Analysis

Polygenic risk scores:

improving the prediction of future disease or added complexity?

INTRODUCTION

'Genomic prevention' is a key pillar of the UK government's current 10-year genomics strategy, and polygenic risk scores (PRS) have been highlighted as an exciting area of development.[1] PRS quantifies the cumulative effects of multiple common genetic variants in an individual; this can be used to predict and stratify disease predisposition for common complex conditions such as cardiovascular disease (CVD), diabetes, and cancer.

The clinical utility of PRS is also currently being evaluated through the UK government's 'Our Future Health' research programme, which will generate PRS for up to 5 million individuals who can opt in to receive their results.[2] PRS could be mainstreamed into clinical practice, with the aim of providing targeted screening and interventions for high-risk populations. Further, PRS is already provided by some direct-to-consumer testing companies and patients may present to their GPs with these results.

However, while PRS aims to enhance disease detection and prevention, there are multiple ethical, legal, and social issues to consider, particularly to reduce harms from overtreatment, overdiagnosis, or false reassurance. Further, given suboptimal utilisation of some currently available risk prediction tools in primary care, such as QRISK,[3] is PRS in fact a tool too far?

POLYGENIC INHERITANCE OF COMMON COMPLEX DISEASES

Inherited diseases can be broadly categorised into either monogenic or polygenic diseases. Monogenic diseases, such as cystic fibrosis, can be traced to variants in a single gene. In contrast, polygenic common complex diseases, such as coronary artery disease (CAD) and obesity, are influenced by multiple genomic variants dispersed across the genome, with each individual variant having a very low effect on disease susceptibility (called penetrance). These variants are termed 'SNPs' — single nucleotide polymorphisms — a change in one nucleotide in the DNA sequence that occurs approximately once in 100–300 base pairs.[4] SNPs are the most common variant in the human genome and by definition are present in over 1% of the population. SNPs act as biological markers that help locate genes that are associated with disease;[4] they may also play a more direct role in disease by affecting the gene's function.[4]

PRS aims to quantify the cumulative effects of SNPs across the genome. SNPs implicated in common complex diseases are derived from genome wide association studies (GWAS). GWAS identify SNPs that are disproportionately present in cases with a complex condition compared with matched controls.

PRS IN DIRECT-TO-CONSUMER TESTING

PRS is already provided by some direct-to-consumer testing companies. A home testing kit costs around 100 GBP and usually involves sending a sample of saliva in the post and completing an online questionnaire answering demographic questions about yourself. For example, the 23andMe Type 2 Diabetes Health Predisposition (T2DM) report tests for 1244 variants associated with a higher risk of developing T2DM.[5]

Direct-to-consumer PRS tests employ SNP microarrays that hybridise with selected SNPs within the genome. The PRS result generated combines detected SNPs with age and self-reported ethnicity as a measure of disease predisposition for T2DM. The result is returned as a centile describing your relative risk of disease compared with other participants within your self-reported ethnic group.[5] However, direct-to-consumer results are reported to have a false positive rate of up to 40%.[6]

CLINICAL APPLICATIONS AND UTILITY OF PRS

PRS in chronic disease

There is potential for PRS to trigger targeted early preventive treatments for chronic diseases such as CVD and type 2 diabetes, which are a major burden on the health system.[1] Individuals stratified as high risk for disease could be offered preventive measures at an earlier age than within current service delivery such as participation in screening, lifestyle counselling, and prophylactic treatments as appropriate.[7] Studies have shown that a high PRS is predictive of myocardial infarction early in life before other risk factors are evident.[8] Although current risk prediction models are more useful in older populations, preventive measures initiated at this later stage may offer less impact on disease progression.

There is also promising research utilising PRS for coronary artery disease (CAD), suggesting that patients in the top decile for polygenic risk regardless of LDL-C levels derive greater clinical benefit from PCSK-9 inhibitors.[9] However, a recent analysis of adults in the US showed that PRS did not improve risk prediction for CAD in a general middle-aged White population, when compared with a 10-year QRISK calculator equivalent.[10] It is unlikely that PRS will be used in isolation in the future, but rather incorporated as another risk factor into current validated risk prediction tools, such as QRISK, utilising existing IT systems and workflows to avoid the need for primary care practitioners to interpret raw PRS data.

Modelling for incorporation of PRS into the current NHS CVD Health Check highlights issues crucial to wider implementation: staffing resources for genotyping and interpretation, data recording within the electronic health record, and the need for decision aids and guidance for clinicians and patients.[11] There are also concerns that incorporating PRS may extend the proportion of the population selected for primary prevention. Current National Institute for Health and Care Excellence (NICE) guidelines recommend starting a statin in all adults with a QRISK at or above 10%; as a consequence all males aged >68 years automatically meet the risk threshold for being recommended statin therapy. The addition of PRS scores may exacerbate the proportion of the population that meet the risk threshold for treatment.

While PRS is not routinely available in the NHS, it is available on a research basis. Our Future Health is about to perform large-scale PRS testing on UK research volunteers, with the aim of discovering new ways to prevent, detect, and treat common complex diseases more accurately, *'so future generations can live healthier lives for longer'*.[2]

In the future, potentially multiple PRS for various conditions could be generated using one microarray that tests for a variety of disease-associated SNPs. 'MultiPRSs' could be used in both primary and secondary prevention settings — to determine disease risk, as well as facilitate more timely interventions for the complications of established disease. One example of this model has recently been trialled in the context of reducing type 2 diabetes complications from renal and CVD: individuals stratified to

the highest risk group were shown to have a 47% reduction in mortality and the greatest absolute risk reduction (number needed to treat [NNT] = 12 to prevent one death from CVD over 5 years).[12]

PRS in cancer

Cancer is a leading cause of death in the UK and there have been several applications of PRS for the commonest cancers.

'CanRisk' is a 'user-friendly' online risk assessment tool that calculates breast cancer risk (https://www.canrisk.org); it incorporates non-genetic factors, including family history, mammographic density, lifestyle and hormonal factors, and genetic factors in the form of high-impact rarer pathogenic variants in cancer predisposing genes (for example, BRCA1) and PRS for 313 SNPs. The tool is now licensed for clinical use within the NHS and is used within routine clinical practice in clinical genetics.

Combining PRS into breast cancer risk algorithms can be used to stratify females into different risk groups. One PRS study found that the top 19% of females met the same 10-year risk for breast cancer at the start of the current UK national screening programme before 40 years, and approximately one in five would never reach this level of 10-year risk.[13] Another study concluded that females in the highest centile of a PRS study had a 30% lifetime risk of breast cancer comparable with rarer monogenic mutations.[14] Another breast cancer risk model that is commonly used in clinical practice is the Tyrer–Cuzick model;[15] incorporation of PRS into this model showed better risk discrimination than high-risk gene panel testing for patients with a cancer diagnosis or actionable findings.[16] PRS could potentially be used to refine the current screening programme, with those with the highest scores being offered earlier mammographic screening and *vice versa*.

In prostate cancer, the BARCODE 1 study is a PRS study looking at SNP profiling in asymptomatic males of European ancestry aged 55–69 years in general practice. A feasibility pilot study showed that the uptake of providing a saliva sample for polygenic risk assessment was 22%. The full study is still in progress and will report on the association of PRS with prostate MRI and biopsy outcomes.[17]

SOME ETHICAL, SOCIAL, AND LEGAL CONSIDERATIONS

PRS shifts the traditional concept of disease

Studies such as BARCODE 1 will determine if PRS as a method of population risk stratification are feasible and will be successful in targeting screening to identify a higher proportion of individuals with clinically significant disease.

However, since PRS results lie on a continuum then it is not always clear, for a specific condition, where to set the threshold to say a disease is present or absent, or someone is low or high risk;[18] any potential for overmedicalisation and iatrogenic harm from overtreatment must be mitigated. Alteration of the CVD risk threshold to guide treatment with statin therapy illustrates the challenge of balancing disease prevention through reduction in incidence with potential side effects from treatment.

Further, there are unanswered questions about whether a high PRS — depending on the condition — would need to prompt warning and testing of family members, since the inheritance of SNPs is not as predictable as for monogenic conditions.[18]

Psychological impact of testing

Fears around the potential psychological harms of disease predisposition tests have yet to be established, with studies so far showing either no or a positive effect on behaviour change.[18,19] The reporting strategy is a key determinant in how patients receive and process their results.[18]

Concerns about how reporting could safely and effectively be conducted on a larger population screening level have been raised.[7] The current system for pre- and post-test counselling for monogenic testing often involves consultations with a trained genetic counsellor. For wider implementation of PRS, counselling will likely be performed by other healthcare professionals, highlighting the need for evidence-based guidelines and training, particularly of the primary care workforce.[7,18]

Healthy lifestyle promotion is important universally

While PRS has the potential to be used in a whole variety of contexts, in the case of lifestyle promotion, concerns have been raised that PRS could serve as a distraction.[20]

PRS results may also reinforce the concept of 'genetic determinism', particularly in the absence of adequate and effective counselling. While a high PRS may invite conversations around lifestyle modification and preventive therapeutic treatments, low PRS results could offer false reassurance. When genetic testing for CVD had been introduced in primary care, there were concerns that a low PRS, despite increased risk related to other risk factors, may lead to false reassurance.[21] This may lead to individuals failing to adopt lifestyle risk-reducing behaviour and/or take up screening.

Scores are not predictive for all ethnicities

Existing knowledge of the genome is largely based on studies of White European populations and there remains a lack of diversity in genomic databases.[22] Large-scale research projects such as Our Future Health are working to address this imbalance by 'oversampling' from Black, Asian, and minority ethnic groups.[1] However, knowledge of disease-associated SNPs in non-White European populations is still lagging behind.[22]

Risk allele frequencies vary across different populations and so results derived from studies in one population are therefore not always clinically useful universally.[22] For example, one recent study looking at PRS models in risk prediction for breast cancer across ethnic groups found that effect sizes were smaller for females of African ancestry.[23] For PRS to be applicable across diverse groups then a focus on collecting ethnically diverse datasets is vital for research to be clinically relevant for all.

Incomplete risk estimate

GWAS provide a measure of SNPs associated with disease — they do not prove causation but rather correlation.[18] Understanding of the role SNPs play in common complex disease is evolving and scores may not adequately reflect potential gene–environment interactions and epigenetic modifications (changes in gene expression) that may impact disease risk.[24]

There are also concerns that PRS based on common variants alone may considerably skew disease risk estimates.[7] Individuals with a family history of disease, such as breast cancer, may also carry rarer highly penetrant pathogenic variants potentially excluded from PRS testing,[7] underlining the importance of taking a thorough family history.

Genetic discrimination

There is an agreed Code of Practice specifically regarding predictive genetic testing: there are warnings that, in the absence of a robust ethical framework, these results could potentially be used by employers and insurance companies to discriminate against individuals.[25]

The Association of British Insurers advises that the majority of predictive test results should not currently be taken into account by insurers;[26] however, there is potential for

this advice to change in the future.

CONCLUSION

PRS is a risk prediction tool currently in development; it aims to identify individuals at high genetic risk of disease. The clinical validity and utility of PRS in various contexts are still being evaluated through large-scale prospective trials that up until now have been lacking.[7]

There are a number of challenges to future implementation of PRS: increasing ethnic diversity in datasets; development of evidence-based guidelines and a robust ethical framework; campaigns to improve genomic literacy; and investment in education, training, and resources in primary care. In addition, we need to better understand the potential effects on our patients including clinical and psychological outcomes, overdiagnosis, and medicalisation. These are all high hurdles to overcome to ensure PRS are not just 'one tool too many', meaning it may be some time before research becomes mainstream practice.

Aya Ayoub,
(ORCID: 0000-0003-4380-0932) GP and Chair, Royal College of General Practitioners (RCGP) Junior International Committee, RCGP; Imperial College London, London.

Jana McHugh,
(ORCID: 0000-0002-8210-6188) Clinical Research Fellow, Institute of Cancer Research, London.

Judith Hayward,
(ORCID: 0000-0001-7137-1261) GP and RCGP Clinical Champion for Genomics in Primary Care, RCGP, London; Yorkshire Regional Genetics Service.

Imran Rafi,
(ORCID: 0000-0001-9673-9675) GP and Reader in Primary Care and Genomics, Institute for Medical and Biomedical Education, St George's University of London, London.

Nadeem Qureshi,
(ORCID: 0000-0003-4909-0644) GP and Clinical Professor of Primary Care, Primary Care Stratified Medicine Research Group, University of Nottingham, Nottingham.

Provenance
Freely submitted; externally peer reviewed.

Competing interests
Judith Hayward and Imran Rafi both receive funding from Congenica/InnovateUK as PIs for the HAPPY research trial. Nadeem Qureshi was funded by the National Institute for Health and Care Research for a study to improve identification of familial hypercholesterolaemia in primary care, which included polygenic hypercholesterolaemia.

Acknowledgements
We are grateful for helpful discussions with Ros Eeles, The Institute of Cancer Research and Royal Marsden NHS Foundation Trust.

DOI: https://doi.org/10.3399/bjgp22X720437

ADDRESS FOR CORRESPONDENCE

Aya Ayoub
Imperial College London, National Heart and Lung Institute, London SW7 2AZ, UK.

Email: aya.ayoub20@imperial.ac.uk

REFERENCES

1. Department of Health and Social Care (DHSC). *Genome UK: the future of healthcare*. London: DHSC, 2020. https://www.gov.uk/government/publications/genome-uk-the-future-of-healthcare (accessed 13 Jul 2022).
2. Our Future Health. Research programme. 2022. https://ourfuturehealth.org.uk/research-programme (accessed 13 Jul 2022).
3. Finnikin S, Ryan R, Marshall T. Statin initiations and QRISK2 scoring in UK general practice: a THIN database study. *Br J Gen Pract* 2017; DOI: https://doi.org/10.3399/bjgp17X693485.
4. Robert F, Pelletier J. Exploring the impact of single-nucleotide polymorphisms on translation. *Front Genet* 2018; **9:** 507.
5. 23andMe. Let's talk about type 2 diabetes & genetics. 2021. https://www.23andme.com/topics/health-predispositions/type-2-diabetes (accessed 13 Jul 2022).
6. Rafi I, Hayward J, Lucassen A. *Position statement on direct to consumer genomic testing*. 2019. https://www.rcgp.org.uk/policy/rcgp-policy-areas/genomic-position-statement.aspx (accessed 13 Jul 2022).
7. Polygenic Risk Score Task Force of the International Common Disease Alliance. Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. *Nat Med* 2021; **27(11):** 1876–1884.
8. Isgut M, Sun J, Quyyumi AA, Gibson G. Highly elevated polygenic risk scores are better predictors of myocardial infarction risk early in life than later. *Genome Med* 2021; **13(1):** 13.
9. Damask A, Steg PG, Schwartz GG, *et al*. Patients with high genome-wide polygenic risk scores for coronary artery disease may receive greater clinical benefit from alirocumab treatment in the ODYSSEY OUTCOMES trial. *Circulation* 2020; **141(8):** 624–636.
10. Mosley JD, Gupta DK, Tan J, *et al*. Predictive accuracy of a polygenic risk score compared with a clinical risk score for incident coronary heart disease. *JAMA* 2020; **323(7):** 627–635.
11. Brigden T, Sanderson S, Janus J, *et al*. *Implementing polygenic scores for cardiovascular disease into NHS Health Checks*. Cambridge: PHG Foundation, 2021.
12. Tremblay J, Haloui M, Attaoua R, *et al*. Polygenic risk scores predict diabetes complications and their response to intensive blood pressure and glucose control. *Diabetologia* 2021; **64(9):** 2012–2025.
13. Mavaddat N, Michailidou K, Dennis J, *et al*. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet* 2019; **104(1):** 21–34.
14. Mars N, Widén E, Kerminen S, *et al*. The role of polygenic risk and susceptibility genes in breast cancer over the course of life. *Nat Commun* 2020; **11(1):** 6383.
15. Himes DO, Luthy BE, Root AE, Gammon A. Breast cancer risk assessment: calculating lifetime risk using the Tyrer–Cuzick model. *J Nurse Pract* 2016; **12(9):** 581–592.
16. Evans DGR, van Veen EM, Harkness EF, *et al*. Breast cancer risk stratification in women of screening age: incremental effects of adding mammographic density, polygenic risk, and a gene panel. *Genet Med* 2022; **24(7):** 1485–1494.
17. Eeles RA, ni Raghallaigh H, The BARCODE1 Study Group. BARCODE 1: a pilot study investigating the use of genetic profiling to identify men in the general population with the highest risk of prostate cancer to invite for targeted screening. *J Clin Oncol* 2020; **38(15 suppl):** 1505.
18. Lewis ACF, Green RC. Polygenic risk scores in the clinic: new perspectives needed on familiar ethical issues. *Genome Med* 2021; **13(1):** 14.
19. Roberts JS. Assessing the psychological impact of genetic susceptibility testing. *Hastings Cent Rep* 2019; **49 Suppl 1(Suppl 1):** S38–S43.
20. Science Media Centre. Expert reaction to Matt Hancock's speech on genetic testing. 2019. https://www.sciencemediacentre.org/expert-reaction-to-matt-hancocks-speech-on-genetic-testing (accessed 13 Jul 2022).
21. Middlemass JB, Yazdani MF, Kai J, *et al*. Introducing genetic testing for cardiovascular disease in primary care: a qualitative study. *Br J Gen Pract* 2014; DOI: https://doi.org/10.3399/bjgp14X679714.
22. Manolio TA. Using the data we have: improving diversity in genomic research. *Am J Hum Genet* 2019; **105(2):** 233–236.
23. Liu C, Zeinomar N, Chung WK, *et al*. Generalizability of polygenic risk scores for breast cancer among women with European, African, and Latinx ancestry. *JAMA Netw Open* 2021; **4(8):** e2119084.
24. Domingue BW, Trejo S, Armstrong-Carter E, Tucker-Drob E. Interactions between polygenic scores and environments: methodological and conceptual challenges. *Sociol Sci* 2020; **7(19):** 465–486.
25. Chapman CR, Mehta KS, Parent B, Caplan AL. Genetic discrimination: emerging ethical challenges in the context of advancing technology. *J Law Biosci* 2020; **7(1):** lsz016.
26. HM Government, Association of British Insurers. *Code on genetic testing and insurance*. London: ABI, 2018. https://www.abi.org.uk/globalassets/files/publications/public/genetics/code-on-genetic-testing-and-insurance-final.pdf (accessed 13 Jul 2022).