

PERSPECTIVE

Extending precision medicine tools to populations at high risk of type 2 diabetes

Shivani Misra ^{1,2}, Jose C. Florez ^{3,4,5*}

1 Division of Metabolism, Digestion & Reproduction, Faculty of Medicine, Imperial College London, London, United Kingdom, **2** Department of Diabetes, Endocrinology & Metabolism, Imperial College Healthcare NHS Trust, London, United Kingdom, **3** Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, United States of America, **4** Programs in Metabolism and Medical & Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **5** Department of Medicine, Harvard Medical School, Boston, Massachusetts, United States of America

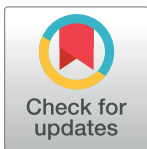
* jcflorez@mgh.harvard.edu

People of South Asian ancestry carry a 3-fold higher risk of developing type 2 diabetes (T2D) than white European individuals [1], with the disease typically manifesting a decade earlier [2] and at a leaner body mass index (BMI) [3]. The South Asian population is often considered as a uniform group, but significant heterogeneity in the prevalence of T2D and its phenotype manifestations across south Asia exists, with a higher prevalence in those from Bangladeshi and Pakistani communities [4]. Genome-wide association studies (GWAS) have not fully explained the excess risk observed in South Asian individuals [5,6], and attention has turned to strategies through which genetic information may be leveraged for clinical benefit, such as generating an aggregate of weighted single nucleotide polymorphisms (SNPs) that capture the overall genetic burden for a trait into a polygenic score (PS) (sometimes described as a polygenic risk score) [7]. However, constructing a PS remains challenging in populations that are underrepresented in GWAS.

In the accompanying article in *PLOS Medicine* [8], Hodgson and colleagues investigate the use of a PS to predict T2D in the Genes & Health (G&H) cohort, addressing a key knowledge gap in the applicability of such tools in underrepresented ethnicities. G&H is a pioneering community-based cohort of approximately 48,000 participants of predominantly British Bangladeshi and Pakistani heritage combining genetic and longitudinal electronic healthcare record data. They first assessed the transferability of known T2D genetic risk loci in G&H and constructed a PS using variants from a multi-ancestry GWAS, adjusting the scores for Pakistani and Bangladeshi individuals and selecting the one with the highest odds for prediction. This score was then integrated with 3 versions of a clinical model (QDiabetes) to predict T2D onset over 10 years in 13,642 individuals diabetes free at baseline. The authors show that incorporation of a PS with QDiabetes provided better discrimination of progression to T2D, especially in those developing T2D under 40 years of age and in women with a history of gestational diabetes. Finally, they incorporated the PS into cluster analyses of baseline routine clinical characteristics, replicating clusters defined in European populations and identifying a cluster resembling a subgroup of severe insulin deficiency. This study significantly advances the field on the transferability of PSs, reproducibility of T2D clusters, and clinical translation of these findings to precision medicine for diabetes.

Transferability of PSs and T2D clusters to other ethnicities

The generation of ethnicity-specific T2D PS for every subethnic group remains aspirational, as the large sample sizes needed to do this robustly are prohibitive (the genetic ancestry of the



OPEN ACCESS

Citation: Misra S, Florez JC (2022) Extending precision medicine tools to populations at high risk of type 2 diabetes. *PLoS Med* 19(5): e1003989. <https://doi.org/10.1371/journal.pmed.1003989>

Published: May 19, 2022

Copyright: © 2022 Misra, Florez. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: BMI, body mass index; G&H, Genes & Health; GWAS, genome-wide association studies; PS, polygenic score; SNP, single nucleotide polymorphism; T2D, type 2 diabetes.

Provenance: Commissioned; not externally peer reviewed.

Indian subcontinent, for example, is more diverse than the whole of Europe [9]). Thus, strategies to utilise existing scores derived from other populations, or leveraging multi-ancestry GWAS, have predominated. Attempts to apply a PS derived in a population of one ancestry to another ethnic group have shown variable performance [10,11]. Indeed, a lower percentage of GWAS loci than expected were replicated in the present study, and several loci previously identified in another South Asian-specific GWAS also failed to replicate [5]; although this may be related to power, it also demonstrates the heterogeneity within South Asian populations, arguably more relevant in cultures where consanguineous unions are prevalent.

Previously described data-driven diabetes clusters [12] have been partially reproducible in other ethnicities including Indian cohorts [13]. Despite the absence of specific biomarker data (C-peptide and pancreatic autoantibodies), the present study was able to identify a probable insulin-deficient cluster on the basis of a high PS reflecting variants associated with beta-cell dysfunction and leaner BMI. It remains unclear whether data-driven clustering approaches stratifying the T2D population offer benefit over and above routine care.

What is the clinical benefit in a higher-risk group?

Precision medicine strategies such as the use of a PS in T2D have predominantly focused on European populations, even though South Asian populations account for a significant proportion of all T2D cases. A strength of Hodgson and colleagues' study is the identification of a specific subset of individuals where the PS may add value; early-onset T2D progresses rapidly to insulin treatment and carries a higher risk of microvascular and macrovascular complications and earlier mortality. The present study poses an interesting research question: Could identifying people earlier in life through an elevated PS attenuate progression to T2D through lifestyle interventions or more aggressive cardiovascular risk reduction?

Future directions

The study by Hodgson and colleagues has some limitations: Because South Asians are under-represented in GWAS, power to discover and select relevant SNPs was finite. Given the unique cohort assembled, there was also no opportunity to replicate prediction and clustering in another cohort. The linked medical records yielded incomplete clinical data to robustly define and replicate previous phenotype clusters and conclusively demonstrate diabetes subtypes. Finally, the risk of undiagnosed diabetes in some control participants cannot be excluded.

Despite these acknowledged limitations, the findings show a clear pathway for future investigations. First, larger GWAS are needed in Pakistani, Bangladeshi, and other South Asian and non-European populations, to enhance the robustness of PS and their portability [4]. Additionally, more investigations are needed to understand the effect of PS in clinical practice. This should occur individually (e.g., qualitative analysis of the effects of genetic information on patient behaviour and lifestyle) and at a clinical pathway level (i.e., whether the PS changes management over and above routine clinical metrics, particularly in resource-limited health-care settings). Finally, the application of process-specific PS (scores that cluster SNPs based on physiological associations) has huge potential to interrogate the biology underlying T2D [14].

Hodgson and colleagues provide compelling evidence that precision medicine tools can be applied meaningfully in hitherto underrepresented populations, and building resources such as the G&H cohort is a valuable endeavour. They provide knowledge upon which clinical studies that test utility of PS in South Asians can be designed. Given the higher prevalence of T2D in South Asians, such studies are critical. However, their study also illustrates the challenges of studying underrepresented ethnicities. Overcoming such challenges may yield not only genetic discovery, but also therapeutic opportunities for vulnerable populations.

References

1. Tillin T, Forouhi NG, McKeigue PM, Chaturvedi N. Southall And Brent REvisited: Cohort profile of SABRE, a UK population-based comparison of cardiovascular disease and diabetes in people of European, Indian Asian and African Caribbean origins. *Int J Epidemiol*. 2012; 41:33–42. <https://doi.org/10.1093/ije/dyq175> PMID: 21044979
2. Wright AK, Welsh P, Gill JMR, Kontopantelis E, Emsley R, Buchan I, et al. Age-, sex- and ethnicity-related differences in body weight, blood pressure, HbA1c and lipid levels at the diagnosis of type 2 diabetes relative to people without diabetes. *Diabetologia*. 2020; 63:1542–53. <https://doi.org/10.1007/s00125-020-05169-6> PMID: 32435821
3. Paul SK, Owusu Adjah ES, Samanta M, Patel K, Bellary S, Hanif W, et al. Comparison of body mass index at diagnosis of diabetes in a multi-ethnic population: A case-control study with matched non-diabetic controls. *Diabetes Obes Metab*. 2017; 19:1014–23. <https://doi.org/10.1111/dom.12915> PMID: 28211609
4. Hills AP, Arena R, Khunti K, Yajnik CS, Jayawardena R, Henry CJ, et al. Epidemiology and determinants of type 2 diabetes in south Asia. *Lancet Diabetes Endocrinol*. 2018; 6: 966–978. [https://doi.org/10.1016/S2213-8587\(18\)30204-3](https://doi.org/10.1016/S2213-8587(18)30204-3) PMID: 30287102
5. Kooner JS, Saleheen D, Sim X, Sehmi J, Zhang W, Frossard P, et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat Genet*. 2011; 43:984–9. <https://doi.org/10.1038/ng.921> PMID: 21874001
6. Sohani ZN, Deng WQ, Pare G, Meyre D, Gerstein HC, Anand SS. Does genetic heterogeneity account for the divergent risk of type 2 diabetes in South Asian and white European populations? *Diabetologia*. 2014; 57:2270–81. <https://doi.org/10.1007/s00125-014-3354-1> PMID: 25145545
7. Udler MS, McCarthy MI, Florez JC, Mahajan A. Genetic Risk Scores for Diabetes Diagnosis and Precision Medicine. *Endocr Rev*. 2019; 40:1500–20. <https://doi.org/10.1210/er.2019-00088> PMID: 31322649
8. Hodgson S, Huang QQ, Sallah N, Genes & Health Research Team, Griffiths CJ, Newman WG, et al. Integrating polygenic risk scores in the prediction of type 2 diabetes risk and subtypes in British Pakistanis and Bangladeshis: A population-based cohort study. *PLoS Med*. 2022. <https://doi.org/10.1371/journal.pmed.1003981>
9. Xing J, Watkins WS, Hu Y, Huff CD, Sabo A, Muzny DM, et al. Genetic diversity in India and the inference of Eurasian population expansion. *Genome Biol*. 2010; 11:R113. <https://doi.org/10.1186/gb-2010-11-11-r113> PMID: 21106085
10. Multhaupt ML, Kita R, Krock B, Eriksson N, Fintanillas P, Asilbeyan S, et al. The science behind 23andMe's type 2 diabetes report. 2019 [cited 2022 Mar 15]. Available from: https://permalinks.23andme.com/pdf/23_19-Type2Diabetes_March2019.pdf
11. Chikowore T, Ekoru K, Vujkovi M, Gill D, Pirie F, Young E, et al. Polygenic Prediction of Type 2 Diabetes in Africa. *Diabetes Care*. 2022; 45:717–23. <https://doi.org/10.2337/dc21-0365> PMID: 35015074
12. Ahlqvist E, Storm P, Käräjämäki A, Martinell M, Dorkhan M, Carlsson A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol*. 2018; 6:361–9. [https://doi.org/10.1016/S2213-8587\(18\)30051-2](https://doi.org/10.1016/S2213-8587(18)30051-2) PMID: 29503172
13. Anjana RM, Baskar V, Nair ATN, Jebarani S, Siddiqui MK, Pradeepa R, et al. Novel subgroups of type 2 diabetes and their association with microvascular outcomes in an Asian Indian population: a data-driven cluster analysis: the INSPIRED study. *BMJ Open Diabetes Res Care*. 2020; 8: e001506. <https://doi.org/10.1136/bmjdr-2020-001506> PMID: 32816869
14. Udler MS, Kim J, von Grotthuss M, Bonàs-Guarch S, Cole JB, Chiou J, et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. *PLoS Med*. 2018; 15:e1002654–4. <https://doi.org/10.1371/journal.pmed.1002654> PMID: 30240442