en genómica traslacional.<sup>34</sup>



# RETOS

Los modelos de predicción de riesgo son una herramienta dirigida a conocer el riesgo a desarrollar diferentes enfermedades en individuos y poblaciones. Estos modelos pueden, emplearse para implementar estrategias preventivas y de salud pública para evitar el desarrollo de enfermedades, así como en la toma de decisiones clínicas en base al riesgo. Tradicionalmente, los modelos se han basado en factores de riesgo convencionales, como la edad, el sexo, determinadas variables clínicas o la historia familiar. Sin embargo, gracias al desarrollo de tecnologías y herramientas de Medicina Personalizada de Precisión como las ciencias ómicas, en los últimos años se ha avanzado considerablemente en el estudio de las enfermedades, pudiendo obtener cada vez más información sobre su complejidad.

En este contexto, la integración de la información obtenida a partir de estas ciencias en el marco de la Medicina Personalizada de Precisión permitirá identificar los sujetos más susceptibles de beneficiarse de cada estrategia o intervención y con ello, dirigir y optimizar los recursos sanitarios. Uno de los grandes avances en este sentido, ha sido la incorporación de la información genética a los modelos de predicción de riesgo a través de los PRS, si bien, aunque su desarrollo se encuentra en estadios iniciales de generación de evidencia, se prevé que esta integración mejore sustancialmente la medicina preventiva y la salud pública.<sup>5</sup>

Por ello, de cara a avanzar en la incorporación de modelos de predicción y en el uso de PRS como estrategias de medicina preventiva y la salud pública en el futuro, es necesario abordar retos de diferente naturaleza.

## RETOS ANALÍTICOS

Es previsible que el desarrollo de modelos de predicción mejore con el tiempo gracias al advenimiento de

nuevas tecnologías que permitan obtener, analizar e integrar cada vez más información a estos modelos, como, por ejemplo, la información derivada del desarrollo de las ciencias ómicas (microbioma, exposoma, etc.). En este sentido, de cara a conseguir una mejora de los modelos de predicción e integrarlos en la práctica clínica, se debe hacer frente a una serie de retos de carácter analítico que afectan principalmente a la generación y validación de los modelos.

- **La necesidad de profundizar en el conocimiento sobre el genoma en su totalidad.** Aunque se han producido grandes avances en el campo de la genómica, la integración de la información genética en los modelos de predicción de riesgo, a la hora de su aplicación clínica, es todavía un reto por el camino que queda por recorrer para la comprensión total de cómo funciona el genoma humano. Concretamente, es necesario elucidar las relaciones que existen entre el genotipo y los eventos clínicos (el fenotipo), ya que es una barrera para poder establecer asociaciones precisas entre la información genética de los individuos y los estados de salud y enfermedad, que es la base para la integración de los PRS en la predicción de riesgo.
- **La falta de datos de todas las poblaciones.** Actualmente los datos poblacionales a partir de los cuales se desarrollan los modelos proceden de determinados subgrupos de población, principalmente de algunas poblaciones europeas de origen caucásico. Esto supone una limitación sobre la validez de los modelos, ya que en muchos casos la información no se puede extrapolar de una población a otra y, por tanto, no se puede generalizar su aplicación a toda la población mundial.
- **La necesidad de ensayos clínicos y estudios piloto sobre los modelos desarrollados.** Si bien es cierto que se está avanzando en la identificación de

SNPs para construir PRS que contribuyan a mejorar los modelos de predicción de riesgo, las estimaciones se están realizando a partir de estudios retrospectivos. En este sentido, es necesario estudiar su aplicabilidad en la práctica clínica a través de ensayos clínicos prospectivos y aleatorizados para que los PRS puedan incorporarse de forma rutinaria a la práctica clínica.

- **La falta de diversidad de las muestras empleadas en los estudios prospectivos.** En los bancos de muestras que se emplean para la realización de estudios prospectivos, existe una mayor representación de muestras de pacientes con patologías más prevalentes, lo que limita la utilidad y validez en el estudio de enfermedades poco frecuentes. Ante esto, es necesario contar con tamaños muestrales mayores en los estudios prospectivos.

acceso a test y pruebas genéticas fuera del entorno clínico. De hecho, en los últimos años, han proliferado las empresas que ofrecen a los ciudadanos test basados en PRS que dan información que puede tener un alto impacto sobre su vida y su salud.<sup>10,11</sup> Las personas que acceden a este tipo de test, no son informados de forma correcta sobre el significado de los resultados, pudiendo generar unas expectativas demasiado altas sobre la información que se puede obtener a partir del estudio de los genes.

- **La falta de regulación en el campo de los PRS.** Actualmente la normativa no está acompañada a los avances tecnológicos existiendo la necesidad de regular este campo no sólo en el ámbito clínico, también en el ámbito social, ya que en el futuro puede suponer un mal uso de la información.

---

## RETOS DE IMPLEMENTACIÓN

Además de los retos analíticos, existen una serie de desafíos y barreras para implementar en la práctica clínica habitual los modelos de predicción de riesgo de enfermedades y los PRS que hayan sobrepasado la fase de desarrollo e investigación.

- **La variabilidad en el desarrollo de los modelos por falta de estándares para su obtención y validación.** En el desarrollo de los modelos de predicción y la incorporación de los PRS a estos modelos, existe una elevada variación dependiendo de la metodología utilizada. Esto se debe a que el proceso de obtención de la información genética y la construcción de los modelos no está estandarizado. En este sentido, aunque existen iniciativas dirigidas a estandarizar los PRS, su aplicación no se ha extendido, de manera que sigue existiendo una variabilidad importante en los modelos que limita la validez de los mismos.
- **La gestión de las expectativas de la población sobre los resultados de los tests genómicos.** El análisis e interpretación de los modelos de predicción y los PRS es complejo y requiere de un profesional cualificado para la correcta explicación de sus resultados. Sin embargo, la población general tiene

---

## RETOS ÉTICOS

La generación de los modelos de predicción de riesgo y el manejo de información genética de los pacientes está sujeta a una serie de consideraciones éticas con el objetivo de asegurar la seguridad de su traslación a la práctica clínica debido al potencial impacto que tiene el manejo de dicha información sobre los propios pacientes.

- **El impacto psicosocial de los resultados sobre la salud de los individuos.** Cada persona puede responder de manera diferente ante los resultados de un test genético o de una predicción de riesgo. Algunas personas pueden reaccionar negativamente al conocer estos resultados, lo que puede repercutir en su salud. Por ello, resulta especialmente importante la consideración hacia las distintas sensibilidades por parte de los profesionales sanitarios responsables de informar de estos resultados.
- **Las implicaciones que tiene informar al paciente y/o individuo de los resultados de modelos predictivos de enfermedades en situaciones en las que no existen soluciones disponibles para su abordaje.** Los modelos de predicción de riesgo tienen como objetivo contribuir a mejorar las estrategias de prevención y abordaje de las distintas enfermedades, para lo que los resultados deben ser



---

correctamente interpretados. Sin embargo, en algunos casos la utilidad de los resultados puede resultar limitada o incluso ser contraproducente si no existe un tratamiento disponible o un abordaje preventivo.

- **Las barreras de accesibilidad a los modelos de predicción y PRS.** En relación a lo comentado en el apartado de retos analíticos, hay una falta de datos en población no europea y no caucásica, que hace que los modelos de predicción no sean aplicables a todas las poblaciones sin una validación previa en dicha población. Por ello, la aplicación de los modelos de predicción de riesgo generados hasta ahora puede generar desigualdades en el campo de la salud dependiendo del origen de cada población.
- **El riesgo de limitar las intervenciones en función de resultados de PRS.** A pesar de que los modelos de predicción de riesgo permiten identificar sujetos más susceptibles de beneficiarse de determinadas intervenciones preventivas, diagnósticas y/o terapéuticas contribuyendo a optimizar recursos sanitarios, no debería privarse a la población general de dichas intervenciones si también pueden repercutir positivamente sobre su salud.





# CONCLUSIONES Y RECOMENDACIONES

---

El campo de la predicción de riesgos es a día de hoy una parte inherente de la práctica clínica. El sistema de salud busca predecir los eventos relacionados con la salud de la población a fin de poder anticiparse y reducir la probabilidad de que ocurran o, si esto no es posible, reducir la gravedad de los mismos. En este sentido, los modelos de predicción de riesgo se han posicionado como una herramienta predictiva útil que permite implementar estrategias preventivas y optimizar las estrategias actuales de diagnóstico y pronóstico.

En este contexto, han aparecido nuevas tecnologías y fuentes de información en el marco de la Medicina Personalizada de Precisión tales como la genómica, que han supuesto un cambio de paradigma en la forma de estudiar y entender las enfermedades. De este modo, se prevé que la incorporación a los modelos de predicción de la información procedente de estos avances permitirá optimizarlos y avanzar en sus aplicaciones, integrando la Medicina Predictiva como un componente esencial de la Medicina Personalizada de Precisión. Para ello, de cara a avanzar en la implementación de los modelos a la práctica, se deben tener en cuenta una serie de consideraciones.

## RECOMENDACIONES

---

- **Impulsar el estudio en ciencias ómicas.** Aunque es previsible que la genómica tenga un papel crucial en el desarrollo de modelos de predicción más avanzados con la incorporación de los PRS, es importante conocer toda la información posible sobre los individuos que pueda afectar a su salud. En este

sentido, el exposoma y otros factores condicionantes de salud o enfermedad modificables, entre otras ómicas podría contribuir a mejorar el desarrollo de modelos de predicción.

- **Crear bancos de información integrados de información sobre condicionantes de salud y enfermedad.** La generación de modelos de predicción óptimos requiere del uso de información procedente de diferentes fuentes y tecnologías. La creación de registros de datos en los que disponer y compartir información genética y de otros determinantes de la salud resultaría útil para el desarrollo de modelos de predicción. Aunque ya existen iniciativas en este sentido en otros países, como el UK biobank o la iniciativa All of Us en EE.UU., en España no se dispone de este tipo de recurso.
- **Generar modelos validados en las diferentes poblaciones.** En relación con el punto anterior, de cara a estimar el riesgo de desarrollar un determinado evento de salud en una población concreta, es imprescindible disponer de modelos que hayan sido elaborados a partir de datos de dicha población sobre la que se va a aplicar o, en su defecto, que hayan sido validados en dicha población, verificando la calibración y discriminación de los mismos. Esto, es esencial, no sólo para la eficacia de los modelos, sino también para favorecer la equidad en su uso y aplicación en diferentes poblaciones.
- **Establecer estándares para la generación de modelos de predicción de riesgos y PRS.** Actualmente existe una gran variabilidad sobre los parámetros y metodologías que se utilizan en la construcción de



modelos de predicción de riesgos. Es recomendable impulsar iniciativas dirigidas a establecer criterios bien definidos a fin de reducir la variabilidad y facilitar su implementación. Además, la definición de estándares de todo el proceso de desarrollo de modelos de predicción, y en concreto de PRS, contribuirá a su futura traslación a la práctica clínica

- **Evaluar la eficiencia de los modelos en el contexto sanitario pertinente.** Con el objetivo de incorporar los modelos de predicción de riesgos a la práctica clínica real en España y contribuir a la sostenibilidad del sistema, es importante generar datos económicos que permitan evidenciar las ventajas que supone su uso desde el punto de vista de la eficiencia en el entorno sanitario actual y su potencial incorporación

en la cartera común de servicios del Sistema Nacional de Salud.

- **Aumentar la formación en el campo de la predicción de riesgos.** La responsabilidad de los profesionales sanitarios en la solicitud, interpretación y transmisión de los resultados a los pacientes implica necesariamente su formación en el funcionamiento de los modelos de predicción de riesgo, sus limitaciones e implicaciones.
- **Concienciar a la población sobre las implicaciones del uso de la información genética.** Es importante no generar excesivas expectativas en la población sobre la capacidad predictiva de los modelos predictivos y transmitir de manera correcta sus resultados.





## BIBLIOGRAFÍA

1. Fundación Instituto Roche. Propuesta de Competencias En Medicina Personalizada de Precisión de Los Profesionales Sanitarios.; 2022. [www.institutoroche.es](http://www.institutoroche.es)
2. Khera A v., Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet.* 2018;50(9):1219-1224. doi:10.1038/s41588-018-0183-z
3. Chatterjee N, Shi J, García-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet.* 2016;17(7):392-406. doi:10.1038/nrg.2016.27
4. Lewis CM, Vassos E. Polygenic risk scores: From research tools to clinical instruments. *Genome Med.* 2020;12(1). doi:10.1186/s13073-020-00742-5
5. Slunecka JL, van der Zee MD, Beck JJ, et al. Implementation and implications for polygenic risk scores in healthcare. *Hum Genomics.* 2021;15(1). doi:10.1186/s40246-021-00339-y
6. Chen Z, Schunkert H. Genetics of coronary artery disease in the post-GWAS era. *J Intern Med.* 2021;290(5):980-992. doi:10.1111/joim.13362
7. National Human Genome Research Institute. Polygenic Risk Scores. Accessed September 28, 2022. <https://www.genome.gov/Health/Genomics-and-Medicine/Polygenic-risk-scores#three>
8. María Brión, José Ramón González-Juanatey, Ángel Carracedo. La puntuación de riesgo poligénico como factor clave en los modelos de predicción clínica cardiovascular. *Rev Esp Cardiol.* 2020;73(8):608-610. doi:10.1016/j.recesp.2019.08.006
9. Illumina. Considerations for polygenic risk score development. Published online 2021. Accessed October 6, 2022. <https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/polygenic-risk-score-application-note-m-gl-00103/polygenic-risk-score-application-note-m-gl-00103.pdf>
10. Wand H, Lambert SA, Tamburro C, et al. Improving reporting standards for polygenic scores in risk prediction studies. *Nature.* 2021;591(7849):211-219. doi:10.1038/s41586-021-03243-6
11. Janssens ACJW, Ioannidis JPA, Bedrosian S, et al. Strengthening the reporting of genetic risk prediction studies (GRIPS): Explanation and elaboration. *Eur J Clin Invest.* 2011;41(9):1010-1035. doi:10.1111/j.1365-2362.2011.02493.x
12. Privé F, Aschard H, Carmi S, et al. Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. *Am J Hum Genet.* 2022;109(1):12-23. doi:10.1016/j.ajhg.2021.11.008
13. European Commission. European “1+Million Genomes” Initiative. Published May 25, 2022. Accessed July 18, 2022. <https://digital-strategy.ec.europa.eu/en/policies/1-million-genomes>
14. Beyond 1 Million Genomes. Published 2020. Accessed July 18, 2022. <https://b1mg-project.eu/>
15. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet.* 2018;19(9):581-590. doi:10.1038/s41576-018-0018-x
16. Fundación Instituto Roche. MEDICINA PERSONALIZADA DE PRECISIÓN EN ESPAÑA: MAPA DE COMUNIDADES.; 2019. [www.institutoroche.es](http://www.institutoroche.es)
17. Costantino JP, Gail MH, Pee D, et al. Validation Studies for Models Projecting the Risk of Invasive and Total Breast Cancer Incidence. <https://academic.oup.com/jnci/article/91/18/1541/2606608>
18. Paquin M, Fasolino T, Steck MB. Breast cancer risk prediction models: Challenges in clinical application. *Clin J Oncol Nurs.* 2019;23(3):256-259. doi:10.1188/19.CJON.256-259
19. Pastor-Barriuso R, Ascunce N, Ederra M, et al. Recalibration of the Gail model for predicting invasive breast cancer risk in Spanish women: A population-based cohort study. *Breast Cancer Res Treat.* 2013;138(1):249-259. doi:10.1007/s10549-013-2428-y
20. BCAC - Breast Cancer Association Consortium. Published 2005. Accessed July 11, 2022. <https://bcac.ccge.medschl.cam.ac.uk/>
21. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast Cancer and

- Breast Cancer Subtypes. *Am J Hum Genet.* 2019;104(1):21-34. doi:10.1016/j.ajhg.2018.11.002
22. Centre for Cancer Genetic Epidemiology. BOADICEA. Accessed July 19, 2022. <https://ccge.medschl.cam.ac.uk/boadicea/>
  23. MyPeBS - My Personal Breast Screening. Accessed July 20, 2022. <https://www.mypebs.eu/es/>
  24. Hsu L, Jeon J, Brenner H, et al. A model to determine colorectal cancer risk using common genetic susceptibility loci. *Gastroenterology.* 2015;148(7):1330-1339.e14. doi:10.1053/j.gastro.2015.02.010
  25. Garcia-Closas M, Rothman N, Figueroa JD, et al. Common genetic polymorphisms modify the effect of smoking on absolute risk of bladder cancer. *Cancer Res.* 2013;73(7):2211-2220. doi:10.1158/0008-5472.CAN-12-2388
  26. SCORE2 working group, ESC Cardiovascular risk collaboration. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. *Eur Heart J.* 2021;42:2439-2454. doi:10.1093/eurheartj/ehab310
  27. National Heart Institute. Framingham Heart Study. Accessed September 6, 2022. <https://www.framinghamheartstudy.org/>
  28. Deloukas P, Kanoni S, Willenborg C, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet.* 2013;45(1):25-33. doi:10.1038/ng.2480
  29. Tanya Brigden A, Sanderson S, Janus J, et al. Implementing Polygenic Scores for Cardiovascular Disease into NHS Health Checks 2 Implementing Polygenic Scores into NHS Health Checks.; 2021. [www.phgfoundation.org](http://www.phgfoundation.org)
  30. Polygenic score pilot for heart disease begins. Genomics Education Programme. Published 2022. Accessed July 19, 2022. <https://www.genomicseducation.hee.nhs.uk/blog/polygenic-score-pilot-for-heart-disease-begins/#:~:text=Polygenic%20scores,aged%20between%2045%20and%2055.>
  31. Forgetta V, Keller-Baruch J, Forest M, et al. Development of a polygenic risk score to improve screening for fracture risk: A genetic risk prediction study. *PLoS Med.* 2020;17(7). doi:10.1371/journal.pmed.1003152
  32. Niemi MEK, Karjalainen J, Liao RG, et al. Mapping the human genetic architecture of COVID-19. *Nature.* 2021;600(7889):472-477. doi:10.1038/s41586-021-03767-x
  33. Horowitz JE, Kosmicki JA, Damask A, Sharma D, Roberts GHL. Genome-wide analysis in 756,646 individuals provides first genetic evidence that ACE2 expression influences COVID-19 risk and yields genetic risk scores predictive of severe disease. medRxiv. Published online 2021. doi:10.1101/2020.12.14.20248176
  34. Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet.* 2022;54(4):412-436. doi:10.1038/s41588-022-01024-z







