Cardiovascular disease: The rise of the genetic risk score

Joshua W. Knowles, Euan A. Ashley*

Center for Inherited Cardiovascular Disease, Stanford University, Stanford, California, United States of America

* euan@stanford.edu

Summary points

• Use of risk factors for decision-making in cardiovascular disease has a long history in medicine.
• Early attempts to augment traditional risk factors with genetic risk scores were hampered by too little understanding of the genetic basis of complex cardiovascular disease.
• Newer studies based on hundreds of thousands of people and millions of genetic variants indicate that genetic risk scores can now outperform traditional risk factors in risk prediction.
• We propose the time has come to incorporate genetic risk scores into clinical practice.
• Studies should focus on the most appropriate way to do this to maximize benefit for our patients.

“[However,] epidemiologic information has accumulated which now allows the physician to recognize certain characteristics of increased risk in patients he [sic] sees in his practice. Some of these characteristics have been convincingly demonstrated, others are still under investigation. More precise identification will undoubtedly be possible in the future.”—William Kannel, Director, Framingham Heart Study [1]

In a classic paper [1], Kannel reported the early results of the longitudinal Framingham Heart Study, demonstrating the identification of factors (for which he coined the term “risk factors”) that “proceed the development of overt coronary heart disease in humans”. Later, the Framingham Risk Score was formalized to include age, sex, diabetes, smoking status, total cholesterol, high-density lipoprotein (HDL) cholesterol, and blood pressure [2], providing a framework for cardiovascular disease (CVD) risk assessment to which all others are compared. Intuitively, the burden of these risk factors accumulates over time (e.g., pack years of smoking or years of hypertension), and some newer risk models allow input of risk factor data from multiple time points [3].

While repeated measurement of CVD biomarkers such as total cholesterol may improve risk prediction, lifelong exposure to CVD risk factors is better captured by genetic susceptibility [4]. Thus, the quest to improve risk prediction for CVD has naturally come to focus on the development of genetic risk scores [5]. This has only been possible because of robust, replicable
findings from genome-wide association studies (GWAS) in extremely large cohorts [6,7]. Early genetic risk scores, based on relatively few single-nucleotide variants, showed a consistent ability to identify those in the highest strata of risk [8,9], with some improvement in “reclassification” of risk. This interest in risk prediction led to an increased focus on the tools for judging utility with a return to prominence of metrics like the C statistic and the proposal of newer metrics, such as the integrated discrimination improvement (IDI) and the net reclassification index (NRI), specifically aimed at judging the merit of adding new factors (i.e., genetic markers) to existing scores [10]. Although the focus of many hundreds of articles, these newer metrics have been criticized for too highly rating poorly fitted risk models and for showing improvement in models with a new biomarker that adds no new information [11–13]. Around 2009, there was also criticism of the common variant studies for failing to find “missing” heritability [14], and the lack of robust risk prediction from discovered variants fed into an overall narrative that genomics was underperforming relative to its hype [15].

Yet it was clear that this was chiefly a problem of study size. While human clinical trials have historically recruited hundreds or thousands of individuals, the genomics community realized that studies with hundreds of thousands to millions of participants would be required to provide the power necessary to fuel discovery of the larger proportion of heritability. This realization ushered in a new era of data sharing. Today, as a result of large-scale collaboration, meta-analysis, and the emergence of national projects such as the United Kingdom Biobank [16,17], there are GWAS of common variants drawing on more than 1 million individuals [18]. Such studies, as modeling would predict [19], are beginning to demonstrate that genetic factors provide robust and powerful risk estimation across diseases that is additive to traditional risk factors [20–22]. Indeed, as the idea that rare variation (synthetic or otherwise) could explain much of the missing heritability of common disease fell in favor [23] and out again [24], the realization dawned that still-too-small studies and overzealous correction of multiple testing had left significant signal in the noise. This stimulated the idea of using a much broader array of variants in a polygenic score. Khera and colleagues [21] used 6.6 million variants, while Inouye et al. [22] used 1.7 million variants as predictors, and both studies demonstrate the ability to identify a group in the upper echelon of genetic risk with a hazard of greater than 4. In particular, the meta genetic risk score (metaGRS) had a higher C-index for incident coronary artery disease than any single traditional risk factor, including smoking, diabetes, hypertension, and body mass index [22]. In that study, the addition of the genetic score to a combination of conventional risk factors increased the C-index by 3.7%. Drawing on multiple interacting mechanisms, it is not surprising that much of the signal of genetic risk scores overlaps traditional risk factors and mechanisms. But the potential of genetic approaches is emphasized by a component of independence, illustrated by the ability to improve on a C-index derived from conventional risk factors alone.

Thus, despite early criticism, most recent genetic risk scores have demonstrated significant improvements in performance for risk prediction in CVD [25]. Given these advantages, it is reasonable to ask whether such scores have the potential to significantly improve multimorbidity assessment for diseases where risk assessment has been routine, especially as the costs of genome-wide genotyping now fall below US$100 per person. Indeed, because genotyping chips survey common variants across the entire genome, reflecting risk for hundreds of conditions besides CVD, it is possible to simultaneously predict risk of multiple diseases with a single “test.” Cardiometabolic scores [26,27] can be combined [28], or estimates can be made, for dozens of diseases, including—as we reported [29]—from whole-genome sequencing.

A critical aspect of the utility of any predictive score is its impact on clinical management. Since Kannel’s coining of the term, risk prediction has been leveraged for management decisions in medicine. Recent guidelines on hypertension [30] and hypercholesterolemia [31,32]
emphasize the role of risk estimation in therapeutic decision-making, particularly for patients with intermediate risk. Yet, whereas cholesterol levels can be lowered through therapy and individuals can stop smoking, what is the specific “answer” to a high genetic risk score? Khera and colleagues [20] provided one answer in a study demonstrating that lifestyle factors are capable of abrogating genetic risk, elegantly underlining the universality of the benefits of diet and exercise while providing a defense for the concern that patients who discover they are at high genetic risk will view that deterministically and be less inclined to lifestyle change (something that has always remained hypothetical [33]). Another recent study has shown that genetic risk for high blood pressure can be mitigated by a healthy lifestyle [34]. Additional data are needed to address the converse concern that individuals shown to have a “protective” genetic background will feel less inclined to maintain a healthy lifestyle. In this regard, the best outcomes are in those individuals that have both a favorable genetic susceptibility and healthy lifestyle [20].

So if polygenic risk scores now outperform traditional risk factors in univariate prediction, augment the C statistics of traditional risk factors taken as a whole, can be implemented for minimal cost, and are targets for intervention, is it not time to incorporate them into clinical practice?

Despite the increasingly well-demonstrated value of the genetic risk scores, few studies have focused on the practical aspects of incorporating scores into clinical practice. Although the benefit of delivering traditional risk factors to physicians and patients has never itself been tested in a randomized controlled trial, the traditional risk score, based on data already gathered, is effectively free to the healthcare system. While there remains an additional cost for genetic scores, albeit modest, it is reasonable to require an outcome benefit to be demonstrated before arguing for adding to medical expenditure. In a small pilot randomized controlled study, we showed the feasibility of delivery of a genetic risk score in a clinical environment [35–37]. While we did not demonstrate that the score led to an improvement in patient adherence to guideline-based therapeutic advice, others have shown that the incorporation of a genetic risk score into clinical care may increase statin usage (mostly through increased statin prescriptions) [38]. We would note that similar challenges in changing behavior despite improving risk prediction have been reported in studies of coronary calcium, carotid ultrasound, and coronary computed tomography (CT) scans [39–41]. However, as we become more sophisticated in delivery of information to “activate” positive behavioral changes, these results are expected to improve. Digital approaches may offer one avenue for improvement: for example, there are now smartphone studies of cardiovascular risk that incorporate genotype data, as well as studies focused specifically on returning genetic risk scores to participants [42–44].

In an additional wrinkle, if the genetic risk score could be calculated from preexisting data, the cost to the healthcare system would be zero, and few would argue that we should not look to refine traditional scores with genetic data. The highly computable nature of genotype data makes for straightforward implementation and future refinement of genetic risk scores when more data become available [45]. Indeed, the ability to create scores across multiple diseases was attractive for direct-to-consumer genetic testing companies who started offering such estimates for multiple diseases and traits many years ago. Early versions received technical criticism based on the small numbers of variants used and the variation between providers in the creation and interpretation of scores. However, this technical criticism was secondary to more general uncertainty over the direct-to-consumer model [46]. Today, with increasing interest from the public and increasing acceptance—at least in the United States from the Food and Drug Administration (FDA)—of consumer-focused tests, the environment is primed for delivery and testing of multimodal risk scores for millions of individuals through direct-to-
consumer services utilizing laboratories accredited under the Centers for Medicare and Medicaid Services Clinical Laboratory Improvement Amendments (CLIA) standard. Healthcare systems and academic clinicians should work together with these companies to ensure standards and transparency in the safe and effective translation of these data for the public good [47].

We believe there are strong reasons to now consider incorporation of genetic risk scores into clinical practice. But questions remain. Since genetic information is viewed as more sensitive than that of other risk factors and since genetic risk does not result from an individual choice, some countries have chosen to separately protect genetic information from discrimination by health insurers or employers. The US Genetic Information Non-Discrimination Act of 2008 [48] includes both of those protections but excludes protection from life insurance discrimination. As such, before testing for a genetic risk score, individuals should receive education beyond that which a treating physician or nurse might be comfortable delivering. The scale of common disease means that the genetic counselor workforce could not meet the demand of delivering counseling for common disease risk scores. Brief video education has, however, been shown to be engaging and compelling, even for much more complex concepts in genetics [49]. Decision support would also be required for physicians and nurses incorporating scores into clinical management. Another challenge that has existed since the earliest use of risk factors in clinical medicine is that of unmeasured factors. A good prognostic score produces a prediction that, on a population level, has acceptable test characteristics. It cannot, however, speak to unmeasured factors in the individual. In the genetic era, this is most relevant for rare variation. It is possible, for example, for an individual to have a common variant risk score that places them in the lowest quintile for risk, but for that individual also to harbor a rare variant of large magnitude in, for example, the gene LDLR that overwhelms the common risk and places them instead in the highest quintile—or even gives them a Mendelian disease. The most proximal answer to this issue is education: unmeasured factors are not a challenge specific to genetic risk scores. The more distant answer—though one with its own nuanced challenge—is to use genome sequencing, a future imagined almost a decade ago [28,29,50].

In conclusion, through collaboration and data sharing, genetic studies of common diseases now allow genetic risk scores that predict future diseases better than traditional risk factors. As millions around the world already have this data in hand, and as the cost of generating this data falls further towards the cost of a daily cup of coffee for just one week, we propose that the time has finally come to build for the testing and incorporation of genetic risk scores into clinical practice.

References

1. Kannel WB. Factors of Risk in the Development of Coronary Heart Disease—Six-Year Follow-up Experience: The Framingham Study. Ann Intern Med. 1961; 55: 33. PMID: 13751193
2. Wilson PW, D’Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998; 97: 1837–1847. PMID: 9603539
3. 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: 350–352. https://doi.org/10.1016/j.jacc.2015.10.068 PMID: 26796407
4. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet. 2012; 380: 572–580. https://doi.org/10.1016/S0140-6736(12)60312-2 PMID: 22607825
5. Morrison AC, Bare LA, Chambliss LE, Ellis SG, Malloy M, Kane JP, et al. Prediction of coronary heart disease risk using a genetic risk score: the Atherosclerosis Risk in Communities Study. Am J Epidemiol. 2007; 166: 28–35. https://doi.org/10.1093/aje/kwm060 PMID: 17443022
6. Liu C, Kraja AT, Smith JA, Brody JA, Franceschini N, Bis JC, et al. Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. Nat Genet. 2016; 48: 1162–1170. https://doi.org/10.1038/ng.3660 PMID: 27618448
7. Preuss M, König IR, Thompson JR, Erdmann J, Absher D, Assimes TL, et al. Design of the Coronary ARtery Disease Genome-Wide Replication And Meta-Analy sis (CARDIoGRAM) StudyClinical Perspec- tive: A Genome-Wide Association Meta-analysis Involving More Than 22 000 Cases and 60 000 Con- trols. Circulation: Genomic and Precision Medicine. 2010; 3: 475–483.

8. Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. Lancet. 2010; 376: 1393–1400. https://doi.org/10.1016/S0140-6736(10)61267-6 PMID: 20971364

9. Ganna A, Magnusson PKE, Pedersen NL, de Faire U, Reilly M, Arnlov J, et al. Multilocus genetic risk scores for coronary heart disease prediction. Arterioscler Thromb Vasc Biol. 2013; 33: 2267–2272. https://doi.org/10.1161/ATVBAHA.113.301218 PMID: 23685553

10. Pencina MJ, D’Agostino RB Sr, D’Agostino RB Jr, Vasan RS. Evaluating the added predicti ve ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008; 27: 112–157.

11. Pepe MS, Fan J, Feng Z, Gerds T, Hilden J. The Net Reclassification Index (NRI): a Misleading Mea- sure of Prediction Improvement Even with Independent Test Data Sets. Stat Biosci. 2015; 7: 282–295. https://doi.org/10.1007/s12561-014-9118-0 PMID: 26504496

12. Hilden J, Gerds TA. A note on the evaluation of novel biomarkers: do not rely on integrated discrimina- tion improvement and net reclassification index. Stat Med. 2014; 33: 3405–3414. https://doi.org/10.1002/sim.5804 PMID: 23553436

13. Kerr KF, Janes H. First things first: risk model performance metrics should reflect the clinical application. Stat Med. 2017; 36: 4503–4508. https://doi.org/10.1002/sim.7414 PMID: 29156498

14. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing herita- bility of complex diseases. Nature. 2009; 461: 747–753. https://doi.org/10.1038/nature08494 PMID: 19812666

15. A Decade Later, Human Genome Project Yields Few New Cures. The New York Times. 12 Jun 2010. Available from: http://www.nytimes.com/2010/06/13/health/research/13genome.html?pagewanted=all. Accessed 19 Feb 2018.

16. Collins R. What makes UK Biobank special? Lancet. 2012; 379: 1173–1174. https://doi.org/10.1016/ S0140-6736(12)60404-8 PMID: 22463865

17. Sudlow C, Gallagher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLoS Med. 2015; 12(3): e1001779. https://doi.org/10.1371/journal.pmed.1001779 PMID: 25826379

18. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over one million people identifies 535 novel loci for blood pressure [Internet]. bioRxiv. 2017. p. 198234. https://doi.org/10.1101/198234

19. 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 and implications for the future [Internet]. bioRxiv. 2017. p. 175406. https://doi.org/10.1101/175406

20. Khera AV, Emdin CA, Drake I, Natarajan P, Bick AG, Cook NR, et al. Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary Disease. N Engl J Med. 2016; 375: 2349–2358. https://doi.org/10.1056/NEJMoa1605086 PMID: 27959714

21. Khera AV, Chaffin M, Aragam K, Emdin CA, Klarin D, Haas M, et al. Genome-wide polygenic score to identify a monogenic risk-equivalent for coronary disease [Internet]. bioRxiv. 2017. p. 218388. https://doi.org/10.1101/218388

22. Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, et al. Genomic risk prediction of coronary artery disease in nearly 500,000 adults: implications for early screening and primary prevention [Internet]. bioRxiv. 2018. https://doi.org/10.1101/250712

23. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol. 2010; 8(1): e1000294. https://doi.org/10.1371/journal.pbio.1000294 PMID: 20126254

24. 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 PMID: 27398621

25. Assimes TL, Roberts R. Genetics: Implications for Prevention and Management of Coronary Artery Disease. J Am Coll Cardiol. 2016; 68: 2797–2818. https://doi.org/10.1016/j.jacc.2016.10.039 PMID: 28007143

26. Fontaine-Bisson B, Renström F, Rolandsson O, MAGIC, Payne F, Hallmans G, et al. Evaluating the discriminative power of multi-trait genetic risk scores for type 2 diabetes in a northern Swedish population. Diabetologia. 2010; 53: 2155–2162. https://doi.org/10.1007/s00125-010-1792-y PMID: 20571754
27. de Miguel-Yanes JM, Shrader P, Pencina MJ, Fox CS, Manning AK, Grant RW, et al. Genetic risk reclassification for type 2 diabetes by age below or above 50 years using 40 type 2 diabetes risk single nucleotide polymorphisms. Diabetes Care. 2011; 34: 121–125. https://doi.org/10.2337/dc10-1265 PMID: 20889853

28. 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 PMID: 24618965

29. Ashley EA, Butte AJ, Wheeler MT, Chen R, Klein TE, Dewey FE, et al. Clinical assessment incorporating a personal genome. Lancet. 2010; 375: 1525–1535. https://doi.org/10.1016/S0140-6736(10)60452-7 PMID: 20435227

30. Whelton PK, Carey RM, Aronow WS, Casey DE Jr, Collins KJ, Denison Himmelfarb C, et al. 2017 ACC/AHA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2017; https://doi.org/10.1016/j.jacc.2017.11.006 PMID: 29146535

31. Nayor M, Vasan RS. Recent Update to the US Cholesterol Treatment Guidelines: A Comparison With International Guidelines. Circulation. 2016; 133: 1795–1806. https://doi.org/10.1161/CIRCULATIONAHA.116.021407 PMID: 27143546

32. Stone NJ, Robinson JG, Lichtenstein AH, Merz CNB, Blum CB, Eckel RH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014; 63: 2889–2934. https://doi.org/10.1016/j.jacc.2013.11.002 PMID: 24399223

33. Bloss CS, Schork NJ, Topol EJ. Effect of direct-to-consumer genomewide profiling to assess disease risk. N Engl J Med. 2011; 364: 524–534. https://doi.org/10.1056/NEJMoa101893 PMID: 21226570

34. Pazzoli R, Dehghan A, Evangelou E, Warren H, Gao H, Caulfield M, et al. Genetic Predisposition to Genetic Risk in Coronary Heart Disease: a policy statement from the American Heart Association. Circulation. 2012; 126: 142–157. https://doi.org/10.1161/CIR.0b013e31825b07f8 PMID: 22645291

35. Knowles JW, Assimes TL, Kiernan M, Pavlovic A, Goldstein B, Yank V, et al. Randomized trial of personalized genomics for preventive cardiology: design and challenges. Circ Cardiovasc Genet. 2012; 5: 368–376. https://doi.org/10.1161/CIRCGENETICS.112.962746 PMID: 22715281

36. Knowles JW, Assimes TL, Klieman M, Pavlovic A, Goldstein B, Yank V, et al. Randomized trial of personal genomics for preventive cardiology: design and challenges. Circ Cardiovasc Genet. 2012; 5: 368–376. https://doi.org/10.1161/CIRCGENETICS.112.962746 PMID: 22715281

37. Johnson HM, Einerstorfer J, Porcaro CE, Aeschlimann SE, Stein JH. Long-term effects of carotid screening on patient outcomes and behaviors. Arch Intern Med. 2011; 171: 589–591. https://doi.org/10.1001/archinternmed.2011.90 PMID: 21444853

38. O’Malley PG, Feuerstein IM, Taylor AJ. Impact of electron beam tomography, with or without case management, on motivation, behavioral change, and cardiovascular risk profile: a randomized controlled trial. JAMA. 2003; 289: 2215–2223. https://doi.org/10.1001/jama.289.17.2215 PMID: 12734132

39. Whelton PK, Carey RM, Aronow WS, Casey DE Jr, Collins KJ, Denison Himmelfarb C, et al. 2017 ACC/AHA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2017; https://doi.org/10.1016/j.jacc.2017.11.006 PMID: 29146535

40. MyGeneRank | Unlock your genetic risk [Internet]. [cited 19 Feb 2018]. Available from: https://mygenerank.scripps.edu/
46. Ng PC, Murray SS, Levy S, Venter JC. An agenda for personalized medicine. Nature. 2009; 461: 724–726. https://doi.org/10.1038/461724a PMID: 19812653

47. Regalado A. 23andMe to Share DNA Data with Researchers Using Apple iPhone. MIT Technology Review. 21 Mar 2016. Available from: https://www.technologyreview.com/s/601082/23andme-to-share-dna-data-with-researchers-using-apple-iphone/. Accessed 20 Feb 2018.

48. Hudson KL, Holohan MK. Keeping pace with the times—the Genetic Information Nondiscrimination Act of 2008. N Engl J Med 2008; 358:2661–2663 Available from: http://www.nejm.org/doi/full/10.1056/NEJMp0803964 https://doi.org/10.1056/NEJMp0803964 PMID: 18565857

49. Kraft SA, Constantine M, Magnus D, Porter KM, Lee SS-J, Green M, et al. A randomized study of multimedia informational aids for research on medical practices: Implications for informed consent. Clin Trials. 2017; 14: 94–102. https://doi.org/10.1177/1740774516669352 PMID: 27625314

50. Dewey FE, Chen R, Cordero SP, Ormond KE, Caleshu C, Karczewski KJ, et al. Phased Whole-Genome Genetic Risk in a Family Quartet Using a Major Allele Reference Sequence. PLoS Genet. 2011; 7(9): e1002280. https://doi.org/10.1371/journal.pgen.1002280 PMID: 21935354