Prospects for Genetic Risk Prediction in Hypertension

Sandosh Padmanabhan

See related article, pp 987–994

Prediction is very difficult, especially about the future.
—Niels Bohr (1885–1962)

Cardiovascular disease (CVD) risk prediction models using conventional risk factors (male sex, hypertension, cholesterol, smoking, diabetes mellitus, and family history) are the essence of current cardiovascular practice.1 Risk estimates from these models are used to identify individuals at high risk of developing CVD (generally 10-year risk estimates) and are usually performed in the fourth or fifth decade of life, when traditional modifiable risk factors become manifest.2 There are 4 major issues with this practice: (1) it is estimated that ≈15% to 20% of patients who develop CVD have none of the traditional risk factors and would be labeled low risk by current prediction algorithms;3 (2) there is increasing evidence that atherosclerosis begins in childhood and is associated with the same traditional cardiovascular (CV) risk factors;4,5 (3) the increasing life-expectancy of the population has made 10-year risk, estimated at age 50 years, less clinically relevant when the actual CVD risks of these individuals are much higher across their remaining lifespan of a further 40 to 50 years;6 and (4) even after correction for traditional risk factors, patients have a high residual risk of CV events. These issues highlight the need to improve current CVD prediction algorithms that will allow risk reduction measures to be implemented at all levels of intervention, from primordial through to secondary prevention.

Genetic variants have attracted interest as markers that may enhance risk prediction models. There are multiple reasons: genotyping is now accurate, rapid, and cheap; variants only need to be measured once as they are fixed at conception and, therefore, may be relevant for younger individuals before conventional risk factors, such as hyperlipidemia, hypertension, and diabetes mellitus, become apparent and before exposure to many environmental risk factors; genome-wide association studies (GWAS) have uncovered several single-nucleotide polymorphisms (SNPs) that are robustly associated with CVD risk factors (lipids, blood pressure [BP]) and CVD (myocardial infarction, coronary heart disease; http://www.genome.gov/gwastudies/). It is likely that SNPs associated with intermediate traits, such as lipids or BP, may predict lipid production or BP over longer periods than a single point laboratory/clinical measurement, and hence may have better predictive potential than clinical measurements. However, the major challenge has been that SNPs arising from GWAS studies have individually small effect sizes, and cumulatively these SNPs explain very little of the population variance of the trait studied. For example, it has been shown that 45 validated genetic variants explain ≈10.6% of the additive genetic variance of coronary artery disease, and 29 GWAS SNPs explain ≈0.9% of the total variance of BP.7–9

The first step in the assessment of a novel risk score incorporating genetic information is that it should show a significant statistical association with the outcome of interest. However, this significance has little relationship to its clinical use, and the second step is required to determine its ability to discriminate future cases from noncases—the calibration of the model, the model fit, the informativeness of the model for the outcome of interest, risk reclassification, and clinical use.10 A minimum of 3 criteria other than statistical significance have been proposed in the evaluation of new biomarkers: discrimination, calibration, and reclassification.11 Recently, recommendations on how to report evaluations of risk prediction models that include genetic variants have been published.12 A risk score’s discriminatory ability is measured by its c-statistic, which represents the area under the receiver operating characteristic curve (plotting sensitivity in relation to 1-specificity; AUC). The c-statistic is dependent on heritability, the genetic variance explained by the genetic variants, the prevalence of the disease condition, and the minor allele frequency in the population.13 Although the AUC or c-statistic is an important measure for clinical validity, it does not differentiate between the accuracy with which the genomic profile predicts the true genetic risk and true genetic risk predicts the disease status. Calibration compares the predicted risk with the observed risk in groups of individuals with varying baseline risk (ie, assess the ability of a risk prediction model to predict accurately the absolute level of risk that is subsequently observed).14 Treatment decisions are dependent on estimates of predicted risk and thus well-calibrated models will have a major impact on clinical management. Reclassification tests whether the addition of a new risk marker results in a substantial proportion of individuals being moved (reclassified) across a predefined treatment threshold.15 To quantify the appropriateness and amount of overall reclassification, Pencina et al10 have proposed 2 indices, the net reclassification improvement and the integrative discrimination index. The net reclassification improvement is calculated by determining the net number of correctly and incorrectly classified subjects among those who develop events and those who do not. The integrative discrimination index indicates how far individuals move on average along the continuum of predicted risk.10
If the integrative discrimination index is small, even in the presence of significant net reclassification improvement, then the change in predicted risk will be small on average (those at higher baseline risk will likely change more than those at lower baseline risk).22 The difference between calibration and reclassification is that calibration gives equal weight to all combinations of predicted and observed risk, and therefore may not be very useful for the most common treatment decision scenario involving individuals who fall into an intermediate risk category.

Common risk variants (SNPs) have effect sizes too small to be used individually as risk predictors, so profiles based on many associated genetic variants are used for predictions of genetic risk. Genetic risk scores are calculated from a panel of validated SNPs using a weighted mean of the number of copies of the risk alleles (reported effect sizes from reference studies are used as weights per copy of the risk allele) or just a simple sum of the risk allele count across all the SNPs in the panel. In theory, the SNPs used to calculate genetic risk scores do not have to be the causative variants as long as they are in high linkage disequilibrium with the causative variants. Two studies looked at genetic risk scores derived from a large panel of >101 GWAS SNPs for CVD and major CVD risk factors, but did not show any association with incident CVD events or any further improvement in discrimination or risk reclassification,7,18 highlighting the limited clinical use of genetic cardiovascular risk prediction currently. Simulation studies assuming mean relative risks of 1.1 to 1.2 for CVD risk variants estimate that 100 genetic variants would explain 1.0% to 9.1% of the variance of CVD and provide c-statistics 0.75,19 which would be similar to but not much better than current prediction models. Hence, to achieve higher levels of discrimination using relative risks more appropriate to GWAS SNPs, the number of SNPs required will increase several fold. In this context, Havulinna et al20 constructed genetic risk scores using 32 bona fide BP GWAS SNPs and showed that this score was strongly associated with BP, hypertension, and incident CVD independent of traditional risk factors, including BP, but with no improvement in reclassification. However, the results of this study support a causal role for BP on CVD, which has hitherto been demonstrated only by randomized clinical trials of antihypertensive therapy.21 Havulinna et al20 highlight the challenges posed by BP variability limiting the precision measurement of the phenotype and suggest that BP genotypes are better surrogates as they can be precisely measured, are unchanging over time, and capture a fixed component of lifetime BP exposure. There is supporting evidence that genetic risk scores using lipid and diabetes mellitus GWAS SNPs predict CVD and diabetes mellitus, respectively, more consistently over time, in contrast to predictive models using clinical biomarkers, which tend to be powerful predictors only in the short term.22,23 Predicting long-term CV risk from a genetic risk score constructed from BP SNPs (or SNPs associated with any other CV risk factor) is limited by the informativeness of the score (eg, a BP genetic risk score will not capture the multiple influences like time-dependent effects, environmental factors, effect of interventions, and competing risks that can influence CV outcomes). Thus, the risk prediction from a risk score using SNPs from multiple risk factors will only inform on risk within a narrow dimension and hence it is not surprising that the results from the multitude of genetic risk score studies are similarly modest.

The future of genetic risk prediction in hypertension can be considered in 2 areas. First, there is the necessity to identify more validated SNPs so that most of the genetic variance is accounted for; second, determining the context in which genetic risk prediction will be clinically relevant. The SNPs included by Havulinna et al20 likely explain the highest fraction of the BP genetic variance per SNP considering all SNPs (discovered and undiscovered) associated with BP. Ongoing larger meta-analyses and future sequencing studies will certainly identify additional risk variants, and these will be SNPs with even smaller effect sizes and also possibly low-frequency variants of large effect size. Increasing the total number of validated SNPs and low-frequency variants should cumulatively explain a greater proportion of the total variance—this will provide a major boost to the accuracy of risk prediction.

The current boom in validated SNPs for BP/HTN9 has placed us only at the start of a multi-step process starting with marker discovery to profile evaluation and finally to an assessment of the net benefit of treatment or prevention strategies guided by genetic profiles.24 A key distinction between genetic risk prediction and classical risk prediction is that the genetic sequence of an individual of a person is constant throughout their life. This has implications for the context in which BP genetic risk prediction would be clinically useful, though genetic markers need be measured only once in the lifetime. A BP genetic risk score will not capture the time-dependent effects on BP from environmental factors, effect of treatment, or competing risks. There is evidence that BP variability is itself an independent predictor of risk and this will not be captured by genetic risk scores created from SNPs associated with BP. Thus, a BP genetic risk score can only reflect risk within a narrow spectrum which uniquely captures lifetime BP exposure. Given the unvarying nature of the genotype, genetic risk prediction may be more aligned with lifetime risk prediction rather than the 10-year CV risk estimates provided by different risk calculators. Data from the National Health and Nutrition Examination Survey 2003 to 2006 reveal that a majority (56%) of US adults, or 87000000 people, have a low 10-year but high lifetime predicted risk for CVD.6 In the Coronary Artery Risk Development in Young Adults Study,9 90% of participants, aged 32 to 47 years, had a 10-year predicted risk of 10%, and half of them had high predicted lifetime risk (39%). As compared with subjects with low 10-year and low lifetime predicted risk, those with low 10-year and high lifetime predicted risk have a significantly greater burden of carotid artery intima-media thickness, coronary artery calcification, and greater progression of subclinical atherosclerosis.4 Data from the Bogalusa Heart Study25 suggest that GWAS SNPs for cardiovascular risk factors identified in adult cross-sectional studies show age-independent effects in children, indicating that traditional risk factors associate with genetic variants early in life. There is strong epidemiological evidence that traditional risk factors like hypertension have a strong and consistent influence on lifetime risk of CV disease, and this effect has been constant.
over decades. Primary prevention statin trials, such as the Anglo-Scandinavian Cardiac Outcomes Trial (ASCOT) and Justification for the Use of statins in Primary prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER), show that there is value in treating everybody with high lifetime risk more aggressively. As traditional risk factors become manifest usually after the fourth decade of life, genetic risk prediction may offer the prospect for early targeting of subjects for primordial prevention strategies. The value of any risk predictor is a balance among predictive power, the cost and invasiveness of the prediction procedure, and the cost and effectiveness of the interventions available. A plethora of available early lifestyle and pharmacological interventions, the high prevalence of the hypertension (implying a high positive predictive value for genetic prediction), would support large-scale genetic screening in predicting hypertension and CVD. Improving the predictive accuracy by identifying all the genetic variants influencing the trait and developing clinical/public health infrastructure to implement genome-driven genetic screening in predicting hypertension and CVD. predictive value for genetic prediction), would support large-scale genetic screening in predicting hypertension and CVD. Improving the predictive accuracy by identifying all the genetic variants influencing the trait and developing clinical/public health infrastructure to implement genome-driven prevention strategies are the 2 key hurdles to be surmounted for genetic risk prediction for CVDs to become a reality.

Disclosures
None.

References