Genetics in Medicine (2022) 24, 549–551

ELSEVIER

Genetics in Medicine

www.journals.elsevier.com/genetics-in-medicine

INVITED COMMENTARY

# Diversity in cancer genomics research is a matter of equity and scientific discovery

Tuya Pal[1,*]

[1]Division of Genetic Medicine, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN

ARTICLE INFO

*Article history:*
Received 8 September 2021
Received in revised form
17 November 2021
Accepted 17 November 2021
Available online 20 November 2021

Promoting diversity to advance genomics research is an ethical issue to ensure that populations beyond those of European descent benefit equally from the advances as highlighted in the study by Wang et al.[1] They show that polygenic risk scores (PRSs) constructed from single-nucleotide variations (SNVs, formerly singe nucleotide polymorphisms [SNPs]) increase the discriminatory ability to predict cancer risks in individuals of European but not of African ancestry. Their findings underscore the potential for PRS to widen existing health disparities because of the paucity of data in populations of non-European descent, thereby reducing opportunities for precision prevention. Beyond being a matter of health equity, diversity in genomics research has tremendous potential to improve human health across all populations by giving us more power to make novel scientific discoveries.

Thousands of genome-wide association studies (GWAS) have been carried out since the completion of the human genome project in 2003 to identify hundreds of common SNVs associated with cancer risk. Although individual SNVs lack clinical utility with a small effect size, combining multiple SNVs to generate PRS has shown great promise in stratifying risks. However, almost 80% of GWAS have been conducted in European descent individuals, who only make up 16% of the world's population.[2-4] In contrast, GWAS have only included 10% Asian individuals, 2% Black individuals, and 1% Hispanic individuals who make up almost 60%, 17%, and 8.5% of the world's population, respectively. Moreover, even though individuals of African ancestry make up approximately 2% of GWAS participants, their data have produced about 7% of the genetic associations.[5] In contrast, European descent individuals make up 78% of GWAS participants but have only contributed to 54% of discoveries. These results show that individuals of African ancestry, who retain the most genetic diversity because humans originated from Africa, disproportionately contribute to the discovery of genetic associations and yet remain a tremendously understudied population. In fact, a recent study of 929 diverse human genomes identified previously unknown common and rare DNA variations exclusive to major geographical regions, representing an important step in expanding genomics to populations underrepresented in GWAS.[6] Consequently, the lack of diversity in GWAS limits our understanding of the genetics underlying complex traits, recognizing that all humans regardless of race have the same biology. Rather, the genetic variants each person carries can determine how and when that biology manifests, which is why we must leverage genetics across all populations to benefit human health as a whole.

*Correspondence and requests for materials should be addressed to Tuya Pal, Division of Genetic Medicine, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, 1500 21st Ave S, Suite 2810, Nashville, TN 37212. *E-mail address:* tuya.pal@vumc.org

doi: https://doi.org/10.1016/j.gim.2021.11.016

GWAS-identified SNVs include both those that contribute to a disease and those that may coincidentally correlate with disease variants. These patterns of correlations differ across populations and lessen the predictive power of GWAS from one population to the next. Consequently, using discoveries about disease biology from GWAS conducted primarily among European descent individuals to predict risks in other populations can lead to misinformation. In fact, accuracy to predict disease risk using data generated primarily from GWAS conducted in European descent individuals is a little less than half among East Asian individuals and drops to a quarter among Black individuals relative to predictive accuracy in European descent individuals.[3] Yet, data from studies among European descent individuals can still be useful for studies in other populations as shown by a study of approximately 50,000 individuals of non-European descent, in which many genetic variants mirrored those from studies strictly among European descent individuals; however, dozens of new associations were also identified that would not have been possible in a single population study.[7]

Although there have been many GWAS focused on cancers, the study by Wang et al[1] is among the few to compare PRS across diverse populations to quantify the lower predictive capability of PRS in populations of African ancestry than in that of European ancestry. They evaluated the discriminatory performance of PRS constructed from GWAS-identified variants to predict cancer risk among participants of European vs African ancestry across 15 major cancers in a large, racially diverse academic medical center biobank. They reported cancer-specific PRS models to have high discriminatory ability to distinguish cancer cases from cancer-free controls across both populations for breast, colorectal, and prostate cancer. Yet, when factoring in age, sex, and principal components, the discriminatory ability of PRS is small to nonexistent in individuals of African ancestry as compared with individuals of European ancestry.

Their findings are consistent with a recent study of patients with breast cancer, through which Du et al[8] reported that PRS did not improve risk prediction among women of African ancestry compared with European ancestry. In contrast, a recent study focused on prostate cancer reported promising results with similar risk prediction capabilities among individuals of European and African ancestry.[9] Specifically, the odds ratio for prostate cancer (based on those in the top PRS decile to those at average risk) was 3.89 (95% CI = 3.24-4.68) in men of European ancestry compared with 3.81 (95% CI = 1.48-10.19) in men of African ancestry. Although results that PRS may identify men of African ancestry at high risk for prostate cancer are encouraging, the substantially larger CI among men of African ancestry than that among men of European ancestry underscores the lower precision in risk prediction. We expect that risk prediction capability in populations of African ancestry will improve as size of GWAS for cancers approaches that of populations of European ancestry, enabling the construction of African ancestry–specific PRS with risk predictive capabilities comparable to those available for European ancestry populations. Simultaneously, there is emerging data to suggest that the proportion of prostate cancers explained by PRS differ across populations, suggesting differences in genetic architecture. Specifically, the effect of PRS on prostate cancer risk was reported to be higher in populations of African ancestry than in populations of European ancestry,[10] with Black individuals having more of their risks explained through GWAS.[11,12] These studies also highlight the tremendous potential for GWAS across diverse populations to improve both predictive ability and accuracy through generation of transethnic risk scores while discovering new risk variants. Moreover, evolutionary history is reported to contribute to the high rates of prostate cancer in men of African ancestry, through natural selection of adaptive alleles that may raise or protect against risk.[10] Thus, genetic factors alongside social and environmental factors, may explain some of the disparities in the higher rates of prostate cancer observed among men of African ancestry, and it further highlights the need for inclusiveness and diversity in GWAS. In fact, over the last few years, PRSs were being offered by certain commercial testing laboratories for patients of European descent, and yet, in light of the widespread recognition that use of PRS solely in populations of European descent may exacerbate disparities, multiple laboratories increasingly revised their testing approach to only offer PRS when they can offer it across all ancestries.

The study by Wang et al[1] serves to highlight the importance of GWAS representation across populations to ensure that all populations may benefit from these advances. Study results show the powerful capability of incorporating PRS into cancer risk prediction models to guide personalized prevention strategies while also underscoring the need to expand PRS clinical utility studies beyond European populations. Their findings are a call to action to emphasize the critical need for further large-scale studies to identify ancestry-specific genetic factors in populations of non-European descent to incorporate PRS into cancer risk assessment. This can only be accomplished through a concerted effort to increase the diversity of participants in genetic studies. Because we strive for diversity, we must also consider cultural competency in the implementation of genomic medicine, which is crucial to increase participation.[13] Simultaneously, it is imperative to broadly consider strategies to make genomics research more equitable through increasing the diversity in the genomics workforce and developing multilevel strategies to reduce barriers to access genomic medicine innovations.[14]

Ultimately, the inclusion of more diverse populations in genetic and genomic studies is both an ethical and scientific issue with tremendous potential to benefit future genetic research and improve human health. It is an ethical issue because the current risk prediction models are most relevant to European individuals, with reduced prediction accuracy among other populations. To implement precision prevention, we need data from populations

underrepresented in genomics studies to get risk prediction answers right, without which existing health inequities based on genomics will continue to widen. It is a scientific issue because without studying diverse populations, we lose the opportunity to make more discoveries, harming our scientific understanding of the genetic foundations of disease in all populations. It is imperative for us to address this disparity to augment our ability to make scientific discoveries while ensuring that genetic technologies may have the opportunity to reduce, rather than widen, existing health disparities.

## Conflict of Interest

The author declares no conflict of interest.

## References


1. Wang L, Desai H, Verma SS, et al. Performance of polygenic risk scores for cancer prediction in a racially diverse academic biobank. *Genet Med*. 2022;24:601–609. http://doi.org/10.1016/j.gim.2021.10.015.
2. Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. *Cell*. 2019;177(4):1080. http://doi.org/10.1016/j.cell.2019.04.032.
3. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Publisher correction: clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2021;53(5):763. http://doi.org/10.1038/s41588-021-00797-z.
4. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. 2016;538(7624):161–164. http://doi.org/10.1038/538161a.
5. Morales J, Welter D, Bowler EH, et al. A standardized framework for representation of ancestry data in genomics studies, with application to the NHGRI-EBI GWAS Catalog. *Genome Biol*. 2018;19(1):21. http://doi.org/10.1186/s13059-018-1396-2.
6. Bergström A, McCarthy SA, Hui R, et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science*. 2020;367(6484):eaay5012. http://doi.org/10.1126/science.aay5012.
7. Wojcik GL, Graff M, Nishimura KK, et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature*. 2019;570(7762):514–518. http://doi.org/10.1038/s41586-019-1310-4.
8. Du Z, Gao G, Adedokun B, et al. Evaluating polygenic risk scores for breast cancer in women of African ancestry. *J Natl Cancer Inst*. 2021;113(9):1168–1176. http://doi.org/10.1093/jnci/djab050.
9. Plym A, Penney KL, Kalia S, et al. Evaluation of a multiethnic polygenic risk score model for prostate cancer. *J Natl Cancer Inst*. Published online April 1, 2021. https://doi.org/10.1093/jnci/djab058
10. Lachance J, Berens AJ, Hansen MEB, Teng AK, Tishkoff SA, Rebbeck TR. Genetic hitchhiking and population bottlenecks contribute to prostate cancer disparities in men of African descent. *Cancer Res*. 2018;78(9):2432–2443. http://doi.org/10.1158/0008-5472.CAN-17-1550.
11. Conti DV, Darst BF, Moss LC, et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. *Nat Genet*. 2021;53(1):65–75. Published correction appears in *Nat Genet*. 2021;53(3):413. https://doi.org/10.1038/s41588-020-00748-0.
12. Harlemon M, Ajayi O, Kachambwa P, et al. A custom genotyping array reveals population-level heterogeneity for the genetic risks of prostate cancer and other cancers in Africa. *Cancer Res*. 2020;80(13):2956–2966. http://doi.org/10.1158/0008-5472.CAN-19-2165.
13. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. *Am J Public Health*. 2014;104(2):e16–e31. http://doi.org/10.2105/AJPH.2013.301706.
14. Jooma S, Hahn MJ, Hindorff LA, Bonham VL. Defining and achieving health equity in genomic medicine. *Ethn Dis*. 2019;29(Suppl 1):173–178. http://doi.org/10.18865/ed.29.S1.173.