Validation of a clinical Blended Genome Exome (cBGE) assay for Prostate Cancer Polygenic and Monogenic Risk in Veterans

Katie Larkin[1], Michael Gatzen[1], Katherine Lafferty[1], Ning Li[1], Candace Patterson[1], Christopher Kachulis[1], Matthew DeFelice[1], Matthew Coole[1], Marissa Gildea[1], Jonna Grimsby[1], Sean Hofherr[1], Morgan E. Danoski[3], Charles A. Brunette[3], Roshan Karunamuni[2], Anna M. Dornisch[2], Tyler Seibert[2], Jason L. Vassy[3], Niall Lennon[1]

[1]Broad Clinical Labs, Burlington, MA, USA; [2]University of California San Diego, La Jolla, CA, USA; [3]VA Boston Healthcare System, Boston, MA, USA

Background

ProGRESS (Prostate Cancer, Genetic Risk, and Equitable Screening Study) is a precision medicine clinical trial within the VA healthcare system.

Traditional PSA-based prostate cancer screening lacks specificity, leading to unnecessary biopsies and overtreatment.

The objective of ProGRESS is to implement a genomics-informed prostate cancer screening strategy to identify individuals who are at high risk for aggressive disease.

U.S. Department of Veterans Affairs
Veterans Health Administration
Boston Healthcare System

Jason Vassy, MD, MPH, MS

The key components of ProGRESS are:

- **Polygenic Risk Score (PRS)** to estimate cumulative genetic risk (utilizing GLIMPSE2 and the **P-CARE model**)
- **Monogenic Testing** of 12 prostate cancer-related genes to identify rare, high-risk pathogenic variants
- Personalized, risk-based screening recommendations

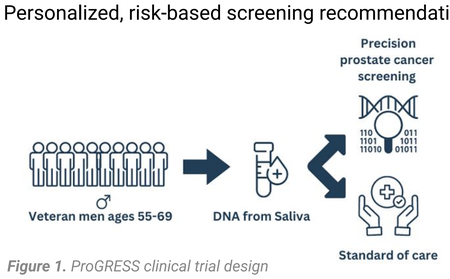

Figure 1. ProGRESS clinical trial design

P-CARE model for Prostate Cancer

Prostate CAncer Integrated Risk Evaluation (P-CARE) is a novel clinical risk model combining:

- A **PRS** - 601-variant polygenic score (PHS601).
- **Family History** of prostate cancer (binary yes or no)
- Genetic principal components (**PCA**) for ancestry-adjusted risk prediction.

P-CARE was developed using data from 585,418 participants in the **Million Veteran Program**.

It was validated using 18,457 samples from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) Consortium, including the ProtecT cohort.[1]

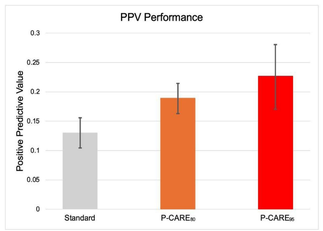

***Figure 2.** In the ProtecT study, a PSA level of ≥3 ng/mL was more likely to indicate clinically significant prostate cancer in men with a higher genetic risk (P-CARE top 20% or top 5%) compared to all men in the study.[1]*

cBGE: A Good Fit for ProGRESS

Clinical Blended Genome Exome (cBGE) was selected as the ideal approach because it:

- Combines genome (2-3x) and exome (>90x) sequencing in a single assay
- Provides high-depth monogenic variant detection and low-pass genome coverage for imputation and PRS
- Is cost-effective, scalable, and minimizes ancestry bias

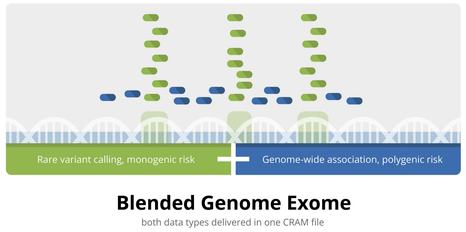

Figure 3. cBGE data types and uses

cBGE was validated to ensure it met the performance criteria for ProGRESS in the following areas:

- Technical performance
- P-CARE risk estimation
- Monogenic variant detection

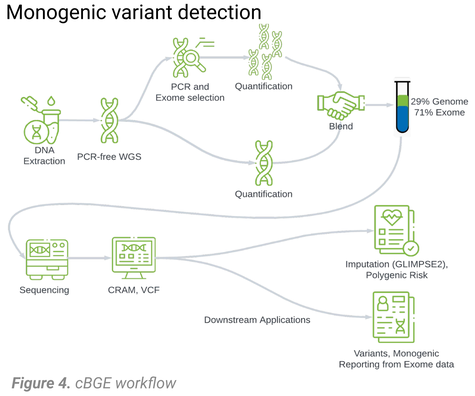

Figure 4. cBGE workflow

Technical Validation

Performance was evaluated using NIST reference data, sequenced on **NovaSeq X with DRAGEN v4.2.7**.

SNPs and Indels (over Exome territory)

HG001 showed high precision (*Figure 5*); HG002 had 99.87-99.89% SNP and 97.35-98.12% indel precision.

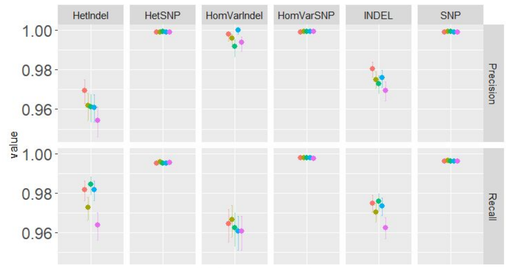

Figure 5. Precision and recall of 5 HG001 replicates

CNVs (over Exome territory)

Deletions ≥3 exons achieved a PPV of 76% and recall of 83%. Duplications achieved a PPV of 87% and recall of 63%.

P-CARE Validation using cBGE

P-CARE validation was conducted by **Michael Gatzen** and **Christopher Kachulis** using cBGE sequencing and GLIMPSE2 imputation, with individuals classified as **low**, **average**, or **high risk** based on P-CARE thresholds.

Cohorts used for analyses:

- 60 samples of diverse ancestry (cBGE and WGS matched)
- 74,331 All of Us (AoU) Research Program v7 samples (with WGS and EHR data, all male)
- BGE replicates (NA12878: 27, NA24385: 7)

Results

PRS/PCA scores calculated from WGS and from BGE data were >99% correlated.

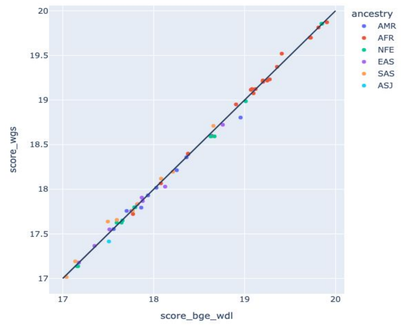

*Figure 7. Scatter plot of **PRS** scores calculated from matched BGE and WGS data. PCA scores were similarly correlated.*

P-CARE clinical validity was established using the AoU cohort. Polygenic scores, genetic principal components, and family history were used to calculate P-CARE values. Cases (n=4,473) and controls (n=69,858) were classified based on prostate cancer diagnoses in electronic health records. Logistic regression, adjusted for age, was used to calculate odds ratios for prostate cancer risk in low and high P-CARE categories relative to the average.

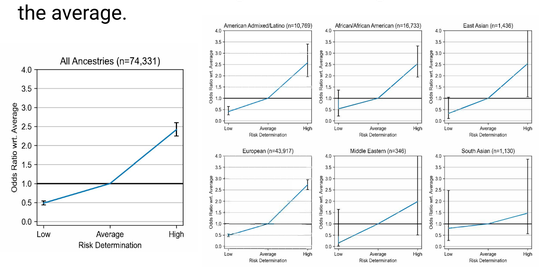

Figure 8. Odds ratios of prostate cancer association for individuals with low or high P-CARE score, as compared to individuals with average P-CARE score, for all ancestries combined, then stratified by ancestries.

The P-CARE score was found to be sufficiently associated with prostate cancer.

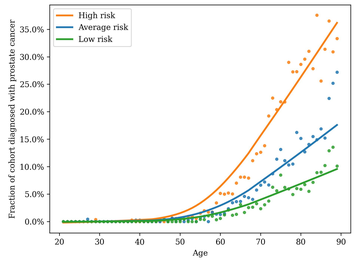

Figure 9. Fraction of cohort diagnosed with prostate cancer by age, separated by high, average, and low P-CARE scores. (LOWESS smoothing intended as a visual guide).

Monogenic Panel Validation using cBGE

12 genes in the hereditary prostate cancer panel were assessed for coverage and callability in cBGE.

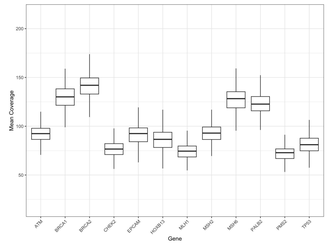

***Figure 10.** Mean coverage of hereditary prostate cancer panel genes in cBGE technical validation cohort*

Each base in each gene interval from the exome (as defined by MANE) was checked. A base was considered undercovered when in at least 80% of samples the base did not achieve ≥20 for depth, base quality, and mapping quality.

>99% of bases (44516 bases of 44689) were sufficiently covered in cBGE data.

Conclusion

cBGE successfully met the validation criteria for ProGRESS, making it a strong fit for precision prostate cancer screening and other screening applications.

By reducing biases related to genomic ancestry and improving access to clinically relevant genetic information, **cBGE broadens the potential for providing personalized care to diverse populations.**

References

1. Vassy, Jason L., et al. "From a Genomic Risk Model to Clinical Trial Implementation in a Learning Health System: The ProGRESS Study." medRxiv, 4 Nov. 2024, p. 2024.11.03.24316516, https://doi.org/10.1101/2024.11.03.24316516.
2. DeFelice, Matthew, et al. "Blended Genome Exome (BGE) as a Cost Efficient Alternative to Deep Whole Genomes or Arrays." *bioRxiv : The Preprint Server for Biology*, Apr. 2024, https://doi.org/10.1101/2024.04.03.587209.
3. Boltz, Toni A., et al. "A Blended Genome and Exome Sequencing Method Captures Genetic Variation in an Unbiased, High-Quality, and Cost-Effective Manner." *bioRxiv*, 9 Sept. 2024, p. 2024.09.06.611689, https://doi.org/10.1101/2024.09.06.611689.
4. https://www.research.va.gov/for_veterans/progress.cfm
5. https://www.research.va.gov/mvp/

Contact

Katie Larkin
Director Clinical Product Development and Strategy
klarkin@broadinstitute.org

Broad Clinical Labs
genomics@broadinstitute.org

Website
broadclinicallabs.org