

# Software für polygene Risiko-Scores: Predict-Modul

Eine leistungsstarke  
Datenanalyselösung  
für die PRS-Forschung

- Optimierte Lösung zur Bestimmung des genotypabhängigen Risikos für ein breites Spektrum an Erkrankungen
- Anwenderfreundliche Benutzeroberfläche zur einfachen Datenanalyse und Erstellung von Befunden zum genetischen Risiko-Score
- Abstammungsspezifische genetische Risiko-Scores zur hochpräzisen Prognose von Erkrankungsrisiken



## Einleitung

Polygenen Erkrankungen wie Typ-2-Diabetes, Brust- und Prostatakrebs, Bluthochdruck, koronarer Herzkrankheit und Alzheimer liegt das Zusammenwirken von Hunderten oder gar Tausenden von Suszeptibilitätsgenen zugrunde. Jede dieser häufigen in Zusammenhang mit Erkrankungen stehenden Varianten trägt neben Umweltfaktoren zum Gesamterkrankungsrisiko bei. Das erschwert die Prognose der Suszeptibilität für Erkrankungen. Polygene Risiko-Scores (PRS) oder genetische Risiko-Scores (GRS) basieren auf der Gesamtzahl der genetischen Varianten, die das individuelle Risiko einer spezifischen Erkrankung erhöhen. Mit diesem vielversprechenden Instrument der Präzisionsmedizin lassen sich die individuelle Erkrankungswahrscheinlichkeit bestimmen, Populationen nach Krankheitsrisiko stratifizieren und Personen ermitteln, die von zusätzlicher Überwachung oder präventiven Maßnahmen profitieren würden.<sup>1,2</sup> Jedoch sind PRS noch ein recht neues Gebiet, das vor der Anwendungsreife weitere Forschung erfordert.

PRS-Studien haben gezeigt, dass es möglich ist, Personen zu ermitteln, bei denen ein hohes genetisch bedingtes Risiko für die Entwicklung komplexer polygener Erkrankungen besteht (Abbildung 1).<sup>3-7</sup>

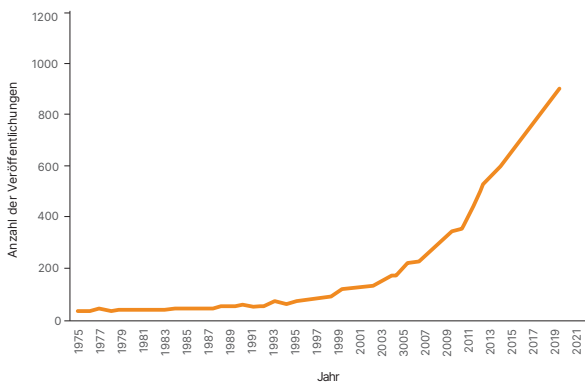


Abbildung 1: Anzahl der Veröffentlichungen zu polygenen Risiko-Scores: Eine Suche mit den Schlagwörtern „PRS“ oder „GRS“ in den Publikationen in der PubMed-Datenbank zeigt einen deutlichen Zuwachs an Studien, bei denen dieses Tool zur Bestimmung des Erkrankungsrisikos zum Einsatz kommt. Stand: Februar 2022.

Forscher können Daten aus genomweiten Assoziationsstudien (GWAS) nutzen, um PRS zu berechnen, die als gewichtete Summe der Allelzahl aller Varianten mit Erkrankungsassoziation definiert sind (Abbildung 2).<sup>8</sup> Aufgrund der großen Anzahl von Varianten, die in großen Populationsstudien identifiziert werden, ist die Erstellung von individuellen PRS jedoch rechenintensiv. Die PRS-Pipeline, von der Eingabe genomischer Daten bis hin zur Qualitätskontrolle und Befunderstellung, erfordert umfassende Kenntnisse im Bereich der Bioinformatik, was einem verbreiteten Einsatz von PRS in der klinischen Forschung entgegensteht.

Das gemeinsam mit Allelica entwickelte Predict-Modul von Illumina vereinfacht den PRS-Prozess, der Berechnung, Analyse und Befunderstellung umfasst, durch eine anwenderfreundliche Benutzeroberfläche zur Generierung individueller PRS in großem Maßstab. Das Predict-Modul steht als Bestandteil eines umfassenden Toolkits zur Verfügung, das zudem den Infinium™ Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip umfasst. Letzterer stellt eine optimierte Lösung zur Prognose des genotypabhängigen Risikos dar, die die Ausweitung der Forschung auf dem noch jungen Gebiet der PRS-Bewertung ermöglicht.

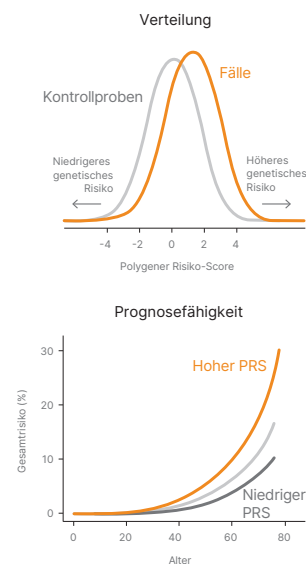


Abbildung 2: Repräsentative Ergebnisse für die Prognose des Erkrankungsrisikos mit PRS: Der obere Bereich zeigt die PRS-Verteilung für Fälle (orangefarbene Linie) und Kontrollproben (graue Linie). Der untere Bereich zeigt die Prognosefähigkeit von PRS für eine einzelne Probe, die ein gegenüber Referenzproben mit einem niedrigen oder durchschnittlichen PRS-Perzentil ein erhöhtes Risiko für eine frühzeitige Erkrankung aufweist.

## Genauere Risikobewertung mit phänotypbasierter PRS-Verteilung

Das Predict-Modul berechnet individuelle PRS und liefert einen Befund zum Erkrankungsrisiko im Verhältnis zu einer Referenzpopulation. Der Vergleich eines numerischen PRS mit einer anhand von Genotypdaten ermittelten Populationsverteilung liefert einen Perzentil-Risiko-Score relativ zur Referenzkohorte, jedoch keine Informationen über das absolute Erkrankungsrisiko. Um den PRS mit dem tatsächlichen Krankheitsrisiko in Beziehung zu setzen, enthalten alle im Predict-Modul verwendeten Populationsverteilungen nicht nur Daten zum Genotyp, sondern auch zum Phänotyp. Das Predict-Modul enthält Scores aus dem [Polygenic Score \(PGS\) Catalog](#) sowie von Allelica entwickelte Risiko-Scores. Das Predict-Modul umfasst eine für jeden PRS spezifische Qualitätskontrolle, die Prognosen von außerordentlicher Qualität ermöglicht.

## Berechnung des Erkrankungsrisikos auf Basis der Abstammung

Die Erblichkeit polygener Erkrankungen ist bei unterschiedlichen Populationen oft verschieden. Daher ist die abstammungsspezifische Risikobewertung entscheidend für die genaue PRS-Bestimmung.<sup>9</sup> Das ausschließlich für die Forschung vorgesehene Predict-Modul umfasst sämtliche im PGS Catalog enthaltenen PRS mit ca. 1.000 vorkompilierten Risikomodellen, die auf über 20 Abstammungsmerkmalen und über 80 PRS für jeden der fünf Kontinente basieren. Somit deckt die Software potenziell pathogene Varianten, die zum Krankheitsrisiko beitragen, umfassend ab.

Das Tool ermöglicht eine feine Kalibrierung der Abstammung auf Basis eines harmonisierten Raums mit ca. 800.000 individuellen Proben für die Hauptkomponentenanalyse (HKA). Zudem können Forscher mit dem Predict-Modul automatische PRS-Befunde mit abstammungsspezifischen Risikoprognosen erstellen. Des Weiteren deckt das Modul ein breites Spektrum an Störungen ab, u. a. in den Bereichen Onkologie, Herz-Kreislauf-Erkrankungen und allgemeines Wohlbefinden.

## Vereinfachter Datenanalyseworkflow

Das Predict-Modul kann Daten verarbeiten, die mit Microarrays oder mit Genomsequenziersystemen generiert wurden. Datendateien von Illumina-Sequenziersystemen werden an die Illumina-Cloudplattform gesendet. Anschließend werden die Scores von Interesse ausgewählt, entweder aus den über 2.600 im Predict-Modul enthaltenen Scores oder aus anwendungsspezifischen vom Labor generierten PRS, und die individuellen PRS für die einzelnen Proben berechnet ([Abbildung 3](#)). Im Anschluss an die Berechnung der Scores erstellt das Predict-Modul personalisierte Befunde mit umfassenden Risikobewertungen, die zur genaueren Prognose des Risikos auch andere Variablen wie Geschlecht und Ethnizität berücksichtigen.

Das Predict-Modul bietet eine Imputationsoption für Sequenzierungsdaten mit geringer Coverage, mit der sich die Anzahl der im Datensatz vorhandenen genetischen Varianten erhöhen lässt. Die Imputations-Engine ermöglicht die Imputation von ca. 10.000 Proben in weniger als drei Stunden mit einer Präzision von über 99 % und hoher Übereinstimmung mit dem Infinium Global Screening Array.



Abbildung 3: Vereinfachter Workflow für die PRS-Bestimmung mit dem Predict-Modul: Mit dem Predict-Modul steht eine umfassende Lösung zur Risikoprognose zur Verfügung, die eine schnelle und genaue Bewertung des Erkrankungsrisikos ermöglicht, ohne dass hierfür umfangreiche bioinformatische Ressourcen erforderlich sind.

## Anwenderfreundliche Benutzeroberfläche

Die Integration der Genotyp-Rohdaten von Illumina-Sequenziersystemen in die intuitive BaseSpace™ Sequence Hub-Oberfläche ermöglicht eine einfache Point-and-Click-Analyse und PRS-Berechnung. Mit dem Predict-Modul lassen sich bis zu 24 PRS in 15 Minuten berechnen. Für jede Probe wird ein standardisierter PRS-Befund erstellt. Dieser umfassende Befund enthält abstammungsspezifische polygene Scores und das relative Lebenszeitrisiko für eine Erkrankung. Dies liefert einen wichtigen Kontext für die PRS-Interpretation (Abbildung 4).

## Sichere, cloudbasierte Datenanalyseplattform

Das Predict-Modul ist in die hinsichtlich Qualität und Leistung optimierte Cloudinfrastruktur von Illumina Connected Analytics (ICA) eingebettet. ICA bietet nicht nur eine skalierbare und sichere Plattform, sondern ermöglicht auch die parallele Verarbeitung von bis zu 1.152 Proben in einer einzigen Analyse zur schnellen, genauen und abstammungsspezifischen PRS-Bestimmung. Die mit dem Predict-Modul generierten Daten lassen sich in andere Tools integrieren, die ebenfalls auf ICA gehostet werden, darunter Analysepipelines für Pharmakogenomik und Genotypisierung.

## Umfassende Lösung für die Prognose des genotypabhängigen Risikos

Das Predict-Modul steht als Bestandteil eines umfassenden PRS-Toolkits zur Verfügung, das den Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip zur hochpräzisen Risikoprognose und standardisierten Befunderstellung umfasst. Der BeadChip umfasst aktualisierte Inhalte mit ca. 160.000 neuen PRS-Markern, die gemeinsam mit Allelica entwickelt wurden und die umfangreiche genomweite Basis von Infinium Global Diversity Array ergänzen. Der umfassende Inhalt des BeadChips wurde sorgfältig so ausgewählt, dass er ein möglichst breites Spektrum wichtiger polygener Erkrankungen abdeckt, darunter Typ-1- und Typ-2-Diabetes, Brust-, Darm- und Prostatakrebs, koronare Herzkrankheit und Alzheimer. Das iScan™ System, die Infinium LCG Assay-Technologie und das integrierte Predict-Modul bilden zusammen eine umfassende, leistungsstarke Lösung für die Prognose des genotypabhängigen Risikos in der PRS-Forschung (Abbildung 5).

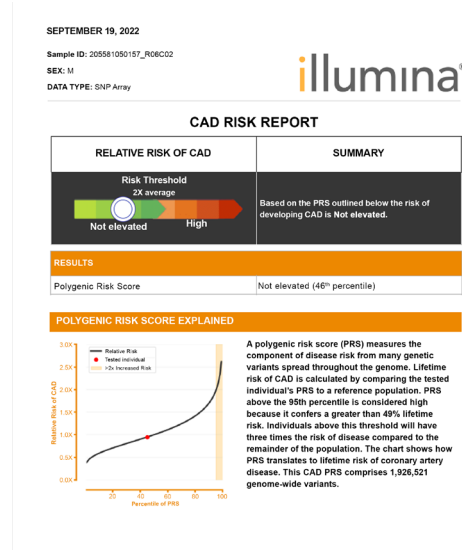
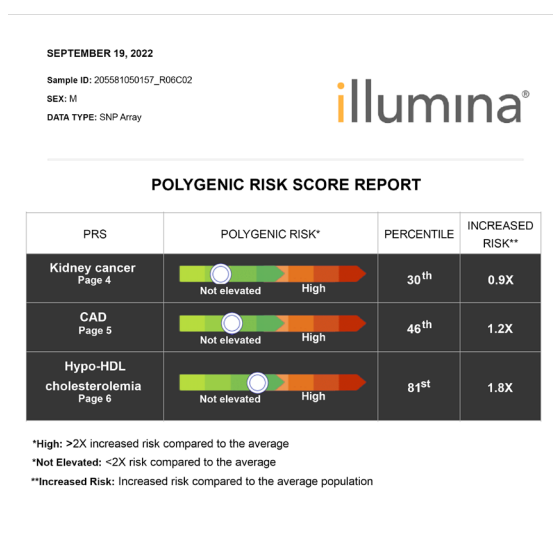


Abbildung 4: Beispielbefund zur Risikoprognose: Das Predict-Modul liefert individuelle Befunde, die das relative Risikoperzentil der einzelnen Proben sowie zusätzliche Informationen zum PRS in einem leicht verständlichen grafischen Format enthalten.



Abbildung 5: Umfassendes Toolkit für die PRS-Forschung: Das Predict-Modul bietet in Kombination mit dem Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip eine optimierte Lösung für die Prognose des genotypabhängigen Risikos für eine hochgradig genaue PRS-Bewertung sowie die standardisierte Befunderstellung.

## Zusammenfassung

Das Predict-Modul vereinfacht mit leistungsstarken Analysealgorithmen und Bioinformatik-Pipelines, die gemeinsam mit Allelica entwickelt wurden, die PRS-Bewertung sowie die entsprechende Befunderstellung. Die anwenderfreundliche BaseSpace Sequence Hub-Benutzeroberfläche ermöglicht eine schnelle, genaue und abstammungsspezifische Bestimmung von PRS für ein breites Spektrum polygener Erkrankungen. Das Predict-Modul berechnet relative Risiko-Scores anhand von Daten aus dem umfangreichen PGS Catalog. Das Softwaremodul steht als Bestandteil eines umfassenden Toolkits zur Verfügung, zu dem auch der Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 BeadChip gehört, der eine optimierte Prognose des genotypabhängigen Risikos für die Forschung im Bereich der Präzisionsmedizin ermöglicht.

## Weitere Informationen

Polygene Risiko-Scores: [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)

## Quellen

1. Dudbridge F. [Power and predictive accuracy of polygenic risk scores](https://doi.org/10.1371/journal.pgen.1003348). *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](https://doi.org/10.1186/s13073-020-00742-5). *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](https://doi.org/10.1038/s41588-018-0183-z). *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](https://doi.org/10.1016/j.ajhg.2018.11.002). *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](https://doi.org/10.1038/s41467-021-21287-0). *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](https://doi.org/10.1001/jama.2019.22241). *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](https://doi.org/10.1186/s13073-022-01074-2). *Genome Med.* 2022;14(1):70. doi:10.1186/s13073-022-01074-2

8. Collier 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

## Bestellinformationen

Produkt	Katalog-Nr.
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



+1.800.809.4566 (USA, gebührenfrei) | +1.858.202.4566 (Tel. außerhalb der USA)  
 techsupport@illumina.com | www.illumina.com

© 2022 Illumina, Inc. Alle Rechte vorbehalten. Alle Marken sind Eigentum von Illumina, Inc. bzw. der jeweiligen Eigentümer. Spezifische Informationen zu Marken finden Sie unter [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).  
 M-GL-01376 DEU v1.0