When genetic burden reaches threshold

Roddy Walsh 1, Rafik Tadros 2, and Connie R. Bezzina 1*

1Department of Clinical and Experimental Cardiology, Heart Centre, Amsterdam Cardiovascular Sciences, Amsterdam UMC, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, Netherlands; and 2Department of Medicine, Cardiovascular Genetics Center, Montreal Heart Institute and Faculty of Medicine, Université de Montréal, 500 Belanger, Montreal, QC H1T 1C8, Canada

Received 6 November 2019; revised 7 January 2020; editorial decision 24 March 2020; accepted 1 April 2020; online publish-ahead-of-print 29 April 2020

Rare cardiac genetic diseases have generally been considered to be broadly Mendelian in nature, with clinical genetic testing for these conditions predicated on the detection of a primary causative rare pathogenic variant that will enable cascade genetic screening in families. However, substantial variability in penetrance and disease severity among carriers of pathogenic variants, as well as the inability to detect rare Mendelian variants in considerable proportions of patients, indicates that more complex etiologies are likely to underlie these diseases. Recent findings have suggested genetic variants across a range of population frequencies and effect sizes may combine, along with non-genetic factors, to determine whether the threshold for expression of disease is reached and the severity of the phenotype. The availability of increasingly large genetically characterized cohorts of patients with rare cardiac diseases is enabling the discovery of common genetic variation that may underlie both variable penetrance in Mendelian diseases and the genetic aetiology of apparently non-Mendelian rare cardiac conditions. It is likely that the genetic architecture of rare cardiac diseases will vary considerably between different conditions as well as between patients with similar phenotypes, ranging from near-Mendelian disease to models more akin to common, complex disease. Uncovering the broad range of genetic factors that predispose patients to rare cardiac diseases offers the promise of improved risk prediction and more focused clinical management in patients and their families.

Keywords
Genetics • Rare cardiac disease • Genetic modifiers • Inherited cardiomyopathies • Ventricular arrhythmias • Genome-wide association studies

Introduction

Heritable cardiovascular diseases have generally been divided into two broad categories. The first category encompasses rare Mendelian genetic diseases such as inherited cardiomyopathies (e.g. hypertrophic cardiomyopathy, HCM) and ventricular arrhythmia syndromes (e.g. long QT syndrome, LQTS) that are usually caused by a rare genetic variant (a so-called ‘mutation’) that has a large risk-increasing effect size. In the second category are common complex diseases such as hypertension and coronary artery disease where environmental factors act in conjunction with a large number of common genetic variants, each with a small risk-increasing effect size. Discoveries into the genetic basis of cardiac diseases, and their application in the clinic, initially focused on the rarer Mendelian conditions. The familial inheritance observed in these diseases and the large effect sizes of the causative rare variants enabled the identification of the major disease genes through linkage studies in large family pedigrees.

Subsequent discoveries demonstrated the genetic heterogeneity of these disorders, with several causative genes identified through linkage and candidate gene studies, though the replication rate for the latter (as described below) has been poor, unless significant rare variant association has been demonstrated by case–control analysis.1 Sequencing of these genes has been recommended for a number of inherited cardiac conditions for several years and has become a standard aspect of clinical management in affected families. The primary benefit of this testing is to identify at risk carriers of the familial pathogenic variant (and non-carriers who are unlikely to develop disease), assuming a penetrant variant is identified that can be predicted with confidence to cause the disease. Clinical genetic testing for these conditions facilitates focused clinical screening of those relatives at

* Corresponding author. Tel: +31 20 5663262, Email: c.r.bezzina@amsterdamumc.nl

© The Author(s) 2020. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
risk of developing disease, has been shown to be cost-effective\(^2\) and can be considered as a success story in the application of genetics into clinical practice.

**Limitations of Mendelian genetic approaches**

The clinical impact of genetic testing for inherited cardiac conditions has been hampered however by two key factors—the substantial proportion of cases where a causative Mendelian variant is still not identified and, where a pathogenic variant is detected, the limited ability to predict clinical outcomes with this information. Diagnostic yields range from 75% to 80% in LQTS to only 20% in Brugada syndrome (BrS).\(^3\) Years of efforts to expand the genetic repertoire of inherited cardiac conditions through candidate gene studies have proved largely fruitless, with contemporary re-evaluation of these studies refuting many of these gene to disease associations. In particular, population genetics databases such as ExAC and gnomAD have shown that some implicated variants have population frequencies incompatible with causing rare diseases, as well as demonstrating that rare variation is collectively common for many genes. For example, all but one gene implicated in BrS (SCNS5A) have now been repudiated by the ClinGen initiative,\(^4\) with similar findings observed for HCM.\(^5,6\) (ClinGen curation for other diseases is ongoing). Similarly, 6.5–13.5% of variants associated with cardiomyopathies are now shown to have population frequencies incompatible with being penetrant variants for such diseases.\(^7\)

Instead, several factors now suggest that the majority of cases where pathogenic variants are not identified (“genotype-negative”) are likely to represent non-Mendelian forms of disease. Most large affected family pedigrees have now been genetically resolved (at least in Europe and North America), with genotype-negative patients more likely to present as sporadic cases and with a much lower family history of disease. In addition, genotype-negative cases can display substantially different phenotypic and clinical characteristics compared to genotype-positive cases—in HCM, this includes distinct left ventricular (LV) morphology and more benign outcomes\(^8\)—suggesting a different genetic aetiology may underlie their disease.

Broad correlations can be observed between pathogenic variant classes (such as variants in a particular gene) and phenotype for several inherited cardiac conditions. For example, differential arrhythmic risk profiles and response to drug therapy are observed amongst variants in the three main LQTS genes.\(^9\) Yet, despite the identification of disease genes and clinical risk factors, predicting disease severity and major complications such as heart failure and sudden cardiac death remains challenging. Even within family pedigrees carrying the same disease-causing variant, incomplete penetrance (variant carriers who do not develop disease) and variable expressivity (a wide range of severity amongst carriers) are common phenomena. Non-genetic factors are known to influence disease risk amongst pathogenic variant carriers in several cardiac diseases, ranging from generic factors such as age and sex to disease-specific modulators (e.g. obesity in HCM\(^10\) and QT prolonging drugs in LQTS\(^11\)). It is also increasingly recognized that additional genetic factors may act to modulate the phenotype in individuals with a primary pathogenic variant and underlie a substantial proportion of the variability in penetrance and disease severity.

**Identification of genetic risk variants in rare cardiac disease**

While all disease-associated genetic variants will lie on a spectrum of phenotype effect size and population frequency, it is useful to distinguish between Mendelian and non-Mendelian variants. The former are primary drivers of disease in affected family pedigrees and can be used for cascade genetic screening to identify at-risk individuals while the latter contribute to disease risk with smaller effect sizes and cannot be used in isolation to define risk. We can consider there to be broadly two types of non-Mendelian genetic risk variants that are detectable with current approaches and study sizes and could potentially contribute to disease risk in rare cardiac diseases. Common variants [usually defined as having a minor allele frequency (MAF) of >5%] identified through genome-wide association studies (GWAS) generally have individually small effect sizes but collectively have been shown to be associated with disease risk across a broad range of disease phenotypes. Intermediate effect variants (MAF < 1–5%) have effect sizes and frequencies between common and Mendelian variants. Preliminary research into genetic risk variants in cardiac disease (described in detail below) suggests these factors may, to varying extents, influence penetrance in individuals with Mendelian genetic defects by pushing the genetic burden towards the threshold of disease\(^12\) (Figure 1). These variants are also likely to contribute to disease risk in individuals without primary Mendelian variants, with a larger burden of disease risk variants expected to be found in such genotype-negative individuals. With some exceptions, the contribution of such variants in determining disease severity remains largely unexplored.

**Common variation and genome-wide association studies**

Genome-wide association studies have had an enormous impact on elucidating the genetic basis of common complex diseases over the last 12 years, with thousands of robust associations identified and sample sizes in some studies now approaching 1 million.\(^13\) With increasing recognition of the genetic complexity of rare Mendelian disease, GWAS are now also starting to be applied to these conditions to identify variation that may underlie both variable penetrance and genotype-negative cases. Two approaches can be used for GWAS in rare genetic disease. The first employs a standard case-control study design including unrelated patients with the genetic disease and population-matched controls, which allows for direct detection of disease-associated variants as well as further stratified analyses based on factors such as pathogenic variant status and disease severity. Rare disease GWAS are limited however by the availability of disease samples, with multi-centre collaborations and meta-analysis required to achieve even moderately powered studies. To complement these efforts, quantitative studies on disease-relevant endphenotypes may be performed using population cohorts such as the UK Biobank. These are powered to detect a larger number of associations and can produce a polygenic risk score (PRS, a weighted
aggregate of associated loci) underlying the endophenotype that can then be tested for association in patients with the rare disease.

The detection of common risk variants in inherited cardiac disease has initially focused on LQTS, given the directly applicable endophenotype (QT interval) readily available in large population studies. The latest QT-interval GWAS, conducted on 76,000 individuals of European descent, identified 35 loci with individually small effects but collectively explaining approximately 10% of QT interval variation in the general population. Studies with LQTS patients have shown that some of these variants can modulate QT interval and risk of arrhythmias and cardiac events—for variants at the NOS1AP locus, KCNQ1 locus, and with a PRS derived from 22 QT interval single polymorphisms (SNPs) in LQTS Type 2 patients (those with a KCNH2 pathogenic variant). These findings highlight the potential role for common variation in explaining phenotypic variability in LQTS patients, though clinical utility remains to be demonstrated.

In contrast, investigations into common variation in cardiomyopathies has so far been largely restricted to case–control studies with moderate sample sizes. For HCM and dilated cardiomyopathy (DCM) relevant endophenotypes related to LV dimensions and function are expensive to measure by cardiac magnetic resonance (CMR), difficult to accurately quantify by echocardiography and are less clearly correlated with disease phenotype. Such data are now becoming available through studies such as the EchoGen consortium and a CMR-scanned subset of the UK Biobank, with LV traits GWAS yielding association signals and PRS that may subsequently be assessed for roles in HCM/DCM disease susceptibility and severity. Indeed, two of the loci identified in the UK Biobank study were...
previously associated with DCM in a case–control GWAS—BAG3 (associated with LV ejection fraction and end-systolic/diastolic volumes) and the CLCNKA/HSPB7 locus (associated with LV ejection fraction). Larger multicentre cardiomyopathy GWAS studies are needed to more fully elucidate the role of common variation in these conditions and assess the utility of these LV endophenotype PRS (Figure 2).

In 2013, a multicentre case–control GWAS in BrS (with 312 cases and 1115 controls) identified three independent association signals at the SCN5A, SCN10A, and HEY2 genes, confirming the central role for the sodium channel in this disorder. In aggregate, these three SNPs alone confer a high relative risk for disease (with an odds ratio >20 for carriers of >4 risk alleles compared to <2), accounting for 7% of the variance in disease susceptibility, though absolute risk is low given the rarity of the disease. The strong association signals and high cumulative effects on relative risk observed in this study, despite its limited sample size, highlight the distinctive genetic architecture of BrS. Rare variants in SCN5A are observed in only 20% of cases and, even in genotype-positive families, they have been shown to be neither necessary nor sufficient to cause disease (with reports of both incomplete penetrance in mutation carriers and non-segregation of putatively pathogenic variants, i.e. non-carriers with a BrS phenotype). These findings point towards a highly polygenic nature of BrS, in contrast to other rare cardiac conditions. More recently, it has been shown that a BrS PRS could potentially be of clinical utility in the diagnostic strategy of BrS.

As well as highlighting the increasing genetic complexity of diseases previously interpreted as principally Mendelian, GWAS and their derived PRS can also now be used to identify individuals in the general population with disease risks equivalent to those conferred by monogenic Mendelian variants. Khera et al. found that 8% of the UK Biobank samples had a PRS that conferred >three-fold risk for coronary artery disease (similar to the risk for carriers of familial hypercholesterolaemia pathogenic variants), with similar risks observed for other common diseases (though at lower percentages). While familial inheritance will be different in such cases, this data could be used to identify individuals for intensive screening and/or preventive intervention. It also illustrates that genetic architecture and risk for many
Intermediate effect variants

Intermediate effect genetic variants lie on a wide spectrum between common variants investigated by GWAS and rare pathogenic Mendelian variants, as measured by both variant frequency in the population and phenotypic effect size in patients. The identification and characterization of such variants is an emerging topic for inherited cardiac conditions and faces substantial challenges in identifying candidate variants (requiring sequencing of large cohorts given their lower frequency) and distinguishing them from both pathogenic Mendelian and benign variants. Once identified, new clinical genetics guidelines will be required for classifying such variants (methods are currently designed for classic Mendelian genetics) and defining actionability when they are detected in patients and their families.

Population genetic resources such as gnomAD have highlighted that many variants previously associated with disease are in fact too common in the population to be penetrant pathogenic variants in rare genetic diseases. While many such variants are likely to be benign and have only been incidentally detected in patients, some may have an effect on disease susceptibility or phenotypic severity. A high burden of proof should be required to define a variant as a potential risk variant, which could include a significant enrichment in case cohorts compared to ethnically matched controls, a demonstrated effect on phenotypic severity within cases carrying a pathogenic variant and well-characterized functional assays.

The classic example of a genetic risk variant in inherited cardiac disease is the p.Asp85Asn missense variant in the KCNE1 gene. With a gnomAD MAF in non-Finnish Europeans of 0.012, the variant is at the lower end of detection frequencies for GWAS and has been associated with prolonged QT interval in the general population, with the largest effect size of any of the associated SNPs (7.42 ms per minor allele). It is enriched in LQTS patients from Japan and USA, where, interestingly, the majority of carriers were genotype-negative for rare pathogenic LQTS variants, suggesting p.Asp85Asn may also contribute substantially to the genetic burden in non-Mendelian LQTS cases. Functional studies have confirmed the deleterious role of this variant by demonstrating significant reductions on repolarizing potassium channel-encoded currents.

As intermediate effect variants are often population-specific, expanding genetic sequencing to non-European populations will be as important as increasing GWAS sample diversity. With a fixed number of variants in the major cardiac genes that are candidate intermediate effect variants, systematic approaches to evaluate their function and effects sizes, using functional assays or population biobanks, may be feasible in the near future.

Prospects for clinical genetics and precision medicine

It is increasingly evident that the genetic architecture of rare cardiac disease is more complex than accounted for by simple Mendelian models. A range of genetic variants of different frequencies and effect sizes may combine to produce an overall genetic burden that, in conjunction with non-genetics factors, may determine whether the threshold of disease is reached in each individual and the severity of disease in patients (Figure 1). While we are still very much in the early stages of discovering these modifying genetic factors and understanding how they affect disease risk, these findings have a number of implications for how we understand the aetiology of these diseases and how they will be applied in clinical practice.

The development of PRS from large GWAS to inform clinical risk prediction is currently an active area of research for common diseases such as coronary artery disease and atrial fibrillation. PRS that identify individuals with high risk of developing disease have the potential for improved prediction over and above current risk factors, but their clinical utility (particularly in non-European populations) remains to be established. For rarer genetic diseases, the integration of the different classes of variants contributing to disease, particularly accounting for the variable effects of primary pathogenic variants, will make clinical adoption even more challenging. Classification guidelines will need to be developed for intermediate effect variants that effectively assess both the likelihood of contribution to the disease phenotype and the estimated effect size. Finally, communicating this complex genetic profile to both clinicians and patients, who are accustomed to receiving deterministic findings from genetic testing, will present additional challenges.

However, more comprehensive assessment of genetic risk for rare cardiac diseases offers enormous potential for improved risk prediction and clinical management in patients and their families. Current Mendelian-based genetics is effective at broadly identifying at-risk individuals but is a blunt tool for individualized risk prediction. For unaffected carriers of cardiomyopathy pathogenic variants, near-life-long periodic clinical screening is performed given the age-dependent penetrance. By more effectively stratifying pathogenic variant carriers according to their broader genetic risk profile, we will be able to identify those at most risk of developing disease and severe cardiac events, facilitating increased monitoring and interventional therapy where appropriate. At the same time, unaffected family members of cardiomyopathy patients, who currently undergo regular clinical screening to detect onset of disease (at considerable cost in healthcare provision), could be discharged if their overall risk was determined to be low.

These new insights into the genetics of rare cardiac diseases also compel us to re-evaluate the aetiology of genotype-negative disease and the clinical management of such cases. The majority of such cases, especially in families with no prior history of disease, are likely to be caused by a range of small to intermediate effect variants and non-genetic factors. Consequently, the risk to family members is likely substantially lower than in pedigrees with penetrant Mendelian variants, such that guidelines for cardiomyopathies might soon consider it feasible to release currently phenotype-negative relatives from ongoing clinical screening. The further development of disease-specific genome-wide risk scores could aid in this decision making by quantifying the risk to relatives through inexpensive genotyping assays. Questions also arise as to how these conditions are defined—are these cases in fact the extreme end of population polygenic risk whose phenotypes converge with Mendelian diseases? Finally, it has become evident that the language and dichotomous classifications that are pervasive in genetic disease (Mendelian/complex disease,
pathogenic/variant/benign variants) are increasingly inadequate for describing genetic cardiac disease over the next few years are likely to reveal ever increasing complexity but yield improvements in risk prediction for patients and their families.

**Funding**

R.W. received support from a post-doctoral grant from the Amsterdam Cardiovascular Sciences. R.T. received support from the Canadian Heart Rhythm Society’s George Mines Award, the European Society of Cardiology research award, and the Philippe and Marvin Casey Cardiology Chair and is currently a clinical research scholar of the Fonds de Recherche du Québec-Santé. C.R.B. acknowledges the support from the Dutch Heart Foundation (CVON Predict 2 project, CVON Concor-genes project), the Netherlands Organization for Scientific Research (VICI fellowship, 016.150.610), and Fondation Leducq.

**Conflict of interest:** none declared.

**References**

1. Hermsen SM, Konrad R, Udupa S, Costain G, Jobling R, Liston E, Jamal SM, Walsh R, Buchan R, Wilk A, John S, Felkin LE, Thomson KL, Chiaw TH, Loong S, Martin A, Haines J, Seidman JG, Seidman CE. Truncations of titin causing dilated cardiomyopathy. Circulation 2017; 135:3765–3776.

2. Wordsworth S, Leal J, Blair E, Legood R, Thomson KL, Taylor J, Watkins H. DNA testing for hypertrophic cardiomyopathy: a cost-effectiveness model. Eur Heart J 2010; 31:926–935.

3. Bezina CR, Lahrouchi N, Priori SG. Genetics of sudden cardiac death. Circ Res 2015; 116:1919–1936.

4. Hosseini SM, Konrad R, Udupa S, Costain G, Jobling R, Liston E, Jamal SM, Walsh R, Buchan R, Wilk A, John S, Felkin LE, Thomson KL, Taylor J, Watkins H. DNA testing for hypertrophic cardiomyopathy: a cost-effectiveness model. Eur Heart J 2010; 31:926–935.
When genetic burden reaches threshold

Arbustini E, Grallert H, Cook SA, Seidman CE, Regitz-Zagrosek V, Cappola TP, Charron P, Cambien F, Villeneuve E. Exome-wide association study reveals novel susceptibility genes to sporadic dilated cardiomyopathy. PLoS One 2017;12:e0172995.

21. Wild PS, Felix JF, Schillert A, Teumer A, Chen M-H, Leening MJG, Völker U, Groißmann V, Brody JA, Irvin MR, Shah SJ, Pramara S, Lieb W, Schmidt R, Stanton AV, Mažahín D, Smith AV, Sundström J, Minelli C, Ruggiero D, Lyttikainen L-P, Tiller D, Smith JG, Manneveau C, Di Tullio MR, Musani SK, Morrison AC, Pers TH, Morley M, Kleber ME, Aragam J, Benjamin EJ, Bis JC, Bispeing A, Broeckel U, Cheng S, Deckers JW, Del Greco M F, Edelmann F, Forngärd M, Franke L, Friedrich N, Harris TB, Hoffer E, Hofman A, Huang J, Hughes AD, Kährönen M, Investigators K, Kruppa J, Lacker J, Lannfelt L, Laskowski R, Launer LJ, Lindgren CM, Loeley C, MacRae CA, Mascioli D, Mayet J, Medenwaldt D, Morris AP, Müller C, Müller-Nurasyid M, Nappo S, Nilsson PM, Nuding S, Nuttle T, Peters A, Pfeuffer A, Pietzner D, Ptramstaller PP, Raitakari OT, Rice KM, Rivadeneira F, Rotter JI, Ruohonen ST, Sacco RL, Samdarshi T, Schmidt H, Sharp AS, Shields DC, Sorice R, Satoosdehnia N, Sinkner BH, Surenrand P, Thom S, Toghihofer AM, Uitterlinden AG, Wachter R, Valérie H, Ziegler A, Munzel T, Marz W, Cappola TP, Hirschhorn JN, Mitchell GF, Smith NL, Fox ER, Dueck ND, Jaddoe VVV, Melander O, Russ M, Lehtimäki T, Ciullo M, Hicks AA, Lind L, Gudnason V, Pleska B, Barron AJ, Zweiker R, Schunkert H, Ingelsson E, Liu K, Arnett DK, Psaty BM, Blankenberg S, Larson MG, Felix SB, Franco OH, Zeller T, Vasan RS, Dörr M. Large-scale genome-wide analysis identifies genetic variants associated with cardiac structure and function. J Clin Invest 2017;127:1798–1812.

22. Ang N, Vargas JD, Yang C, Cabrera CP, Warren HR, Fung K, Tzanis E, Barnes AO, Schwartz PJ, Crotti L, Dagradi F, Guicheney P, Fressart V, Leenhardt A, Pieske B, Barron AJ, Zweiker R, Schunkert H, Ingelsson E, Liu K, Arnett DK, Psaty BM, Blankenberg S, Larson MG, Felix SB, Franco OH, Zeller T, Vasan RS, Dörr M. Large-scale genome-wide analysis identifies genetic variants associated with cardiac structure and function. J Clin Invest 2017;127:1798–1812.

23. Bezzina CR, Barc J, Mizusawa Y, Remme CA, Gourraud JB, Simonet F, Verkerk J, Aung N, Vargas JD, Yang C, Cabrera CP, Warren HR, Fung K, Tzanis E, Barnes AO, Schwartz PJ, Crotti L, Dagradi F, Guicheney P, Fressart V, Leenhardt A, Pieske B, Barron AJ, Zweiker R, Schunkert H, Ingelsson E, Liu K, Arnett DK, Psaty BM, Blankenberg S, Larson MG, Felix SB, Franco OH, Zeller T, Vasan RS, Dörr M. Large-scale genome-wide analysis identifies genetic variants associated with cardiac structure and function. J Clin Invest 2017;127:1798–1812.