

---

## Genome-wide rare variant analysis for thousands of phenotypes in >70,000 exomes

Lavania Sharma<sup>2</sup>, Peter Hulick<sup>1</sup>, Brian Helfand<sup>1</sup>, Elissa Levin<sup>2</sup>, Alexandre Bolze<sup>2</sup>, Keith Dunaway<sup>2</sup>, James Lu<sup>2</sup>, Jianfeng Xu<sup>1</sup>

<sup>1</sup> NorthShore University HealthSystem, Evanston, Illinois

<sup>2</sup> Helix, San Carlos, California

Pathogenic variants explain only a small portion of prostate cancer (PCa), however the utility of disease risk measured by polygenic risk score, has been consistently demonstrated for many common diseases, including PCa. Polygenic tests are currently being developed, offered and used in the clinical space. In collaboration with experts in the field, we have developed a comprehensive clinical polygenic PCa test that looks at SNPs stratified by risk among different ethnicities, namely, East Asian, African American and Non-Hispanic White. The individuals are categorized as average, moderately-high or high risk for developing PCa. Medical management recommendations like earlier and more frequent screenings for PCa are included in the clinical report. These recommendations are based on the individuals' risk, and generally follow the risk-based PCa screening guidelines of the U.S. Preventive Service Task Force. Here we propose best practices for developing a clinical polygenic risk score test in multiple ethnicities as well as provide recommendations on counseling patients that have undergone such tests.

For test development, we used the odds ratio (OR)-weighted and population-standardized genetic risk score (GRS) method. In this method, each SNP is first standardized against the general population and then multiplied for all SNPs. Thus its expected mean in the general population will always be 1.0, regardless of the number of SNPs used in calculation, and its values can be simply interpreted as relative risk to that of the general population. These two important features of population-standardized GRS makes the interpretation and implementation of an individual's risk comprehensible. We also applied two benchmarks, baseline and calibration, to ensure the reliability of GRS values.

This work aims to provide procedures to build a simple and interpretable risk score in multiple ethnicities and provide guidance on medical management. We will also discuss challenges and future work to address the gaps in the process of clinical test development and return of results.