

# Software per il punteggio di rischio poligenico: Predict Module

Una soluzione efficace  
di analisi dei dati per  
la ricerca sul PRS

- Soluzione ottimizzata per la valutazione del rischio a partire dal genotipo su un ampio numero di disordini
- Interfaccia di facile utilizzo che semplifica l'analisi dei dati e la creazione di report per il punteggio di rischio genetico
- Punteggi di rischio genetico con informazioni ancestrali forniscono una predizione altamente accurata del rischio di patologia

**illumina**<sup>®</sup>

## Introduzione

Le patologie poligeniche, come il diabete di tipo 2, il cancro al seno e il cancro alla prostata, l'ipertensione, la patologia coronarica e l'Alzheimer, sono causate da un insieme di effetti provocati da centinaia a migliaia di geni suscettibili. Ciascuna di queste varianti comuni associate alla patologia contribuisce al rischio di patologia complessivo che può essere influenzato anche da fattori ambientali. Ciò rende difficile la predizione della suscettibilità della patologia. I punteggi di rischio poligenico (PRS, Polygenic Risk Score), o punteggi di rischio genetico (GRS, Genetic Risk Score), rappresentano il numero totale di varianti genetiche che aumentano il rischio, per un individuo, di sviluppare una determinata patologia. Questo strumento per medicina di precisione è promettente e può essere utilizzato per stimare la probabilità di un individuo di sviluppare una malattia, stratificare le popolazioni in base al rischio di malattia e di identificare gli individui che potrebbero trarre vantaggio da ulteriore monitoraggio o misure preventive.<sup>1,2</sup> Tuttavia, il PRS è un campo emergente e richiede maggiore ricerca prima di evolversi.

Gli studi sul PRS hanno dimostrato la capacità di identificare gli individui che presentano un alto rischio geneticamente determinato di sviluppare patologie poligeniche complesse (Figura 1).<sup>3-7</sup>

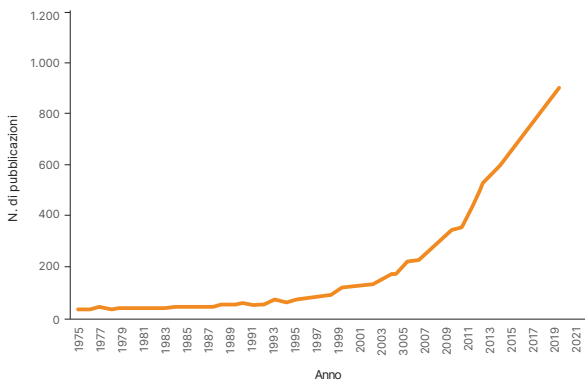


Figura 1: Numero di pubblicazioni con punteggi di rischio poligenico: una ricerca per parola chiave "PRS" o "GRS" nelle pubblicazioni contenute nel database PubMed mostra un marcato aumento di studi che utilizzano questo strumento per determinare il rischio di patologia. Dati al febbraio 2022.

I ricercatori possono calcolare il PRS sfruttando i dati ottenuti dagli studi di associazione sull'intero genoma (GWAS, Genome-Wide Association Study). Tali punteggi sono definiti come una somma ponderata dei conteggi degli alleli sulle varianti associate alla patologia (Figura 2).<sup>8</sup> Tuttavia, dato il vasto numero di varianti identificato negli studi sulla popolazione su larga scala, il PRS per singolo individuo è un calcolo intensivo. Per eseguire la pipeline PRS è necessaria una profonda esperienza in bioinformatica, a partire dall'immissione dei dati genomici fino al controllo qualità e alla creazione di report. Per questo risulta limitato l'utilizzo diffuso del PRS nella ricerca clinica.

Il Predict Module di Illumina, sviluppato in partnership con Allelica, semplifica il processo di calcolo, di analisi e di creazione di report del PRS grazie a un'interfaccia di facile utilizzo che consente di generare correttamente punteggi PRS personalizzati. Il Predict Module fa parte di una serie di strumenti completa che include l'Infinium™ Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip e offre una soluzione ottimizzata di predizione del rischio a partire dal genotipo che aiuta la ricerca ad ampliarsi nel campo emergente della valutazione del PRS.

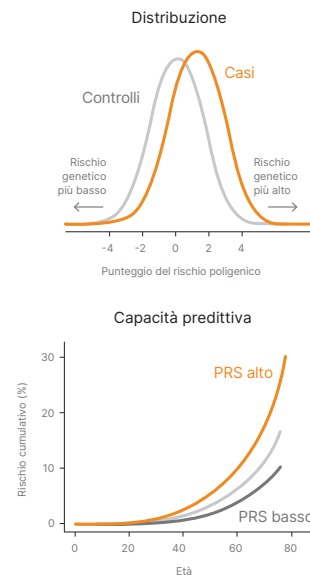


Figura 2: Risultati rappresentativi per la predizione del rischio di patologia utilizzando il PRS: il pannello superiore mostra la distribuzione del PRS per casi (linea arancione) e controlli (linea grigia). Il pannello inferiore mostra la capacità predittiva del PRS per un singolo campione dimostrando un aumentato rischio di sviluppare la patologia in anticipo rispetto ai campioni di riferimento con un percentile di PRS basso o medio.

## Valutazione accurata del rischio con la distribuzione del PRS in base al fenotipo

Il Predict Module calcola i PRS individuali e riporta il rischio di patologia rispetto a una popolazione di riferimento. Sebbene il confronto di un PRS numerico rispetto a una distribuzione della popolazione utilizzando i dati sul genotipo fornisca un punteggio di rischio in percentile rispetto alla coorte di riferimento, non fornisce informazioni sul rischio assoluto di patologia. Per collegare il PRS all'effettivo rischio di patologia, tutte le distribuzioni delle popolazioni utilizzate nel Predict Module contengono, oltre ai dati sul genotipo, i dati sul fenotipo. Il Predict Module include punteggi contenuti nel catalogo [Polygenic Score \(PGS\)](#) e i punteggi di rischio sviluppati da Allelica. Le misure di controllo qualità specifiche per ciascun PRS sono contenute nel Predict Module per mantenere elevate prestazioni predittive.

## Calcolo del rischio di patologia con informazioni ancestrali

L'ereditabilità di patologie poligeniche varia spesso tra popolazioni diverse, il che rende fondamentale la valutazione del rischio con informazioni ancestrali per l'accurata determinazione del PRS.<sup>9</sup> Il Predict Module, che è previsto solo a uso di ricerca, include tutti i PRS contenuti nel catalogo PGS con circa 1.000 modelli di rischio precompilati che comprendono più di 20 tratti su più di 80 PRS per ciascuna delle cinque origini ancestrali continentali. In questo modo, il software fornisce una copertura esaustiva delle varianti potenzialmente patogene che contribuiscono al rischio

di patologia. Questo strumento fornisce la calibrazione particolareggiata dell'origine ancestrale sfruttando uno spazio di analisi delle componenti principali (PCA, Principal Component Analysis) di circa 800.000 campioni individuali. Il Predict Module consente inoltre ai ricercatori di generare report PRS automatizzati con predizione di rischio specifico per l'origine ancestrale e copre un ampio numero di disordini inclusi oncologia, patologia cardiovascolare e benessere generale.

## Flusso di lavoro semplificato per l'analisi dei dati

Il Predict Module è compatibile con i dati di input generati sui microarray o con i sistemi di sequenziamento per l'intero genoma. I file di dati ottenuti dai sistemi di sequenziamento di Illumina vengono inviati alla piattaforma sul cloud di Illumina e vengono selezionati i punteggi di interesse, sia da più di 2.600 punteggi inclusi nel Predict Module sia da PRS personalizzati generati in laboratorio, per calcolare un PRS personalizzato per ogni campione ([Figura 3](#)). Dopo il calcolo di un punteggio, il Predict Module genera report personalizzati con valutazioni complete del rischio che includono altre variabili, come sesso ed etnia, per l'accurata predizione del rischio.

Per i dati di sequenziamento con bassa copertura, il Predict Module offre un'opzione di imputazione per aumentare il numero di varianti genetiche presenti nel set di dati. Con il motore di imputazione, possono essere imputati circa 10.000 campioni in meno di tre ore con una precisione che supera il 99% e con un'elevata concordanza con l'Infinium Global Screening Array.



Figura 3: Flusso di lavoro semplificato per la determinazione del PRS utilizzando il Predict Module: il Predict Module offre una soluzione completa per la predizione del rischio che consente valutazioni veloci e accurate di rischio di patologia senza richiedere grandi risorse bioinformatiche.

## Interfaccia di facile utilizzo

I dati del genotipo non elaborati ottenuti dai sistemi di sequenziamento di Illumina si integrano con l'interfaccia intuitiva di BaseSpace™ Sequence Hub che consente una semplice analisi "punta e clicca" e il calcolo del PRS. Con il Predict Module è possibile calcolare 24 PRS in 15 minuti. Per ogni campione viene generato un report PRS standardizzato. Questo report esaustivo include i punteggi poligenici numerici con informazioni ancestrali e il relativo rischio di contrarre la patologia nel corso della vita. Questo fornisce contesto critico per l'interpretazione del PRS (Figura 4).

## Piattaforma di analisi dei dati sicura basata sul cloud

Il Predict Module è integrato nell'infrastruttura basata sul cloud Illumina Connected Analytics (ICA) che consente di ottimizzare la qualità e le prestazioni. Oltre a fornire una piattaforma scalabile e sicura, ICA consente di calcolare in parallelo fino a 1.152 campioni da valutare in una singola analisi e fornisce la determinazione rapida e accurata del PRS con informazioni sull'origine ancestrale. I dati generati con il Predict Module possono essere integrati in altri strumenti presenti sull'ICA, incluse le pipeline di farmacogenomica e di analisi di genotipizzazione.

## Soluzione completa per la predizione del rischio a partire dal genotipo

Il Predict Module fa parte di una serie di strumenti a elevate prestazioni per il PRS che include l'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip per la predizione altamente accurata del rischio e la creazione di report standardizzata. Il BeadChip offre contenuto aggiornato con circa 160.000 nuovi marker PRS, sviluppati in collaborazione con Allelica, aggiunti alla struttura portante (backbone) sull'intero genoma in continua crescita dell'Infinium Global Diversity Array. Il contenuto completo del BeadChip è stato minuziosamente selezionato per fornire un'ampia copertura dei disordini poligenici chiave, inclusi diabete di tipo 1 e di tipo 2, cancro al seno, al colon e alla prostata, patologie coronariche e Alzheimer. LiScan™ System, la tecnologia dell'Infinium LCG Assay e il Predict Module integrato lavorano assieme per creare una soluzione completa e a elevate prestazioni per la predizione del rischio a partire dal genotipo e supportare la ricerca del PRS (Figura 5).

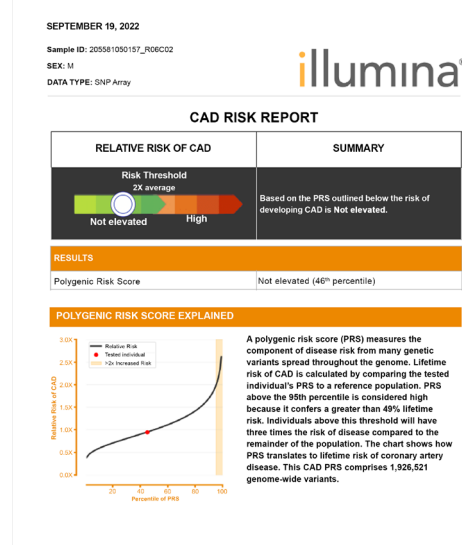
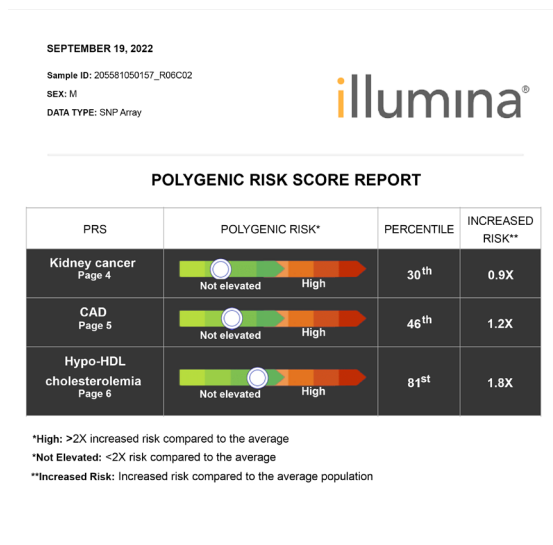


Figura 4: Esempio di report di predizione del rischio: il Predict Module fornisce report personalizzati che contengono il percentile relativo del rischio per ogni campione in un formato di semplice visualizzazione con informazioni sul PRS utilizzato per calcolare il rischio.



Figura 5: Serie di strumenti completa per la ricerca sul PRS: il Predict Module assieme all'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip offre una soluzione ottimizzata per la predizione del rischio a partire dal genotipo per una valutazione accurata del PRS e la creazione di report standardizzata.

## Riepilogo

Il Predict Module utilizza una pipeline efficace di algoritmi di analisi e di bioinformatica sviluppati in partnership con Allelica per semplificare la valutazione e la creazione di report per il PRS. L'interfaccia utente del BaseSpace Sequence Hub è di facile utilizzo e consente di determinare il PRS in modo veloce, accurato e con informazioni sull'origine ancestrale con ampia copertura dei disordini poligenici. Il Predict Module sfrutta i dati contenuti nel catalogo PGS in continua crescita per calcolare i punteggi relativi del rischio. Questo modulo software fa parte di una serie di strumenti completa che include l'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip e offre una soluzione ottimizzata di predizione del rischio a partire dal genotipo per la ricerca nel campo della medicina di precisione.

## Maggiori informazioni

Punteggio di rischio poligenico, [illumina.com/areas-of-interest/complex-disease-genomics/polygenic-risk-scores](https://illumina.com/areas-of-interest/complex-disease-genomics/polygenic-risk-scores).

Illumina Connected Analytics, [illumina.com/products/by-type/informatics-products/connected-analytics](https://illumina.com/products/by-type/informatics-products/connected-analytics).

## Bibliografia

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

## Informazioni per gli ordini

Prodotto	N. di catalogo
Predict Module	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



Numero verde 1.800.809.4566 (U.S.A.) | Tel. +1.858.202.4566  
 techsupport@illumina.com | www.illumina.com

© 2022 Illumina, Inc. Tutti i diritti riservati. Tutti i marchi di fabbrica sono di proprietà di Illumina, Inc. o dei rispettivi proprietari. Per informazioni specifiche sui marchi di fabbrica, consultare la pagina Web [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).  
 M-GL-01376 ITA v1.0