ESC European Society of Cardiology

European Heart Journal (2022) **00**, 1–11
https://doi.org/10.1093/eurheartj/ehac648

STATE OF THE ART REVIEW
Genetics

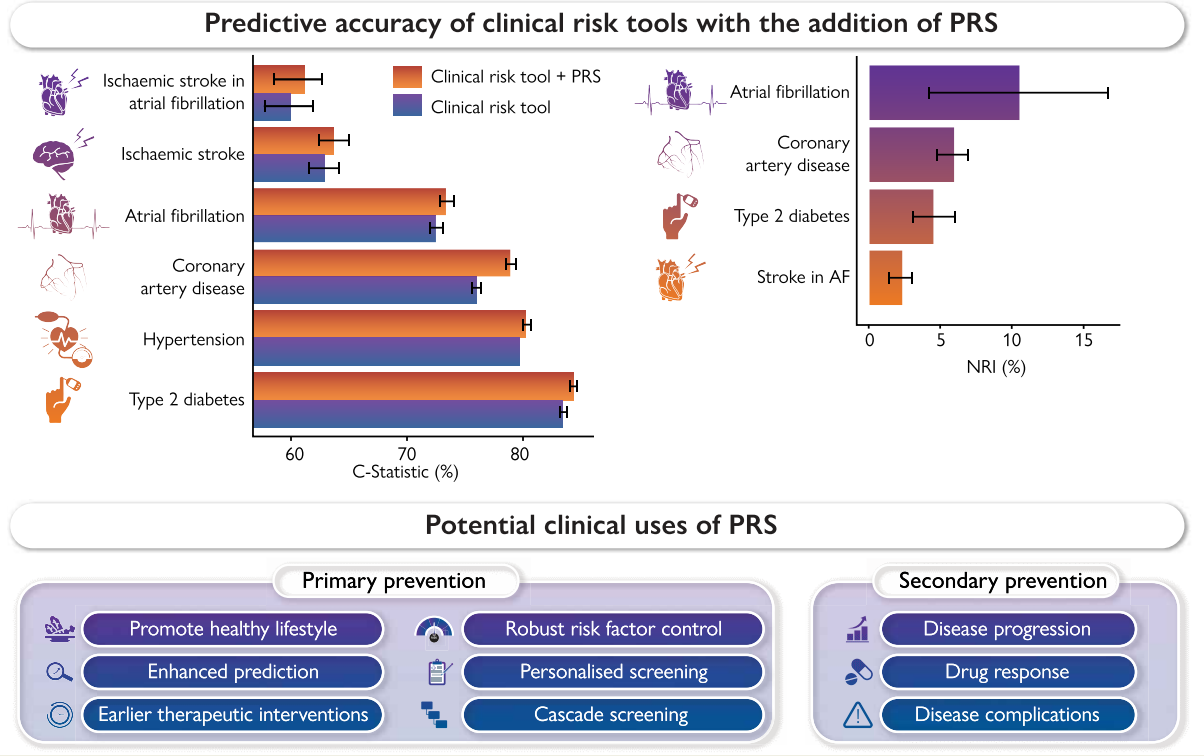
Polygenic risk scores for the prediction of cardiometabolic disease

Jack W. O'Sullivan [1,2], Euan A. Ashley [1,2,3], and Perry M. Elliott [4,5]*

[1]Stanford Center for Inherited Cardiovascular Disease, Stanford University School of Medicine, Stanford, CA, USA; [2]Division of Cardiology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA; [3]Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA; [4]UCL Institute of Cardiovascular Science, Gower Street, London WC1E 6BT, UK; and [5]St. Bartholomew's Hospital, W Smithfield, London EC1A 7BE, UK

Received 8 December 2021; revised 28 August 2022; accepted 27 October 2022

Graphical Abstract



Summary of predictive metrics for polygenic risk scores (PRS), when they are added to existing clinical risk tools. Also discussed are the potential clinical uses for PRS.

Abstract

Cardiometabolic diseases contribute more to global morbidity and mortality than any other group of disorders. Polygenic risk scores (PRSs), the weighted summation of individually small-effect genetic variants, represent an advance in our ability to predict the development and complications of cardiometabolic diseases. This article reviews the evidence supporting the use of PRS in seven common cardiometabolic diseases: coronary artery disease (CAD), stroke, hypertension, heart failure and cardiomyopathies, obesity, atrial fibrillation (AF), and type 2 diabetes mellitus (T2DM). Data suggest that PRS for CAD, AF, and T2DM consistently improves prediction when incorporated into existing clinical risk tools. In other areas such as ischaemic stroke and hypertension, clinical application appears premature but emerging evidence suggests that the study of larger and more diverse populations coupled with more granular phenotyping will propel the translation of PRS into practical clinical prediction tools.

Keywords Polygenic risk scores • Genetics • Precision Medicine

* Corresponding author. Tel: + 44 020 7679 6500, Email: perry.elliott@ucl.ac.uk

Introduction

The heritability of cardiometabolic diseases is well recognized.[1] Over the last ~40 years, the mechanisms underpinning this epidemiological observation have begun to be appreciated.[2] A major breakthrough came in 1985[3] when Lehrman *et al.*[3] identified a deletion in the low-density lipoprotein receptor (LDLR) gene in a patient with homozygous familial hypercholesterolaemia. This was one of the first descriptions of cardiometabolic disease caused by a single genetic variant that disrupts or truncates a protein conferring a substantial risk of developing a disease (i.e. a large effect size). However, most cardiovascular disease is not caused by rare, monogenic variants. Twin studies have confirmed the genetic susceptibility for common cardiometabolic disease traits in the absence of single gene variants.[4] For example, monozygotic twins are at double the risk of developing atrial fibrillation (AF) compared with dizygotic twins [hazard ratio (HR) 2.0, 95% confidence interval (CI) 1.3–3.0].[5]

Genome-wide association studies (GWAS) examine the association of up to millions of single variations in DNA [for example 'A' instead of 'C', known as single nucleotide variants (SNVs)] and their association with a particular disease.[6–8] SNVs are common and individually have a small effect on disease risk, but when combined across the genome give significant insight into an individual's cardiometabolic risk.[9] This summation of risk from SNVs across the genome is the basis of a polygenic risk score (PRS).

This state-of-the-art review describes the evidence pertaining to PRS for the prediction of coronary artery disease (CAD), obesity, type 2 diabetes mellitus (T2DM), hypertension, AF, stroke, and heart failure/cardiomyopathies. It considers the data concerning prediction from PRS alone, the added predictive benefit of PRS when combined with clinical risk factors, the polygenic modulation of monogenic variants, and the clinical utility of disease-specific PRS (*Graphical Abstract*). Throughout the review follows the 2020 NHLBI guidelines on the use and reporting of race, ethnicity, and ancestry.[10]

Methods used to construct polygenic risk scores

Polygenic risk scores can be conceptualized as the summation of SNVs into a single score. These SNVs are found throughout the genome and each individual SNV has a specific risk of developing a disease. A PRS combines these individual risks into a single score. The individual risk from each SNV is generally small, but their summation into a PRS significantly stratifies risk between individuals. Important to this process of PRS construction is the phenomenon of linkage disequilibrium (LD). Linkage disequilibrium is the nonrandom association of alleles (e.g. SNVs) at different genomic loci (positions in the genome). Typically, these SNVs are located in close proximity to each other and are inherited together.[11] SNVs that are inherited together and highly correlated with each other (i.e. in linkage disequilibrium) are often both associated with a disease of interest (e.g. CAD, from a GWAS). However, this association does not reveal which, if any, of these SNVs, are truly associated with a disease of interest. It has been shown that inclusion of all SNVs in LD decreases the predictive performance of PRS[12] and so SNVs that are in LD need to be either removed, or their effect sizes modified to more accurately reflect the true effect size of these variants. Methods include LD clumping, and LD pruning, both of which remove one of a pair of SNVs that are identified to be highly correlated.[13] The SNV removed is either the one with a higher *P*-value (clumping), or at random (pruning).[13] Newer methods do not exclude SNVs but reweight their effect size using difference approaches that utilize a Bayesian framework [LDpred,[14] Bayesian Sparse Linear Mixed Models (BSLMM),[15] AnnoPred,[16] LDpred-funct,[17] PRS-CS,[18] Algorithm by Newcombe *et al.*[19] and PleioPred[20]], penalized regression (Lassosum[21]). Recent evidence suggests that Bayesian approaches lead to more accurate measures.[22]

Polygenic risk scores are often presented as percentiles across a group; for example: the top 25th, 10th, 5th, or 1st percentile of PRS. These PRS percentiles are then compared with the remainder of the population and a risk of disease is calculated.[23] For example, a 2018 study showed that the odds ratio (OR) for individuals with an AF PRS in the highest 1% compared to the remaining 99% was 4.63 (95% CI 3.96–5.39).[23]

Many studies construct risk models using PRS to predict incident or prevalent disease (typically controlling for age, sex and genetic measures of ancestry). The accuracy of these models are often expressed as a C-statistic or area under the (receiver operating) curve (AUC). These parameters are rank-correlation probabilities (0–100) that can be broadly defined as the probability that a randomly selected participant who does have the disease has a higher predicted probability than that of a randomly selected participant who does not have the disease.[24] The higher the C-statistic (or AUC), generally, the better the prediction model.

Cardiometabolic polygenic risk scores

Coronary artery disease

Numerous studies have compared the predictive accuracy of PRS with established clinical risk factors for CAD.[23,25–27] Different approaches have been examined including: (i) the predictive ability of PRS compared to individual conventional clinical risk factors (e.g. PRS vs. hypertension), (ii) the predictive ability of PRS compared to aggregated conventional clinical risk factors [i.e. PRS vs. risk factors summed into a clinical risk tool, such as the American College of Cardiology (ACC)/American Heart Association (AHA) atherosclerotic cardiovascular disease (ASCVD) risk score], and (iii) the predictive ability when PRS are integrated into clinical risk tools (i.e. PRS integrated into the ACC/AHA ASCVD risk score vs. the conventional ACC/AHA ASCVD risk score).

In comparison to individual clinical risk factors, many studies have shown PRS to be of similar or enhanced accuracy in predicting CAD.[25–27] For example, a 2018 UK Biobank (UKBB) study of European participants showed that a PRS had greater predictive accuracy than any clinical risk factor: PRS C-statistic: 0.623 (95% CI 0.615–0.631), compared with a range of ~0.550 to 0.594 for clinical risk factors [smoking, T2DM, self-reported family history of heart disease, body mass index (BMI), hypertension, high cholesterol].[26] Similar results have been replicated in other cohorts of European participants (for example the Malmö Diet and Cancer Study).[28] Trans-ancestral PRS derived from a combination of European and non-European data have been shown to be similarly predictive of CAD, and PRS derived from European cohorts are significant predictors even across non-European populations.[25,29] Although ancestral-specific scores for non-European populations outperform European derived PRS in non-European populations.[30,31]

When compared to combined clinical risk factors, PRS has largely been shown to be less predictive of CAD. In the UKBB, a collection of clinical risk factors summed as a clinical risk tool was found to have a C-statistic of 0.670 (0.663–0.678), compared with 0.623 (0.615–0.631) for PRS alone.[26] Similar results were seen in the Malmö Diet and Cancer Study, C-statistics for clinical risk tools (summed individual risk factors): 0.776 (0.737–0.815) compared with 0.759 (0.724–0.794) for PRS.[28] However, studies in other biobanks (for example: FinnGenn) have shown higher C-statistics for PRS alone vs. clinical risk tools for CAD: 0.832 (0.828–0.836) vs. 0.823 (0.819–0.827), respectively.[9,26]

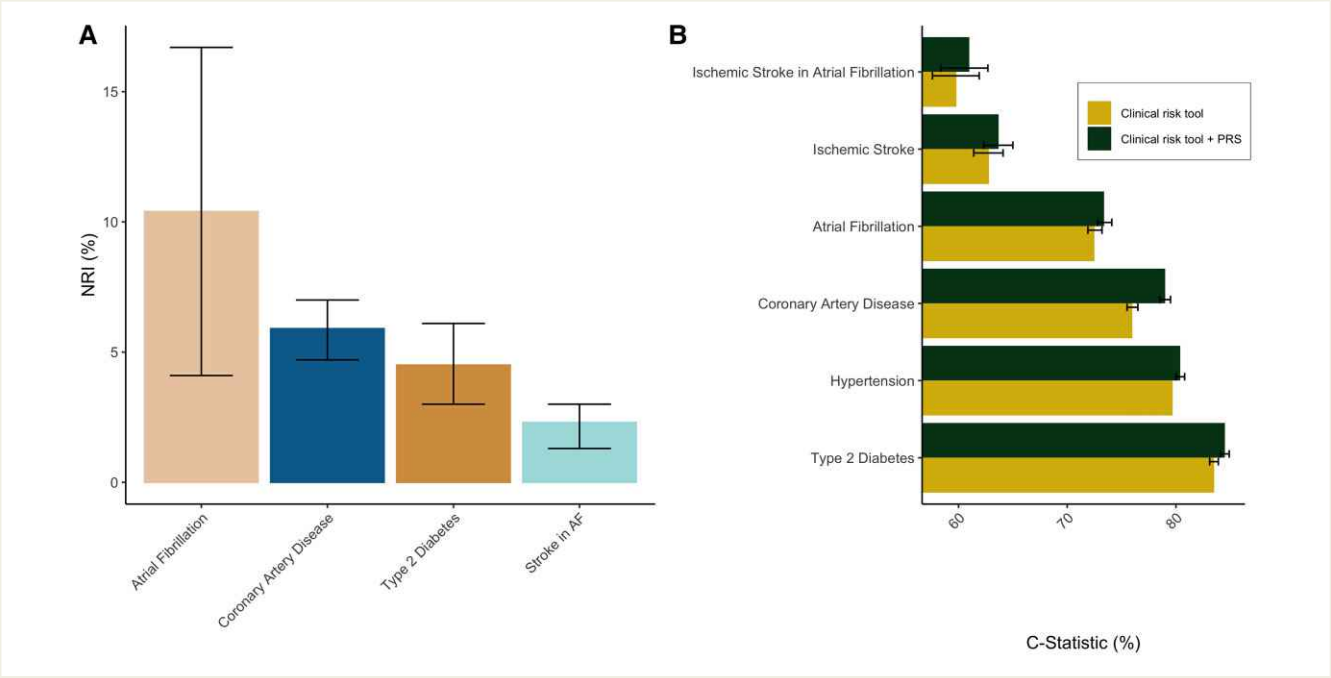


Figure 1 Improvement in prediction with the integration of polygenic risk scores into current, guideline-recommended risk tools. Plot *A* presents Net Reclassification Improvement, Net Reclassification Improvement data were not available for all diseases included in this review. Plot *B* presents C-statistic: both for the guideline-recommended clinical risk tool and for risk tools that integrates polygenic risk score into the respective guideline-recommended risk tool. The guideline-recommended risk tool for each disease is as follows: Atrial Fibrillation: CHARGE-AF using a risk threshold of >5% over 5 years, variables included in CHARGE-AF: age, height, weight, systolic blood pressure, diastolic blood pressure, smoking status, blood-pressure-lowering medication, diabetes, heart failure and history of myocardial infarction; coronary artery disease: AHA/ACC PCE: American Heart Association/American College of Cardiology Pooled Cohort Equation using a 7.5% risk threshold over 10 years, including the following variables: Age, diabetes, sex, race, smoker, total cholesterol, HDL, systolic blood pressure, treatment for hypertension; Type 2 Diabetes was the American Diabetes Association risk score using a 33% risk threshold over 10 years and including the following variables age, sex, body mass index, history of stroke or CHD, parental history of diabetes, systolic blood pressure, diastolic blood pressure, HDL and triglycerides; Ischaemic Stroke in Atrial Fibrillation: CHA_2DS_2-VASc, variables included: Congestive heart failure, hypertension, Age (>65, >75), diabetes, vascular disease, prior stroke and sex; Hypertension: Framingham hypertension risk score, variables included: age, sex, systolic blood pressure, diastolic blood pressure, body mass index, and current smoking, as well as diabetes; Ischaemic Stroke: Combination of family history, blood pressure, body mass index, smoking status, and diabetes. Data was obtained from the following primary sources: atrial fibrillation,[32] coronary artery disease,[27] type 2 diabetes mellitus,[32] ischaemic stroke in atrial fibrillation,[47] hypertension,[64] and ischaemic stroke[46] only C-statistics for coronary artery disease were statistically compared. Confidence intervals for C-statistic for clinical risk tool for hypertension were not reported in primary data.

Across a number of different biobanks, the addition of PRS to clinical risk tools has shown improvement in risk prediction for CAD.[25–27,32,33] The risk conferred by clinical markers and PRS is largely uncorrelated,[27,28] suggesting possibly distinct mechanisms. The most common clinical risk scores used for CAD are the ACC/AHA ASCVD risk score and the UK's QRISK score. For example, in the UKBB, net reclassification index (NRI) using a 7.5% 10-year risk threshold improved with the addition of PRS; NRI: 5.9% (4.7–7.0; *Figure 1*), and this increased to 15.4% (11.6%–19.3%) for younger subgroups.[27] Similar results were appreciated in younger participants (<55 years) in the FinnGenn cohort: NRI: 3.9 (1.6–6.2).[32] A significant improvement in predictive performance was also shown in more diverse cohorts [Atherosclerosis Risk in Communities (ARIC) cohort, the Multi-Ethnic Study of Atherosclerosis (MESA): NRI = 3.0% (1.7–4.3)], although only when cohorts were meta-analyzed together to increase sample size.[25,34] These improvements translate into opportunities to offer preventative therapy (i.e. statins) to individuals not identified via current guidelines. A 2020 analysis showed that an additional 4% of the population may be considered high-risk and offered a statin if PRS is incorporated into risk stratification.[35] This improved reclassification of risk with the inclusion of PRS into conventional clinical risk models is seen across sexes, ages and, increasingly, ancestry groups (*Table 1*), with greatest benefit derived in younger patients before the emergence of clinical risk factors.[23,26,27,32] Further, evidence suggests the interrogation of PRS into the ACC/AHA ASCVD risk score for CAD is overall more cost-effective than the ACC/AHA risk score alone, and results in a lower number of ASCVD events.[40]

There are some data to suggest that a high CAD polygenic risk confers a similar risk as monogenic variants in LDLR, APOB, and PCSK9.[41–44] For example, in the UKBB, familial hypercholesterolaemia (FH) carriers were at an 3.2-fold (1.7–6.0) increased risk of CAD, compared to 2.3-fold (2.1–2.5) for non-carriers in the top 20% PRS.[41] These results appear consistent across numerous ancestry groups.[44] Furthermore, PRS appears to modulate monogenic risk.[41] In the UKBB, FH carriers in the lowest quintile PRS had only a 1.30-fold (0.39–4.32) increased risk.[41] Whereas, FH carriers in highest quintile PRS were at a 12.61-fold (2.96–53.62) increased risk.[41] Similar results showed the protective effect of a low PRS equating to proportional protection to that

Table 1 Net reclassification improvement across ancestral groups for coronary artery disease

	Net reclassification Improvement (95% CI)
Black/African American/Black Caribbean/Black African	2.5% (0.6–4.3) in Weale *et al.*[25]
South Asian (Indian, Bangladeshi or Pakistani)	8.7% (3.1–14.4) in Weale *et al.*[25] 3.9% (0.9–7.0) in Huang *et al.*[36]
Hispanic[a]	7.5% (−1.4–16.5) in Weale *et al.*[25]
European	2.7% (1.1–4.2) in Weale *et al.*[25]
Chinese	3.5% (1.2–6.0) in Lu *et al.*[31]

Net reclassification improvement (NRI) of a risk tool with the incorporation of PRS. Weale *et al.* compared AHA/ACC PCE + addition of PRS with the AHA/ACC PCE alone. This PRS was constructed using data largely from the CARDIoGRAMplusC4D consortium,[6] which consisted exclusively of participants of European Ancestry and was validated in the MESA, ARIC and UK Biobank cohorts. [25] Data from Huang et al (which compared QRISK + addition of PRS compared with QRISK alone[36] obtained from the Genes & Health cohort and validated in the same cohort[36]) and Lu et al (which compared a Chinese ASCVD risk score, with similar covariates to QRISK and the AHA/ACC ASCVD, but used a Chinese ASCVD risk score with a risk threshold of 4.5%)[31] obtained data from a variety of Chinese and European GWAS studies,[6,37–39] and was validated in three cohorts in the China-PAR project: the International Collaborative Study of Cardiovascular Disease in Asia (InterASIA), the China Multi-Center Collaborative Study of Cardiovascular Epidemiology (ChinaMUCA-1998), and Community Intervention of Metabolic Syndrome in China and Chinese Family Health Study (CIMIC).

[a]NRI was not significant in Hispanics although sample size was substantially smaller and the PRS effect sizes itself significant and of comparable size to those seen in individuals of Europeans.[45]

from PCSK9 loss of function (LOF) variants.[43] These results suggest that future risk tools should include both PRS and monogenic risk.

Stroke

There is a consistent signal for predictive benefit of PRS for stroke, although this varies with stroke type.[45–48] Most stroke PRS are derived from the MEGASTROKE consortium, which conducted a GWAS on 67 000 participants who had suffered a stroke (cases) and 454 000 controls.[49] The MEGASTROKE consortium identified SNVs associated with any stroke type and also four stroke subtypes–ischaemic stroke (IS), large artery stroke, cardioembolic stroke, and small vessel stroke– and included participants from across the world (European, East Asian, African, South Asian, mixed Asian, and Latin American).[49]

The initial PRS built from the MEGASTROKE data showed that those in the top third of PRS risk had a 35% increased risk of incident stroke (stroke of any kind) compared with those at low genetic risk (bottom third) [HR 1.35 (1.21–1.50)].[45] This risk appeared to be independent of lifestyle behaviours (smoking, diet, BMI, and physical exercise).[45] Further work in a largely European population in the UK Biobank showed that PRS that are specific to stroke subtypes (such as IS) are more predictive [C-statistic: 0.580 (0.566–0.593)] than all clinical risk factors individually (self-reported family history of stroke, smoking, diabetes, LDL cholesterol), aside from hypertension and systolic blood pressure (SBP) [C = 0.590 (0.577–0.603); C = 0.584 (0.570–0.598), respectively].[46] Notably, the combination of clinical risk factors was found to have a larger C-statistic than the IS PRS individually [C = 0.628 (0.614–0.641)].[46] The addition of PRS to the combination of clinical risk factors is of unclear benefit. While the addition of PRS did lead to the highest C-statistic: C = 0.637 (0.623–0.650), compared with clinical risk factors combined: C = 0.628 (0.614–0.641), this improvement was not statistically significant.[46] Similar results have been shown when PRS are examined within randomized controlled trials (RCTs). A 2021 analysis of the ASPREE trial of older adults (>75 years) showed that PRS alone was a better predictor than most clinical risk factors, however, with the addition of PRS to a clinical tool, the C-statistic was increased non-significantly [AUC improved from 66.6% (62.2–71.1) to 68.5 (64.0–73.0) $P = 0.095$].[50] Lastly, a 2021 analysis of participants of five RCTs (ENGAGE AF-TIMI 48, SOLID-TIMI 52, SAVOR-TIMI 53, PEGASUS-TIMI 54, and FOURIER) showed that participants with a PRS in the highest third were at greater risk of IS even after adjustment for clinical risk factors [HR: 1.27 (1.04–1.53)—results presented for primary prevention cohort].[48]

Studies have also explored the role of PRS in the prediction of IS in patients with AF.[48] The risk of stroke in AF is currently determined using a clinical risk tool (CHA_2DS_2-VASc[51] but despite the heritability of IS (~40%),[52] the CHA_2DS_2-VASc tool does not include any index of genetic susceptibility, such as family history. Evidence from GWAS shows a shared genetic aetiology between AF and IS (e.g. loci at the *PITX2* and *ZFHX3* genes are associated with risk of AF and IS).[49,53] In addition, several AF PRS are predictive of stroke,[54–56] and prophylactic oral anticoagulants have been shown to be effective in high risk patients.[57] A recent study in the UKBB showed that a cardioembolic PRS was predictive of IS in participants with AF, and also showed the addition of PRS to the CHA_2DS_2-VASc tool improved prediction with a NRI of 2.3% (1.3%–3.0%) (using a 4% 10-year risk threshold) (*Figure 1*).[47] Most of this reclassification improvement reflected the downgrading of IS risk [NRI for non-cases: 2.3% (95% CI 0.6%–5.4%)] suggesting participants with low PRS are at lower risk than their CHA_2DS_2-VASc risk score suggests.[47]

Hypertension

GWAS have identified many genomic loci associated with the development of hypertension across different ancestry groups.[58,59] PRS based on these data show a consistent association between a high PRS and the development of hypertension,[60–63] but results regarding the predictive benefit of PRS over established clinical risk factors are mixed.[60–63] More recently, Bayesian methods for PRS construction and validation on larger samples have shown improved performance compared to clinical risk factors.[64,65] For example, a 2021 analysis of 218 754 individuals (55 917 of whom had hypertension) found a HR of 1.42 (1.41–1.43) per 1 standard deviation (SD) increase in PRS for SBP and a HR of 1.41 (1.40–1.42) for diastolic blood pressure (DBP) (defined as 'SBP ≥140 mmHg and DBP ≥90 mmHg').[64] This same study showed an association of PRS with the age of onset of hypertension.[64] Participants in the top 2.5th percentile of PRS developed SBP and DBP hypertension 10.6 and 10.5 years earlier than those in the 20th–80th PRS percentile.[64] The addition of a SBP and DBP PRS to the Framingham hypertension risk score[66] improved the predictive performance of the model modestly [C-statistic 79.7–80.4 (80.0–80.8)][64] and was predictive of cardiovascular events.[64] Encouragingly, similar results are seen in non-European ancestral groups.[65]

Heart failure and cardiomyopathies

Phenotype heterogeneity, changing disease definitions, and a diverse aetiology have posed challenges for GWAS in heart failure cohorts. Studies have identified genomic loci associated with heart failure with

reduced ejection fraction (HFrEF), and preserved ejection fraction (HFpEF),[7] but have been insufficient to generate useful PRS in either scenario.[67] For example, a heart failure PRS in the UKBB showed a HR of only 1.01 per SD ($P = 0.38$),[67] although this improved when PRS was combined with a collective biomarker PRS for heart failure [35 biomarkers including albumin, total cholesterol, LDL cholesterol, sodium in urine, and glycated haemoglobin (HbA1c)]; HR 1.08 per SD ($P < 2 \times 10^{-16}$).[67] This relatively poor performance of generic heart failure PRS is at least in part explained by the non-selective phenotype coding of 'heart failure' in many genetic biobanks. For example, most heart failure participants in the UKBB are classified as 'heart failure unspecified' (ICD10-I509), 'heart failure' (ICD10-I500), or left ventricular failure unspecified' (ICD10—I501). Conversely, PRS studies that have used manual curation have shown more promising results.[68] For example, a dilated cardiomyopathy (DCM) PRS, which used cardiac magnetic resonance imaging to classify DCM based on the body surface area (BSA) indexed for left ventricular ejection function (LVEF), left ventricular end-systolic volume (LVESV), and stroke volume (SV), showed an OR of 1.51 per SD increase in LVESV polygenic score ($P = 8.5 \times 10^{-34}$).[68]

PRS for more phenotypically distinct cardiomyopathies have shown more encouraging results.[69] A 2021 analysis showed that some of the variability in hypertrophic cardiomyopathy (HCM) disease course can be explained by PRS.[69] Participants in the lowest 20th percentile of PRS risk had half the odds of developing HCM as those in the middle 60% PRS [OR = 0.53 (0.45–0.63)], whereas those in the top 20th percentile risk were at greater than a twofold increased odds of HCM [OR = 2.30 (2.02–2.62)].[69] These results were also expressed in PRS-predicted risk of left ventricular hypertrophy: a 1 SD unit increase in PRS 'conferred a 0.71 ± 0.35 mm increase in maximum left ventricular wall thickness ($P = 0.048$) in carriers of *MYBPC3* truncating variants ($n = 232$) and a 0.73 ± 0.36 mm increase ($P = 0.037$) in carriers of *MYH7* missense variants ($n = 186$)'.[69] All of these results speak to the value of meticulously defining phenotypes to allow maximum benefit of heart failure PRS.

Weight, body mass Index, and obesity

Increased BMI and obesity are well recognized risk factors for cardiometabolic disease,[70] and several single gene variants are associated with an increased risk of obesity.[71] While these variants confer a large risk, they are rare and account for only a small proportion of people with obesity (between 1.75% and 4%).[72–74] GWAS have revealed the contribution of common, polygenic variation to obesity risk,[75] and have facilitated the creation of obesity PRS, which have subsequently been shown to accurately identify those at increased susceptibility.[76,77] For example, a 2019 PRS validated across four cohorts (UKBB, Partners Healthcare, Framingham Offspring/CARDIA, and Avon Longitudinal Study of Parents and Children) showed that 43% of participants in the highest 10% of PRS risk were obese (BMI ≥ 30 kg/m^2), compared to just 9.5% in the lowest 10% risk.[76] This was similar for severe obesity (BMI >40 kg/m^2), where 5.6% of participants in the highest 10th percentile of PRS were severely obese compared to 0.2% of the lowest 10th percentile risk; a 25-fold gradient in risk of severe obesity ($P <$ 0.0001). In absolute terms, those in the highest 10% of PRS risk had an average 4.8 kg/m^2 higher BMI than those in the lowest 10% risk (BMI: 30.0 kg/m^2 vs. 25.2 kg/m^2, $P < 0.0001$).[76] A high obesity PRS also corresponded to an increased risk of associated traits. For example, individuals with the highest 10% risk were at a 28% increased risk of CAD, a 72% increased risk for diabetes mellitus, a 38% increased risk for hypertension, a 34% increased risk for congestive heart failure, a 23% increased risk for IS, and a 41% increased risk for venous thromboembolism.[76] People with very high PRS have a similar risk of obesity as those with rare monogenic forms. For example, in one study, carriers of the monogenic variant *MC4R* have, on average, a 4.1 kg/m^2 higher BMI compared to non-carriers. Those in the top 1.6% PRS risk have, on average, a 4.1 kg/m^2 higher BMI compared to the remaining 98.4% of the population. This is particularly noteworthy as only 0.14% of the general population carries pathogenic *MC4R* variants, whereas around 1.5% of the general population has a PRS with a comparable effect size.[76]

Despite the association between obesity PRS and BMI, the ability of PRS to predict BMI beyond clinical risk factors is less clear.[77] A 2020 study of 2517 individuals in the CARDIA cohort showed that PRS explained between 11.9% and 13.6% of variation in BMI, compared with the strongest clinical predictor (52.2% for BMI at baseline).[77] Some have suggested that serial BMI measurements are more predictive than PRS, but the discriminative ability of PRS from an earlier age (before the emergence of an increased BMI) is likely to remain valuable.[76,77] Lastly, it should be noted that healthy lifestyle habits (i.e. regular exercise) can mitigate some of the effects of a high obesity PRS; up to 17% of participants in the top 10% PRS risk group had a normal BMI[76] and high fitness level (measured by treadmill test) has a mitigating effect on the development of a high BMI, even in the presence of a high obesity PRS.[77]

Atrial fibrillation

PRS have been shown to predict AF[23,32,78–84] and, as with many PRS, increasingly sophisticated methods have led to improved performance.[32,78,79] The predictive ability of PRS seems to be comparable to conventional clinical risk factors. In a 2020 study of European participants in the FinnGenn cohort, the CHARGE-AF risk score (comprising age, height, weight, SBP, DBP, smoking status, blood pressure-lowering medication, diabetes, heart failure and history of myocardial infarction) had a C-statistic of 0.725 (0.719–0.732), compared with 0.751 (0.744–0.757) for PRS.[32]

Further evidence has supported the role of PRS for predicting AF in the absence of clinical risk factors. A 2021 study showed those in the top 10th percentile of PRS risk had an OR of 5.70 (2.60 to 13.95) for developing lone AF, compared to the bottom 90th percentile of PRS.[85] Similarly, a 2020 study in the UKBB showed that those in the top 2.5th percentile PRS risk developed AF almost 7 years before those in the 20–80th percentile.[32] The same study showed of those who developed AF before age 60 years, 27.9% were at high PRS risk (>5% 5-year risk of developing AF determined by PRS model only), whereas only 4.9% of these participants with early AF were deemed high risk by the common clinical risk tool, CHARGE-AF.[32] Similarly, even in patients with monogenic AF, polygenic risk appears to modulate this risk.[82]

The inclusion of PRS with a clinical risk model led to an increase in the predictive ability. In European populations discrimination was modestly improved [C-statistic improved from 0.725 (0.719–0.732) to 0.734 (0.728–0.741)], but reclassification was significantly better [NRI using a 5% risk threshold over 5 years: 10% (4.2% to 15.7%)].[32] Similar improvements in prediction were seen when a PRS was added to a clinical risk model across four TIMI RCTs ($n = 36\,662$), with results consistent across participants of European and non-European ancestry.[84] Discrimination was even better appreciated in a Japanese population: AUC improved from 0.72 (0.67–0.74) to 0.84 (0.80–0.86) for clinical risk score and clinical risk score with the inclusion of PRS,

Table 2 Potential clinical utility of cardiometabolic PRS

Primary prevention	Promote a healthy lifestyle	Communicating a patient's risk may promote a healthier diet and increased exercise, although data examining this are mixed.
	Enhanced prediction	For a number of cardiometabolic diseases, the addition of PRS enhances the prediction of classification.
	Earlier therapeutic interventions	Patients at high risk may benefit from early interventions, such as those at high risk of CAD.
	Robust risk factor control	Many cardiometabolic diseases are risk factors for other cardiometabolic diseases, such as hypertension for coronary artery disease. Identification of those at high risk could facilitate more robust risk factor control, e.g. lower blood pressure target for those at high risk of CAD.
	Personalized screening	Those at higher PRS may benefit from earlier routine screening to facilitate earlier intervention, such as regular HbA1c screening. Although long-term data is required.
	Cascade screening	For patients who have a PRS that is similar to that inferred by monogenic risk. This may mean families may benefit from screening, as is recommended for monogenic diseases
Secondary prevention	Drug response	For some cardiometabolic diseases, there is an array of medication options. PRS may indicate which medication or which combination of medicines is most beneficial.
	Disease progression	At diagnosis, the trajectory of disease is largely unknown, PRS could indicate which participants will develop mild or severe disease.
	Disease complications	Prediction of complications from cardiometabolic conditions, such as ischaemic stroke in patients with atrial fibrillation.

respectively.[78] The augmented predictive performance of AF when PRS is added to clinical risk models is encouraging and PRS does appear to be an effective strategy for early detection of early-onset AF and its complications. However, the earlier prediction of AF may have less clinical actionability than other cardiometabolic diseases.

Type 2 diabetes mellitus

Many GWAS have confirmed the polygenic nature of T2DM,[86,87] and PRS have shown promise in disease prediction.[88] Early studies showed an apparent risk gradient between PRS and development of T2DM; a 2018 UKBB study showed that participants within the top 2.5% of PRS were at 3.4-fold increased risk compared to the average (median) PRS, and 9.4-fold increased risk compared with participants with the bottom 2.5% risk.[88]

In comparison to clinical risk factors, studies have shown mixed results regarding the predictive ability of PRS alone. A 2018 study showed comparable AUC (both PRS and the collection of BMI, age and sex had a reported AUC of 66%, no 95%CI reported).[88] Conversely, a 2020 study showed the combination of clinical risk factors (age, sex, BMI, history of stroke or CHD, parental history of diabetes, SBP, DBP, HDL and triglycerides) was more predictive for the development of T2DM than PRS alone [C-statistic: 0.835 (0.831–0.839) compared with 0.763 (0.758–0.767), respectively].[32] This was supported by a 2021 study that showed the combination of clinical risk factors was more predictive of T2DM than PRS [C-statistic: 0.709 (0.696–0.722) for PRS, compared with 0.839 (0.829–0.849) for the combination of clinical risk factors].[89]

Nevertheless, the addition of PRS to clinical risk models has shown promise. The integration of PRS into the American Diabetes Association (ADA) clinical risk model tool showed a NRI (using 33% risk threshold over 10 years) of 4.5% (3.0–6.1), but a modest improvement in C-statistic 0.84 (0.83–0.84) with clinical risk factors only to 0.85 (0.84–0.85) with the inclusion of PRS.[32] Similarly, a 2021 study showed the addition of PRS to a clinical risk score (consisting of sex, age, family history, BMI, SBP, serum glucose levels, serum HDL-cholesterol, and serum triglycerides) improved NRI [22.5% (17.4%–28.0%) for continuous NRI, and 6% (2.0%–10.9%) for categorical NRI].[89]

Although there are no studies that have sequenced patients to examine the polygenic modulation of monogenic T2DM, there is indirect evidence that suggests this occurs,[90] via the varying clinical penetrance of monogenic variants. Similarly, the highest PRS appears to confer a similarly high risk as monogenic variants. In a 2020 study of European participants, those with a PRS in the top 2.5th percentile (compared to participants with a PRS in the middle 20–80th percentile) were at 3.5-fold increased risk of developing T2DM whereas those with a PRS in the bottom 2.5th percentile risk had an ~80% reduction in lifetime T2DM risk.[32]

Clinical utility of polygenic risk scores

A number of clinical uses for PRS have been suggested.[13] In this section, the use cases are divided into primary and secondary prevention to mirror current, clinical risk stratification (*Table 2*). Many of the described examples are supported only by observational data. Furthermore, as clinical implementation is considered, the recalibration of PRS to the target population is essential. Recalibration is the adaptation of risk models to account for differences between the population in which the risk model was derived and the target population (in which the PRS is being deployed). Like all risk models, PRS are derived from research cohorts. These cohorts are likely to vary from clinical cohorts in the healthcare systems where PRS may be used. To ensure PRS models remain accurate they should be recalibrated using the genetic, and clinical risk factor profile of the target population. This has been shown to ensure accuracy of clinical risk tools upon clinical integration.[91]

Primary prevention

For primary prevention, earlier identification of high-risk individuals could facilitate earlier interventions. For most cardiometabolic diseases, these are lifestyle modification (diet and exercise), however for others (CAD and stroke) pharmacological interventions may also be indicated. A protective effect of exercise has been shown for almost all cardiometabolic diseases,[92] but the impact of communicating genetic risk on positive behaviour change is inconsistent.[93,94] Recent evidence suggests the communication of genetic risk via an interactive web tool does promote behaviour change; in a cohort of more than 7000 participants, 42.6% of participants that were identified at high risk of an myocardial infarction or stroke had made at least one positive behaviour change (weight loss, smoking cessation, or online health coaching).[94] Similar 2021 data showed a greater reduction in the cases of T2DM in those with a PRS in the highest third of T2DM PRS who successfully completed lifestyle interventions compared with those in the middle and lowest third of T2DM PRS.[95] For both T2DM and CAD, it appears positive lifestyle changes have protective benefits across all strata of PRS risk, but the greatest protective benefits are seen in those with the highest PRS.[22]

Regarding pharmacological interventions, it appears PRS may have a role in enhanced primary prevention of myocardial infarctions and strokes (ASCVD). Contemporary data for adult patients suggests the addition of state-of-the-art PRS to current primary prevention ASCVD risk tools improves prediction (*Figure 1*).[25,27,32] This will facilitate the identification of more patients that will benefit from medications. A building evidence base of cost-effectiveness studies has suggested that the integration of PRS into ASCVD risk tools leads to overall lower costs to health care systems and society, as well as more precise identification of those at risk and, ultimately a reduced number of ASCVD events.[40,96] Also, acquisition of genetic data is rapidly falling, both the older, but still common, genotyping technique (estimated at $30USD, £25)[40,96] and the newer low-pass whole genome sequencing are decreasing in price even while the technology improves. It is anticipated that the cost-effectiveness of PRS will continue to improve.[40,96]

Beyond early interventions, identification of people at high-risk could facilitate more effective screening. Examples where this might be the case include wearable devices for AF detection, regular cardiac imaging for heart failure, home blood pressure monitoring for hypertension, and targeted HbA1c checks for type 2 diabetes. Family screening may also be beneficial in individuals with a high PRS.

Secondary prevention

The clinical utility for PRS in secondary prevention is more limited, as the presence of disease typically signifies high-risk and leads to the initiation of medications and interventions. However, there may be some utility in predicting disease complications and disease progression. Following a diagnosis, the trajectory of disease can be difficult to determine and PRS could conceivably have a role in identifying individuals who require more intensive therapies; for example, multiple medications or higher doses in hypertension and tailoring of therapy in T2DM where PRS may assist in predicting macro- and microvascular complications.

Next steps

There are many opportunities to further PRS research. Most of the current limitations stem from a lack of diversity within biobanks and enrichment of non-European participants is an urgent priority. There is also a paucity of data for people aged <40 and older than 80. While the clinical benefit in those over 80 will likely be modest, it is expected that predictive benefit will be greatest in the young,[27,32] as PRS is appreciable from birth, long before the emergence of clinical risk factors. Evidence to support this come from the FinnGen biobank where individuals in the top 2.5 percentile PRS developed CAD >4 years before those in the 20–80th percentile PRS [−4.35 (−4.84 to −3.86)], and >6 years before those in the lowest 2.5 percentile PRS.[32] Furthermore, in the UKBB, almost 10% of participants who had at least a 3-fold increased risk of CAD could not be identified by clinical risk factors alone.[23]

Concerns surrounding the portability of PRS across diverse ancestry groups are well-recognized,[97] and increasing the ancestral diversity of participants in studies and biobanks is critical. However, data produced from biobanks will take time. In the meantime, emerging evidence suggests some portability of PRS across ancestral groups[25] (*Table 1*). For example, using data from Atherosclerosis Risk in Communities cohort, the Multi-Ethnic Study of Atherosclerosis, UKBB, and Chinese Biobanks show that the improved prediction with the addition of PRS into the AHA/ACC PCE was seen in individuals identified as 'Europeans', 'black/African American/black Caribbean/black African', 'Chinese' and 'South Asians (Indian, Bangladeshi or Pakistani)' (*Table 1*). It is important to note that all of these studies ensured recalibration of PRS to the target population, typically by comparing the observed vs. expected events. This is essential to ensure any enhanced predictive performance is meaningful improvement, rather than a poorly calibrated model. Despite predictive benefits with the addition of PRS to clinical risk tools across ancestry groups, benefit still remains unequal. [98]

Portability tools and methods should be used to ensure equity of PRS across ancestral groups, but the ultimate solution remains the inclusion of more diverse groups into studies and biobanks. Perhaps the best example of the power of diversity of genetic studies is the 2021 analysis of 1.65 million individuals that showed improved polygenic prediction of lipids with diverse participants.[99] With the 2022 release of All of Us genetic data, of which >50% of participants identify as 'non-white', portability and equality of PRS should improve.[100]

A small number of RCTs have examined PRS (*Figure 2*), but most have been post-hoc analyses. For example, Mega et al examined the effect of statins stratified by genetic risk in an observational study (the Malmo Diet and Cancer Study) and four RCTs (two primary prevention: JUPITER and ASCOT, and two secondary prevention: CARE and PROVE IT-TIMI 22).[101] The meta-analysis of these trials showed greater relative benefit from statins in higher risk genetic groups: relative reductions in CAD across genetic risk quintiles: intermediate quintiles (2–4) [29% (16%–41%)] and highest quintile [48% (29%–63%)] and also greater absolute benefit for the participants in high genetic risk [three-fold decrease in the number needed to treat to prevent one CAD event ($P = 0.01$)].[101] Similar results were seen in analyses of other statin trials, across WOSCOPS, ASCOT, and JUPITER primary prevention trials, where participants in the highest quintile of genetic risk had a greater relative risk reduction (46% compared with 26% for the rest of the cohort, P for heterogeneity = 0.05).[102] Similarly, absolute risk reduction was greater: 3.6% (95% CI 2.0–5.1) vs. 1.3% (95% CI 0.6–1.9).[102]

Newer PRS, with larger numbers of included SNVs, have also been shown to be of value in trials of PCSK9 inhibitors. For instance, the ODYSSEY OUTCOMES investigators showed that a PCSK9 inhibitor (alirocumab) had greater absolute (6.0% vs. 1.5%) and relative risk reduction in participants with high PRS (>90th percentile, HR:

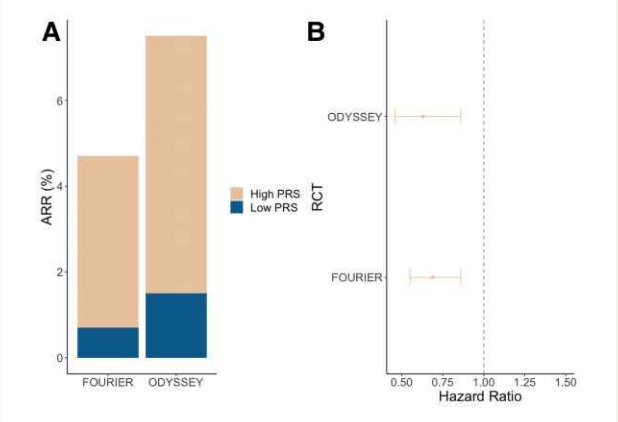


Figure 2 Derived from randomized clinical trials: the effect of PCSK9 inhibitors (*A* and *B*) on major vascular events (coronary heart death, myocardial infarction, coronary revascularization, stroke) stratified by polygenic risk scores. Panel *A* shows the absolute risk reduction by polygenic risk score group in the FOURIER trial [absolute risk reduction: 4% vs 0.7%, $P = 0.004$, hazard ratio: 0.69 (0.55–0.86) for high polygenic risk score][104] and ODYSSEY OUTCOMES trial (absolute risk reduction: 6% vs. 1.5%, interaction $P = 0.04$).[103] High polygenic risk score was defined as top 20% in FOURIER, and top 10% in ODYSSEY OUTCOMES. Panel *B* shows the hazard ratios for PCSK9 inhibitors in the high risk polygenic risk score groups in each trial: FOURIER: hazard ratio: 0.69 (0.55–0.86), and ODYSSEY OUTCOMES: 0.63 (0.46–0.86).

0.63 (0.46–0.86)) compared with those with low PRS [≤90th percentile, HR: 0.87 (0.78–0.98)] (*Figure 2*).[103] Similarly, results from the FOURIER trial showed that evolocumab had a greater absolute (4%) and relative risk reduction [HR: 0.69 (0.55–0.86)] in participants with high genetic risk compared with participants at high clinical risk (*Figure 2*).[104] Similar results have been seen in RCTs for IS.[48,50]

Future RCTs will add valuable data as most current measures of predictive performance are from observational data and expressed as NRI and C-statistic, which have been criticized as misleading, and overly reliant on risk thresholds.[105,106] There is now a consensus regarding best methodological practice for PRS papers,[107] most notably: an independent training and test cohort of participants, methodological approaches to account for LD, calibration assessment, and quantification of accuracy metrics (such as C-statistic), particularly against current clinical practice.

As clinical utility is considered, there is likely to be varying thresholds and metrics that stakeholders will consider sufficient to merit intervention. To some degree, this theme has already emerged in academic literature. For example, a 2020 study of PRS for CAD concluded that a NRI of 4.4% (3.5%–5.3%) was 'modest' and did not warrant clinical implementation.[33] Others have disagreed.[108] Thresholds of improvement that constitutes clinical improvement will likely be disease-specific.

It should be noted that many of the studies reported in this review use the same datasets. For instance, most stroke PRS reported in this state-of-the-art review are constructed from the MEGASTROKE consortium, and many PRS are validated in the UKBB. With the emergence of further biobanks (All of Us, Japanese Biobank, Chinese Biobank, FinnGen, and Our Future Health), this limitation should diminish with time.

Conclusions

Of the diseases reviewed, PRS for CAD, AF, and T2DM appear to have the most consistent evidence for clinical utility. A lack of diversity among participants in GWAS and PRS studies is an important limitation that needs to be addressed through development of ancestrally-diverse biobanks. Improved clinical phenotyping is required to improve stratification of broad phenotypes such as heart failure. Nevertheless, as the science continues to improve, it seems likely that the inclusion of PRS into risk tools will benefit clinical patient care.

Conflict of interest: JOS is an advisor to Google Health AI. E.A. is founder of Personalis, DeepCell, Svexa; advisor to Pacific Biosciences, SequencBio, and Apple and non-executive director of AstraZeneca. PE reports consulting fees from Pfizer, Sanofi Genzyme, Sarepta, Freeline, Bristol Myers Squibb.

Data availability

All data in this manuscript is publicly available via published research.

References

1. Vinkhuyzen AAE, Wray NR, Yang J, Goddard ME, Visscher PM. Estimation and partition of heritability in human populations using whole-genome analysis methods. *Annu Rev Genet* 2013;**47**:75–95. https://doi.org/10.1146/annurev-genet-111212-133258
2. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, *et al.* Genome-wide association studies. *Nat Rev Methods Primers* 2021;**1**:59. https://doi.org/10.1038/s43586-021-00056-9
3. Lehrman MA, Schneider WJ, Südhof TC, Brown MS, Goldstein JL, Russell DW. Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science* 1985;**227**:140–146. https://doi.org/10.1126/science.3155573
4. Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med* 1994;**330**:1041–1046. https://doi.org/10.1056/NEJM199404143301503
5. Christophersen IE, Ravn LS, Budtz-Joergensen E, Skytthe A, Haunsoe S, Svendsen JH, *et al.* Familial aggregation of atrial fibrillation: a study in Danish twins. *Circ Arrhythm Electrophysiol* 2009;**2**:378–383. https://doi.org/10.1161/CIRCEP.108.786665
6. Nikpay M, Goel A, Won H-H, Hall LM, Willenborg C, Kanoni S, *et al.* A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;**47**:1121–1130. https://doi.org/10.1038/ng.3396
7. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, *et al.* Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. *Nat Commun* 2020;**11**:163. https://doi.org/10.1038/s41467-019-13690-5
8. Dewey FE, Grove ME, Pan C, Goldstein BA, Bernstein JA, Chaib H, *et al.* Clinical interpretation and implications of whole-genome sequencing. *JAMA* 2014;**311**:1035–1045. https://doi.org/10.1001/jama.2014.1717
9. O'Sullivan JW, Raghavan S, Marquez-Luna C, Luzum JA, Damrauer SM, Ashley EA, *et al.* Polygenic risk scores for cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* 2022;**146**:e93–e118.
10. NHLBI Trans-Omics for Precision Medicine. Recommendations on the use and reporting of race, ethnicity, and ancestry in the NHLBI Trans-Omics for Precision Medicine (TOPMed) program. https://www.nhlbiwgs.org/guidelines-use-and-reporting-race-ethnicity-and-ancestry-topmed (11 March 2021)
11. Slatkin M. Linkage disequilibrium–understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* 2008;**9**:477–485. https://doi.org/10.1038/nrg2361
12. Vilhjálmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, *et al.* Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet* 2015;**97**:576–592. https://doi.org/10.1016/j.ajhg.2015.09.001
13. Lambert SA, Abraham G, Inouye M. Towards clinical utility of polygenic risk scores. *Hum Mol Genet* 2019;**28**:R133–R142. https://doi.org/10.1093/hmg/ddz187
14. Privé F, Arbel J, Vilhjálmsson BJ. LDpred2: better, faster, stronger. *Bioinformatics* 2021;**36**:5424–5431. https://doi.org/10.1093/bioinformatics/btaa1029
15. Zhou X, Carbonetto P, Stephens M. Polygenic modeling with Bayesian sparse linear mixed models. *PLoS Genet* 2013;**9**:e1003264. https://doi.org/10.1371/journal.pgen.1003264
16. Hu Y, Lu Q, Powles R, Yao X, Yang C, Fang F, *et al.* Leveraging functional annotations in genetic risk prediction for human complex diseases. *PLoS Comput Biol* 2017;**13**:e1005589. https://doi.org/10.1371/journal.pcbi.1005589

17. Márquez-Luna C, Gazal S, Loh P-R, Kim SS, Furlotte N, Auton A, *et al.* Incorporating functional priors improves polygenic prediction accuracy in UK biobank and 23andMe data sets. *Nat Commun* 2021;**12**:6052. https://doi.org/10.1038/s41467-021-25171-9
18. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun* 2019;**10**:1776. https://doi.org/10.1038/s41467-019-09718-5
19. Newcombe PJ, Nelson CP, Samani NJ, Dudbridge F. A flexible and parallelizable approach to genome-wide polygenic risk scores. *Genet Epidemiol* 2019;**43**:730–741. https://doi.org/10.1002/gepi.22245
20. Hu Y, Lu Q, Liu W, Zhang Y, Li M, Zhao H. Joint modeling of genetically correlated diseases and functional annotations increases accuracy of polygenic risk prediction. *PLoS Genet* 2017;**13**:e1006836. https://doi.org/10.1371/journal.pgen.1006836
21. Mak TSH, Porsch RM, Choi SW, Zhou X, Sham PC. Polygenic scores via penalized regression on summary statistics. *Genet Epidemiol* 2017;**41**:469–480. https://doi.org/10.1002/gepi.22050
22. Ye Y, Chen X, Han J, Jiang W, Natarajan P, Zhao H. Interactions between enhanced polygenic risk scores and lifestyle for cardiovascular disease, diabetes, and lipid levels. *Circ Genom Precis Med* 2021;**14**:e003128.
23. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;**50**:1219–1224. https://doi.org/10.1038/s41588-018-0183-z
24. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 2011;**30**:1105–1117. https://doi.org/10.1002/sim.4154
25. Weale ME, Riveros-Mckay F, Selzam S, Seth P, Moore R, Tarran WA, *et al.* Validation of an integrated risk tool, including polygenic risk score, for atherosclerotic cardiovascular disease in multiple ethnicities and ancestries. *Am J Cardiol* 2021;**148**:157–164. https://doi.org/10.1016/j.amjcard.2021.02.032
26. Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, *et al.* Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *J Am Coll Cardiol* 2018;**72**:1883–1893. https://doi.org/10.1016/j.jacc.2018.07.079
27. Riveros-Mckay F, Weale ME, Moore R, Selzam S, Krapohl E, Michael SR, *et al.* An integrated polygenic tool substantially enhances coronary artery disease prediction. *Circ Genom Precis Med* 2021;**14**:e003304. https://doi.org/10.1161/CIRCGEN.120.003304
28. Hindy G, Aragam KG, Ng K, Chaffin M, Lotta LA, Baras A, *et al.* Genome-Wide polygenic score, clinical risk factors, and long-term trajectories of coronary artery disease. *Arterioscler Thromb Vasc Biol* 2020;**40**:2738–2746. https://doi.org/10.1161/ATVBAHA.120.314856
29. Dikilitas O, Schaid DJ, Kosel ML, Carroll RJ, Chute CG, Denny JA, *et al.* Predictive utility of polygenic risk scores for coronary heart disease in three Major racial and ethnic groups. *Am J Hum Genet* 2020;**106**:707–716. https://doi.org/10.1016/j.ajhg.2020.04.002
30. Kamiza AB, Toure SM, Vujkovic M, Machipisa T, Soremekun OS, Kintu C, *et al.* Transferability of genetic risk scores in African populations. *Nat Med* 2022;**28**:1163–1166. https://doi.org/10.1038/s41591-022-01835-x
31. Lu X, Liu Z, Cui Q, Liu F, Li J, Niu X, *et al.* A polygenic risk score improves risk stratification of coronary artery disease: a large-scale prospective Chinese cohort study. *Eur Heart J* 2022;**43**:1702–1711. https://doi.org/10.1093/eurheartj/ehac093
32. Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, *et al.* Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med* 2020;**26**:549–557. https://doi.org/10.1038/s41591-020-0800-0
33. Elliott J, Bodinier B, Bond TA, Chadeau-Hyam M, Evangelou E, Moons KGM, *et al.* Predictive accuracy of a polygenic risk score–enhanced prediction model vs a clinical risk score for coronary artery disease. *JAMA* 2020;**323**:636–645. https://doi.org/10.1001/jama.2019.22241
34. Mosley JD, Gupta DK, Tan J, Yao J, Wells QS, Shaffer CM, *et al.* Predictive accuracy of a polygenic risk score compared with a clinical risk score for incident coronary heart disease. *JAMA* 2020;**323**:627–635. https://doi.org/10.1001/jama.2019.21782
35. Aragam KG, Dobbyn A, Judy R, Chaffin M, Chaudhary K, Hindy G, *et al.* Limitations of contemporary guidelines for managing patients at high genetic risk of coronary artery disease. *J Am Coll Cardiol* 2020;**75**:2769–2780. https://doi.org/10.1016/j.jacc.2020.04.027
36. Huang QQ, Sallah N, Dunca D, Trivedi B, Hunt KA, Hodgson S, *et al.* Transferability of genetic loci and polygenic scores for cardiometabolic traits in British Pakistani and Bangladeshi individuals. *Nat Commun* 2022;**13**:4664. https://doi.org/10.1038/s41467-022-32095-5
37. Lu X, Wang L, Chen S, He L, Yang X, Shi Y, *et al.* Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. *Nat Genet* 2012;**44**:890–894. https://doi.org/10.1038/ng.2337
38. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, *et al.* Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet* 2017;**49**:1385–1391. https://doi.org/10.1038/ng.3913
39. Howson JMM, Zhao W, Barnes DR, Ho W-K, Young R, Paul DS, *et al.* Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. *Nat Genet* 2017;**49**:1113–1119. https://doi.org/10.1038/ng.3874
40. Mujwara D, Henno G, Vernon ST, Peng S, Di Domenico P, Schroeder B, *et al.* Integrating a polygenic risk score for coronary artery disease as a risk-enhancing factor in the pooled cohort equation: a cost-effectiveness analysis study. *J Am Heart Assoc* 2022;**11**:e025236. https://doi.org/10.1161/JAHA.121.025236
41. Fahed AC, Wang M, Homburger JR, Patel AP, Bick AG, Neben CL, *et al.* Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. *Nat Commun* 2020;**11**:3635. https://doi.org/10.1038/s41467-020-17374-3
42. Trinder M, Francis GA, Brunham LR. Association of monogenic vs polygenic hypercholesterolemia with risk of atherosclerotic cardiovascular disease. *JAMA Cardiol* 2020;**5**:390–399. https://doi.org/10.1001/jamacardio.2019.5954
43. Oetjens MT, Kelly MA, Sturm AC, Martin CL, Ledbetter DH. Quantifying the polygenic contribution to variable expressivity in eleven rare genetic disorders. *Nat Commun* 2019;**10**:4897. https://doi.org/10.1038/s41467-019-12869-0
44. Khera AV, Chaffin M, Zekavat SM, Collins RL, Roselli C, Natarajan P, *et al.* Whole-Genome sequencing to characterize monogenic and polygenic contributions in patients hospitalized with early-onset myocardial infarction. *Circulation* 2019;**139**:1593–1602. https://doi.org/10.1161/CIRCULATIONAHA.118.035658
45. Rutten-Jacobs LC, Larsson SC, Malik R, Rannikmäe K, MEGASTROKE consortium. International stroke genetics consortium, et al. Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK biobank participants. *BMJ* 2018;**363**:k4168. https://doi.org/10.1136/bmj.k4168
46. Abraham G, Malik R, Yonova-Doing E, Salim A, Wang T, Danesh J, *et al.* Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat Commun* 2019;**10**:5819. https://doi.org/10.1038/s41467-019-13848-1
47. O'Sullivan JW, Shcherbina A, Justesen JM, Turakhia M, Perez M, Wand H, *et al.* Combining clinical and polygenic risk improves stroke prediction among individuals with atrial fibrillation. *Circ Genom Precis Med* 2021;**14**:e003168. https://doi.org/10.1161/CIRCGEN.120.003168
48. Marston NA, Patel PN, Kamanu FK, Nordio F, Melloni GM, Roselli C, *et al.* Clinical application of a novel genetic risk score for ischemic stroke in patients with cardiometabolic disease. *Circulation* 2021;**143**:470–478. https://doi.org/10.1161/CIRCULATIONAHA.120.051927
49. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* 2018;**50**:524–537. https://doi.org/10.1038/s41588-018-0058-3
50. Neumann JT, Riaz M, Bakshi A, Polekhina G, Thao LTP, Nelson MR, *et al.* Predictive performance of a polygenic risk score for incident ischemic stroke in a healthy older population. *Stroke* 2021;**52**:2882–2891. https://doi.org/10.1161/STROKEAHA.120.033670
51. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, *et al.* 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines and the heart rhythm society. *J Am Coll Cardiol* 2019;**74**:104–132. https://doi.org/10.1016/j.jacc.2019.01.011
52. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NLM, Jackson C, *et al.* Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke* 2012;**43**:3161–3167. https://doi.org/10.1161/STROKEAHA.112.665760
53. Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Bragi Walters G, Thorgeirsson G, *et al.* A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet* 2009:876–878. https://doi.org/10.1038/ng.417
54. Tada H, Shiffman D, Smith JG, Sjögren M, Lubitz SA, Ellinor PT, *et al.* Twelve–single nucleotide polymorphism genetic risk score identifies individuals at increased risk for future atrial fibrillation and stroke. *Stroke* 2014;**45**:2856–2862. https://doi.org/10.1161/STROKEAHA.114.006072
55. Lubitz SA, Yin X, Lin HJ, Kolek M, Gustav Smith J, Trompet S, *et al.* Genetic risk prediction of atrial fibrillation. *Circulation* 2017;**135**:1311–1320. https://doi.org/10.1161/CIRCULATIONAHA.116.024143
56. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, *et al.* Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat Genet* 2018;**50**:1234–1239. https://doi.org/10.1038/s41588-018-0171-3
57. López-López JA, Sterne JAC, Thom HHZ, Higgins JPT, Hingorani AD, Okoli GN, *et al.* Oral anticoagulants for prevention of stroke in atrial fibrillation: systematic review, network meta-analysis, and cost effectiveness analysis. *BMJ* 2017;**359**:j5058. https://doi.org/10.1136/bmj.j5058
58. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, *et al.* Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet* 2017;**49**:403–415. https://doi.org/10.1038/ng.3768
59. International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, *et al.* Genetic variants in novel

pathways influence blood pressure and cardiovascular disease risk. *Nature* 2011;**478**:103–109. https://doi.org/10.1038/nature10405
60. Niiranen TJ, Havulinna AS, Langén VL, Salomaa V, Jula AM. Prediction of blood pressure and blood pressure change with a genetic risk score. *J Clin Hypertens (Greenwich)* 2016;**18**:181–186. https://doi.org/10.1111/jch.12702
61. Fava C, Sjögren M, Montagnana M, Danese E, Almgren P, Engström G, *et al.* Prediction of blood pressure changes over time and incidence of hypertension by a genetic risk score in Swedes. *Hypertension* 2013;**61**:319–326. https://doi.org/10.1161/HYPERTENSIONAHA.112.202655
62. Juhola J, Oikonen M, Magnussen CG, Mikkilä V, Siitonen N, Jokinen E, *et al.* Childhood physical, environmental, and genetic predictors of adult hypertension: the cardiovascular risk in young Finns study. *Circulation* 2012;**126**:402–409. https://doi.org/10.1161/CIRCULATIONAHA.111.085977
63. Havulinna AS, Kettunen J, Ukkola O, Osmond C, Eriksson JG, Kesäniemi YA, *et al.* A blood pressure genetic risk score is a significant predictor of incident cardiovascular events in 32,669 individuals. *Hypertension* 2013;**61**:987–994. https://doi.org/10.1161/HYPERTENSIONAHA.111.00649
64. Vaura F, Kauko A, Suvila K, Havulinna AS, Mars N, Salomaa V, *et al.* Polygenic risk scores predict hypertension onset and cardiovascular risk. *Hypertension* 2021;**77**:1119–1127. https://doi.org/10.1161/HYPERTENSIONAHA.120.16471
65. Lim N-K, Lee J-Y, Lee J-Y, Park H-Y, Cho M-C. The role of genetic risk score in predicting the risk of hypertension in the Korean population: Korean genome and epidemiology study. *PLoS One* 2015;**10**:e0131603.
66. Parikh NI, Pencina MJ, Wang TJ, Benjamin EJ, Lanier KJ, Levy D, *et al.* A risk score for predicting near-term incidence of hypertension: the Framingham heart study. *Ann Intern Med* 2008;**148**:102–110. https://doi.org/10.7326/0003-4819-148-2-200801150-00005
67. Sinnott-Armstrong N, Tanigawa Y, Amar D, Mars N, Benner C, Aguirre M, *et al.* Genetics of 35 blood and urine biomarkers in the UK biobank. *Nat Genet* 2021;**53**:185–194. https://doi.org/10.1038/s41588-020-00757-z
68. Pirruccello JP, Bick A, Wang M, Chaffin M, Friedman S, Yao J, *et al.* Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nat Commun* 2020;**11**:2254. https://doi.org/10.1038/s41467-020-15823-7
69. Harper AR, Goel A, Grace C, Thomson KL, Petersen SE, Xu X, *et al.* Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. *Nat Genet* 2021;**53**:135–142. https://doi.org/10.1038/s41588-020-00764-0
70. Buscot M-J, Thomson RJ, Juonala M, Sabin MA, Burgner DP, Lehtimäki T, *et al.* Distinct child-to-adult body mass index trajectories are associated with different levels of adult cardiometabolic risk. *Eur Heart J* 2018;**39**:2263–2270. https://doi.org/10.1093/eurheartj/ehy161
71. Barsh GS, Farooqi IS, O'Rahilly S. Genetics of body-weight regulation. *Nature* 2000;**404**:644–651. https://doi.org/10.1038/35007519
72. Larsen LH, Echwald SM, Sørensen TIA, Andersen T, Wulff BS, Pedersen O. Prevalence of mutations and functional analyses of melanocortin 4 receptor variants identified among 750 men with Juvenile-onset obesity. *J Clin Endocrinol Metab* 2005;**90**:219–224. https://doi.org/10.1210/jc.2004-0497
73. Stutzmann F, Tan K, Vatin V, Dina C, Jouret B, Tichet J, *et al.* Prevalence of melanocortin-4 receptor deficiency in Europeans and their age-dependent penetrance in multigenerational pedigrees. *Diabetes* 2008;**57**:2511–2518. https://doi.org/10.2337/db08-0153
74. Vaisse C, Clement K, Durand E, Hercberg S, Guy-Grand B, Froguel P. Melanocortin-4 receptor mutations are a frequent and heterogeneous cause of morbid obesity. *J Clin Invest* 2000;**106**:253–262. https://doi.org/10.1172/JCI9238
75. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;**518**:197–206. https://doi.org/10.1038/nature14177
76. Khera AV, Chaffin M, Wade KH, Zahid S, Brancale J, Xia R, *et al.* Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell* 2019;**177**:587–596.e9. https://doi.org/10.1016/j.cell.2019.03.028
77. Murthy VL, Xia R, Baldridge AS, Carnethon MR, Sidney S, Bouchard C, *et al.* Polygenic risk, fitness, and obesity in the coronary artery risk development in young adults (CARDIA) study. *JAMA Cardiol* 2020;**5**:40–48.
78. Okubo Y, Nakano Y, Ochi H, Onohara Y, Tokuyama T, Motoda C, *et al.* Predicting atrial fibrillation using a combination of genetic risk score and clinical risk factors. *Heart Rhythm* 2020;**17**:699–705. https://doi.org/10.1016/j.hrthm.2020.01.006
79. Kloosterman M, Santema BT, Roselli C, Nelson CP, Koekemoer A, Romaine SPR, *et al.* Genetic risk and atrial fibrillation in patients with heart failure. *Eur J Heart Fail* 2020;**22**:519–527. https://doi.org/10.1002/ejhf.1735
80. Homburger JR, Neben CL, Mishne G, Zhou AY, Kathiresan S, Khera AV. Low coverage whole genome sequencing enables accurate assessment of common variants and calculation of genome-wide polygenic scores. *Genome Med* 2019;**11**:74. https://doi.org/10.1186/s13073-019-0682-2
81. Weng L-C, Preis SR, Hulme OL, Larson MG, Choi SH, Wang B, *et al.* Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. *Circulation* 2018;**137**:1027–1038. https://doi.org/10.1161/CIRCULATIONAHA.117.031431
82. Choi SH, Jurgens SJ, Weng L-C, Pirruccello JP, Roselli C, Chaffin M, *et al.* Monogenic and polygenic contributions to atrial fibrillation risk. *Circ Res* 2020;**126**:200–209. https://doi.org/10.1161/CIRCRESAHA.119.315686
83. Shoemaker MB, Husser D, Roselli C, Al Jazairi M, Chrispin J, Kühne M, *et al.* Genetic susceptibility for atrial fibrillation in patients undergoing atrial fibrillation ablation. *Circ Arrhythm Electrophysiol* 2020;**13**:e007676. https://doi.org/10.1161/CIRCEP.119.007676
84. Marston NA, Garfinkel AC, Kamanu FK, Melloni GM, Roselli C, Jarolim P, *et al.* A polygenic risk score predicts atrial fibrillation in cardiovascular disease. *Eur Heart J* 2022.
85. Lazarte J, Dron JS, McIntyre AD, Skanes AC, Gula LJ, Tang AS, *et al.* Role of common genetic variation in lone atrial fibrillation. *Circ Genom Precis Med* 2021;**14**:e003179. https://doi.org/10.1161/CIRCGEN.120.003179
86. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, *et al.* The genetic architecture of type 2 diabetes. *Nature* 2016;**536**:41–47. https://doi.org/10.1038/nature18642
87. Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, *et al.* Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat Genet* 2020;**52**:680–691. https://doi.org/10.1038/s41588-020-0637-y
88. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, *et al.* Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;**50**:1505–1513. https://doi.org/10.1038/s41588-018-0241-6
89. He Y, Lakhani CM, Rasooly D, Manrai AK, Tzoulaki I, Patel CJ. Comparisons of polyexposure, polygenic, and clinical risk scores in risk prediction of type 2 diabetes. *Diabetes Care* 2021;**44**:935–943. https://doi.org/10.2337/dc20-2049
90. Hani EH, Boutin P, Durand E, Inoue H, Permutt MA, Velho G, *et al.* Missense mutations in the pancreatic islet beta cell inwardly rectifying K+ channel gene (KIR6.2/BIR): a meta-analysis suggests a role in the polygenic basis of type II diabetes mellitus in Caucasians. *Diabetologia* 1998;**41**:1511–1515. https://doi.org/10.1007/s001250051098
91. Pennells L, Kaptoge S, Wood A, Sweeting M, Zhao X, White I, *et al.* Equalization of four cardiovascular risk algorithms after systematic recalibration: individual-participant meta-analysis of 86 prospective studies. *Eur Heart J* 2019;**40**:621–631. https://doi.org/10.1093/eurheartj/ehy653
92. Khurshid S, Weng L-C, Nauffal V, Pirruccello JP, Venn RA, Al-Alusi MA, *et al.* Accelerometer-derived and genetically inferred physical activity and human disease. *medRxiv* August 8, 2021. https://doi.org/10.1101/2021.08.05.21261586.
93. Hollands GJ, French DP, Griffin SJ, Prevost AT, Sutton S, King S, *et al.* The impact of communicating genetic risks of disease on risk-reducing health behaviour: systematic review with meta-analysis. *BMJ* 2016;**352**:i1102. https://doi.org/10.1136/bmj.i1102
94. Widén E, Junna N, Ruotsalainen S, Surakka I, Mars N, Ripatti P, *et al.* 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. https://doi.org/10.1161/CIRCGEN.121.003459
95. Raghavan S, Jablonski K, Delahanty LM, Maruthur NM, Leong A, Franks PW, *et al.* Interaction of diabetes genetic risk and successful lifestyle modification in the diabetes prevention programme. *Diabetes Obes Metab* 2021;**23**:1030–1040. https://doi.org/10.1111/dom.14309
96. Sun L, Pennells L, Kaptoge S, Nelson CP, Ritchie SC, Abraham G, *et al.* Polygenic risk scores in cardiovascular risk prediction: a cohort study and modelling analyses. *PLoS Med* 2021;**18**:e1003498. https://doi.org/10.1371/journal.pmed.1003498
97. 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**:584–591. https://doi.org/10.1038/s41588-019-0379-x
98. Tcheandjieu C, Zhu X, Hilliard AT, Clarke SL, Napolioni V, Ma S, *et al.* Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. *Nat Med* 2022;**28**:1679–1692. https://doi.org/10.1038/s41591-022-01891-3
99. Graham SE, Clarke SL, Wu K-HH, Kanoni S, Zajac GJM, Ramdas S, *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* 2021;**600**:675–679. https://doi.org/10.1038/s41586-021-04064-3
100. Data snapshots. https://www.researchallofus.org/data-tools/data-snapshots/(24 August 2022)
101. Mega JL, Stitziel NO, Smith JG, Chasman DI, Caulfield M, Devlin JJ, *et al.* Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet* 2015;**385**:2264–2271. https://doi.org/10.1016/S0140-6736(14)61730-X
102. Natarajan P, Young R, Stitziel NO, Padmanabhan S, Baber U, Mehran R, *et al.* Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation* 2017;**135**:2091–2101. https://doi.org/10.1161/CIRCULATIONAHA.116.024436
103. Damask A, Steg PG, Schwartz GG, Szarek M, Hagström E, Badimon L, *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**:624–636. https://doi.org/10.1161/CIRCULATIONAHA.119.044434

104. Marston NA, Kamanu FK, Nordio F, Gurmu Y, Roselli C, Sever PS, *et al.* Predicting benefit from evolocumab therapy in patients with atherosclerotic disease using a genetic risk score: results from the FOURIER trial. *Circulation* 2020;**141**:616–623. https://doi.org/10.1161/CIRCULATIONAHA.119.043805
105. Pepe MS, Fan J, Feng Z, Gerds T, Hilden J. The net reclassification Index (NRI): a misleading measure of prediction improvement even with independent test data sets. *Stat Biosci* 2015;**7**:282–295. https://doi.org/10.1007/s12561-014-9118-0
106. Cook NR. Quantifying the added value of new biomarkers: how and how not. *Diagn Progn Res* 2018;**2**:14. https://doi.org/10.1186/s41512-018-0037-2
107. Wand H, Lambert SA, Tamburro C, Iacocca MA, O'Sullivan JW, Sillari C, *et al.* Improving reporting standards for polygenic scores in risk prediction studies. *Nature* 2021;**591**:211–219. https://doi.org/10.1038/s41586-021-03243-6
108. Levin MG, Rader DJ. Polygenic risk scores and coronary artery disease: ready for prime time? *Circulation* 2020;**141**:637–640. https://doi.org/10.1161/CIRCULATIONAHA.119.044770