The future of polygenic risk scores in direct-to-consumer genomics

Yuhao Sun[1,2,3], John Vines[2], Albert Tenesa[1,4]

Corresponding authors: Yuhao Sun (yuhao.sun@ed.ac.uk), Albert Tenesa (albert.tenesa@ed.ac.uk)

[1] *Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian, UK*

[2] *Institute for Design Informatics, School of Informatics, University of Edinburgh, Edinburgh, UK*

[3] *Edinburgh Futures Institute, University of Edinburgh, Edinburgh, UK*

[4] *MRC Human Genetics Unit at the MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK*

In March 2025, direct-to-consumer (DTC) genomics company 23andMe filed for Chapter 11 bankruptcy protection after years of declining demand for ancestry kits and reputational harm from an October 2023 data breach affecting nearly seven million customers (https://www.reuters.com/business/healthcare-pharmaceuticals/dna-testing-firm-23andme-files-chapter-11-bankruptcy-sell-itself-2025-03-24/). Although the company subsequently announced plans to restructure and continue its operations under new financial arrangements, the bankruptcy proceedings raised concerns over the long-term handling of over 15 million users' genomic profiles and prompted a U.S. House committee hearing on data and national-security implications (https://www.congress.gov/event/119th-congress/house-event/118362). Since its founding in 2006, 23andMe established its reputation through saliva-based ancestry reports and a vision of expanding into health-related predictions. In recent years, the company introduced reports based on polygenic risk scores (PRS) models for conditions, such as coronary artery disease (CAD) and type 2 diabetes (https://permalinks.23andme.com/pdf/23_21-PRSMethodology_May2020.pdf).

PRS compute a weighted sum of thousands of genetic variants, which are effect sizes derived from genome-wide association studies (GWAS), to stratify disease risk[1]. Numerous studies have demonstrated clinical utility; for example, the *BOADICEA* model integrates a 313-variant PRS to improve breast cancer risk stratification and guide prevention[2]. Well-calibrated PRS for CAD have similarly been shown to identify high-risk individuals for early lifestyle and pharmacological interventions[3]. However, the broader performance of PRS remains dependent on key methodological and demographic factors. Their predictive accuracy often diminishes in populations underrepresented in GWAS, raising concerns that current PRS may exacerbate health disparities[4]. Additionally, small effect sizes and gene-environment interactions further complicate the translation from statistical association to meaningful individual risk estimates[4]. In clinical settings, ambiguities around PRS can be addressed and contextualised through professional guidance. By contrast, in DTC settings, motivated primarily by curiosity or a desire for health empowerment[5], many users approach PRS without sufficient context to weigh its strengths and limitations. Once they receive point estimates devoid of uncertainty bounds or professional interpretation, users are left ill-equipped to understand the nuances of the implications. Therefore, we ask to what extent can consumers responsibly navigate these complex probabilistic scores without clinical support? In this correspondence, without questioning the scientific basis or clinical potential of PRS, we highlight challenges that are specific to DTC settings and advocate for greater transparency, contextualised communication, and sustained consumer support to ensure their responsible use.

Clinical trials have shown that PRS can inform positive health behaviours when provided alongside professional guidance and support tools[6,7]. For example, the *MI-GENES* randomised trial demonstrated that disclosing a CAD PRS alongside conventional risk estimates, with genetic counselling and shared decision-making, led to higher statin initiation and significantly lower low-density lipoprotein cholesterol levels at 6 months[6]. However, such engagement may be absent or inconsistently provided in DTC contexts. In our recent mixed-methods study[8] exploring public perceptions of PRS and prospective DTC use, UK-based participants engaged with speculative DTC-PRS scenarios and struggled to interpret predictive outputs. Many could grasp concepts of PRS and welcomed the idea of early intervention, but felt ungrounded: unable to appreciate or understand from the cohort context or calibration, and uncertain about next steps. In the absence of such clarity, some anticipated confusion, anxiety, and diminished trust in the information[8]. As recently discussed, these difficulties reflect broader gaps in public understanding, where statistical uncertainty and genetic determinism remain poorly addressed[7]. We do not suggest that PRS inherently cause psychological or behavioural harms; however, when delivered

without transparent communication, interpretive support, and sustained follow-up in DTC settings, they can leave users uncertain and reduce their perceived usefulness.

Beyond individual reactions, we argue that PRS in DTC settings operate within a complex socio-technical system where algorithmic opacity, commercial incentives, and delayed feedback loops collide. Small fluctuations in score can shift risk categories. However, lifestyle interventions based on those scores yield benefits only years later, by which point the initial emotional impact may have faded, leaving behind uncertainty or disappointment about how useful or actionable the prediction really was. Moreover, foresight-driven commercial narratives compress the multifactorial and evolving nature of individual health trajectories into a single static metric, obscuring how well-being unfolds across social, emotional, and environmental dimensions. This trend extends beyond legacy players: *impute.me*, originally an "open-source, non-profit" PRS platform, had processed 28,651 genome uploads and 3.1 million analytical queries by mid-2019[9]. As of the time of writing, the source code originally referenced in the publication is no longer available in the linked repository; the *impute.me* website now redirects to a commercial provider offering whole-genome sequencing and PRS analysis for a considerable fee. Transitions from research-oriented to profit-driven platforms risk undermining core scientific values – limiting access, reducing transparency, and removing codebases that once supported reproducibility – thereby exacerbating disparities in genomic health equity.

Therefore, we advocate a reflective re-examination of PRS deployment in DTC genomics as part of its wider socio-technical system. First, DTC-PRS service providers should ensure transparent data governance by publicly sharing analytical workflows, validation metrics, and data-use policies. Second, consent and onboarding materials should go beyond legal jargon to explain limitations, uncertainties, and psychological implications in clear and accessible language. Third, consistent with prior recommendations on risk communication[10], providers should prioritise absolute risk, clearly presented using reference groups, defined timeframes, practical scenarios, and visual aids tailored to users with varying levels of numeracy. Relatedly, human-computer interaction (HCI) research – echoing recent calls for co-design and digital decision support in PRS implementation[7,8] – could further strengthen these efforts by informing interface layout, explanatory text, and tool development. Finally, multi-stakeholder governance models are essential to prioritise ethical accountability and consumer empowerment over short-term profit. To ensure feasibility, these recommendations should align with existing data-protection frameworks and clarify how providers can embed consent, transparency, and breach-notification requirements within legal obligations. For example, the European Union's General Data Protection Regulation (GDPR) and the U.S. Health Insurance Portability and Accountability Act (HIPAA). By reframing our approach from 'predicting the future' to 'navigating

uncertainty' in partnership, we may transform a sector-wide challenge into an opportunity to build public trust and promote more humane, equitable, and accountable engagement with PRS in DTC genomics.

Competing interests

The authors declare no competing interests.

Acknowledgements

AT is supported by awards from the Biotechnology and Biological Sciences Research Council (BBS/E/D/10002070, BBS/E/D/30002275, BBS/E/RL/230001A).

References

1. Khera, A. V. et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat. Genet. 50, 1219-1224 (2018).
2. Lee, A. et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. Genetics in Medicine 21, 1708-1718 (2019).
3. Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of polygenic risk scores. Nat. Rev. Genet. 19, 581-590 (2018).
4. Martin, A. R. et al. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat. Genet. 51, 584-591 (2019).
5. Peck, L. et al. Why do people seek out polygenic risk scores for complex disorders, and how do they understand and react to results? Eur. J. Hum. Genet. 30, 81-87 (2021).
6. Kullo, I. J. et al. Incorporating a genetic risk score into coronary heart disease risk estimates. Circulation 133, 1181-1188 (2016).
7. Sanderson, S. C. & Inouye, M. Psychological and behavioural considerations for integrating polygenic risk scores for disease into clinical practice. Nat. Hum. Behav. 9, 1098-1106 (2025).
8. Sun, Y., Tenesa, A. & Vines, J. Human-precision medicine interaction: public perceptions of polygenic risk score for genetic health prediction. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems 1-26 (2025) doi:10.1145/3706598.3713567.
9. Folkersen, L. et al. Impute.me: an open-source, non-profit tool for using data from direct-to-consumer genetic testing to calculate and interpret polygenic risk scores. Front. Genet. 11, 578 (2020).
10. Spiegelhalter, D. Risk and uncertainty communication. Annu. Rev. Stat. Appl. 4, 31-60 (2017).