

# Polygenic Risk Scores

Clinical Value of Polygenic Risk  
November 2019

## Highlights

A genetic component of disease has been identified across many conditions, cancers and syndromes. As we continue to uncover the genomic mechanisms of human disease, our multifaceted perspective continues to expand and evolve.

Herein is a brief introduction to polygenic risk and its potential value to the future of preventative care.



A genetic component of disease has been identified across many conditions, cancers and syndromes. As we continue to uncover the genomic mechanisms of human disease our multifaceted perspective continues to expand and evolve.

Some of these mechanisms are as simple as a monogenic disease like Tay-Sachs, which follows classical Mendelian genetics and has an autosomal recessive pattern. Other more complex diseases like cancers, diabetes, and heart disease are clearly multifactorial and continue to unravel complex interactions at every level of the genomic cascade. This ranges from epigenetics, polygenic risk, transcriptional genetics to proteomics, and is often further influenced by environmental factors. Researchers are not yet capable of connecting all the genomic pieces of the puzzle, but are inching their way forward. Polygenic risk, although just one piece of the bigger picture, can provide clinical utility when considering population-level risk stratification as it relates to disease susceptibility because it allows an individual's risk of disease to be identified. Polygenic risk tests have the potential to inform many aspects of medical practice including: disease risk stratification, stratified screening strategies, early diagnosis, therapeutic management, chemoprevention, and behavior modification for risk reduction. GeneType risk assessments are a series of tests that capitalize on this current understanding of genetic risk as it relates to complex disease, utilizing the most advanced methodologies to calculate individual risk scores.

### What is a Polygenic Risk Score?

The individual variants that make up the Polygenic Risk Score (PRS) were originally identified from large genome-wide association study (GWAS) datasets. Each disease will have several variants significantly associated with the occurrence of said disease. Each variant will

**Genome-Wide Association Studies [GWAS]** - Large population level studies which enable scientists to identify genes and genetic markers involved in human disease. This method searches the genome for SNPs that occur more frequently in people with a particular disease than in people without the disease. Each study can look at hundreds or many thousands of SNPs at the same time. Researchers use data from this type of study to pinpoint genetic variations that may contribute to a person's risk of developing a certain disease.

be assigned a strength-of-association to the disease in question. Polygenic risk is determined by a mathematical summation of risk from each of these individual variants across a genome. Although the PRS can be estimated as relative risk compared to a control population, there is greater clinical utility in an absolute risk score derived by incorporating competing disease and mortality across a unit of time. Ultimately, the PRS can then be used as an independent risk factor of disease in the same way that a clinical factor such as age, gender, or family history would be used in a risk assessment model.

**Genomic** - pertaining to function of genetics from structure to relationship between genetic events.

**Genetic** - pertaining to a gene.

**Variant** - Single Nucleotide Polymorphism (SNP), an alteration in DNA that may be a common or rare event.

**Monogenic Disease** - a disease or condition associated with a single genetic variant, this pathogenic variant is often referred to as a mutation.

**Polygenic Disease** - a disease or condition associated with multiple (tens to thousands) genetic variants. Although complete mechanisms of action are yet unknown, the variants in isolation are not disease-causing. However, together the variants may support the disease or condition in the context of external environmental factors.

**Polygenic Risk** - combination of risk variants associated with a certain disease or condition.

**Polygenic Risk Score [PRS]** - a number associated with a person's disease risk based on the aggregated effects of individual risk variants through a multiplicative algorithm.

The number of DNA variants in a PRS can vary from tens to many thousands. The number of incorporated variants varies between diseases due to the limitations in cohort sizes of the initial GWAS datasets. Furthermore, there may be a theoretical limit to the contribution of PRS in a multifactorial disease model, as such, there will be a point when the addition of variants does not provide additional discrimination. However, in the context of complex diseases such as breast cancer, each sub-classification, whether based on immunohistochemical status, molecular phenotype, recurrence score, etc. can be thought of as an independent disease model. Thus, the number of contributory variants will expand in parallel to the availability of GWAS datasets.

## Learning Datasets

One of the most limiting factors in the development of clinically useful PRS is the lack of study cohorts across populations. The majority of GWAS have been conducted in Caucasian individuals of European descent. The paucity of cohort data from other ethnicities/races is of significance, particularly in the context of polygenic risk, which relies

on disease-associated variants that are known to differ between ethnicities.

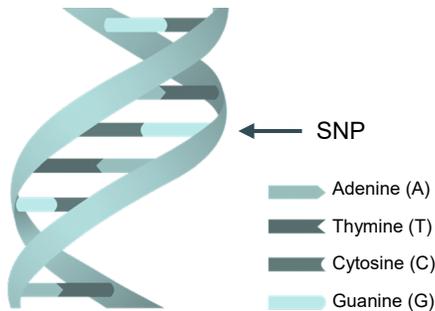
For example, GeneType for Breast Cancer genotypes a handful of variants that are specific to African Americans and independent from both Caucasian or Hispanic variants for breast cancer. Furthermore, the strength-of-association of each risk allele varies in a Caucasian female compared to an African American female. The PRS for both African American and Hispanic women were validated in ethnic-specific cohorts (Allman et al., 2015). As the research community continues to expand multi-ethnic cohorts, PRS will continue to improve in its capacity to stratify at-risk individuals (Spaeth et al., 2018).

## Clinical Utility: A Present-Day Case Study

The current retrospective data supporting clinical utility of PRS is strong across some diseases. Although not yet included in practice guidelines for any major medical organization, for breast cancer risk assessment, there is a significant amount of data supporting the risk variants used to calculate PRS in Caucasian women (Dite et al., 2016, Mavaddat et al., 2015, Rudolph et al., 2018, Mavaddat et al., 2019, Kuchenbaecker et al., 2014, Kuchenbaecker et al., 2017).

### Single Nucleotide Polymorphisms [SNP] -

Frequently called SNPs (pronounced "snips"), are the most common type of genetic variation among people. Each SNP represents a difference in a single DNA building block, called a nucleotide. For example, a SNP may replace the nucleotide cytosine (C) with the nucleotide thymine (T) in a certain stretch of DNA.



SNPs occur normally throughout a person's DNA. Most commonly, these variations are found in the DNA between genes. They can act as biological markers, helping scientists locate genes that are associated with disease. When SNPs occur within a gene or in a regulatory region near a gene, they may play a more direct role in disease by affecting the gene's function.

Most SNPs have no effect on health or development. Some of these genetic differences, however, have proven to be very important in the study of human health. Researchers have found SNPs that may help predict an individual's response to certain drugs, susceptibility to environmental factors such as toxins, and risk of developing particular diseases.

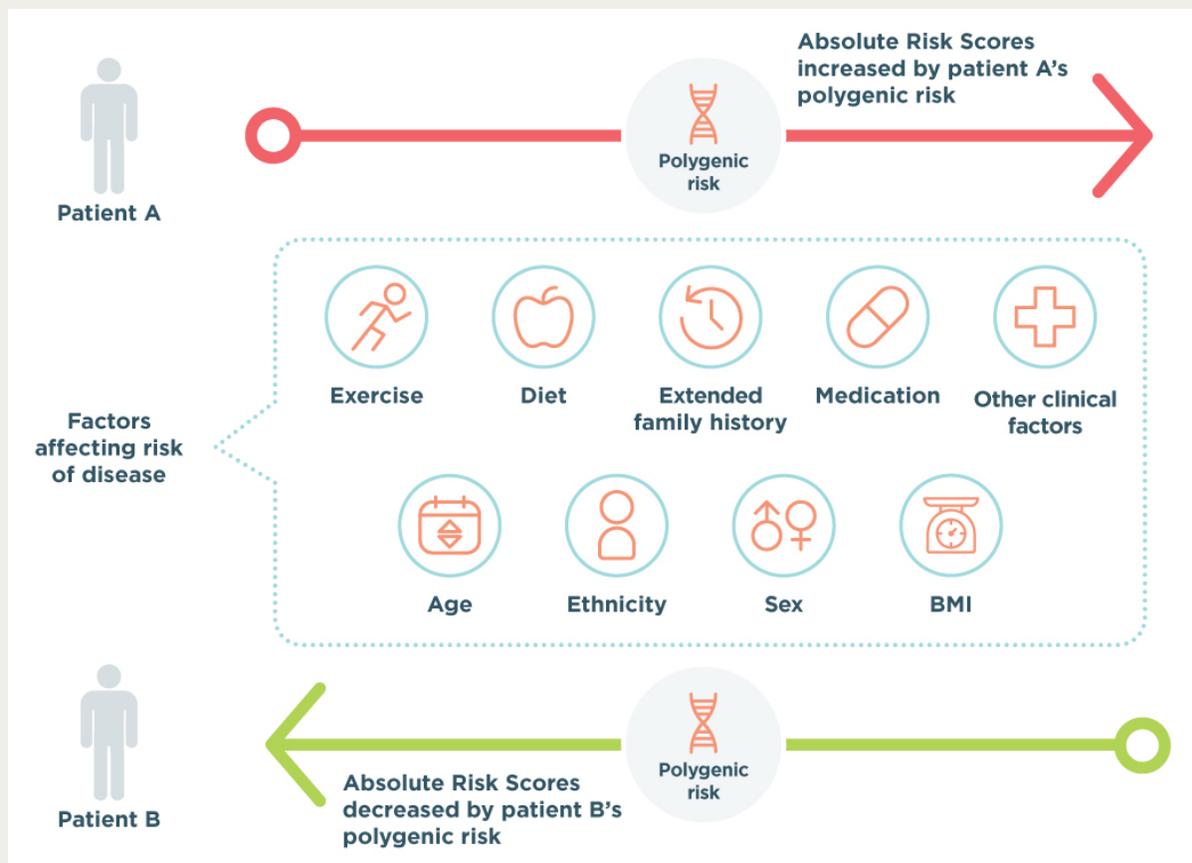
Prospective cohort studies consistently provide evidence that polygenic risk scores for breast cancer provide more accurate risk prediction than family history alone and that polygenic risk scores can aid in risk-stratified screening and prevention.

Li et al. (2017) analyzed 4,365 women from the Australian Breast Cancer Family Registry and the KConFab familial breast cancer cohorts, with 2,599 prospectively analyzed over 7.4 years. During this follow-up period, 205 breast cancers occurred. The hazard ratios for continuous PRS per standard deviation was 1.38 (95% CI: 1.22–1.56), and the upper vs. lower quintile hazard ratio was 3.18 (95% CI: 1.84–5.23). Based on a threshold of 20% lifetime risk, the authors concluded that the clinical management of 23% of women could be altered by incorporating PRS into their risk assessment. This conclusion provided support for the use of breast cancer-associated SNPs in risk assessment within family breast cancer cohorts by showing more accurate risk prediction than family history alone. More importantly, this can influence recommendations for cancer screening and prevention modalities for high-risk women.

In a large prospective study, Naslund-Koch et al (2017); approximately 35,000 women from the Danish general population were followed in Danish health registries for up to 21 years after blood sampling. Patients were genotyped for 72 breast cancer associated SNPs. These alleles were not associated with incidence of other cancers, but each breast cancer risk allele was associated with a 4% increase of breast cancer incidence and with 5% increased breast cancer mortality. After including a breast cancer allele sum in the risk assessment, 25% of women currently being offered screening mammography had an absolute 5-year risk below the cut-off for a 50-year-old woman. The study provides strong evidence that a polygenic risk score can aid in risk-stratified breast cancer screening.

Van Veen et al. (2018) analyzed 9,363 women in a prospective cohort between 2009-2017 in the UK. Of these, 466 were found to have breast cancer (271 prevalent; 195 incident). A combined risk assessment stratified the cohort such that those at 5% or greater 10-year risk were 5-fold more likely to develop a high stage cancer than those at less than 2% 10-year risk. Furthermore, compared to only 18% of the control population at greater than 5% 10-year risk, 30% of the cases were stratified to this moderate/high risk category consisting of 35% interval-detected cancers and 42% stage 2+ cancers. In contrast, a third of the cohort had 10-year risk scores less than 2% and accounted for only 18% of breast cancer cases (17% interval-detected and 15% stage 2+). The conclusion being that a combined risk is likely to aid risk-stratified screening and prevention strategies.

The challenge in breast cancer prevention lies in the widespread adoption at the clinician level through increased awareness of the benefits of risk assessment. Lack of awareness and implementation of basic risk assessment in the primary care setting is a larger issue within our US healthcare system due to time and cost restraints. This may be, in part, alleviated through the incorporation of PRS into routine clinical care.



**Figure 1:** Polygenic Risk Scores can modify the disease risk of two seemingly identical patients. Importantly, two patients may be the same age and ethnicity and have the same clinical risk factors including family history, medication use, BMI and their diet and exercise may be at similar levels. However, a patient's polygenic risk is not visible. Patient A and Patient B may have drastically different polygenic risk scores, thus modifying their absolute risk scores. Integration of polygenic risk into a patient's absolute risk score will push Patient A into an *increased-risk* category whereas Patient B will remain in an *average risk* category.

### Clinical Utility: The Future

The clinical value of PRS will be further exploited as we continue into the age of big-data. New polygenic risk variants are being published weekly with an increased capacity to handle computational and statistically challenging, heterogeneous, multivariate datasets. Our clinical challenge continues to be the introduction of this novel risk factor into mainstream medical practice. There is substantial retrospective and emerging prospective clinical data to support clinical decision-making based on PRS in several different diseases (such as breast cancer). But, prospective clinical data is financially more burdensome when assessing disease risk and prevention in a general population because you must assess much larger sample sizes over an extended period. For preventative medicine to become a reality, not only do we need to stratify an at-risk population, we need to apply effective screening and preventative metrics. The future of PRS will lie in its ability to identify the at-risk population and likewise suggest the most efficacious screening regimen and preventative therapy option.

## Advancing with Caution

While PRS may be intertwined in the future of preventative medicine, it will not be a predictor of disease risk in isolation of other clinical and environmental factors. As such, clinicians will play a significant role in the expansion and adoption of PRS into routine clinical care. Clinicians must be able to trust the scientific and clinical evidence supporting the use of PRS tests. In a fast-paced world of biotech start-ups, sometimes it is difficult to distinguish between the “next best thing” and the “real deal.” With numerous consumer-facing companies that dabble in hobby-level genetic and ancestry testing, it is important to differentiate the tests built on strong clinical evidence. We realize the field of genetics is constantly growing and changing and we strive to provide comprehensive reviews of the clinical data that support the utility of PRS across the GeneType test portfolio. We support clinician training and education efforts in order to advance the clinical value of PRS in routine clinical practice.

## References

- Allman R, Dite GS, Hopper JL, Gordon O, Starlard-Davenport A, Chlebowski R, *et al.* SNPs and breast cancer risk prediction for African American and Hispanic women. *Breast Cancer Res Treat.* 2015;154(3):583-9.
- Dite GS, MacInnis RJ, Bickerstaffe A, Dowty JG, Allman R, Apicella C, *et al.* Breast Cancer Risk Prediction Using Clinical Models and 77 Independent Risk-Associated SNPs for Women Aged Under 50 Years: Australian Breast Cancer Family Registry. *Cancer Epidemiol Biomarkers Prev.* 2016;25(2):359-65.
- Kuchenbaecker KB, Neuhausen SL, Robson M, Barrowdale D, *et al.* Associations of common breast cancer susceptibility alleles with risk of breast cancer subtypes in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res.* 2014;16(6):3416. doi: 10.1186/s13058-014-0492-9.
- Kuchenbaecker KB, McGuffog L, Barrowdale D, Lee A, *et al.* Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. *J Natl Cancer Inst.* 2017;109(7). doi: 10.1093/jnci/djw302.
- Li, H, Feng B, Miron A, Chen X, *et al.* Breast cancer risk prediction using a polygenic risk score in the familial setting: a prospective study from the Breast Cancer Family Registry and kConFab. *Genet Med.* 2017; 19(1): 30–35. doi:10.1038/gim.2016.43
- Mavaddat N, Pharoah PD, Michailidou K, *et al.* Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* 2015;107(5):djv036.
- Mavaddat N, Michailidou K, Dennis J, Lush M, *et al.* Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *American Journal of Human Genetics.* 2019. 104, 21–34.
- Naslund-Koch C, Nordestgaard BG & Bojesen SE. Common breast cancer risk alleles and risk assessment: a study on 35 441 individuals from the Danish general population. *Annals of Oncology* 2017; 28: 175–181,
- Rudolph A, Song M, Brook MN, Milne RL, *et al.* Joint associations of a polygenic risk score and environmental risk factors for breast cancer in the Breast Cancer Association Consortium. *International Journal of Epidemiology*, 2018, 526–536.
- Spaeth E, Starlard-Davenport A, Allman R. Bridging the Data Gap in Breast Cancer Risk Assessment to Enable Widespread Clinical Implementation across the Multiethnic Landscape of the US. *J Cancer Treat & Diagnosis.* (2018);2(4):1-6.
- van Veen EM, Brentnall AR, Byers H, Harkness EF, *et al.* Use of Single-Nucleotide Polymorphisms and Mammographic Density Plus Classic Risk Factors for Breast Cancer Risk Prediction. *JAMA Oncol.* 2018; 4(4): 476-482.