

STATE-OF-THE-ART REVIEW

Genetic Risk Stratification

A Paradigm Shift in Prevention of Coronary Artery Disease



Robert Roberts, MD,^a Chih Chao Chang, MD,^a Trevor Hadley, MD^b

HIGHLIGHTS

- CAD is a pandemic that can be prevented. Conventional risk factors are inadequate to detect who is at risk early in the asymptomatic stage.
- Genetic risk for CAD can be determined at birth, and those at highest genetic risk have been shown to respond to lifestyle changes and statin therapy with a 40% to 50% reduction in cardiac events.
- Genetic risk stratification for CAD should be brought to the bedside in an attempt to prevent this pandemic disease.

SUMMARY

Coronary artery disease (CAD) is a pandemic disease that is highly preventable as shown by secondary prevention. Primary prevention is preferred knowing that 50% of the population can expect a cardiac event in their lifetime. Risk stratification for primary prevention using the American Heart Association/American College of Cardiology predicted 10-year risk based on conventional risk factors for CAD is less than optimal. Conventional risk factors such as hypertension, cholesterol, and age are age-dependent and not present until the sixth or seventh decade of life. The genetic risk score (GRS), which is estimated from the recently discovered genetic variants predisposed to CAD, offers a potential solution to this dilemma. The GRS, which is derived from genotyping the population with a microarray containing these genetic risk variants, has indicated that genetic risk stratification based on the GRS is superior to that of conventional risk factors in detecting those at high risk and who would benefit most from statin therapy. Studies performed in >1 million individuals confirmed genetic risk stratification is superior and primarily independent of conventional risk factors. Prospective clinical trials based on risk stratification for CAD using the GRS have shown lifestyle changes, physical activity, and statin therapy are associated with 40% to 50% reduction in cardiac events in the high genetic risk group (20%). Genetic risk stratification has the advantage of being innate to an individual's DNA, and because DNA does not change in a lifetime, it is independent of age. Genetic risk stratification is inexpensive and can be performed worldwide, providing risk analysis at any age and thus has the potential to revolutionize primary prevention. (J Am Coll Cardiol Basic Trans Science 2021;6:287-304)

© 2021 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

The approach to any common disease is from the perspective of detection, prevention, and specific management once the event occurs. Coronary artery disease (CAD), which occurs commonly throughout the world, has long been recognized to be associated with factors that can be prevented. Risk factors for disease are usually categorized into 3 classes: genetic susceptibility;

From the ^aDepartment of Medicine, Dignity Health at St. Joseph's Hospital and Medical Center, Phoenix, Arizona, USA; and the ^bBaylor College of Medicine, Houston, Texas, USA.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the [Author Center](#).

Manuscript received July 28, 2020; revised manuscript received September 8, 2020, accepted September 13, 2020.

ABBREVIATIONS AND ACRONYMS

ACC = American College of
Cardiology

AHA = American Heart
Association

ANRIL = antisense non-coding
RNA in the INK4 Locus

bp = base pair

CAD = coronary artery disease

GRS = genetic risk score

GWAS = genome-wide
association study

LDL-C = low-density
lipoprotein cholesterol

MR = Mendelian randomization

SNP = single nucleotide
polymorphism

environmental exposure; and lifestyle factors. The environmental and lifestyle components of this triad for CAD have been well-documented since the 1960s, starting with the Framingham study (1). Guidelines have been standardized for their detection and reduction of their associated risk for CAD (2).

Several randomized placebo-controlled clinical trials have confirmed that a reduction in these risk factors for CAD is associated with a corresponding reduction in cardiac events and mortality (3-5). This was comprehensively reviewed by The Cholesterol Treatment Trialist, which did a meta-analysis that involved >170,000 individuals and analyzed the safety and efficacy of statin therapy (6,7). In the United States, over the past 30 years, reduction of risk from environmental and lifestyle risk factors has been the predominant reason for a 50% reduction in cardiac events and mortality (8,9).

Epidemiologists have claimed for some time, based on family history (10) and studies on identical twins (11), that genetic risk factors account for 40% to 60% of the predisposition for CAD (12). Discovery and isolation of these genetic variants predisposed to CAD were enabled by the development of new technologies in the early part of this century (13-17). In this review, we summarize these large studies, their research implications for the pathogenesis of CAD, and recent attempts to incorporate these genetic variants into a risk score to risk stratify for primary and secondary prevention of CAD. The application of a risk score based on an individual's DNA has several advantages over conventional risk factors because it is independent of age and can be determined at birth or anytime thereafter. The advantages and limitations of a genetic risk score (GRS) for CAD are summarized. The quality and accuracy of GRSs are still in evolution but could be enhanced by clinical application (Central Illustration).

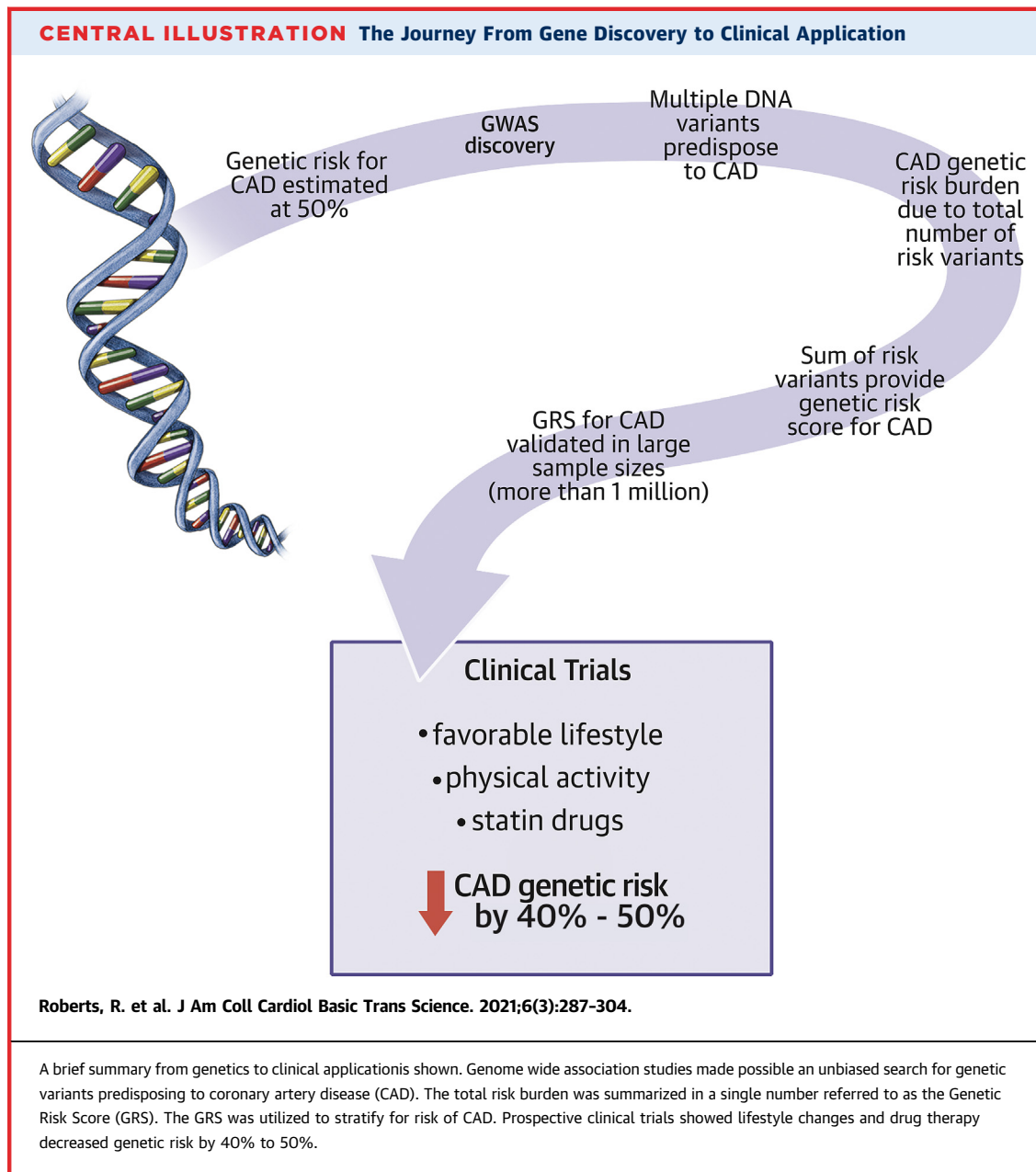
RARE SINGLE GENE DISORDERS

Variations in the DNA sequence that are disease related occur in a continuum from rare to common. However, because of the earlier clinical detection of the phenotypes and detection through pedigrees, we refer to inherited diseases as those that are rare versus those that are common. Rare disorders by definition occur in <1% of the population and are usually associated with high penetrance. These genes were discovered through a technique referred to as genetic linkage analysis (18). This requires a pedigree

of 2 or 3 generations of several individuals affected with the disease. A DNA marker, detected in those affected with the disease (more often than by chance), indicates it is in close physical proximity to the gene responsible for the disease. The DNA region is then cloned and sequenced to identify the responsible mutation, a process referred to as positional cloning. Using this approach, genes responsible for multiple rare cardiovascular inherited disorders were discovered, including familial hypertrophic cardiomyopathy (19,20), and Wolff-Parkinson-White syndrome (21). The last 2 decades of the 20th century were a golden era for these rare diseases and provided tremendous insight into their pathogenesis and management. It is claimed of the 8,315 rare inherited disorders, a responsible gene has been discovered for >5,482 (22). Expression of the responsible mutant gene as a transgene in animal models for rare cardiovascular disorders [e.g., familial hypertrophic cardiomyopathy (23) or Wolff-Parkinson-White syndrome (24)] induces the expected human phenotype in the offspring. Thus, these disorders are often referred to as single-gene disorders because the mutant gene is both necessary and sufficient to induce the phenotype.

POLYGENIC DISORDERS: GENOME-WIDE ASSOCIATION STUDIES

The predisposition of common disorders, such as CAD and diabetes type 2, are all due to multiple DNA risk variants and are referred to as polygenic disorders. This is in keeping with the hypothesis that the DNA risk variants transmitting predisposition for polygenic disorders are multiple, with each having only minimal risk. Identification of these DNA risk variants requires a different approach than that of collecting pedigrees and analyzing for genetic linkage. The preferred unbiased approach is that of a case-control association study based on an analysis of markers distributed throughout the genome (25). This is, in principle, simple; the frequency of DNA markers is determined by genotyping individuals known to have the disease and genotyping control subjects. Those DNA markers that occur more frequently in cases than controls would be designated in a DNA sequence that contains an increased risk for the disease. The marker itself is unlikely to be the cause of increased risk but is in close physical proximity to the causative sequence. The approach requires DNA markers evenly distributed throughout the human genome at intervals of not more than 3,000 to 6,000 base pairs (bp). This requires 500,000 to 1 million markers to meet this requirement.



Several advancements in technology occurred to make the case-control association study approach practical. In 2001, the human genome (13,17) was sequenced, followed by the HapMap Project (16), which made available millions of informative polymorphic DNA markers spanning the human genome. These markers are single nucleotide polymorphisms (SNPs) that are fairly evenly distributed throughout the genome at approximately 1 SNP/1,000 bp. The number of SNPs per human genome is fairly consistent at approximately 5 million (16,26,27).

SNPs were encrypted on computerized microarrays and used as the template for platforms designed to perform high-throughput genotyping (15). The first of these studies was performed for age-related macular degeneration (28). Genome-wide association studies (GWASs) have become a major international focus over the past decade. The National Human Genome Research Institute GWAS catalog indicates >5,687 GWASs have been performed, involving >71,000 variant trait associations and published in >3,000 publications (29). These studies

have greatly enhanced the architecture of genetic predisposition for CAD and many other polygenic diseases.

Genetic predisposition, or lack of predisposition for a disease, is determined by the DNA sequences transmitted to offspring at conception. Genes are classically defined as DNA sequences that code for proteins. This definition was developed before we knew much regarding the remainder of the sequences in the total genome. It is now expected that >90% of the genome is functional and influences the expression directly or indirectly of genes upstream and downstream, including across other chromosomes. Sequences that code for proteins account for >1% of the DNA in the human genome. All genes have multiple forms, referred to as alleles, as do DNA sequences in the nonprotein-coding regions. These alleles differ from each other and their parent sequence, usually by only a single nucleotide substitution mutation. The frequency of these alleles in the population varies from rare to common. Despite the minor sequence difference between alleles, its function is often altered, which for disease-related alleles could lead to being neutral or associated with an increased or decreased predisposition. Each individual has a maximum of 2 copies of each allele, 1 from each parent. The particular allele occupying any chromosomal locus of an individual is determined at conception. The use of the term allele in this review refers to alternative forms of any DNA sequence, which may or may not be a part of a protein-coding sequence. Those DNA sequences shown to be associated with increased risk for a disease are referred to as genetic risk variants.

GENETIC RISK VARIANTS FOR CAD

The first genetic risk variant for CAD, 9p21, was discovered in 2007, simultaneously and independently by 2 groups (30,31), followed shortly by confirmation from 2 other groups (32,33). Over the next 2 years, several groups throughout the world confirmed the association between the 9p21 genetic risk variant and CAD (34). The reported features of the 9p21 genetic risk variant were similar throughout the world, except in Africans, who had the 9p21 variant is not a risk factor (35-37). The 9p21 genetic risk variant was shown to be common, being present in approximately 75% of the population. The risk for CAD is mediated independently of conventional risk factors for CAD. Subsequent efforts by groups such as the Myocardial Infarction Genetics Consortium (38) identified 2 novel genetic risk variants for CAD, followed by a variant at 3q22 found by Erdmann *et al.* (39), the SLC22A3 variant

found by Trégouët *et al.* (40), and the 12q24 variant found by Gudbjartsson *et al.* (41). The Coronary Artery Disease Genetics Consortium (42) identified 5 novel genetic risk variants for CAD. The characteristics of the genetic variant 9p21, along with the subsequent 10 genetic risk variants predisposed to CAD, were consistent with the hypothesis that common polygenic disorders are due to multiple commonly occurring alleles, each associated with only minimal risk.

The associated minimal risk of each variant strongly indicated there were many more, and the search would have to be intensified. It would require a greater density of DNA markers and massive sample sizes. This led to the formation of an international consortium, with >80,000 cases and control subjects, the CARDIoGRAM (Coronary Artery Disease Genome-Wide Replication, and Meta-analysis) Consortium (43). The GWAS approach used millions of SNPs as markers with increased sample sizes. This required statistical correction, which with 1 million SNPs and a Bonferroni correction, demanded a *p* value of 0.00000005 that became known as genome-wide significance (44). In addition, replication in an independent population was required.

Over the next 10 years, international consortium and individual efforts discovered 171 genetic risk variants that are of genome-wide significance and have been replicated in an independent population. A total of 163 of these genetic risk variants for CAD were carefully tabulated in the review by Erdmann *et al.* (45). The additional 8 variants are summarized in Table 1. It is expected there are many more to be identified. There are many other genetic risk variants for CAD that are not of genome-wide significance but are highly significant in that they have <5% false positives as determined by the false discovery rate (39,45-47).

NEED FOR GENETIC RISK STRATIFICATION

CAD is a slow-growing disease that usually takes decades to manifest clinical features. This process is influenced by several known environmental and lifestyle risk factors. Decreasing environmental or lifestyle risk has been shown to be extremely effective, particularly for secondary prevention. Secondary prevention is activated when an individual has a clinical event (e.g., myocardial infarction or angina). In contrast, primary prevention, by definition, does not have such defining signals. Unfortunately, the conventional risk factors for CAD are age-dependent and are only detected late in life, which is less than optimal for primary prevention. The potential to markedly attenuate this disease, if not eliminate it, is robust with the currently existing therapeutic

TABLE 1 Eight Newly Discovered Genetic Risk Variants

	Position	Lead SNP	RAF	OR	Gene(s) at Locus	Ref. #
1	6p21.3	rs3869109	G (0.55)	1.16	HLC-C,HLA-B, HCG27	(140)
2	chr2: 19809058	rs2123536	T (0.39)	1.25	TTC32-WDR35	(141)
3	chr4: 156730909	rs1842896	T (0.76)	1.23	GUCY1A3	(141)
4	chr6: 32449331	rs9268402	G (0.59)	1.17	C6orf10	(141)
5	chr12: 88605319	rs7136259	T (0.39)	1.21	ATP2B1	(141)
6	6p24.1	rs6903956	A (0.08)	1.51	C6orf105	(142)
7	chr3: 17099388	rs4618210	A (0.45)	0.91	PLCL2	(143)
8	chr19: 2111529	rs3803915	A (0.21)	0.89	AP3D1-DOT1L-SF3A2	(143)

Since the paper by Erdmann et al. (45), there has been 8 newly discovered genetic risk variants predisposed to coronary artery disease.
 OR = odd's ratio; RAF = relative allele frequency; SNP = single nucleotide polymorphism.

armamentarium. Modification of risk factors clearly would have to be initiated early in life, and, in some cases, even before adolescence. In women with hormonal protection during the pre-menopausal stage, it is perhaps possible to delay primary prevention until the fourth or fifth decade, and still prevent the development of clinical sequelae (46,47).

The predominant culprit in the pathogenesis of coronary atherosclerosis is the plasma low-density lipoprotein cholesterol (LDL-C) and the duration of exposure. Every additional 10 years of exposure to plasma LDL-C is associated with a doubling of morbidity and mortality (48). The Cardiology Practice Guidelines have established a plasma LDL-C of ≤ 70 mg/dl as the target for primary or secondary prevention of CAD. However, the plasma LDL-C of a woman in her fourth decade averages 121 mg/dl, whereas a man in his fourth decade averages 146 mg/dl, which is nearly twice that of the desired target (49). Should we treat everyone with statin therapy? Primary prevention based on conventional risk factors requires treatment with statin therapy in 100 individuals to prevent a single cardiac event (50). Only one-half are expected to benefit because only 50% will develop a cardiac event.

To avoid the cost of treating everyone and inappropriately treating one-half of the population who will not benefit beckons the issues to risk stratify. The Cardiology Practice Guidelines were set up as an attempt to avoid treating everyone with the hope of selecting those who would most benefit from preventive therapy. Using the conventional risk factors, namely, age, cholesterol, diabetes, obesity, blood pressure, sedentary way of life, and smoking, a simple calculation was devised to be applicable to anyone from age 40 to 75 years, taking into account that everyone at age 40 years will have increased plasma LDL-C. The American College of Cardiology/The American Heart Association (AHA/ACC) adopted the method of the Pooled

Cohort Equation, which calculates risk for 10 years based on these risk factors (51). Anyone with a 10-year risk of $\geq 7.5\%$ are candidates for preventive therapies.

A 40-year-old woman with the risk profile of Caucasian, blood pressure of 120/70 mm Hg, total cholesterol of 240, LDL-C 180 mg/dl, high-density lipoprotein (HDL) 30 mg/dl, non-smoker, and without a history of diabetes has a 10-year risk of 1.3% based on the Pooled Cohort Equation. This is far below the level of 7.5% required to recommend preventive therapy. Nevertheless, a plasma LDL-C of 180 mg/dl is >2 -fold higher than the desired level, and despite the lack of other risk factors, it is likely to contribute to CAD. It appears to be a missed opportunity for primary prevention of CAD in women with these clinical findings. It is well-documented that the incidence of heart disease is the same in women as in men but is delayed by approximately 10 years. If we are to take advantage of the opportunity for primary prevention in pre-menopausal women, it would be greatly enhanced with a method of risk detection that is independent of age.

Mendelian randomization (MR) has shown that individuals with a minimal decrease in LDL-C due to genotypes inherited at conception are associated with a 50% reduction in cardiac risk per mmol/l reduction in LDL-C (52). This is approximately 3-fold greater than the benefit observed in clinical trials that enroll individuals with an average age of 60 years. To be effective, primary prevention in men has to occur in their 20s or 30s for risk stratification for early intervention (53). Because CAD is now a pandemic disease, attempts to halt its spread will require earlier primary prevention in both men and women, preferably based on a risk assessment that is independent of age. Risk assessment for CAD, based on genetic risk, is independent of age because DNA does not change in an individuals' lifetime. This is discussed in further detail in subsequent sections.

A TREASURE TROVE FOR NEW DRUGS AND RESEARCH-FUNCTIONAL ANALYSIS OF GENETIC RISK VARIANTS

Genetics have been most favorable for CAD (54). Families with familial hypercholesterolemia provided strong evidence for cholesterol being a causative agent (55). The development of the first statin drug and its lowering of cholesterol in patients with familial hypercholesterolemia was associated with a corresponding reduction in cardiac mortality and morbidity. The statin class of drugs is currently our main therapy in both primary and secondary prevention of CAD. Approximately a decade ago, a new armamentarium in the prevention of CAD was introduced; PCSK9 inhibitors were stimulated yet again by a genetic observation. The discovery of a gain-of-function mutation in the *PCSK9* gene was observed to be associated with increased risk for myocardial infarction (56). This was followed by the discovery of a loss-of-function in the *PCSK9* gene that was associated with decreased risk for myocardial infarction (57). PCSK9 inhibitors are now part of the routine clinical management and prevention of CAD (58).

Angiopoietin-like 3 is known to inhibit lipoprotein and endothelial lipase, which leads to increased levels of triglyceride and other lipids (59-61). Individuals inheriting a loss of function in the gene encoding angiopoietin-like 3 experience a 41% lower risk of CAD due to low plasma levels of LDL-C and triglycerides (62,63). Inhibition of angiopoietin-like 3 by pharmacological means is also associated with lower plasma levels of LDL-C independent of the LDL-C receptor. This led to development of an antibody, evinacumab, which inhibits angiopoietin-like 3. It has undergone clinical trials, and phase III trials performed recently in individuals with homozygous familial hypercholesterolemia showed a 41% reduction in LDL-C (64). Thus, a new mechanism has been exploited to add to the armamentarium of drug therapy to decrease LDL-C independent of cholesterol (statins) or removal by the LDL receptor (PCSK9 inhibitors). Genetics has again contributed significantly to the future prevention and management of CAD.

More than one-half of the genetic risk variants do not mediate their risk through known conventional risk factors (34). This would indicate that several factors, yet to be discovered, are involved with the pathogenesis of CAD. This is inspirational fodder for the investigator. Elucidation of the mechanisms whereby they mediate their risk enable the development of drugs based on novel targets. The thrust of research in this field is still primitive, in part, because

of the lack of tools to pursue such a task. First, the lead SNP signifies a risk region that contains thousands of bp of which only ≥ 1 bp are the causative risk mutation. The efforts to map and sequence these regions are expensive, at a ratio of approximately 30 to 1 over the cost of genotyping. Even after massive sequencing, it will be a prohibitive task. This is made more difficult because no 1 risk variant is necessary or sufficient to induce the phenotype of CAD; therefore, researchers cannot depend on a conventional knockout or transgenic animal models to confirm or disprove causality. Furthermore, >80% of the risk variants for CAD, as for other polygenic diseases, are located in nonprotein-coding sequences (34) that are inherently unfriendly to functional analysis. Their location in nonprotein-coding regions relegates risk variants to a regulatory role, namely, influencing the expression of protein-coding sequences that can be hundreds of thousands of bp upstream or downstream (cis-acting) or even on other chromosomes (trans-acting). This task is further complicated by the recent observation that suggested that only approximately one-third of causal genes are the nearest genes to the GWAS hit (65,66). One approach is to express these sequences in various in vitro experiments to seek a clue to their function. Investigators are pursuing means to place the risk variant sequences and their proximal genes into biological pathways in the hopes of determining their function (45).

Despite the difficulties, research is extremely active in pursuing the function of genetic risk variants, as can be exemplified by discussion of a few of these variants located at loci 9p21, 1p13, 15q24, and 6p24. The 9p21 risk variant for CAD was the first genetic risk variant to be discovered for CAD and has undergone extensive research. Multiple studies confirmed 9p21 mediates its risk for CAD, independent of known risk factors. Fine mapping of the 9p21 risk variant region showed that it is confined to 58,000 bp (67). This 58,000 bp does not contain any protein-coding sequences but does overlap with a long noncoding RNA of 126,000 bp referred to as an antisense noncoding RNA in the *INK4* Locus (*ANRIL*) (68,69). *ANRIL* has no known function, but its expression induces the expression of 3 nearby well-known genes, *CDKN2a*, *CDKN2b*, and *ARF*, which encode for cyclin-dependent kinase inhibitors. An attractive hypothesis is that the 9p21 risk variant reduces the expression of *ANRIL* and the kinase inhibitors that enable the growth of vascular atheroma. Several studies provided support for this hypothesis (70,71). Unfortunately, other studies showed lack of

correlation between the 9p21 risk variant and expression of cyclin-dependent kinase inhibitors (72-74).

Visel et al. (75) pursued studies of the 9p21 variant in a mouse model. This corresponds to a 70,000 bp region in the mouse. This region was knocked out, and the resulting phenotype showed excessive proliferation of smooth muscle cells and several neoplasms. However, none of the vessels showed fatty streaks or any evidence of coronary atherosclerosis. The absence of atherosclerosis might be, in part, because the mouse was resistant to atherosclerosis or the lack of the 9p21 risk variant, because there was only a 50% sequence homology between the human 9p21 58,000 bp and the 70,000 bp region of the mouse. Studies suggested that the 9p21 risk variant did not appear in the genome until that of higher primates (67), and thus, would not be expected to give an atherosclerotic phenotype in the mouse.

Harismendy et al. (76) performed a series of studies that concluded that the 9p21 risk variant mediated its risk for CAD through interferon- γ . However, we subsequently showed that the effects of interferon- γ on CDKN2a/b were independent of the 9p21 risk variant (77), which was confirmed recently by Erridge et al. (78). Recent studies involved the deletion of the 58-kb locus-induced pluripotent stem cells derived from individuals homozygous for the nonrisk haplotype. Expression of this region in the smooth muscle cells suggested the 9p21 risk variants were related to adhesion, contraction, and proliferation (79).

Although the mechanism of the 9p21 risk variant remains unknown, clinical observation suggests that it acts through the vessel wall and predisposes to the development of atheroma. This was confirmed by the deCode group, which found that the 9p21 risk variant was also associated with an increased risk for abdominal aortic aneurysms and intracranial aneurysms (80). There was no association between the 9p21 risk variant and myocardial infarction (81,82). This strongly indicated 9p21 was not associated with plaque rupture or thrombosis. Several studies indicated the 9p21 risk variant was associated with progression of coronary atherosclerosis (81,83,84). The 9p21 risk variant was predisposed to atheroma in the coronaries and ballooning in the aortic aneurysm, which suggested a defect in the vessel wall and led to these 2 different abnormalities.

The *SORT1* gene located near the 1p13 locus is known to increase the risk for CAD approximately 15% to 20% per copy and causes a 10% to 15% increase in plasma LDL-C. *SORT1* is highly expressed in the liver and primary hepatocytes (85,86). Studies have been

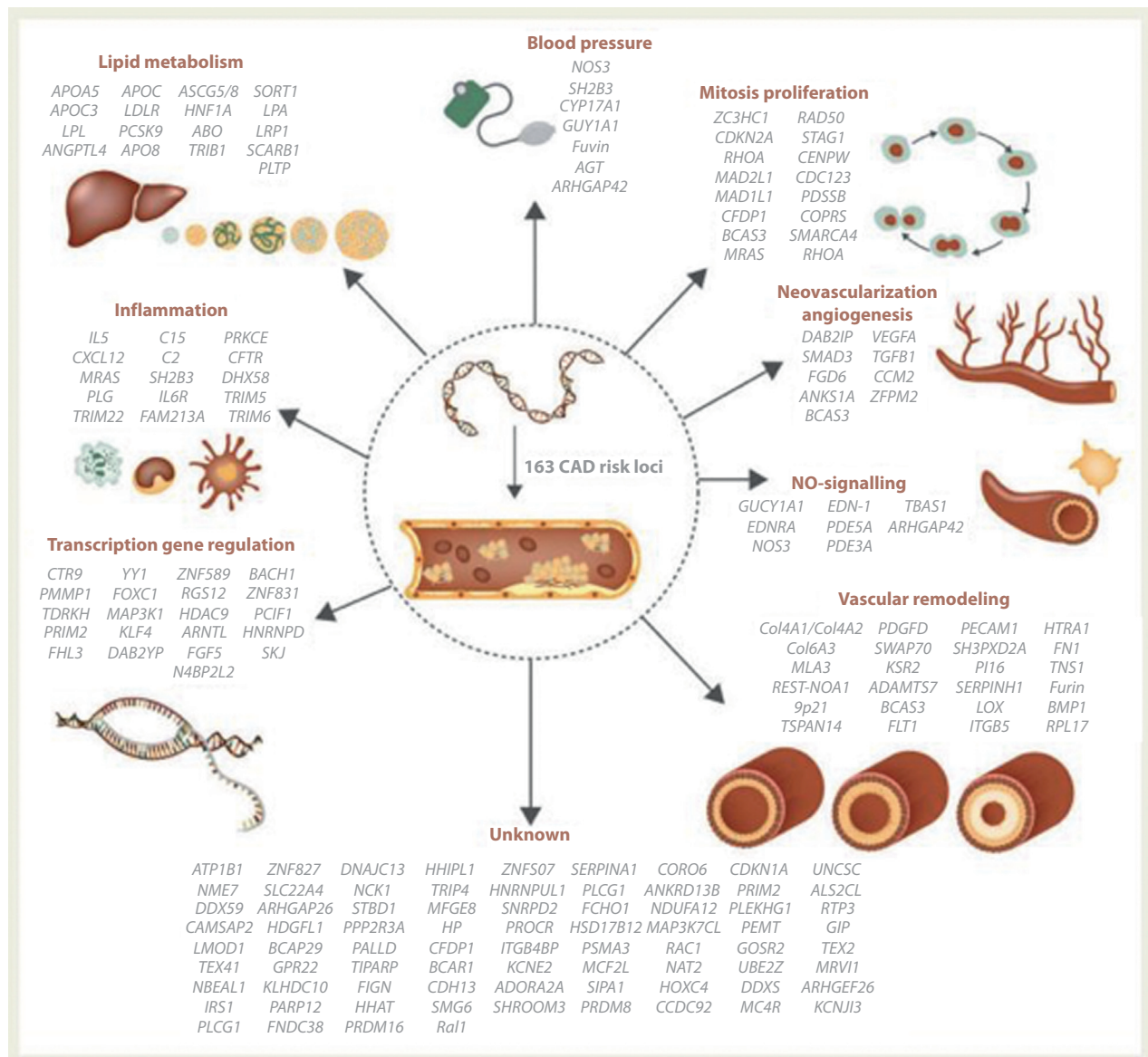
controversial, showing increased expression of very LDL in the mouse, and other studies showing decreased very LDL secretion (87-89). Because of these confusing results, *SORT1*, as a risk variant for CAD, has received less attention.

A third risk variant for CAD located on chromosome 6p24 has at least 2 candidate causal genes, *PHACTR1* (phosphatase and actin regulator 1) and *EDN1* (endothelin 1). Fine mapping of this region strongly suggest endothelin 1 is more likely to be the candidate because it is known to be involved with vascular biology. Nevertheless, studies show that *PHACTR1* also has vascular properties, but conflicting results have been observed in comparing *PHACTR1* versus *EDN1* on expression of vascular tissue. It remains to be determined which of these 2 genes are likely to be the causal candidate for CAD predisposition (90,91).

The risk variant for CAD on chromosome 15q25 harbors the gene *ADAMTS7* (A disintegrin and metalloproteinase with thrombospondin motifs-7). This variant is associated with reduced protein maturation, proteinase activity, and also migration of vascular smooth muscle cells in vitro (92). Knockout experiments in the mouse show reduced atherosclerosis and decreased cell proliferation in response to vascular injury (93,94). These results need to be confirmed and are still in process.

In a recent review, Erdmann et al. (45) compiled a listing of the known most proximal genes of the SNP that show peak association. This is shown in **Figure 1**. Several variants suggest inflammation is a risk for CAD, including the major histocompatibility complex variant and others shown in the diagram. A recent gene-tissue expression study showed expression of HLA-C (Major Histocompatibility Complex, Class I, C) in the coronary artery (95), which lent further support to a role for chronic inflammation in the development of atherosclerosis. This is in keeping with other experimental and clinical evidence collected over the years that signifies a role for inflammation (96). The recent clinical trial CANTOS (Canakinumab Anti-Inflammatory Thrombosis Outcomes Study) (96) confirmed that interleukin-1 β is involved as 1 of the causative factors of coronary atherosclerosis.

The availability of multiple genetic risk variants has enabled MR to be used to determine whether associations are causative (97). Using genetic variants that increase plasma HDL cholesterol as the variable in MR studies, we cast considerable doubt on the dogma that HDL is protective of CAD (98). This was confirmed in several other studies (99-104). In contrast, MR studies that used genetic variants that influenced plasma triglyceride levels or plasma LDL-C strongly indicated that they were causative

FIGURE 1 Suggested Pathways for Involvement of Genetic Risk Variants

Various organs or pathways associated with 163 mapped genetic risk loci for CAD are shown in the figure. Reprinted with permission by Erdmann et al. (45). CAD = coronary artery disease.

for CAD (98). Several favorable and potential targets for the development of drugs to treat CAD and myocardial infarction were excluded by MR studies and included fibrinogen, folic acid, uric acid, and C-reactive protein (97). In selecting a target for the development of novel drugs, it is strongly recommended, when possible, to perform MR to confirm that the target is causative of the disease. The association of biomarkers with a disease does not necessarily reflect causality. A good example of this

is plasma C-reactive protein, a marker for inflammation, that increases in concentration in the presence of inflammation. Inflammation is 1 of the causative factors for CAD, thus, increased C-reactive protein reflects the tendency for increased CAD. Nevertheless, this immune response has nothing to do with causing coronary atherosclerosis. Confirmation by MR of a biomarker or drug target as non-causative can save millions of dollars and years spent on a large clinical trial. A confounding factor

for MR may be pleiotropy, in which the genetic variant affects multiple traits (97); nevertheless, some techniques can minimize limitations (e.g., the latent causal variable model) (105).

DEVELOPMENT OF A GRS

Perhaps the most well-known approach to modeling the contribution of common genetic variations predisposed to CAD is risk stratification based on the GRS. In brief, the CAD GRS is derived from a list of common genetic variants that have risk-weighted associations with CAD. The weight of each risk variant is multiplied by the number of variants at that site (0, 1, or 2), and the final score is simply the sum of the weighted dosage for each risk variant included in the GRS. This sum is typically represented as a natural log for the purposes of distributing individuals into percentiles for risk stratification.

The potential use of CAD GRS for clinical risk stratification has fueled interest in the methods behind GRS development (50,106,107). Although there is no standardized procedure for developing a CAD GRS for clinical use, establishing a CAD GRS for clinical use can be broken down into 3 basic steps: 1) selection of the CAD risk variants to be included in the microarray; 2) genotyping the individual's DNA using the microarray; and 3) formation of an informatics pipeline to generate the GRS.

The CAD GRS can be estimated from variants using weights for CAD previously established by GWASs. Recently discovered genetic risk variants can be added by imputation. As discussed earlier, the most recent CAD GWAS brought the number of CAD risk variants with genome-wide significance to 171 (45). Previous studies described the application of CAD GRS derived from these variants (106,108). Recently, methods were used to develop the GRS based on millions of variants by relaxing the thresholds for genome-wide significance and by using software that accounted for linkage disequilibrium and estimated the proportion of SNPs with non-zero weights (109). One such genome-wide CAD GRS included 6.6 million variants and demonstrated superior performance to the CAD GRS that used less CAD risk variants, although it was genome-wide significant (5×10^{-8}) (109). In both cases, the risk variants, along with their respective weights derived from GWASs, are readily available (108-110).

A detailed description of the procedure for generating a novel CAD GRS from a CAD GWAS has been previously described (111). However, as CAD GRS moves closer toward the reality of clinical implementation, it is reasonable to expect that a

standardized list of variants, shown to demonstrate the best performance for the intended use, will be established and made readily available. However, before the clinical implementation of CAD GRS, a number of issues must be addressed; namely, that the current CAD GRS showed variable performance when applied to populations of non-European ancestry. This issue of disparities in accuracy of the CAD GRS is largely due to most of the risk variants being derived from GWASs performed in populations of predominantly European descent and the lack of well-powered studies conducted in globally diverse populations (112). Further work is needed to ensure that progress made in CAD GRS development does not increase disparities in the treatment of CAD.

Once the method of score generation has been determined, a procedure for genotyping patients and then processing that information into a GRS must be developed. If the CAD GRS being calculated has a limited number of risk variants, such as a CAD GRS that contains 50 SNPs, then calculating the GRS is simple. However, using the 6.6 million SNP genome-wide CAD GRS will require the use of larger microarrays and the use of software that can perform genotype imputation. Briefly, genotype imputation is a computational procedure that uses statistical inference to approximate the identity of variants not covered by the array being used. A number of software programs exist for genotype imputation. A few of the commonly used ones include IMPUTE2, BEAGLE, and Minimac 4 (113). The choice of software will depend upon the needs of the study, because each program will differ slightly in the time it takes to process each sample, the central processing unit it requires, and its average imputation quality.

For the purposes of setting up a CAD GRS, it will be necessary to ensure sufficient imputation accuracy of all variants to be included in the GRS, which will depend on the variants included in the array, the imputation software selected, and the reference panel used by the imputation software. The reference panel enables imputation software to infer the identity of genetic variants untyped by the array used to genotype the patient (16,27). Older reference panels, such as those from the HapMap project or 1000 Genome Project, have struggled to generate acceptable imputation accuracy in ethnically diverse samples, particularly those of African ancestry (114,115). However, the newest reference panel developed by the TOPMed (Trans-Omics for Precision Medicine) group was derived from 60,039 individuals of diverse ancestry who underwent deep whole-genome sequencing and demonstrated marked progress in

this area (115). Notably, the TOPMed reference panel demonstrated an improvement of 0.26 to 0.92 in the average imputation quality in African ancestry genomes (115). These improvements mark an important step toward closing the gap in the variable performance of CAD GRS in multiethnic populations.

Once a patient's sample has been received and genotyped, the genetic data will need to move through an informatics pipeline that can process the data into a CAD GRS. In short, from start to finish, a pipeline for a genome-wide CAD GRS consists of: 1) genotyping of the sample; 2) imputation; and 3) score calculation. During the set up of the pipeline, a series of quality control checks need to be performed to ensure that results are accurate and reliable. Because the anticipated use of a CAD GRS for risk stratification, further work will also be necessary to evaluate the validity of the CAD GRS before it can be brought to the clinic (116). Standardization of CAD GRS development and delivery will need to occur before clinical implementation.

VALIDATION OF A GRS

Development of a GRS was 1 of the goals derived from the GWAS search for risk variants predisposed to CAD. Risk stratification based on genetic risk was not age-dependent and could be obtained as early as at birth. This would be a major advantage for primary prevention of CAD as discussed previously (see the section on Need for Genetic Risk Stratification). The initial efforts using only the 9p21 risk variant were discouraging (117). Predictability of cardiac events became more promising as more risk variants became available, and the total burden of genetic risk for CAD was shown to be proportional to the number of risk variants (34). A GRS based on 12 risk variants predisposed to CAD (118) was encouraging, although the benefit compared with conventional risks factors was clinically insignificant. The sustained discovery effort provided many more DNA risk variants, which enabled the development of a much improved GRS.

This was tested by Mega *et al.* (106) in 2015 using a large study with a sample size of 48,421 individuals. This clinical trial consisted of 2 primary prevention trials (JUPITER [Justification for the Use of Statins in Prevention: An Intervention Trial Evaluating Rosuvastatin] and ASCOT [Anglo-Scandinavian Cardiac Outcomes Trial]) and 2 secondary prevention trials (CARE [Cholesterol and Recurrent Events] and PROVE-IT-TIMI [the Pravastatin or Atorvastatin Evaluation and Infection Therapy-Thrombolysis in Myocardial Infarction]). These trials were completed

and showed that statin therapy was effective for primary and secondary prevention (106). Baseline blood samples were available and were genotyped for 27 genetic risk variants predisposed to CAD. The GRS predicted cardiac events independent of traditional risk factors, including family history. Furthermore, it predicted who would benefit most from statin therapy. The number needed to treat to prevent 1 cardiac event was decreased 3-fold in the primary prevention trials. The risk gradient across the GRS was significant: to prevent 1 cardiac event in 10 years, 66 subjects with low genetic risk needed to be treated, 42 subjects in those with intermediate risk needed to be treated, and 25 subjects with high genetic risk needed to be treated in JUPITER, and 57, 47, and 20 subjects needed to be treated, respectively, in ASCOT (106). Thus, the GRS not only provided risk stratification regarding prognosis but also detected those who benefitted most from the proven preventive drug for CAD, namely, statins (119). In a similar randomized clinical trial termed WOSCOPS (West of Scotland Coronary Prevention Study) (120), 57 genetic risk variants were genotyped in a sample size of 10,456 individuals. The results showed that the high genetic risk group had a reduction of cardiac events of 44% versus 24% in the remaining groups. The number needed to treat to prevent 1 cardiac event in the high-risk group was 13 subjects versus 38 subjects in all other groups. These results confirmed the greater power of genetic risk stratification other than that of traditional risk factors. It is unusual for a single biomarker to have both predictive and prognostic properties (121).

Using a microarray that contained 49,310 risk variants derived from CARDIoGRAMplusC4D consortium, Abraham *et al.* (122) risk stratified 5 prospective cohorts, 3 from the FINRISK (a large Finnish population survey on risk factors on chronic, non-communicable diseases) group and 2 from the Framingham Heart Study for a combined sample size of 16,082 subjects. The GRS was strongly associated with incident cardiac events and essentially independent of conventional risk factors, including family history. Integration of the GRS with the Framingham Risk Score or the ACC/AHA Pooled Cohort Equations improved the 10-year risk prediction for CAD events.

Abraham *et al.* (122) showed 49,310 risk variants had greater power to risk stratify CAD than the previous efforts that used 27 or 57 genetic risk variants predisposed to CAD. We have been consistently reminded that the genetic risk variants of known genome-wide significance only account for a small

fraction of the expected inherited predisposition for CAD. If a less statistically stringent correction was used, namely, the false discovery rate, hundreds of other genetic risk variants could be associated with predisposition for CAD. Abraham et al. (122) used 49,310 genetic risk variants and accounted for 27% of inherited predisposition, a much larger proportion than using the genome-wide significant risk variants predisposed to CAD. There was a concern from the beginning that requiring $5 \cdot 10^{-8}$, as proposed by the Bonferroni correction, might be too stringent. This led to the use of other less stringent statistical correction methods to select genetic risk variants predisposed to CAD.

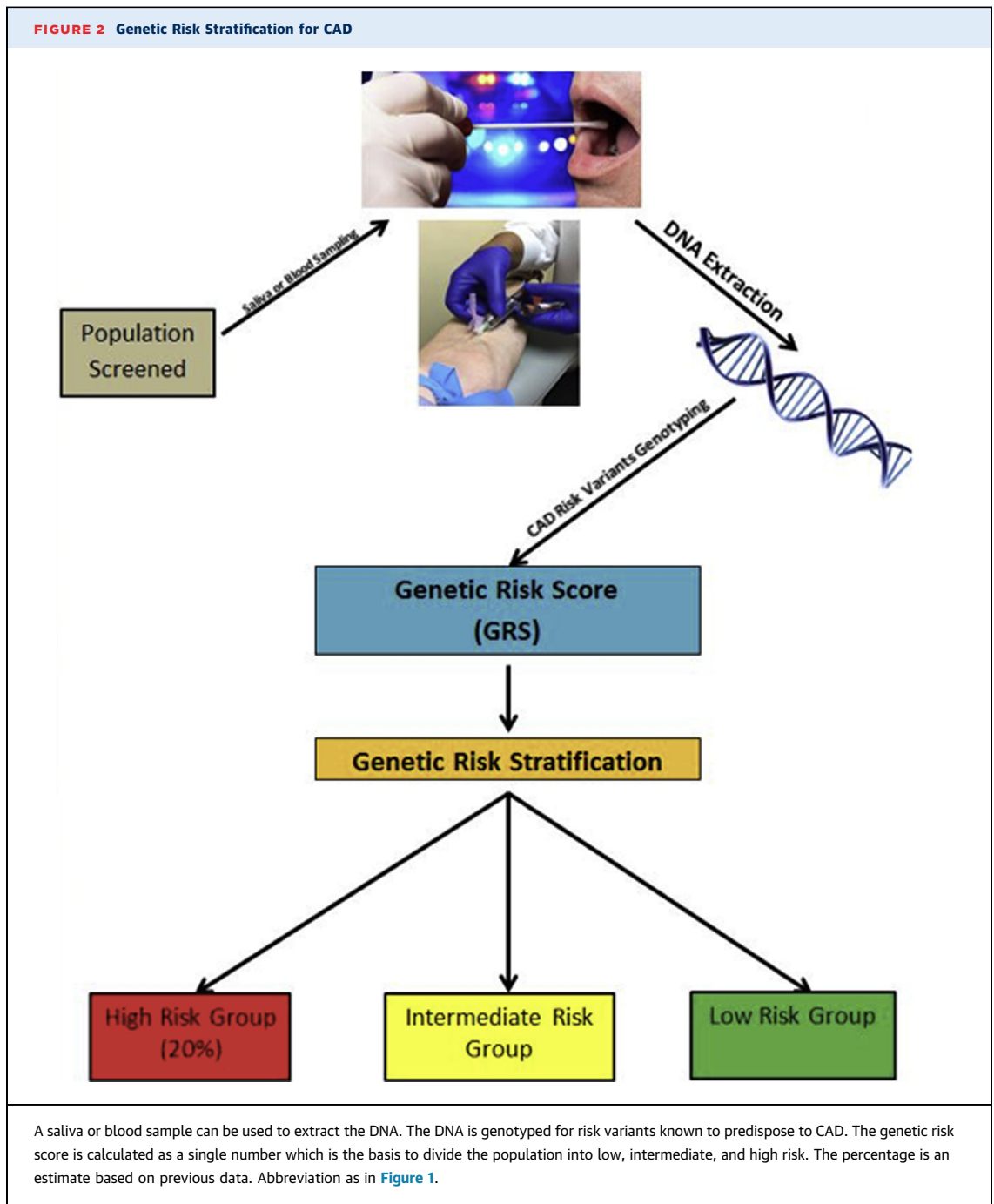
Inouye et al. (123) developed a microarray with 1.7 million SNPs, which represented risk variants of genome-wide significance and others with a false discovery rate of <5%. They genotyped a population of 500,000 from the UK Biobank and observed that the top 20% with the highest GRS had a 4-fold increased risk of CAD. The GRS was relatively independent of conventional risk factors. The microarray of 1.7 million SNPs was superior to the previous microarray, which contained 49,310 variants.

Khera et al. (124) used a computerized algorithm, LDpred (109) to select 6.6 million risk variants predisposed to CAD. A pruning analysis was performed to assure all risk variants were in linkage disequilibrium to avoid redundancy (125). The total number of genetic risk variants contributing to predisposition of CAD might well be only a few thousand. However, the statistical programs were such that weakly associated genetic risk variants contributed to the GRS, whereas those that contributed zero risk did not dilute the predictive power (125). Khera et al. (124) genotyped a population of 288,978 from the UK Biobank and observed the top 8% with the highest GRS had a 3-fold increased risk of CAD. The top 0.5% had a 5-fold increased risk for CAD. Phenotypic analysis showed conventional risk factors would not detect most individuals with a high genetic risk for CAD. There were only 20% with hypercholesterolemia, 35% with a family history, and 28% with hypertension. They concluded, as did previous studies, that the GRS was superior to and relatively independent of conventional risk factors.

Although multiple studies, as discussed, have shown the GRS to be superior to conventional risk factors for risk stratification of CAD, there are 2 studies with contrasting results. Both studies were performed with a microarray using >6 million genetic risk variants predisposed to CAD. Mosley et al. (126) genotyped 2 American populations of European descent in the ARIC (Atherosclerosis Risk In

Community) ($n = 4,847$) and the MESA (Multi-Ethnic Study of Atherosclerosis) ($n = 2,390$) cohorts. They found a close association in the ARIC cohort between the GRS and coronary heart disease, with an odds ratio of 1.89. Those in the top 20% had a 2.89 increased risk for coronary heart disease versus the bottom 80% percentile. Analysis of incident cardiac events in the follow-up of ARIC and MESA showed good correlation, with genetic risk in the top 20%, but the odds ratios were less (1.54 and 1.63, respectively). Results of the risk for CAD estimated from the Pooled Cohort Equations added to the GRS did not significantly improve prediction of cardiac events or change the C-statistic in either the ARIC or MESA population. The second study by Elliott et al. (127) genotyped a cohort of 352,660 individuals from the UK Biobank. The GRS showed a strong association with coronary heart disease, but when added to the result of the Pooled Cohort Equations, showed only a modest improvement in the C statistic, from 0.76 to 0.78. Although this was statistically significant, it was probably not of great clinical significance. The conclusion from both studies was that the GRS was equivalent to the risk score obtained from conventional risk factors but added no additional benefit in this group of individuals with a mean age of 60 years. It was difficult to reconcile these results with the multiple other studies that showed superiority of GRS over conventional risk factors. Nevertheless, Elliott et al. (127) recognized the GRS might offer an advantage in predicting CAD in younger age groups who had conventional risk factors have not yet been developed. This is the dilemma for early primary prevention of CAD. The authors agreed a study should be performed to assess the predictive value of GRS in this younger age group.

Two recent trials that assessed the effect of PCSK9 inhibitors further confirmed the predictability of the GRS for CAD events. The FOURIER (Further Cardiovascular Outcomes Research With PCSK9 Inhibition in Subjects With Elevated Risk) trial (128) assessed the benefit of a PCSK9 inhibitor, evolocumab, on cardiac events in a randomized, placebo-controlled trial of 11,953 patients who agreed to have their DNA analyzed. The microarray used in the study contained either 27 genetic risk variants predisposed to CAD or 6 million genetic risk variants. There was a strong correlation between individuals with intermediate and high genetic risk for CAD and cardiac events with 1.23 and 1.65 hazard ratios, respectively. Those who received evolocumab therapy had a 13% relative reduction in the group with conventional risk factors without high genetic risk and a 31% reduction in the high genetic risk group with or



without conventional risk factors. Thus, the group with the highest genetic risk had the most benefit from lowering plasma LDL-C due to evolocumab and was independent of conventional risk factors. The microarray with 27 genetic risk variants predisposed to CAD was just as accurate for predicting cardiac events as the microarray with 1 million risk variants. The ODYSSEY (Evaluation of Cardiovascular

Outcomes After an Acute Coronary Syndrome During Treatment With Alirocumab) trial (129), a randomized, placebo-controlled study, assessed the effect of alirocumab on cardiac events in 11,953 individuals. The group with the highest GRS had the highest risk for CAD and a 37% reduction in cardiac events by alirocumab versus a 13% reduction in the group with the lowest GRS. These results also indicated that

the GRS was superior to conventional risk factor stratification in identifying those at greater risk who would benefit most from cholesterol-lowering medications.

LIFESTYLE CHANGES AND DRUG THERAPY REDUCE GENETIC RISK FOR CAD

Tikkanen et al. (130) risk-stratified for CAD using a GRS to assess the effect of physical activity on the risk of cardiac events in a large sample size of 468,095 individuals from the UK Biobank. The physical activity performed was that of the handgrip for 3 s and cardiorespiratory fitness determined by oxygen consumption on a stationary bike. The highest genetic risk group, as determined by the GRS, had the most benefit from cardiorespiratory fitness, exhibiting a 49% lower risk for CAD.

A prospective study was performed by Khera et al. (108) involving a sample size of 55,685 individuals. Individuals with a high GRS for CAD (20%) had a 90% higher risk of cardiac events than those with a lower GRS. The group was risk-stratified to compare a healthy lifestyle (no current smoking, no obesity, regular physical activity, and a healthy diet) to that of an unhealthy lifestyle (at least 2 unfavorable features). The group with the highest GRS and a favorable lifestyle experienced 46% fewer cardiac events than those with an unfavorable lifestyle.

The GRS was evaluated in >1 million individuals. It was shown to be more predictive and prognostic than using traditional risk factors for CAD. Using the GRS showed that lifestyle changes, such as exercise, diet, and drugs such as statins, could markedly reduce genetic risk.

LIMITATIONS OF GRS

THE INFLUENCE OF ETHNICITY AND REGIONALIZATION.

The global application of a GRS assumes the alleles-carrying risk variants have a similar frequency across different ethnic populations. This assumption is a concern because the genetic risk variants were discovered primarily in populations of European descent. A brief overview of human evolution supports a common ancestry for many disease-related variants, but not necessarily for all, particularly for ethnic groups such as Africans. It also provides the framework for adopting a cautious approach in other ethnic groups and the necessity for ongoing research.

In support of global applications is our common African ancestry, which has slowly evolved from the chimpanzee over approximately 6 million years. It is reasonable to assume the dynamics of genetic drift would provide an equilibrium of alleles across the

populations. However, this was interrupted by the large migration out of Africa approximately 100,000 years ago. This provided a bottleneck to the ebb and flow of genetic drift as a proportion of the population migrated into East and South Asia, whereas others migrated to Southern and Northern Europe. On a global basis, separate alleles may be divided into ancestral (transmitted from other primates) and derived (due to new mutations) alleles. Ancestral alleles will be older, having evolved over >100,000 years ago, whereas derived alleles would be ≤100,000 years old. It is well known that certain bottlenecks can change disease-associated alleles, such as the increased risk for cystic fibrosis in Québécois (131) or increased cardiovascular disease in descendants of the HMS Bounty (132). Polygenic diseases such as CAD are due to multiple alleles with minimal effect sizes and the different frequencies at different loci would be expected to distribute toward the mean. Globally, a similar distribution of allele frequencies (133), with the possible expectation of a reduced genetic load in the African population (134), would be anticipated. Genetic drift is expected to be less in those who stayed in Africa because there is less pressure from natural selection compared with those who moved onto the challenge of adapting to a new environment (135).

It remains to be determined whether disease-associated alleles exhibit the same pattern of distribution and frequency as non-disease-related alleles. Disease-associated alleles may influence survival and natural selection. An analysis (136) showed 69.2% of ancestral risk alleles had a higher frequency in Africans, and 64.5% of derived alleles had a higher frequency in non-Africans. Notably, 44% of all risk alleles discovered are ancestral, meaning they originated >100,000 years ago. The mean frequency of ancestral risk alleles is 9.5% higher in African populations, whereas derived risk allele frequencies are 5.4% less in African populations. Analysis of the 1000 Genome Project showed the frequency of derived alleles are similar across the continents, with Africans more likely to be heterozygous for derived alleles and non-Africans more likely to be homozygous for derived alleles (136). In assessing a GRS for a disease, it may have some validity to separate the world's population into African and non-African. Kim et al. (136) analyzed 306 GWAS loci for each continental population in the 1000 Genome Project. The highest risk alleles for disease were observed in Africans for metabolic, morphological, neoplastic, and neurological diseases. The African population had lower frequencies for risk alleles associated with cardiovascular disease. The first genetic risk variant

discovered for CAD, 9p21 (30,31), transmitted an increased relative risk for CAD of approximately 25% per copy in Europeans. In contrast, 9p21 was not associated with any risk for CAD in the African population (35–37). The lower frequency of risk alleles for cardiovascular disease is at odds with the clinical observation of a high incidence and prevalence of CAD in Africans. Thus, as we go forward using genetic risk stratification for CAD and other cardiac disorders, it must be born in mind that the risk for CAD may be misestimated in individuals of African descent. Precise genetic risk stratification for diseases may require knowing what proportion of risk alleles is ancestral versus derived.

More than 3,000 GWAS have been performed in the past decade, and the ancestry of the participants were >80% Europeans, and 14% Asians (137). There are limited data on other ethnic groups. The 9p21 risk variant discovered in Europeans is confirmed to be a risk variant in other ethnic groups, including East and South Asians, but not in Africans. Genetic risk stratifying in Africans for CAD may need an enrichment of risk variants for CAD from the African population. The GRS will have to be assessed in the various ethnic groups to determine its overall accuracy.

DATA EVALUATION AND REPLICATION PERFORMED IN THE SAME POPULATION. Most of the GRSs have been developed with the test set and the validated set selected from the same population (64,65). Any confounding factors, such as ethnicity, have not been detected. It is more desirable for a score to be determined from a test set and validation in an independent population.

COMPREHENSIVE RISK STRATIFICATION COUPLED TO COMPREHENSIVE THERAPY

The recent discovery of genetic risk variants has the potential to predict those at higher risk for ischemic heart disease. The genetic prediction has the same accuracy at birth as it does anytime throughout life. Genetic risk is innate to individuals' DNA and does not change in a lifetime. It is possible in the future, we will discover there are exceptions due to epigenetic changes, whereby expression of certain variants may be altered after birth. This makes it possible to risk stratify for CAD in young males, even as teenagers, to determine who would benefit most from primary preventative measures. Similarly, among the premenopausal female population, genetic risk stratification could be performed and therapy initiated before the onset of CAD. The schema indicated in **Figure 2** would be 1 approach to detect those at greatest risk and who would be more likely to benefit from

primary prevention. Use of the GRS and conventional risk stratification could revolutionize primary prevention in both men and women. It is worthy to note that the mechanisms mediating genetic risk are primarily unknown. Elucidating the mechanisms will lead to new drug targets and ultimately novel therapies specific for genetic risk. This will couple comprehensive risk stratification and comprehensive therapy for primary and secondary prevention of CAD.

A FUTURE PROSPECTIVE

CAD is the number 1 killer in the world and primary and secondary prevention have been shown to be effective. It is also known that approximately 50% of individuals in the United States who live out a normal lifespan will experience a cardiac event. It has been proven that lifestyle changes and cholesterol-lowering therapy are safe and effective as primary preventive measures. CAD originates early in life and slowly accumulates over the decades, with clinical manifestations such as myocardial infarction and sudden cardiac death peaking in the sixth decade in men and the seventh decade in women. To prevent these events, we must initiate prevention early in life. Conventional risk factors for CAD are often not present until the sixth or seventh decade and do not correlate with preclinical coronary atherosclerosis (138). In contrast, the GRS, which can be determined any time after birth, shows good correlation with preclinical CAD. The GRS is a very promising approach. There is much to be done to improve the GRS. We do not know the causal sequences, and the SNPs simply serve as proxies for the risk variant sequence. There are many more risk variants yet to be discovered. There are risk variants specific for ethnic groups, which can be imputed onto existing microarrays, as recently shown by Wang *et al.* (139) for South Asians.

The GRS must be pursued because it offers tremendous hope for a disease that exists throughout the world and can be prevented. Genetic testing and therapy are available everywhere and is inexpensive. Although there are arguments for and against the GRS being used as a clinical tool, we believe the time is now. Plasma cholesterol as a risk factor was recognized in the 1960s and received clinical application as total cholesterol. Clinical application refined it as a risk factor leading to the recognition of LDL-C, HDL-C, and triglycerides. We believe that current use of the GRS in a clinical setting will enhance its development, enable early preventive measures for the 10% to 20% at highest genetic risk, and initiate appropriate primary prevention in the premenopausal woman. Acquisition of more knowledge

and experience will gradually include more of the population at intermediate risk for CAD.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

The authors have received research funding from the Canadian Institutes of Health Research, CIHR no. MOP P82810, the Canada Foundation for Innovation, the CFI, no. 11966, and the Dignity Health

Foundation. The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr. Robert Roberts, Department of Medicine, University of Arizona, 500 West Thomas Road #480, Phoenix, Arizona, 85013, USA.
E-mail: Robert.Roberts@dignityhealth.org.

REFERENCES

1. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J III. Factors of risk in the development of coronary heart disease—six-year follow-up experience: the Framingham study. *Ann Internal Med* 1961;55:33-50.
2. Gluckman TJ, Kovacs RJ, Stone NJ, Damalas D, Mullen JB, Oetgen WJ. The ASCVD risk estimator app: from concept to the current state. *J Am Coll Cardiol* 2016;67:350-2.
3. The Multiple Risk Factor Intervention Trial Group. Statistical design considerations in the NHLI multiple risk factor intervention trial (MRFIT). *J Chronic Dis* 1977;30:261-75.
4. Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994;344:1383-9.
5. Shepherd J, Cobbe SM, Ford I, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Engl J Med* 1995;333:1301-8.
6. Cholesterol Treatment Trialists' (CTT) Collaboration, Baigent C, Blackwell L, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet* 2010;376:1670-81.
7. Collins R, Reith C, Emberson J, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet* 2016;388:2532-61.
8. Lloyd-Jones DM, Larson MG, Beiser A, Levy D. Lifetime risk of developing coronary heart disease. *Lancet* 1999;353:89-92.
9. American Heart Association. Heart and Stroke Statistical Update. Dallas, TX: American Heart Association, 2000.
10. Myers RH, Kiely DK, Cupples LA, Kannel WB. Parental history is an independent risk factor for coronary artery disease: the Framingham study. *Am Heart J* 1990;120:963-9.
11. Marenberg ME, Risch N, Berkman LF, et al. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med* 1994;330:1141-6.
12. Chan L, Boerwinkle E. Gene-environment interactions and gene therapy in atherosclerosis. *Cardiol Rev* 1994;2:130-7.
13. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004;431:931-45.
14. International HapMap Consortium, Frazer KA, Ballinger DG, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449:851-61.
15. Roberts R. New gains in understanding coronary artery disease, interview with Dr Robert Roberts. *Affymetrix Microarray Bull* 2007;3:1-4.
16. Gibbs R, Belmont J, Hardenbol P, et al. The International HapMap Project. *Nature* 2003;426:789-96.
17. Lander E, Linton L, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860-921.
18. Marian AJ, Brugada R, Roberts R. Chapter 55: Mendelian basis of congenital and other cardiovascular disease. In: Fuster V, Harrington RA, Eapen ZJ, et al., editors. *Hurst's The Heart*. 14th Ed. New York: McGraw Hill, 2017.
19. Hejtmancik JF, Brink PA, Towbin J, et al. Localization of gene for familial hypertrophic cardiomyopathy to chromosome 14q1 in a diverse US population. *Circulation* 1991;83:1592-7.
20. Jarcho JA, McKenna W, Pare JA, et al. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. *N Engl J Med* 1989;321:1372-8.
21. Gollob MH, Green MS, Tang AS-L, et al. Identification of a gene responsible for familial Wolff-Parkinson-White syndrome. *N Engl J Med* 2001;344:1823-31.
22. Centers for Mendelian Genomics. Mendelian traits by the numbers. Available at: <http://mendelian.org/mendelian-traits-numbers>. Accessed October 3, 2020.
23. Geisterfer-Lowrance AAT, Christe M, Conner DA, et al. A mouse model for familial hypertrophic cardiomyopathy. *Science* 1996;272:731-4.
24. Sidhu JS, Rajawat YS, Rami TG, et al. Transgenic mouse model of ventricular preexcitation and atrioventricular reentrant tachycardia induced by an AMP-activated protein kinase loss-of-function mutation responsible for Wolff-Parkinson-White syndrome. *Circulation* 2005;111:21-9.
25. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;6:95-108.
26. Altshuler D, Donnelly P. A haplotype map of the human genome. *Nature* 2005;437:1299-320.
27. Auton A, Abecasis G, Altshuler D, et al. A global reference for human genetic variation. *Nature* 2015;526:68-74.
28. Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005;308:385-9.
29. Buniello A, MacArthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2018;47:D1005-12.
30. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;316:1488-91.
31. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 2007;316:1491-3.
32. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661-78.
33. Samani NJ, Erdmann J, Hall AS, et al. Genome wide association analysis of coronary artery disease. *N Engl J Med* 2007;357:443-53.
34. Assimes TL, Roberts R. Genetics: implications for prevention and management of coronary artery disease. *J Am Coll Cardiol* 2016;68:2797-818.
35. Barbalic M, Reiner AP, Wu C, et al. Genome-wide association analysis of incident coronary heart disease (CHD) in African Americans: a short report. *PLoS Genet* 2011;7:e1002199.
36. Franceschini N, Hu Y, Reiner AP, et al. Prospective associations of coronary heart disease loci in African Americans using the MetaboChip: the PAGE study. *PLoS One* 2014;9:e113203.
37. Lettre G, Palmer CD, Young T, et al. Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: the NHLBI CARE project. *PLoS Genet* 2011;7:e1001300.
38. Myocardial Infarction Genetics Consortium, Kathiresan S, Voight BF, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet* 2009;41:334-41.
39. Erdmann J, Grosshennig A, Braund PS, et al. New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat Genet* 2009;41:280-2.
40. Trégouët D-A, König IR, Erdmann J, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat Genet* 2009;41:283-5.
41. Gudbjartsson DF, Bjornsdottir US, Halapi E, et al. Sequence variants affecting eosinophil

- numbers associate with asthma and myocardial infarction. *Nat Genet* 2009;41:342-7.
- 42.** The Coronary Artery Disease (CAD) Genetics Consortium. a genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet* 2011; 43:339-44.
- 43.** Preuss M, König IR, Thompson JR, et al., CARDIoGRAM Consortium. Design of the Coronary ARtery Disease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: a genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ Cardiovasc Genet* 2010;3: 475-83.
- 44.** Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; 273:1516.
- 45.** Erdmann J, Kessler T, Munoz Venegas L, Schunkert H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. *Cardiovasc Res* 2018;114:1241-57.
- 46.** CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet* 2013;45:25-33.
- 47.** Nelson C, Goel A, Butterworth A, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet* 2017;49:1385-91.
- 48.** Navar-Boggan AM, Peterson ED, D'Agostino RBS, Neely B, Sniderman AD, Pencina MJ. Hyperlipidemia in early adulthood increases long-term risk of coronary heart disease. *Circulation* 2015;131:451-8.
- 49.** Swiger KJ, Martin SS, Blaha MJ, et al. Narrowing sex differences in lipoprotein cholesterol subclasses following mid-life: the very large database of lipids (VLDL-10B). *J Am Heart Assoc* 2014;3:e000851.
- 50.** Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet* 2018;19:581-90.
- 51.** American College of Cardiology, American Heart Association. ASCVD Risk Estimator. Available at: https://tools.acc.org/dl/ascvd_risk_estimator/index.html#!/calculate/estimator. Accessed April 20, 2020.
- 52.** Ference BA, Majeed F, Pennumettha R, Flack JM, Brook RD. Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or both: a 2 × 2 factorial Mendelian randomization study. *J Am Coll Cardiol* 2015;65:1552-61.
- 53.** Ference BA, Graham I, Tokgozoglul L, Catapano AL. Impact of lipids on cardiovascular health. *J Am Coll Cardiol* 2018;72:1141.
- 54.** Roberts R. PCSK9 inhibition—a new thrust in the prevention of heart disease: genetics does it again. *Can J Cardiol* 2013;29:899-901.
- 55.** Brown MS, Goldstein JL. Familial hypercholesterolemia: a genetic defect in the low-density lipoprotein receptor. *N Engl J Med* 1976;294: 1386-90.
- 56.** Abifadel M, Varret M, Rabès JP, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 2003;34:154-6.
- 57.** Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet* 2005;37:161-5.
- 58.** Stein EA, Mellis S, Yancopoulos GD, et al. Effect of a monoclonal antibody to PCSK9 on LDL cholesterol. *N Engl J Med* 2012;366:1108-18.
- 59.** Köster A, Chao YB, Mosior M, et al. Transgenic angiotensin-like (ANGPTL)4 overexpression and targeted disruption of *angptl4* and *angptl3*: regulation of triglyceride metabolism. *Endocrinology* 2005;146:4943-50.
- 60.** Mitsuru S, Matsuda M, Hiroaki Y, et al. Angiotensin-like protein3 regulates plasma HDL cholesterol through suppression of endothelial lipase. *Arterioscler Thromb Vasc Biol* 2007;27: 366-72.
- 61.** Fujimoto K, Koishi R, Shimizugawa T, Ando Y. *Angptl3*-null mice show low plasma lipid concentrations by enhanced lipoprotein lipase activity. *Exp Anim* 2006;55:27-34.
- 62.** Dewey FE, Gusarova V, Dunbar RL, et al. Genetic and pharmacologic inactivation of ANGPTL3 and cardiovascular disease. *N Engl J Med* 2017; 377:211-21.
- 63.** Athyros VG, Katsiki N, Dimakopoulou A, Patoulias D, Alataki S, Doumas M. Drugs that mimic the effect of gene mutations for the prevention or the treatment of atherosclerotic disease: from PCSK9 inhibition to ANGPTL3 inactivation. *Curr Pharm Des* 2018;24:3638-46.
- 64.** Raal FJ, Rosenson RS, Reeskamp LF, et al. Evinacumab for homozygous familial hypercholesterolemia. *N Engl J Med* 2020;383:711-20.
- 65.** Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* 2016;48:245-52.
- 66.** Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* 2016; 48:481-7.
- 67.** Jarinova O, Stewart AFR, Roberts R, et al. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler Thromb Vasc Biol* 2009;29:1671-7.
- 68.** Holdt LM, Beutner F, Scholz M, et al. ANRIL expression is associated with atherosclerosis risk at chromosome 9p21. *Arterioscler Thromb Vasc Biol* 2010;30:620-7.
- 69.** Pasmant E, Laurendeau I, Héron D, Vidaud M, Vidaud D, Bièche I. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res* 2007;67:3963-9.
- 70.** Leeper NJ, Radesdana A, Kojima Y, et al. Loss of CDKN2B promotes p53-dependent smooth muscle cell apoptosis and aneurysm formation. *Arterioscler Thromb Vasc Biol* 2013;33:e1-10.
- 71.** Pilbrow AP, Folkersen L, Pearson JF, et al. The chromosome 9p21.3 coronary heart disease risk allele is associated with altered gene expression in normal heart and vascular tissues. *PLoS One* 2012; 7:e39574.
- 72.** Kim JB, Deluna A, Mungrue IN, et al. Effect of 9p21.3 coronary artery disease locus neighboring genes on atherosclerosis in mice. *Circulation* 2012; 126:1896-906.
- 73.** Folkersen L, Kyriakou T, Goel A, et al. Relationship between CAD risk genotype in the chromosome 9p21 locus and gene expression. Identification of eight new ANRIL splice variants. *PLoS One* 2009;4:e7677.
- 74.** Hamsten A, Eriksson P. Quest for genes and mechanisms linking the human chromosome 9p21.3 locus to cardiovascular disease. *Circulation* 2012; 126:1815-7.
- 75.** Visel A, Zhu Y, May D, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* 2010;464:409-12.
- 76.** Harismendy O, Notani D, Song X, et al. 9p21 DNA variants associated with coronary artery disease impair interferon- γ signalling response. *Nature* 2011;470:264-8.
- 77.** Almontashiri NAM, Fan M, Cheng BLM, Chen H-H, Roberts R, Stewart AFR. Interferon- γ activates expression of p15 and p16 regardless of 9p21.3 coronary artery disease risk genotype. *J Am Coll Cardiol* 2013;61:143-7.
- 78.** Erridge C, Gracey J, Braund PS, Samani NJ. The 9p21 locus does not affect risk of coronary artery disease through induction of type 1 interferons. *J Am Coll Cardiol* 2013;62:1376-81.
- 79.** Lo Sardo V, Chubukov P, Ferguson W, et al. Unveiling the role of the most impactful cardiovascular risk locus through haplotype editing. *Cell* 2018;175:1796-810.e20.
- 80.** Helgadottir A, Thorleifsson G, Magnusson KP, et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet* 2008;40:217-24.
- 81.** Dandona S, Stewart AFR, Chen L, et al. Gene dosage of the common variant 9p21 predicts severity of coronary artery disease. *J Am Coll Cardiol* 2010;56:479-86.
- 82.** Reilly MP, Li M, He J, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet* 2011;377:383-92.
- 83.** Ardissino D, Berzuini C, Merlini PA, et al. Influence of 9p21.3 genetic variants on clinical and angiographic outcomes in early-onset myocardial infarction. *J Am Coll Cardiol* 2011;58:426-34.
- 84.** Chan K, Patel RS, Newcombe P, et al. Association between the chromosome 9p21 locus and angiographic coronary artery disease burden: a collaborative meta-analysis. *J Am Coll Cardiol* 2013;61:957-70.
- 85.** Musunuru K, Strong A, Frank-Kamenetsky M, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* 2010; 466:714-9.
- 86.** Wang X, Raghavan A, Peters DT, Pashos EE, Rader DJ, Musunuru K. Interrogation of the

- atherosclerosis-associated SORT1 (Sortilin 1) locus with primary human hepatocytes, induced pluripotent stem cell-hepatocytes, and locus-humanized mice. *Arterioscler Thromb Vasc Biol* 2018;38:76-82.
87. Kjolby M, Andersen OM, Breiderhoff T, et al. Sort1, Encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. *Cell Metab* 2010;12:213-23.
88. Strong A, Ding Q, Edmondson AC, et al. Hepatic sortilin regulates both apolipoprotein B secretion and LDL catabolism. *J Clin Invest* 2012;122:2807-16.
89. Gustafsen C, Kjolby M, Nyegaard M, et al. The hypercholesterolemia-risk gene SORT1 facilitates PCSK9 secretion. *Cell Metab* 2014;19:310-8.
90. Beaudoin M, Gupta RM, Won H-H, et al. Myocardial infarction-associated SNP at 6p24 interferes with MEF2 binding and associates with PHACTR1 expression levels in human coronary arteries. *Arterioscler Thromb Vasc Biol* 2015;35:1472-9.
91. Gupta RM, Hadaya J, Trehan A, et al. A genetic variant associated with five vascular diseases is a distal regulator of endothelin-1 gene expression. *Cell* 2017;170:522-33.e15.
92. Pu X, Xiao Q, Kiehl S, et al. ADAMTS7 cleavage and vascular smooth muscle cell migration is affected by a coronary-artery-disease-associated variant. *Am J Human Genet* 2013;92:366-74.
93. Kessler T, Zhang L, Liu Z, et al. ADAMTS-7 Inhibits re-endothelialization of injured arteries and promotes vascular remodeling through cleavage of thrombospondin-1. *Circulation* 2015;131:1191-201.
94. Bauer RC, Tohyama J, Cui J, et al. Knockout of Adamts7, a novel coronary artery disease locus in humans, reduces atherosclerosis in mice. *Circulation* 2015;131:1202-13.
95. Gamazon ER, Segrè AV, van de Bunt M, et al. Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nat Genet* 2018;50:956-67.
96. Libby P. Interleukin-1 beta as a target for atherosclerosis therapy: biological basis of CANTOS and beyond. *J Am Coll Cardiol* 2017;70:2278-89.
97. Roberts R. Mendelian randomization studies promise to shorten the journey to FDA approval. *J Am Coll Cardiol Basic Trans Science* 2018;3:690-703.
98. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomisation study. *Lancet* 2012;380:572-80.
99. Lincoff AM, Nicholls SJ, Riesenmeyer JS, et al. Evacetrapib and cardiovascular outcomes in high risk vascular disease. *N Engl J Med* 2017;376:1933-42.
100. Kawashiri MA, Tada H, Nomura A, Yamagishi M. Mendelian randomization: its impact on cardiovascular disease. *J Cardiol* 2018;72:307-13.
101. Barter PJ, Caulfield M, Eriksson M, et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med* 2007;357:2109-22.
102. Schwartz GG, Olsson AG, Abt M, et al. Effects of dalcetrapib in patients with a recent acute coronary syndrome. *N Engl J Med* 2012;367:2089-99.
103. HPS3/TIMI55-REVEAL Collaborative Group, Bowman L, Hopewell JC, Chen F, et al. Effects of anacetrapib in patients with atherosclerotic vascular disease. *N Engl J Med* 2017;377:1217-27.
104. Nomura A, Won HH, Khera AV, et al. Protein truncating variants at the cholesteryl ester transfer protein gene and risk for coronary heart disease. *Circ Res* 2017;121:81-8.
105. O'Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. *Nat Genet* 2018;50:1728-34.
106. Mega JL, Stitziel NO, Smith JG, 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;385:2264-71.
107. Roberts R, Campillo A, Schmitt M. Prediction and management of CAD risk based on genetic stratification. *Trends Cardiovasc Med* 2020;30:328-34.
108. Khera AV, Emdin CA, Drake I, et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N Engl J Med* 2016;375:2349-58.
109. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50:1219-24.
110. Khera AV, Chaffin M, Zekavat SM, et al. Whole-genome sequencing to characterize monogenic and polygenic contributions in patients hospitalized with early-onset myocardial infarction. *Circulation* 2019;139:1593-602.
111. Chatterjee N, Shi J, García-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet* 2016;17:392-406.
112. 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;51:584-91.
113. Schurz H, Müller SJ, van Helden PD, et al. Evaluating the accuracy of imputation methods in a five-way admixed population. *Front Genet* 2019;10:34.
114. Hancock DB, Levy JL, Gaddis NC, et al. Assessment of genotype imputation performance using 1000 Genomes in African American studies. *PLoS One* 2012;7:e50610.
115. Taliun D, Harris DN, Kessler MD, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. *bioRxiv* 2019:563866.
116. Janssens ACJW. Validity of polygenic risk scores: are we measuring what we think we are? *Human Mol Genet* 2019;28:R143-50.
117. Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann Intern Med* 2009;150:65-72.
118. Davies RW, Dandona S, Stewart AF, et al. Improved prediction of cardiovascular disease based on a panel of single nucleotide polymorphisms identified through genome-wide association studies. *Circ Cardiovasc Genet* 2010;3:468-74.
119. Tezak Z, Kondratovich MV, Mansfield E. US FDA and personalized medicine: in vitro diagnostic regulatory perspective. *Per Med* 2010;7:517-30.
120. Natarajan P, Young R, Stitziel NO, 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;135:2091-101.
121. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Davies C, Godwin J, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* 2011;378:771-84.
122. Abraham G, Havulinna AS, Bhalala OG, et al. Genomic prediction of coronary heart disease. *Eur Heart J* 2016;37:3267-78.
123. Inoué M, Abraham G, Nelson CP, et al. Genomic risk prediction of coronary artery disease in 480,000 adults. *J Am Coll Cardiol* 2018;72:1883-93.
124. Vilhjálmsdóttir BJ, Yang J, Finucane HK, et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet* 2015;97:576-92.
125. Zhang Y, Qi G, Park J-H, Chatterjee N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. *Nat Genet* 2018;50:1318-26.
126. Mosley JD, Gupta DK, Tan J, et al. Predictive accuracy of a polygenic risk score compared with a clinical risk score for incident coronary heart disease. *JAMA* 2020;323:627-35.
127. Elliott J, Bodinier B, Bond TA, et al. Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease. *JAMA* 2020;323:636-45.
128. Marston NA, Kamanu FK, Nordio F, et al. Predicting benefit from evolocumab therapy in patients with atherosclerotic disease using a genetic risk score. *Circulation* 2020;141:616-23.
129. Damask A, Gabriel SP, Schwartz GG, et al. Patients with high genome-wide polygenic risk scores for coronary artery disease may receive greater clinical benefit from alirocumab treatment in the ODYSSEY OUTCOMES trial. *Circulation* 2020;141:624-36.
130. Tikkanen E, Gustafsson S, Ingelsson E. Associations of fitness, physical activity, strength, and genetic risk with cardiovascular disease: longitudinal analyses in the UK Biobank study. *Circulation* 2018;137:2583-91.
131. Laberge A-M, Michaud J, Richter A, et al. Population history and its impact on medical genetics in Quebec. *Clin Genet* 2005;68:287-301.
132. Benton MC, Stuart S, Bellis C, et al. "Mutiny on the Bounty": the genetic history of Norfolk Island reveals extreme gender-biased admixture.

- Investig Genet 2015;6:11 [published correction appears in Investig Genet 2015;6:12].
- 133.** Lohmueller KE. The distribution of deleterious genetic variation in human populations. *Curr Opin Genet Devel* 2014;29:139-46.
- 134.** Henn BM, Botigue LR, Peischl S, et al. Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proc Natl Acad Sci U S A* 2016;113:E440-9.
- 135.** Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc Natl Acad Sci U S A* 2005;102:15942.
- 136.** Kim MS, Patel KP, Teng AK, Berens AJ, Lachance J. Genetic disease risks can be misestimated across global populations. *Genome Biol* 2018;19:179.
- 137.** Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature* 2016;538:161-4.
- 138.** Baber U, Mehran R, Sartori S, Schoos MM, Sillesen H, Muntendam P, et al. Prevalence, impact, and predictive value of detecting sub-clinical coronary and carotid atherosclerosis in asymptomatic adults: the BiImage study. *J Am Coll Card* 2015;65:1065-74.
- 139.** Wang M, Menon R, Mishra S, et al. Validation of a genome-wide polygenic score for coronary artery disease in South Asians. *J Am Coll Cardiol* 2020;76:703-14.
- 140.** Davies RW, Wells GA, Stewart AF, et al. A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. *Circ Cardiovasc Genet* 2012;5:217-25.
- 141.** Lu X, Wang L, Chen S, et al. Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. *Nat Genet* 2012;44:890-4.
- 142.** Wang F, Xu C-Q, He Q, et al. Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. *Nat Genet* 2011;43:345-9.
- 143.** Hirokawa M, Morita H, Tajima T, et al. A genome-wide association study identifies PLCL2 and AP3D1-DOT1L-SF3A2 as new susceptibility loci for myocardial infarction in Japanese. *Eur J Hum Genet* 2015;23:374-80.

KEY WORDS cardiovascular genetics, coronary artery disease, genetic risk score for CAD, genome-wide association studies, prevention of CAD