

Software de puntuación de riesgo poligénico: módulo Predict

Una potente solución de análisis de datos para la investigación de la PRS

- Una solución de evaluación de la relación genotipo-riesgo optimizada con una amplia cobertura de trastornos
- Su interfaz fácil de usar simplifica el análisis de datos y la generación de informes sobre la puntuación de riesgo genético
- Las puntuaciones de riesgo genético basadas en la ascendencia proporcionan una predicción muy precisa del riesgo de enfermedad



Introducción

Las enfermedades poligénicas, como la diabetes tipo 2, los cánceres de mama y de próstata, la hipertensión, la arteriopatía coronaria y la enfermedad de Alzheimer, están causadas por los efectos combinados de cientos o miles de genes de susceptibilidad. Cada una de estas variantes comunes asociadas a la enfermedad contribuye al riesgo global de la misma, que también puede ser influenciado por factores ambientales, lo que hace que la predicción de la susceptibilidad a la enfermedad suponga un reto. Las puntuaciones de riesgo poligénico (PRS, Polygenic Risk Scores), o las puntuaciones de riesgo genético (GRS, Genetic Risk Scores), representan el número total de variantes genéticas que aumentan el riesgo individual de desarrollar una enfermedad específica. Esta novedosa herramienta de la medicina de precisión se puede emplear para estimar la probabilidad de que una persona desarrolle una enfermedad, estratificar las poblaciones según el riesgo de enfermedad e identificar a las personas que se beneficiarían de una supervisión adicional o de intervenciones preventivas.^{1,2} Sin embargo, la PRS es un campo en constante evolución que requiere más investigación antes de llegar a considerarla una herramienta consolidada.

Los estudios de PRS han demostrado la capacidad de identificar personas con un alto riesgo genético de desarrollar enfermedades poligénicas complejas (Figura 1).³⁻⁷

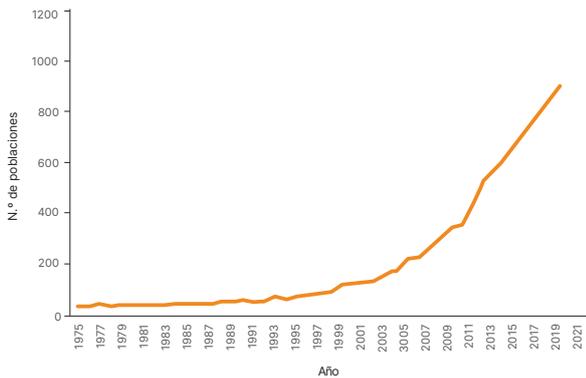


Figura 1: Número de publicaciones con puntuaciones de riesgo poligénico. Una búsqueda de palabras clave de "PRS" o "GRS" en las publicaciones de la base de datos PubMed muestra un notable aumento de los estudios que emplean esta herramienta para determinar el riesgo de enfermedad. Actualizado a febrero de 2022.

Los investigadores pueden hacer uso de los datos procedentes de los estudios de asociación de todo el genoma (GWAS, Genome-Wide Association Studies) para calcular la PRS, definida como la suma ponderada de los recuentos de alelos de las variantes asociadas a la enfermedad (Figura 2).⁸ Sin embargo, debido al gran número de variantes identificadas en los estudios de poblaciones a gran escala, la generación de la PRS por persona es un procedimiento con un gran consumo de recursos de computación. Se requieren amplios conocimientos de bioinformática para llevar a cabo el proceso de la PRS, desde la asignación de los datos genómicos hasta el control de calidad y la generación de informes, lo que limita el uso generalizado de la PRS en la investigación clínica.

El módulo Predict de Illumina, desarrollado en colaboración con Allelica, simplifica el proceso de cálculo, análisis y generación de informes de la PRS con una interfaz fácil de usar para generar PRS individualizadas a escala. El módulo Predict está disponible como parte de un completo conjunto de herramientas que incluye Infinium™ Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip, que brinda una solución de predicción de la relación genotipo-riesgo optimizada para ayudar a ampliar la investigación en el campo en constante evolución de la evaluación de la PRS.

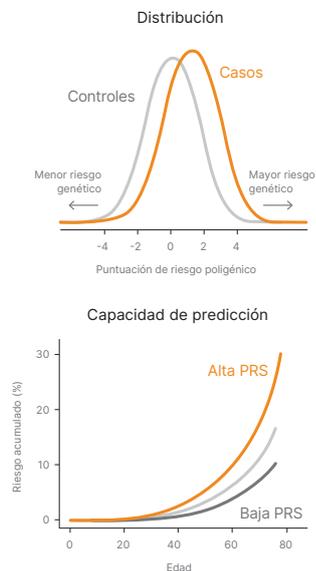


Figura 2: Resultados representativos de la predicción del riesgo de enfermedad usando la PRS. El gráfico superior muestra las distribuciones de la PRS de los casos (línea de color naranja) y los controles (línea de color gris). El gráfico inferior muestra la capacidad de predicción de la PRS de una muestra individual que demuestra un mayor riesgo de desarrollar la enfermedad antes que las muestras de referencia con un percentil de PRS bajo o promedio.

Evaluación precisa del riesgo con la distribución de la PRS basada en el fenotipo

El módulo Predict calcula las puntuaciones individuales de PRS y genera informes sobre el riesgo de enfermedad en relación con una población de referencia. Si bien la comparación de una PRS numérica con una distribución de población que cuente con datos del genotipo proporciona un percentil de la puntuación de riesgo en relación con la cohorte de referencia, esta no proporciona información sobre el riesgo absoluto de enfermedad. A fin de relacionar la PRS con el riesgo real de enfermedad, todas las distribuciones de población que se usan en el módulo Predict contienen datos del fenotipo, además del genotipo. El módulo Predict incluye puntuaciones del [catálogo de puntuación poligénica \(PGS, Polygenic Score\)](#) y puntuaciones de riesgo desarrolladas por Allelica. Las medidas de control de calidad específicas de cada PRS se incorporan al módulo Predict con el fin de mantener un alto rendimiento predictivo.

Cálculo del riesgo de enfermedad basado en la ascendencia

La posibilidad de heredar enfermedades poligénicas suele ser variable en diversas poblaciones, lo que hace que la evaluación del riesgo basada en la ascendencia resulte fundamental para la determinación precisa de las PRS.⁹ El módulo Predict, destinado para uso exclusivo en investigación, incluye todas las PRS incluidas en el catálogo de PGS con aproximadamente 1000 modelos de riesgo precompilados que comprenden más de 20 rasgos y más de 80 PRS para cada una de las cinco ascendencias continentales. Como resultado, el software proporciona una cobertura completa de las variantes potencialmente patógenas que contribuyen

al riesgo de enfermedad. Esta herramienta ofrece una calibración de la ascendencia a escala precisa haciendo uso de un espacio armonizado de análisis de componentes principales (PCA, Principal Component Analysis) de aproximadamente 800 000 muestras individuales. El módulo Predict también facilita a los investigadores la generación de informes de PRS de forma automatizada con predicciones de riesgo específicas para cada ascendencia y ofrece una amplia cobertura de trastornos, como la oncología, las enfermedades cardiovasculares y el bienestar general.

Flujo de trabajo de análisis de datos simplificado

El módulo Predict es compatible con los datos de entrada generados en microarrays o sistemas de secuenciación del genoma completo. Los archivos de datos de los sistemas de secuenciación de Illumina se envían a la plataforma en la nube de Illumina y se seleccionan las puntuaciones de interés, ya sea entre las más de 2600 puntuaciones incluidas en el módulo Predict o las generadas por el laboratorio, con el fin de calcular una PRS individualizada para cada muestra ([Figura 3](#)). Una vez calculada la puntuación, el módulo Predict genera informes personalizados con evaluaciones de riesgo completas que incluyen otras variables, como el sexo y el origen étnico, para una predicción de riesgo precisa.

En el caso de los datos de secuenciación de baja cobertura, el módulo Predict ofrece una opción de asignación para aumentar el número de variantes genéticas presentes en el conjunto de datos. Gracias al mecanismo de asignación, se pueden asignar aproximadamente 10 000 muestras en menos de tres horas con una precisión superior al 99 % y una alta concordancia con Infinium Global Screening Array.



Figura 3: Flujo de trabajo simplificado para la determinación de la PRS mediante el módulo Predict. El módulo Predict ofrece una solución integral de predicción del riesgo para realizar evaluaciones rápidas y precisas del riesgo de enfermedad sin necesidad de contar con amplios recursos bioinformáticos.

Interfaz fácil de usar

Los datos de genotipo sin procesar de los sistemas de secuenciación de Illumina se integran en la intuitiva interfaz de BaseSpace™ Sequence Hub, lo que permite un fácil análisis y cálculo de la PRS con tan solo hacer clic. Con el módulo Predict se pueden calcular hasta 24 PRS en 15 minutos. Se genera un informe estandarizado de PRS para cada muestra. Este completo informe incluye puntuaciones numéricas poligénicas basadas en la ascendencia y el riesgo relativo de enfermedad a lo largo de la vida, lo que proporciona un contexto fundamental para la interpretación de la PRS (Figura 4).

Plataforma de análisis de datos basada en la nube segura

El módulo Predict está integrado en la infraestructura basada en la nube de Illumina Connected Analytics (ICA), que está optimizada para la calidad y el rendimiento. Además de proporcionar una plataforma escalable y segura, ICA posibilita el cálculo en paralelo de hasta 1152 muestras que se evaluarán en un solo análisis, para una determinación rápida, precisa y basada en la ascendencia de la PRS. Los datos generados con el módulo Predict se pueden integrar con otras herramientas también alojadas en ICA, incluyendo procesos de análisis de farmacogenómica y genotipado.

Solución integral de la predicción de la relación genotipo-riesgo

El módulo Predict está disponible como parte de un kit de herramientas de PRS de alto rendimiento que incluye Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip para una predicción de riesgo de alta precisión y una generación de informes estandarizada. El BeadChip presenta un contenido actualizado con aproximadamente 160 000 nuevos marcadores de PRS, desarrollados en colaboración con Allelica, que se añaden a la amplia estructura de todo el genoma de Infinium Global Diversity Array. El amplio contenido del BeadChip se seleccionó minuciosamente para ofrecer una amplia cobertura de los principales trastornos poligénicos, como la diabetes tipo 1 y 2, los cánceres de mama, colon y próstata, la arteriopatía coronaria y la enfermedad de Alzheimer. iScan™ System, la tecnología Infinium LCG Assay y el módulo Predict integrado trabajan en conjunto para crear una solución de predicción de la relación genotipo-riesgo completa y de alto rendimiento para respaldar la investigación de la SRP (Figura 5).

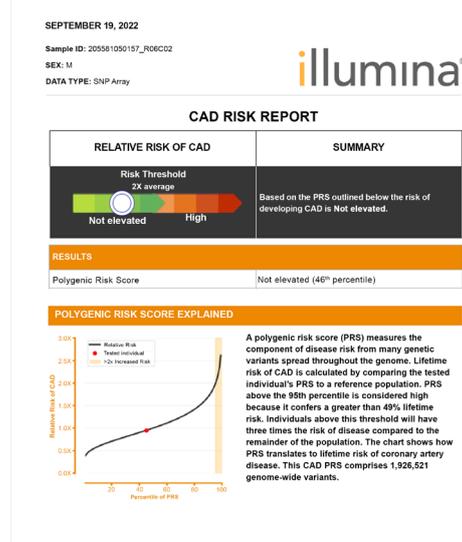
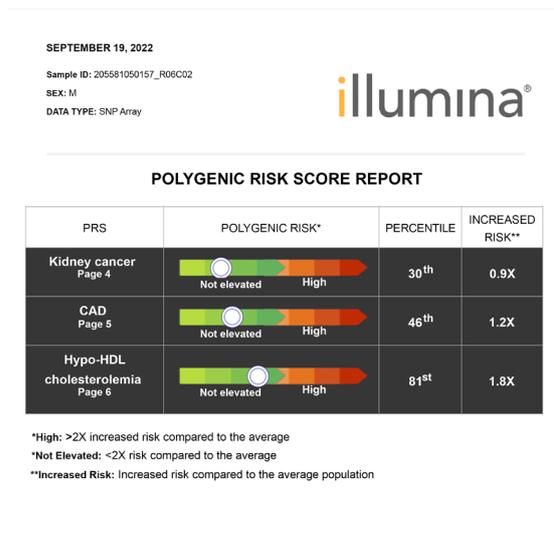


Figura 4: Ejemplo de informe de predicción de riesgo. El módulo Predict proporciona informes individualizados que transmiten el percentil de riesgo relativo de cada muestra en un formato fácil de visualizar con información sobre la PRS que se usó para el cálculo del riesgo.



Figura 5: Completo conjunto de herramientas para la investigación del PRS. El módulo Predict, combinado con Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip, brinda una solución de predicción de la relación genotipo-riesgo optimizada para una evaluación de gran precisión de la PRS y una generación de informes estandarizada.

Resumen

El módulo Predict usa potentes algoritmos de análisis y procesos bioinformáticos desarrollados en colaboración con Allelica para simplificar la evaluación y la generación de informes de las PRS. La interfaz de usuario de BaseSpace Sequence Hub es fácil de usar y ofrece una determinación rápida, precisa y basada en la ascendencia con una amplia cobertura de trastornos poligénicos. El módulo Predict se beneficia de los datos del amplio catálogo de PGS para calcular las puntuaciones de riesgo relativas. Este módulo de software está disponible como parte de un completo kit de herramientas que incluye Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip, que brinda una solución de predicción de la relación genotipo-riesgo optimizada para la investigación en la medicina de precisión.

Información adicional

Puntuaciones de riesgo poligénico, illumina.com/areas-of-interest/complex-disease-genomics/polygenic-risk-scores

Illumina Connected Analytics, illumina.com/products/by-type/informatics-products/connected-analytics

Bibliografía

1. Dudbridge F. [Power and predictive accuracy of polygenic risk scores](#). *PLoS Genet*. 2013;9(3):e1003348. doi:10.1371/journal.pgen.1003348
2. Lewis CM, Vassos E. [Polygenic risk scores: from research tools to clinical instruments](#). *Genome Med*. 2020;12(1):44. doi:10.1186/s13073-020-00742-5
3. Khera AV, 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
4. 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
5. Huynh-Le MP, Fan CC, Karunamuni R, et al. [Polygenic hazard score is associated with prostate cancer in multi-ethnic populations](#). *Nat Commun*. 2021;12(1):1236. doi:10.1038/s41467-021-21287-0
6. Elliott J, Bodinier B, Bond TA, et al. [Predictive Accuracy of a Polygenic Risk Score-Enhanced Prediction Model vs a Clinical Risk Score for Coronary Artery Disease](#). *JAMA*. 2020;323(7):636-645. doi:10.1001/jama.2019.22241
7. Ge T, Irvin MR, Patki A, et al. [Development and validation of a trans-ancestry polygenic risk score for type 2 diabetes in diverse populations](#). *Genome Med*. 2022;14(1):70. doi:10.1186/s13073-022-01074-2

8. Collister JA, Liu X, Clifton L. [Calculating polygenic risk scores \(PRS\) in UK Biobank: A practical guide for epidemiologists.](#) *Front Genet.* 2022;13:818574. doi:10.3389/fgene.2022.818574.
9. Duncan L, Shen H, Gelaye B, et al. [Analysis of polygenic risk score usage and performance in diverse human populations.](#) *Nat Commun.* 2019;10(1):3328. doi:10.1038/s41467-019-11112-0

Datos para realizar pedidos

Producto	N.º de catálogo
Módulo Predict	20086666
Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip (16 samples)	20090683
Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip (48 samples)	20090684
Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip (96 samples)	20090685
Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip (384 samples)	20090686



1 800 809 4566 (llamada gratuita, EE. UU.) | tel.: +1 858 202 4566
 techsupport@illumina.com | www.illumina.com

© 2022 Illumina, Inc. Todos los derechos reservados. Todas las marcas comerciales pertenecen a Illumina, Inc. o a sus respectivos propietarios. Si desea consultar información específica sobre las marcas comerciales, consulte www.illumina.com/company/legal.html.
 M-GL-01376 ESP v1.0