Check for updates

Realistic expectations are key to realising the benefits of polygenic scores

We must not let enthusiasm around polygenic scores allow us to forget other factors that are bigger, more modifiable, and relevant for everyone, argue **Amit Sud, Rachel Horton, and colleagues**

Amit Sud, [1,2] Rachel H Horton, [3,4,5] Aroon D Hingorani, [6,7,8,9] Ioanna Tzoulaki, [10,11,12] Clare Turnbull, [1,13] Richard S Houlston, [1] Anneke Lucassen [4,5,14]

Key messages

- Polygenic scores will always be limited in their ability to predict disease, as much of a person's disease risk is determined by factors that polygenic scores cannot measure

- If we do not effectively communicate this limitation, we risk overemphasising the role of polygenic scores, which could undermine current effective screening programmes

- The enthusiasm around polygenic scores must not distract from efforts to tackle modifiable risk factors for disease

Polygenic scores look at thousands of genetic variants across a person's genome to estimate their risk of developing a specific disease. Each individual genetic variant has a small effect on a person's disease risk, but by looking at all the variants together, something clinically meaningful might be said about their overall risk of developing a disease. This is in contrast to monogenic variants, such cancer predisposing *BRCA* variants, where a variant in a single gene has a very marked effect on a person's disease risk. Polygenic scores can, in theory, be developed for any disease for which genetics influences risk, but the two areas in which their use has most widely been described are cancer and coronary artery disease. We focus on these in our article.

Enthusiasm surrounds government reports on polygenic scores, with the *Genome UK* report describing them as offering a "step change" in screening for disease.[1] The UK NHS will offer risk information based on polygenic scores to five million people as part of *Our Future Health*, which is set to become the UK's largest health research programme.[2][3] Such information is expected to inform clinical decisions including access to screening.[4]

Amid the hope that polygenic scores will "change the whole paradigm of healthcare,"[5] we should recognise that these scores are limited in their potential to predict disease. If we do not set our expectations accordingly, they could harm rather than help.

Polygenic scores will always be limited in their ability to predict disease

Polygenic scores offer the possibility of assessing a person's genetic risk for multiple diseases simultaneously, at any point in their life course. But they do not consider the effects of environmental or poorly understood non-genetic factors that contribute to most common diseases. Thus, polygenic scores will always remain one of many risk factors and will never reach a point where they can accurately predict who will and will not develop disease.[6]

As with any screening tool, understanding the sensitivity and specificity of polygenic scores is essential to evaluate their clinical utility. A 2022 preprint evaluating polygenic scores in disease prevention indicates that, with specificity set at 95% (meaning that 5% of people who will not develop the disease will have a high polygenic score), the typical sensitivity for a polygenic score is 10-15% (meaning that only 10-15% of people who will develop the disease will have a high polygenic score)[7]—for example, a polygenic score developed to detect women at >17% lifetime risk of breast cancer has a sensitivity of 39% (it will identify 39% of the women who will go on to develop breast cancer, but miss 61% of them) and a specificity of 78% (22% of women who will not go onto develop breast cancer will be classified as having a "high risk score").[7][8] Increasing the sensitivity of a polygenic score reduces the specificity, and vice versa (fig 1).
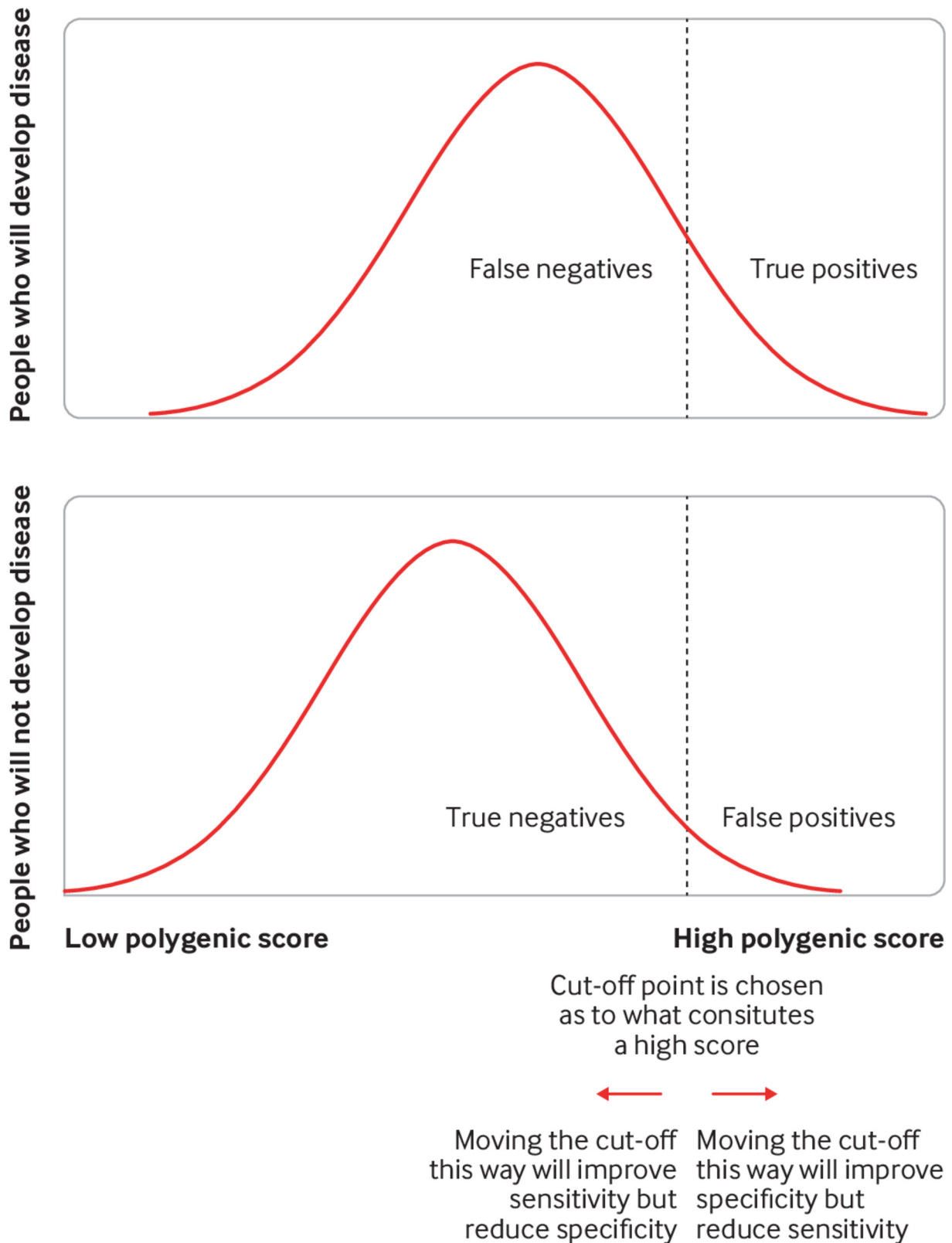
Fig 1 | Choosing thresholds for declaring a "high" polygenic score

It is tempting to imagine that there will be a transformative improvement in the predictive ability of polygenic scores through the discovery of more genetic risk variants. But modelling shows that, even in the theoretical scenario that all common genetic risk variants are identified and used in a polygenic score, they will still be limited in their ability to differentiate between those who will and will not develop disease. Zhang et al calculated the maximum predictive ability achievable with polygenic scores for various

cancers and found that they hit a ceiling. In the case of breast cancer, for example, with a specificity set at 95%, the best achievable sensitivity would be 19% (4% better than current scores). At the extremes of the distribution, that is for a small number of people, they have the scope to be clinically useful. With theoretical best possible polygenic scores, for example, people in the highest 1% of polygenic scores for breast cancer would have a relative risk of four times the average risk; for prostate cancer, five times; for colorectal cancer, 3.5 times.[9]

Balancing the benefits and harms of polygenic scores in clinical practice

Given the intrinsically limited predictive abilities of polygenic scores, a popular approach in research studies has been to integrate such scores into existing prediction models that also consider other risk factors, aiming to give a more holistic overview of disease risk. Using this strategy, polygenic scores stand to slightly improve risk prediction.

Generating an integrated risk tool by adding a polygenic score for coronary artery disease to the pooled cohort equation or QRISK score (clinical models for estimating a person's risk of an atherosclerotic cardiovascular disease event over the next 10 years) improves predictive accuracy by 3-4%.[10 11] This integrated risk tool has been lauded as "substantially enhanc[ing]" coronary artery disease prediction and has been piloted in a collaboration between the NHS and the healthcare company Genomics in 836 general practice patients in the HEART study.[12] If this integrated risk tool were to be used, assuming that everyone exceeding the specified risk cut-off receives and takes a statin (which results in a 20% relative risk reduction), 8713 people would need to undergo integrated risk testing to prevent one additional coronary artery disease event. A comparable effect could be achieved by lowering the current 10 year risk threshold for offering statin treatment in the UK from 10% to around 7.5%. A recent cost effectiveness analysis of polygenic scores in coronary artery disease prevention indicated an incremental cost effectiveness ratio of around $140 000 per quality adjusted life year.[13] This analysis costed polygenic scoring at $70 per person, accounting for technical analytical costs, but not budgeting for other downstream costs (such as the cost of appointments with healthcare professionals for people to discuss their scores). It also assumed 100% adherence to statin treatment.

Many hope that polygenic scores will improve cancer screening programmes through early or more frequent screening for those at higher polygenic risk. It has been proposed, for example, that annual mammography should be offered to women aged 40-50 with polygenic scores that indicate they are at moderate or high risk of breast cancer.[3] This has the potential to detect 1700 more cancers, but at the cost of 5722 false positive results and with 4112 cancers still being missed.[8 14 -16]Figure 2 uses 100 person diagrams to indicate how polygenic scores might perform for cancer detection in people not currently offered cancer screening in the UK.
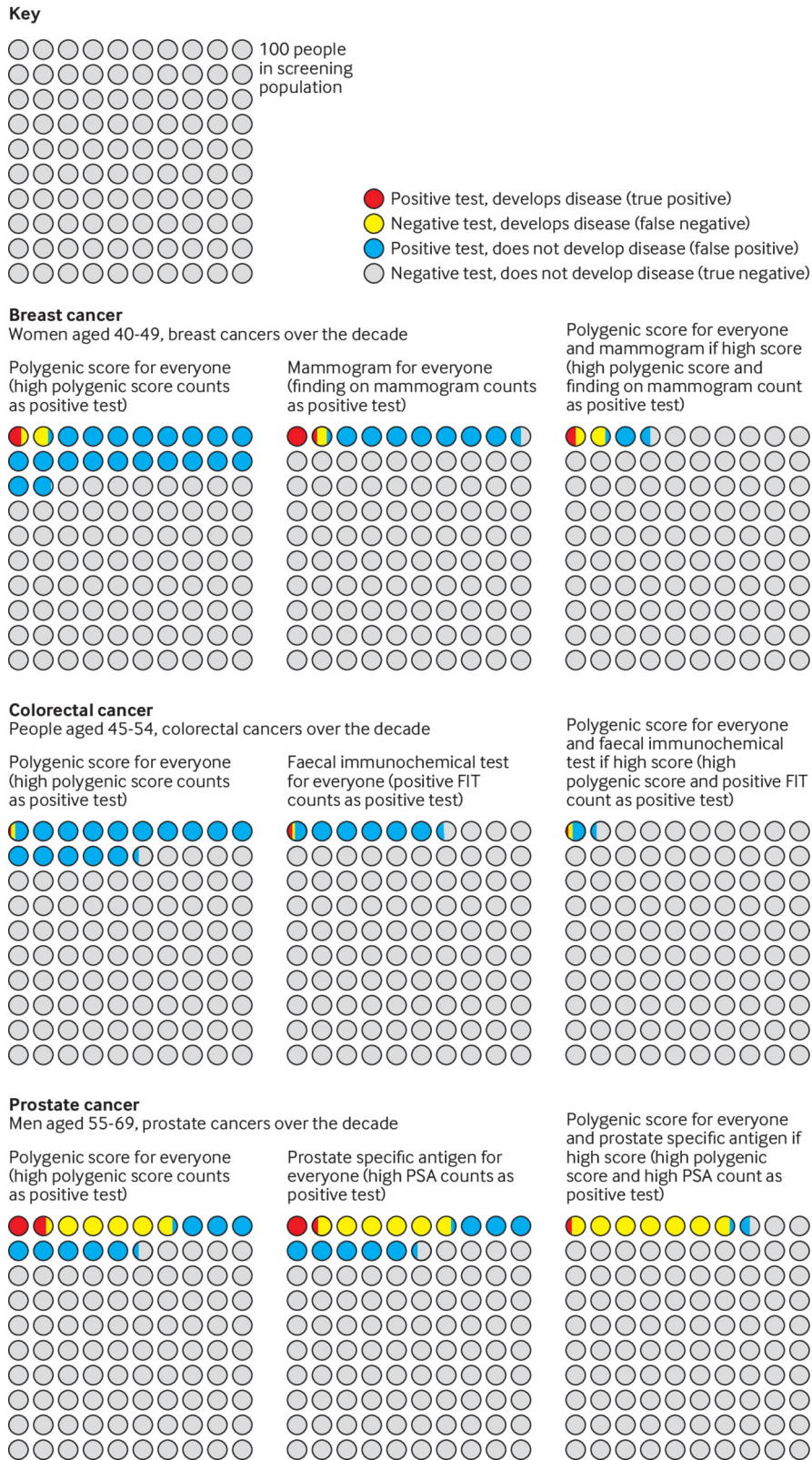
Fig 2 | Polygenic scoring for cancer screening. This figure draws on cancer registration rates from the Office for National Statistics,[14] with data on polygenic score performance from Jia et al 2020,[8] mammography from Pisano et al 2005,[16] faecal immunochemical testing from Lee et al 2014,[17] and prostate specific antigen from Thompson et al 2005.[18] It assumes 100% uptake, no interval cancers, and portability of polygenic scores across all ancestries.

Incorporating polygenic scores into existing screening programmes is not without risk. A study looking population screening for colorectal cancer found that adding a polygenic score to faecal immunochemical testing did not improve the diagnostic accuracy but did add complexity and cost.[19] In situations such as this, where existing screening is already effective and cheap but not well taken up, polygenic scores could worsen outcomes if people take "low risk" scores as a reason to disengage from screening altogether.[20] Although we do not know how likely this would be if polygenic scores were offered at scale, there are indications from a cohort study that provided personalised breast cancer risk estimates: of 127 000 women invited to participate, 46% accepted risk estimation, and attendance at the first screening appointment was slightly reduced among women estimated to be at "below average" risk.[21]

Furthermore, polygenic scores cannot tackle overdiagnosis, a major harm of screening.[22] Most polygenic scores for cancer are based on variants associated with incidence, not mortality, which compromises their usefulness for diseases like prostate cancer, which many men die with rather than from.[23] When existing screening has limitations (such as prostate specific antigen testing for prostate cancer), the limited positive predictive value of a polygenic score adds little diagnostic accuracy and might increase the number of people who will not develop cancer (false positives) but who will, nonetheless, be offered invasive confirmatory investigations. Ambitions to introduce widespread polygenic scoring for prostate cancer would require unprecedented investment in diagnostic imaging, such as magnetic resonance imaging, which is a constrained resource in the UK.[24]

What might people expect from polygenic scores?

Without wider conversations between the public, researchers, healthcare providers, and policy makers around their limitations, polygenic scores are vulnerable to misinterpretation. Unlike other factors that might subtly nudge a person's risk one way or another, people might read more into "genetic" tests.[25 26] In 2019, for example, the then UK health secretary Matt Hancock told the Royal Society that having a polygenic score for prostate cancer "may have saved my life" and he would ensure that he did not "miss any screening appointments in the future" after being told that he had a 15% risk of developing prostate cancer by age 75, neglecting to mention that the background population risk is 13% and that there is currently no screening programme for prostate cancer in the UK.[27]

Absolute risk estimates account both for the relative risk often reported with polygenic scores and the underlying population disease risk—for example, people in the top 5% of polygenic scores for breast cancer have a lifetime risk of 19% (compared with a population risk of 11.8%). For prostate cancer, this is 22.2% (population risk 12.7%), for colorectal cancer 6.9% (population risk 4.6%). For less common conditions, the effect on absolute risk is often more modest. People in the top 5% of polygenic scores for ovarian cancer, for example, have a lifetime risk of 2.1%, compared with a population risk of 1.6%.[7 9]

Risk is notoriously difficult to communicate, and supporting people in understanding the results of their polygenic scores could put major strain on the health service. People with "high risk" scores might want to discuss their results with a clinician, even when their absolute risk is still small, and further costs might accrue depending on decisions around future screening. Conversely, people who do not have "high risk" polygenic scores might be less likely to seek medical attention for concerning symptoms, or their clinicians might be less inclined to investigate. This is concerning, as most people who develop disease will not have a high polygenic score.

Proponents could argue that being informed of high polygenic risk might prompt helpful lifestyle changes, although evidence is lacking. A 2016 meta analysis found that people tend not to change their lifestyle on the basis of genetic results, and a study of nearly 1000 blood donors found that provision of polygenic risk information for coronary heart disease did not affect objectively measured levels of physical activity and other health related behaviours.[28 29] An observational follow-up study of 7342 people in their 50s found that communication of high cardiovascular risk (based on a composite score incorporating both traditional and polygenic risk) was associated with better health behaviour, but, because everyone in the study had polygenic testing, we cannot know whether people offered a traditional risk score alone would have made similar changes.[30]

Some evidence indicates that polygenic scores have the potential to be misunderstood and cause distress. A survey of 227 people accessing polygenic scores online without counselling for a wide variety of diseases (including some without clear preventive or treatment options) found that only 25.6% answered all questions relating to understanding and interpretation of polygenic scores correctly, but 60.8% experienced some degree of negative reaction (upset, anxious, or sad on the "feelings about genomic testing results" scale) after receiving their results. A lower understanding of polygenic scores was associated with a negative psychological reaction.[31] Research exploring how best to communicate these scores and what they do and do not mean will therefore be essential if polygenic scores are to be widely adopted clinically.

A further concern is that, in the future, insurers might seek to use polygenic scores to determine eligibility, given the prospect of widespread polygenic score use increasing information asymmetry between insurers and their customers.[32 33] Since 2001, the UK insurance industry has followed a code setting limitations on the use of health related genomic information in determining eligibility for insurance, but this code is voluntary and designed to cover single gene risk factors such as *BRCA* variants.

Non-genetic risk factors need greater attention

The development and use of polygenic scores is attracting money and attention, but, for most common diseases, unglamorous but well established risk factors like smoking, obesity, and socioeconomic deprivation matter more than a person's genetic background. Childhood postcode, for example, is probably as good a predictor of risk for most common diseases as most polygenic scores.[34]

We need to invest in tackling lifestyle risk factors for disease through, for example, "stop smoking" initiatives and policies that make it easier to afford to make healthy diet and exercise choices. We also need to work to remove barriers to accessing existing effective screening and treatments. Although polygenic scores have the potential to subtly improve our ability to predict who will and will not develop disease, most disease will occur in people who do not have a high score. We argue that enthusiasm around polygenic scores should not detract from efforts to tackle big, modifiable environmental risk factors, which have generalisable and population-wide utility—we need to get the cake tasting better before we work too hard on the icing.

A further point to note is that the variants used in polygenic scores are established in genome-wide association studies, and over 95% of participants in these studies are European (https://gwasdiversitymonitor.com/).[35] These scores will typically have lower predictive accuracy when applied to people with non-European ancestry. Although, as we argue above, polygenic scores at best only slightly

improve each person's risk prediction, the use of polygenic scores is set to benefit people with European ancestry more than anyone else. *Our Future Health* seeks to address this by trying to recruit a diverse range of people to the study, but it is important to remember that at present, where polygenic scores work, they may widen gaps.

In summary, polygenic scores have the potential to slightly improve risk prediction for common diseases, but the benefits of using them will be modest. Wider discussion regarding the limitations of polygenic scores is essential, along with robust research that examines their clinical utility in the real world. This is necessary to ensure that excessive focus on genetic risks does not divert time, money, and attention away from other far greater contributors to disease. Contrary to what many people might expect given usual deterministic discourses around genomics, a high polygenic score will generally have a rather underwhelming impact on absolute risk and both clinicians and the public need to know this.

AUTHOR AFFILIATIONS

1
Division of Genetics and Epidemiology, Institute of Cancer Research, London, UK

2
Haemato-oncology Unit, Royal Marsden Hospital NHS Foundation Trust, Sutton, UK

3
Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, UK

4
Clinical Ethics, Law, and Society Group, Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK

5
Centre for Personalised Medicine, University of Oxford, Oxford, UK

6
Institute of Cardiovascular Science, Faculty of Population Health, University College London, UK

7
University College London British Heart Foundation Research Accelerator Centre, London, UK

8
Health Data Research UK, London, UK

9
University College London, National Institute of Health Research Biomedical Research Centre, London, UK

10
Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, UK

11
National Institute for Health Research, Imperial Biomedical Research Centre, Imperial College London, UK

12
Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece

13
Cancer Genetics Unit, Royal Marsden NHS Foundation Trust, London, UK

14
Data Health and Society, NIHR Biomedical Research Centre, Southampton, UK

1 Department of Health and Social Care, Department for Business, Energy and Industrial Strategy, Office for Life Sciences, Lord Bethell of Romford. Genome UK: the future of healthcare. https://www.gov.uk/government/publications/genome-uk-the-future-of-healthcare, 2020.

2 Our Future Health. Our future health study protocol. 2022. https://ourfuturehealth.org.uk/wp-content/uploads/Our-Future-Health-protocol-for-website-2022.pdf

3 Genomics. Genomics PLC announces genetic risk scoring partnership with Our Future Health to test new personalised disease prevention approaches, 2022. https://www.genomic-splc.com/news/genomics-plc-announces-genetic-risk-scoring-partnership-with-our-future-health-to-test-new-personalised-disease-prevention-approaches/#:~:text=Ge-nomics%20plc%20have%20today%20announced,disease%2C%20stroke%2C%20and%20cancer

4 Davis N. Millions invited to take part in UK scheme to diagnose diseases earlier. *The Guardian* 2022. https://www.theguardian.com/science/2022/oct/24/uk-scheme-to-diagnose-diseases-earlier-cancer-obesity-mental-health

5 Our Future Health. Dr. Raghib Ali: "We can change the whole paradigm of healthcare." 2022. https://ourfuturehealth.org.uk/dr-raghib-ali-we-can-change-the-whole-paradigm-of-health-care/2022

6 Wald NJ, Old R. The illusion of polygenic disease risk prediction. *Genet Med* 2019;21:-7. doi: 10.1038/s41436-018-0418-5 pmid: 30635622

7 Hingorani AD, Gratton J, Finan C, et al. Performance of polygenic risk scores in screening, prediction, and risk stratification. *medRxiv* 2022:2022.02.18.22271049. https://www.medrxiv.org/content/10.1101/2022.02.18.22271049v2

8 Jia G, Lu Y, Wen W, et al. Evaluating the utility of polygenic risk scores in identifying high-risk individuals for eight common cancers. *JNCI Cancer Spectr* 2020;4:pkaa021.

9 Zhang YD, Hurson AN, Zhang H, etalBreast Cancer Association Consortium (BCAC)Barrett's and Esophageal Adenocarcinoma Consortium (BEACON)Colon Cancer Family Registry (CCFR), etal. Assessment of polygenic architecture and risk prediction based on common variants across fourteen cancers. *Nat Commun* 2020;11:. doi: 10.1038/s41467-020-16483-3 pmid: 32620889

10 Elliott J, Bodinier B, Bond TA, etal. Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease. *JAMA* 2020;323:-45. doi: 10.1001/jama.2019.22241 pmid: 32068818

11 Riveros-Mckay F, Weale ME, Moore R, etal. Integrated polygenic tool substantially enhances coronary artery disease prediction. *Circ Genom Precis Med* 2021;14:e003304. doi: 10.1161/CIRCGEN.120.003304 pmid: 33651632

12 Genomics. Genomics PLC announces successful world-first pilot using improved genomic risk assessment in cardiovascular disease prevention in the NHS. 2022. https://www.genomic-splc.com/news/successful-world-first-pilot-using-improved-genomic-risk-assessment-in-cardio-vascular-disease-prevention-in-the-nhs/

13 Kiflen M, Le A, Mao S, et al. Cost-effectiveness of polygenic risk scores to guide statin therapy for cardiovascular disease prevention. *Circ Genom Precis Med* 2022:101161CIRCGEN121003423.

14 National Office for Statistics. Cancer registration statistics, England: 2017. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddis-eases/bulletins/cancerregistrationstatisticsengland/2017

15 National Office for Statistics. Population of the UK by country of birth and nationality. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmi-gration/datasets/populationoftheunitedkingdombycountryofbirthandnationality2021

16 Pisano ED, Gatsonis C, Hendrick E, etalDigital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005;353:-83. doi: 10.1056/NEJMoa052911 pmid: 16169887

17 Lee JK, Liles EG, Bent S, Levin TR, Corley DA. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. *Ann Intern Med* 2014;160:. doi: 10.7326/M13-1484 pmid: 24658694

18 Thompson IM, Ankerst DP, Chi C, etal. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ml or lower. *JAMA* 2005;294:-70. doi: 10.1001/jama.294.1.66 pmid: 15998892

19 Niedermaier T, Guo F, Weigl K, Hoffmeister M, Brenner H. Combined performance of fecal immunochemical tests and a genetic risk score for advanced neoplasia detection. *Cancer Prev Res (Phila)* 2022;15:-52. doi: 10.1158/1940-6207.CAPR-21-0552 pmid: 35679356

20 Duncan DE. *Biobanks are on the cusp of translating big data into new medicine.* NeoLife, 2022. https://neo.life/2022/08/biobanks-are-on-the-cusp-of-translating-big-data-into-new-medicine

21 French DP, McWilliams L, Howell A, Evans DG. Does receiving high or low breast cancer risk estimates produce a reduction in subsequent breast cancer screening attendance? Cohort study. *Breast* 2022;64:-9. doi: 10.1016/j.breast.2022.05.001 pmid: 35569186

22 Vickers AJ, Sud A, Bernstein J, Houlston R. Polygenic risk scores to stratify cancer screening should predict mortality not incidence. *NPJ Precis Oncol* 2022;6:. doi: 10.1038/s41698-022-00280-w pmid: 35637246

23 Klein RJ, Vertosick E, Sjoberg D, etal. Prostate cancer polygenic risk score and prediction of lethal prostate cancer. *NPJ Precis Oncol* 2022;6:. doi: 10.1038/s41698-022-00266-8 pmid: 35396534

24 Royal College of Radiologists. NHS must do more to future-proof its MRI capacity, say imaging experts. https://www.rcr.ac.uk/posts/nhs-must-do-more-future-proof-its-mri-capacity-say-imaging-experts2017

25 Ballard LM, Horton RH, Fenwick A, etal. Genome sequencing in healthcare: understanding the UK general public's views and implications for clinical practice. *Eur J Hum Genet* 2019.pmid: 31527856

26 IPSOS MORI. A public dialogue on genomic medicine: time for a new social contract? 2019. https://www.ipsos.com/sites/default/files/ct/publication/documents/2019-04/public-dialogue-on-genomic-medicine-full-report.pdf

27 BBC. Hancock criticised over DNA test "over reaction." https://www.bbc.co.uk/news/health-47652060

28 Hollands GJ, French DP, Griffin SJ, etal. The impact of communicating genetic risks of disease on risk-reducing health behaviour: systematic review with meta-analysis. *BMJ* 2016;352:. doi: 10.1136/bmj.i1102 pmid: 26979548

29 Silarova B, Sharp S, Usher-Smith JA, etal. Effect of communicating phenotypic and genetic risk of coronary heart disease alongside web-based lifestyle advice: the INFORM Randomised Controlled Trial. *Heart* 2019;105:-9. doi: 10.1136/heartjnl-2018-314211 pmid: 30928969

30 Widén E, Junna N, Ruotsalainen S, etal. How communicating polygenic and clinical risk for atherosclerotic cardiovascular disease impacts health behavior: an observational follow-up study. *Circ Genom Precis Med* 2022;15:e003459. doi: 10.1161/CIRCGEN.121.003459 pmid: 35130028

31 Peck L, Borle K, Folkersen L, Austin J. Why do people seek out polygenic risk scores for complex disorders, and how do they understand and react to results?*Eur J Hum Genet* 2022;30:-7. doi: 10.1038/s41431-021-00929-3 pmid: 34276054

32 Government Office for Science. Genomics beyond health. https://www.gov.uk/government/publications/genomics-beyond-health/genomics-beyond-health-full-report-accessible-webpage#data-security-and-public-attitudes-to-genomics, 2022.

33 Re S. Polygenic risk scores: a better cancer predictor for insurers? 2020 https://www.swiss-re.com/reinsurance/life-and-health/l-h-risk-trends/polygenic-risk-scores-better-cancer-predictor-for-insurers.html

34 Belsky DW, Caspi A, Arseneault L, etal. Genetics and the geography of health, behaviour and attainment. *Nat Hum Behav* 2019;3:-86. doi: 10.1038/s41562-019-0562-1 pmid: 30962612

35 Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 2019;51:-91. doi: 10.1038/s41588-019-0379-x pmid: 30926966