Title: Prioritizing the detection of rare pathogenic variants in population screening

Authors:
Paul Lacaze[1*], Ranjit Manchanda[2‡] and Robert C Green[3‡]

*Corresponding author: paul.lacaze@monash.edu
‡ Co-senior authors / These authors contributed equally

Affiliations:
[1] Public Health Genomics, Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia
[2] Wolfson Institute of Population Health, CRUK Barts Centre, Queen Mary University of London, Charterhouse Square, London, United Kingdom; Department of Gynaecological Oncology, St Bartholomew's Hospital, London, United Kingdom; Department of Health Services Research, London School of Hygiene & Tropical Medicine, London, UK; Department of Health Services Research, Faculty of Public Health & Policy, London School of Hygiene & Tropical Medicine, London, United Kingdom.
[3] Mass General Brigham, Broad Institute, Ariadne Labs and Harvard Medical School, Boston, Massachusetts, USA

Standfirst:

Population genomic screening for rare monogenic variants is now supported by substantial evidence of clinical utility and cost-effectiveness, but there is much less evidence to support screening with polygenic risk scores (which do not detect rare variants). Using 'only' polygenic scores in population screening initiatives, while ignoring the detection of higher risk rare monogenic variants, is ill-advised.

Manuscript:

Population genomic screening to identify high-risk individuals from the general population presents a significant opportunity for public health intervention and disease prevention. At this time, prioritizing the detection of rare clinically significant pathogenic variants for penetrant medically actionable monogenic conditions represents an evidence-based, appropriate and justifiable strategy for large-scale population screening. There is currently less evidence to support population screening with polygenic risk scores, which when calculated using genotyping arrays, will miss the identification of rare variant carriers for medically actionable conditions, who are among the highest risk individuals in the general population.

Background

Modern medical genetics has largely emerged in the context of molecular diagnosis of persons affected by inherited conditions. But the potential for using genomic technologies to screen certain segments of the population, or even the general population, has considerable provenance and tremendous appeal.[1] Screening occurs routinely today within affected families for hereditary forms of cancer, cardiovascular disease and other highly penetrant conditions, affording an opportunity for enhanced surveillance and risk-reducing interventions among those carrying gene variants. Building upon this precedent, it has been proposed that population screening to identify high risk individuals from the general population carrying pathogenic or likely pathogenic variants (PVs) in a limited set of genes for highly penetrant inherited conditions presents a significant opportunity for public health benefit.[2,3]

The genes most often mentioned in the context of population genomic screening include high-risk cancer susceptibility genes (CSGs) such as *BRCA1, BRCA2* and *PALB2* for hereditary breast and ovarian cancer (HBOC)[4] and *MLH1*, *MSH2* and *MSH6* for Lynch syndrome (LS),[5] as well as lipid metabolism genes *LDLR*, *APOB* and *PCSK9* for familial hypercholesterolemia (FH)[6]. For those identified at high risk of these conditions, there is consensus that effective risk management interventions are available to prevent disease and reduce risk.[3,7]

Population based *BRCA* testing in Jewish populations[8] has already provided an evidence-based model for clinical implementation of such screening. Israel recently implemented such a program, and the UK National Health Service Cancer Programme team is launching its

population based *BRCA* programme in early 2023. There is also evidence from modelling studies that population based screening for a limited set of high risk monogenic genes will be cost effective from payer and societal perspectives.[9-12] General population screening for targeted detection of rare PVs in this limited set of high risk genes (monogenic screening) therefore represents a tractable strategy with rational precedent for the medical management of individuals who are so identified from the general population.

Polygenic risk scores (PRS), by contrast, are not yet supported by such an evidence-based rationale for population screening. A PRS calculates the collective influence of many common genetic variants on the risk of a particular disease, typically calculated as a weighted sum of trait-associated alleles.[13] Population based PRS estimates of disease risk have been available to the public through direct-to-consumer (DTC) testing companies since 2007, at times attracting fierce criticism for omitting rare monogenic risk variants.[14] Making clinical decisions based on PRS alone, without undertaking monogenic testing, may provide false reassurance of low genetic risk to some, and be potentially harmful in that respect (for a hypothetical case-study, see Box 1 below). In addition to longstanding products from DTC companies, newer commercial vendors are building PRS products, and sometimes promoting these without mention of how monogenic risk variants would be missed.

PRSs have become more sophisticated and there is enthusiasm for implementing PRS-based population screening initiatives within medicine and public health, often without reference to monogenic risks.[15] A number of such population based studies and biobanking initiatives have preferentially undertaken PRS testing, in the absence of monogenic testing for rare variants. A recent example is the UK "Our Future Health" initiative, proposing to screen 5-million individuals using PRSs for common disease.[16] In a cohort of this size, using DNA sequencing for the detection of rare PVs would identify ~50,000 monogenic PV carriers for HBOC, Lynch syndrome and FH alone. Most of these genetically high-risk individuals would be missed by PRS testing, since the underlying technology most often used for calculating PRSs (genotyping microarrays) would not typically detect rare monogenic variants.

The existing evidence base regarding the clinical utility of using PRSs for population screening is less robust than monogenic screening, where clinical benefits within affected families are well established. Validation of predictive models using PRSs, including defining absolute risk thresholds for effective clinical interventions, is still lacking for most disease states,[17]

especially since PRSs will ideally require integration into a combined disease risk model with other risk-factors. Unlike clinical monogenic testing, reporting of PRSs is not yet standardized for most diseases, nor are there well aligned clinical guidelines for action. Nevertheless, there is good evidence and emerging applicability of PRS for risk stratification in some diseases, the prime example being in risk-adapted screening for breast cancer.

Despite the availability of effective interventions for monogenic conditions such as HBOC, LS and FH, these conditions remain chronically underdiagnosed. Over 95% of those carrying PVs in the *BRCA1/2* genes remain unidentified in the general UK population, despite over 25 years of genetic testing based on clinical presentation or family history.[18] Similar circumstances hold in the US[19] and the rates for testing and detection of LH and FH variants are even lower.[5,6] Unfortunately, an estimated 50-80% of PV-carriers for these conditions do not fulfil current clinical genetic testing criteria.[7,8,18] The system for proactive screening in most countries is plagued by restricted access and underutilisation of testing. Finding unaffected rare PV carriers for highly penetrant monogenic conditions in the general population should be an upfront priority of population genomics.

Box-1: Hypothetical case study: A healthy 30-year-old woman participates in a proactive genetic screening study for breast cancer risk, but the study only offers polygenic-score testing. The participant's polygenic-score result comes back "low" for breast cancer (bottom 20%-25% of the polygenic or even modelled breast-cancer risk distribution). On the basis of this result, the participant perceives her breast cancer risk to be low, and is offered only routine breast cancer screening (e.g. mammography) starting at age 50, in line with the general population recommendations (less frequent breast cancer screening is also being considered for the lowest-risk quartile). However, the polygenic-score was not capable of detecting rare pathogenic variants in known breast cancer susceptibility genes, and therefore failed to detect her rare germline *BRCA1* pathogenic variant (inherited from her father, without manifest family history, meaning she didn't qualify for clinical criteria-based genetic testing). Years later at age 42, still under the impression of low genetic-risk, she is unexpectedly diagnosed with invasive breast cancer (detected at a late-stage) before the commencement of any routine breast cancer screening.

Clinical gene panels based on targeted sequencing of known disease-associated genes are currently offered to patients of all genetic ancestries, and the clinical implications of monogenic testing are more likely to be consistent across ancestries. PRS, by contrast, are subject to well recognized ancestry-specific biases.[13] For population genomic screening to maximise opportunities for prevention and cost effectiveness for health systems in the future, this aspect of population scale testing and implementation must be given careful consideration.

Emerging population screening initiatives

Recent population genomic screening studies have prioritized the detection of highly penetrant, rare PVs in medically actionable genes using targeted sequencing. The recently launched Australian 'DNA Screen' national pilot study is offering preventive DNA screening for 10 medically actionable genes to 10,000 adults aged 18-40 years.[7] Likewise, the recently announced PROTECT-C study is offering DNA screening for nine (HBOC and LS) CSGs to over 5000 women in the UK. In PROTECT-C, PRS will be utilized concurrently to provide personalised breast and ovarian cancer risk prediction for risk-adapted breast cancer screening and breast/ovarian cancer prevention, but will not replace monogenic testing. The increasing popularity of PRS is likely due to low-cost, improved accessibility, potential to generate risk prediction on large populations, and the potential for population stratification for risk-adapted screening and prevention. PRS will certainly be an important screening tool.

However, to maximise the preventive potential and public health impact of population genomic screening, initiatives ought to also prioritize monogenic testing for medically actionable conditions before or concurrently with polygenic testing. It is important to prioritise identification of the most genetically high-risk individuals in the general population and provide them access to risk management and preventive care based on current guidelines. Using PRS alone without monogenic testing, long considered a major limitation of low-cost DTC approaches, will miss the most clinically significant genetic risk information associated with rare high-risk PVs for heritable conditions. This would be a missed opportunity for genomic medicine and prevention. Using PRSs 'alone' is therefore misguided as a population screening strategy to maximise precision prevention, and would be akin to "putting the cart before the horse" at a critical inflection point for population genomics.

References:

1 Khoury, M. J., McCabe, L. L. & McCabe, E. R. Population screening in the age of genomic medicine. *N Engl J Med* **348**, 50-58, doi:10.1056/NEJMra013182 (2003).

2 Bean, L. J. H. *et al.* DNA-based screening and personal health: a points to consider statement for individuals and health-care providers from the American College of Medical Genetics and Genomics (ACMG). *Genet Med* **23**, 979-988, doi:10.1038/s41436-020-01083-9 (2021).

3 Murray, M. F., Evans, J. P. & Khoury, M. J. DNA-Based Population Screening: Potential Suitability and Important Knowledge Gaps. *JAMA* **323**, 307-308, doi:10.1001/jama.2019.18640 (2020).

4 Breast Cancer Association, C. *et al.* Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med* **384**, 428-439, doi:10.1056/NEJMoa1913948 (2021).

5 Win, A. K. *et al.* Prevalence and Penetrance of Major Genes and Polygenes for Colorectal Cancer. *Cancer Epidemiol Biomarkers Prev* **26**, 404-412, doi:10.1158/1055-9965.EPI-16-0693 (2017).

6 Watts, G. F. *et al.* Integrated Guidance for Enhancing the Care of Familial Hypercholesterolaemia in Australia. *Heart Lung Circ* **30**, 324-349, doi:10.1016/j.hlc.2020.09.943 (2021).

7 Lacaze, P. A., Tiller, J., Winship, I. & Group, D. N. A. S. I. Population DNA screening for medically actionable disease risk in adults. *Med J Aust* **216**, 278-280, doi:10.5694/mja2.51454 (2022).

8 Manchanda, R. *et al.* Randomised trial of population-based BRCA testing in Ashkenazi Jews: long-term outcomes. *BJOG* **127**, 364-375, doi:10.1111/1471-0528.15905 (2020).

9 Guzauskas, G. F. *et al.* Cost-effectiveness of Population-Wide Genomic Screening for Hereditary Breast and Ovarian Cancer in the United States. *JAMA Netw Open* **3**, e2022874, doi:10.1001/jamanetworkopen.2020.22874 (2020).

10 Manchanda, R. *et al.* Cost-effectiveness of Population-Based BRCA1, BRCA2, RAD51C, RAD51D, BRIP1, PALB2 Mutation Testing in Unselected General Population Women. *J Natl Cancer Inst* **110**, 714-725, doi:10.1093/jnci/djx265 (2018).

11 Zhang, L. *et al.* Population genomic screening of all young adults in a health-care system: a cost-effectiveness analysis. *Genet Med* **21**, 1958-1968, doi:10.1038/s41436-019-0457-6 (2019).

12 Marquina, C. *et al.* Population genomic screening of young adults for familial hypercholesterolaemia: a cost-effectiveness analysis. *Eur Heart J* **43**, 3243-3254, doi:10.1093/eurheartj/ehab770 (2022).

13 Kullo, I. J. *et al.* Polygenic scores in biomedical research. *Nat Rev Genet* **23**, 524-532, doi:10.1038/s41576-022-00470-z (2022).

14 Times, N. Y. Don't Count on 23andMe to Detect Most Breast Cancer Risks, Study Warns. *https://www.nytimes.com/2019/04/16/health/23andme-brca-gene-testing.html (accessed Dec 2022)* (2019).

15 Polygenic Risk Score Task Force of the International Common Disease, A. Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. *Nat Med* **27**, 1876-1884, doi:10.1038/s41591-021-01549-6 (2021).

16 Our Future Health - Protocol. *https://ourfuturehealth.org.uk/research-programme/* (Accessed October 2022).

17 Ding, Y. *et al.* Large uncertainty in individual polygenic risk score estimation impacts PRS-based risk stratification. *Nat Genet* **54**, 30-39, doi:10.1038/s41588-021-00961-5 (2022).

18 Manchanda, R. *et al.* Current detection rates and time-to-detection of all identifiable BRCA carriers in the Greater London population. *J Med Genet* **55**, 538-545, doi:10.1136/jmedgenet-2017-105195 (2018).

19 Manickam, K. *et al.* Exome Sequencing-Based Screening for BRCA1/2 Expected Pathogenic Variants Among Adult Biobank Participants. *JAMA Netw Open* **1**, e182140, doi:10.1001/jamanetworkopen.2018.2140 (2018).