UNIVERSITAT DE BARCELONA

Assessing prognosis in Mendelian and complex heart diseases: clinical and genomic risk prediction models

Paloma Jordà Burgos

Assessing prognosis in Mendelian and complex heart diseases: clinical and genomic risk prediction models

Doctoral thesis dissertation presented by **Paloma Jordà Burgos** to apply for the degree of doctor at the University of Barcelona

Co-directed and tutorized by **Elena Arbelo Lainez**, Universitat de Barcelona, and co-directed by **Rafik Tadros**, Université de Montréal

Doctoral Program in Medicine and Translational Research

School of Medicine and Health Sciences. University of Barcelona.

September, 2023

Acknowledgements

To my mentors, Ana, Elena and Rafik. Thank you for sharing your passion with me, for your trust and advice.

To Josep, initial tutor of the thesis, for all his valuable support.

To all my professors.

To my fellow residents. To all those who have taught me during my residency. To all those who work every day taking care of patients: nurses, doctors, administrative and management staff. To all my colleagues, thank you for educating me, for your collaboration and care, and for accompanying me to accompany people in the process of health and disease.

To the patients, for teaching me so much about life and human resilience.

To those who believe in and defend the public healthcare system.

On a more personal level:

To my family, for their values. In particular, to *la iaia*.

To my friends, my lifelong ones and all those I have met in the journey.

To all those who feed my curiosity, thank you.

Thanks to all the people and resources that have allowed me the privilege of culminating such a beautiful process as a doctoral thesis.

Table of contents

Abbreviations and acronyms

AF Atrial fibrillation

AHA/ACC American Heart Association/American College of Cardiology

ARVC Arrhythmogenic right ventricle cardiomyopathy

BAG3 BCL-2 associated athanogene 3

Bpm Beats per minute

BrS Brugada syndrome

CAD Coronary artery disease

CHA2DS2-VASc Congestive heart failure or left ventricular dysfunction, hypertension, age ≥75 (doubled), diabetes, stroke (doubled)-vascular disease, age 65–74, sex category (female) (score)

CI Confidence interval

CMR Cardiac magnetic resonance

CVD Cardiovascular disease

DCM Dilated cardiomyopathy

DSC2 Desmocollin 2

DSG2 Desmoglein 2

DSP Desmoplakin

ECG Electrocardiogram

EF Ejection fraction

EMD Emerin

ESC European Society of Cardiology

FLNC Filamin C

GWAS Genome-wide association studies

HCM Hypertrophic cardiomyopathy

HF Heart failure

HRS Heart Rhythm Society

ICD Implantable cardioverter defibrillator

ITFC International Task Force consensus statement

KCNE1 Potassium voltage-gated channel subfamily E regulatory subunit 1

KCNH2 Potassium voltage-gated channel subfamily H member 2

KCNQ1 Potassium voltage-gated channel subfamily Q member 1

LGE-CMR Late gadolinium enhancement in the cardiac magnetic resonance

LMNA Lamin A/C

LPA Lipoprotein A

LQTS Long QT syndrome

LV Left ventricle

LVEF Left ventricular ejection fraction

LVOTO Left ventricular outflow tract obstruction

mITFC Modified International Task Force consensus statement

MTAG Multi-trait analysis of genome-wide association studies

MYBPC3 Myosin binding protein C3

MYH7 β-myosin heavy chain

N Sample size

NDLVC Non-dilated left ventricular cardiomyopathy

NGS Next-generation sequencing

NMDA N-methyl-D-aspartate

NNT Number needed to treat

NPR3 Natriuretic peptide receptor 3

NRI Net reclassification index

NSVT Non-sustained ventricular tachycardia

OR Odds ratio

PCSK9 Proprotein convertase subtilisin/kexin type 9

PGS Polygenic scores

PITX2 Paired like homeodomain 2

PKP2 Plakophilin 2

PLN Phospholamban

RBM20 RNA binding motif protein 20

RV Right ventricle

RVEF Right ventricular ejection fraction

SCD Sudden cardiac death

SCN5A Sodium voltage-gated channel alpha subunit 5

S.d Standard deviation

SGLT2 Sodium–glucose co-transporter 2

SNP Single nucleotide polymorphism

SRR Serine racemase

TMEM43 Transmembrane protein 43

TTN Titin

TWI T wave inversion

VA Ventricular arrhythmia

VALOR-HCM A study to evaluate mavacamten in adults with symptomatic obstructive hypertrophic cardiomyopathy who are eligible for septal reduction therapy

VF Ventricular fibrillation

VT Ventricular tachycardia

List of articles in the thesis

Thesis in compendium of publications format which consists of five objectives and three articles.

First article:
Paloma Jordà, Daniel Martínez, Marta Farrero, María Ángeles Marcos, Elena Sandoval, María Ángeles Castel, Daniel Pereda, Eduard Quintana, Elena Arbelo, Aurora Sánchez, Òscar Campuzano, Coloma Tirón de Llano, Ramon Brugada, Josep Brugada, Manuel Castellá, Félix Pérez-Villa, and Ana García-Álvarez. BAG3 Genetic Cardiomyopathy May Overlap Fulminant Myocarditis Clinical Findings. Circ Heart Fail. 2022;15(3):e008443.
Impact factor: 9.7. Q1 in the category CARDIAC & CARDIOVASCULAR SYSTEMS.

Second article:
Paloma Jordà, Yiwei Lai, Amélie Jeuken, Louis Philippe Lemieux Perreault, Elisabeth Goulet, Najim Lahrouchi, CHARGE Inflammation Working Group, Anna Nozza, Michael W Tanck, Peter Guerra, Julia Cadrin-Tourigny, Simon De Denus, Connie R Bezzina, Guillaume Lettre, David Busseuil, Marie-Pierre Dubé, Jean-Claude Tardif, and Rafik Tadros. Multi-trait analysis of genome-wide association studies for the discovery of novel genetic loci and genomic prediction in common cardiovascular diseases.
Under review.

Third article:
Paloma Jordà, Laurens P. Bosman, Alessio Gasperetti, Andrea Mazzanti, Jean-Baptiste Gourraud, Brianna Davies, Tanja Charlotte Frederiksen, Zoraida Moreno Weidmann, Andrea Di Marco, Jason D. Roberts, Ciorsti MacIntyre, Colette Seifer, Antoine Delinière, Wael Alqarawi, Deni Kukavica, Damien Minois, Alessandro Trancuccio, Marine Arnaud, Mattia Targetti, Annamaria Martino, Giada Oliviero, Daniel C. Pipilas, Corrado Carbucicchio, Paolo Compagnucci, Antonio Dello Russo, Iacopo Olivotto, Leonardo Calò, Steven A. Lubitz, Michael J. Cutler, Philippe Chevalier, Elena Arbelo, Silvia Giuliana Priori, Jeffrey S. Healey, Hugh Calkins, Michela Casella, Henrik Kjærulf Jensen, Claudio Tondo, Rafik Tadros, Cynthia A. James, Andrew D. Krahn, and Julia Cadrin-Tourigny. Arrhythmic risk prediction in arrhythmogenic right ventricular cardiomyopathy: external validation of the arrhythmogenic right ventricular cardiomyopathy risk calculator. Eur Heart J. 2022;43(32):3041-52.
Impact factor: 39.3. Q1 in the category CARDIAC & CARDIOVASCULAR SYSTEMS.

Thesis summary in Catalan

Avaluació pronòstica en les malalties cardiovasculars mendelianes i complexes: models de predicció del risc clínic i genòmic

Les malalties cardiovasculars (MCVs) contribueixen de forma important a la morbiditat i mortalitat de la població general. El pronòstic de les MCVs pot ser molt heterogeni en funció de l'entitat, i fins i tot en individus afectats per la mateixa patologia, variant des d'un estat asimptomàtic fins a la presentació d'arrítmies ventriculars, la necessitat de transplantament cardíac o la mort.

La base genètica i l'heretabilitat de les malalties cardiovasculars estan ben demostrades, amb un substrat genètic que abasta des de variants genètiques comunes (SNPs) de baix risc fins a variants genètiques rares de major risc que contribueixen acumulativament a l'expressió fenotípica d'una afectació concreta. En els últims anys, hem avançat en la comprensió dels fonaments genètics de les MCVs, en part gràcies a la millora en les tècniques d'anàlisi genòmic i a la creació de biobancs que contenen informació sobre milers d'individus. Així mateix, s'han realitzat múltiples esforços per millorar la precisió en la predicció de malaltia i l'estratificació pronòstica de les MCVs mitjançant la integració de factors de risc validats de forma ponderada en models de predicció clínica.

En aquesta tesi ens enfoquem en la predicció clínica i l'estratificació pronòstica de les MCVs, bé a través d'informació sobre variants genètiques rares, bé a través de models de predicció clínica a partir de puntuacions poligèniques (PGS, de l'anglès *polygenic scores*), que integren el valor additiu de l'efecte de SNPs de tot el genoma en relació a un fenotip concret, o bé a través de models de predicció clínica que integren marcadors demogràfics i clínics.

A la primera secció d'aquest compendi presentem el pronòstic d'una sèrie d'individus portadors d'una variant rara en el gen *BAG3* (p.Gln88*). Els portadors afectats (71%) van mostrar un fenotip clínic de miocardiopatia dilatada (MCD) amb una alta progressió a insuficiència cardíaca (IC) aguda i xoc cardiogènic a edats joves (57%) que, en alguns casos, anava precedida de símptomes compatibles amb infecció viral. La histopatologia va revelar dany miocàrdic agut i/o infiltrats inflamatoris en absència de marcadors serològics d'infecció viral aguda o càrrega viral elevada en el teixit miocàrdic. En aquest treball concloem que la MCD ocasionada per la variant *BAG3* p.Gln88* s'associa a una

forma de MCD d'alta penetrància, que condueix a l'aparició d'IC aguda i a un patró histopatològic específic, essent aquesta caracterització potencialment útil en la comprensió de les bases moleculars de la malaltia amb la fi de millorar el tractament específic de la mateixa.

A la segona secció, avaluem el substrat poligènic de tres malalties cardiovasculars comunes: la fibril·lació auricular (FA), la insuficiència cardíaca (IC) i la malaltia arterial coronària (MAC). L'objectiu del nostre treball fou millorar les estimacions de l'associació entre SNP i tret dels estudis convencionals d'associació del genoma complert (GWAS) emprant l'anàlisi multi-tret de GWAS (MTAG), que augmenta la potencia estadística d'aquestes associacions mitjançant l'anàlisi conjunta de trets genèticament correlacionats. MTAG va permetre el descobriment de 19, 131 i 141 nous loci genètics associats a la FA, la IC i la MAC. En el treball es descriuen els gens possiblement associats a aquests loci i es destaquen aquells amb vies fisiopatològiques més interessants o bé amb potencial terapèutic. Les estimacions de l'efecte variant-tret millorades a través d'MTAG van ser comparades amb les procedents de GWAS convencionals mitjançant PGS. Els models predictius incloent els PGS derivats dels resultats de MTAG van mostrar una capacitat discriminatòria i predictiva similar als models emprant PGS derivats de GWAS convencionals, si bé en la sub-anàlisi per sexes els PGS derivats de MTAG van mostrar una millor capacitat discriminatòria i predictiva en dones per a la MAC [ΔNagelkerke's R^2 1.735% (Interval de confiança (IC) 95%: 0.609-2.856); Índex de reclassificació neta 0.208 (95% CI: 0.139-0.277)]. Els models incloent els PGS per a la IC, van mostrar una predicció i calibratge deficients que van impedir l'avaluació de la seva capacitat discriminatòria. Aquest treball demostra que MTAG és una eina útil en el descobriment de nous loci genètics associats a les MCVs comunes – FA, IC i MAC - , i si bé no millora la predicció i discriminació de l'estatus de malaltia mitjançant l'ús de PGS en aquestes malalties en la cohort global sí que permet millorar la discriminació i predicció dels PGS per a la MAC en la cohort de sexe femení.

A la tercera secció, presentem la validació externa de la calculadora de risc d'arrítmies ventriculars a la miocardiopatia arritmogènica del ventricle dret (MAVD) en una cohort de 429 individus amb MAVD [103 amb arrítmies ventriculars durant una mitjana de seguiment de 5.02 anys (2.05–7.90)], demostrant que el model publicat al 2019 és

adequat per a la predicció d'arrítmies ventriculars. El model inclou de forma ponderada les següents variables al moment del diagnòstic: sexe, edat, síncope en els darrers 6 mesos, taquicàrdia ventricular no sostinguda prèvia, recompte d'extrasístoles ventriculars, nombre de derivacions anteriors i inferiors amb T invertida i fracció d'ejecció del ventricle dret. En el desenvolupament inicial del model també es va avaluar el valor predictiu de la fracció d'ejecció del ventricle esquerre que va resultar no significatiu a l'anàlisi multivariada. Altres possibles predictors no van ser inclosos en el model ja que no foren significatius en un metanàlisi previ. El model aplicat a la cohort de validació externa va mostrar un calibratge i discriminació adequats [concordança entre el risc predit i observat en les gràfiques de calibratge, pendent de calibratge: 1.01 (IC 95%: 0.99–1.03), Índex C: 0.70 (IC 95%: 0.65–0.75)] i un millor benefici de reclassificació net per un llindar de risc per sota del 35% respecte als altres algoritmes recomanats a les guies de pràctica clínica. Per tant, concloem que la calculadora de risc arrítmic per a la MAVD és una eina vàlida i útil en la estratificació pronòstica d'aquests pacients.

En el seu conjunt, aquesta tesi avalua el paper de la genètica cardiovascular en la comprensió dels mecanismes fisiopatològics subjacents a les MCVs, el valor diagnòstic i pronòstic de la mateixa, així com el rol dels models de predicció de malaltia i estratificació pronòstica en aquestes entitats.

Introduction

1. Prevalence, morbidity, and mortality of cardiovascular diseases.

Cardiovascular diseases (CVDs) have a prevalence in the general population ranging from 0.05% in rare diseases such as Brugada syndrome (BrS) (1) to 7.2% in common diseases such as coronary artery disease (CAD) (2). CVDs encompass a varied spectrum of conditions that affect the heart and the cardiovascular system, by affecting the myocardium, and thus the diastolic and/or systolic function of the heart, as well as the electrical properties of the cardiac tissue which may cause atrial or ventricular arrhythmias (VAs). When VAs occur, and they are sustained and cause hemodynamic instability, its most devastating complication is sudden cardiac death (SCD). SCD represents 50% of cardiovascular deaths, being up to 50% the first manifestation of the disease (3-5). SCD is mainly related to coronary heart disease in the Western world, accounting for up to 75–80% of SCD cases. (6) However, the underlying cardiac diseases associated with SCD vary through lifespan. In the young, primary electric diseases and cardiomyopathies predominate, as well as myocarditis and coronary anomalies. However, half of SCD cases from the fourth decade are related to CAD. (7) The annual incidence rate of SCD is age-dependent, ranging from 2.3 to 21.7 per 100,000 persons in individuals aged 1 to 35 and 36 to 49 years, respectively, and increasing thereafter throughout life. (8) The increase in the annual incidence of SCD across lifespan is due to the increase in acquired heart disease such as coronary heart disease. In the younger strata, although the overall incidence is low, it may be higher in certain individuals with congenital or heritable cardiac diseases with onset in the first decades of life.

Nowadays, one of the therapeutic approaches to CVDs is aimed at prevention: prevention of disease onset in healthy individuals of the general population, i.e., primary prevention, and prevention of progression of the structural abnormalities and the severity of symptoms in case the disease already exists, i.e., secondary prevention.

Disease prevention strategies (primary prevention) of great impact can be promoted from a political and institutional framework, through measures with population impact, such as pollution control, regulation of toxic substances such as alcohol and tobacco, or

promoting healthy lifestyles at the population level as facilitation to resources for physical exercise or through educational programs that favor healthy diet. Likewise, the health system, with the primary health care as the crucial element, carries out interventions at the individual level for the prevention and treatment of individual risk factors for CVD. For example, detecting and treating arterial hypertension or hyperlipidemia, through support programs for smoking cessation, among other disease preventing actions.

Regarding other preventive actions in the realm of CVDs, in patients diagnosed with atrial fibrillation (AF) prophylactic anticoagulant treatment is initiated in certain cases at high-risk of cardioembolic stroke (9). Likewise, in individuals with heart disease at high-risk of SCD, such as in some individuals with heritable or acquired myocardial diseases or some heritable arrhythmic disorders, placement of implantable cardioverter defibrillator (ICD) is recommended in some at-risk individuals to prevent SCD (10).

Secondary prevention and management to mitigate adverse outcomes differ depending on the entity. As an example, in patients with left ventricular (LV) impairment and heart failure (HF), treatment with myocardial remodeling drugs such as beta-blockers, drugs that intervene in the renin-angiotensin-aldosterone and natriuretic peptide pathways, and SGLT2 (Sodium–glucose co-transporter 2) inhibitors, are recommended in order to improve cardiac function and symptoms (11). The instauration of the appropriate treatment in patients affected with HF will not only improve HF-symptoms but also reduce their risk of SCD. (12)

The benefit of detecting individuals at risk for a certain event depends on the ability to detect the event, the potential to intervene (and improve) the course of the disease, and the harm and benefit of both: the event and the suggested intervention. Overall, the intervention is effective when we prevent a serious event in individuals at high susceptibility to suffer it and the harms of the intervention are low. Depending on the pre-test probability of the event, i.e., prevalence, the screening strategy can be recommended at the population level or in sub-groups of interest. Of note, not all individuals with an identifiable condition predisposing to a certain event will develop the event. That is, not all individuals with AF will suffer cardioembolic events and not all

individuals with certain myocardial diseases or heritable cardiac disorders are at risk of SCD. Therefore, an effective preventive strategy will require to identify predisposing conditions and count on risk stratification tools with adequate discrimination and predictive accuracy.

As an example, individuals with AF have a substantial increased risk of cardioembolic stroke. (13) In approximately 20% of the individuals that suffer a stroke in association with AF, AF is first diagnosed at the time of the stroke or shortly thereafter. (14) The use of screening strategies to detect AF through tools such as the 24-hour electrocardiogram or devices capable of detecting cardiac electrical signals or irregular beats for cardioembolic risk assessment and the initiation of prophylactic treatment of cardioembolic events has insufficient evidence. (15) As mentioned above, an effective preventive strategy requires appropriate identification of at-risk individuals, and CHA_2DS_2-VASc cardioembolic stroke risk stratification [congestive HF, hypertension, age ≥75 years (doubled), diabetes, stroke/transient ischemic attack/thromboembolism (doubled), vascular disease (prior myocardial infarction, peripheral artery disease, or aortic plaque), age 65-74 years, female sex category] has been developed from populations of patients with clinically diagnosed AF, not screen-detected individuals in which AF may be subclinical and/or low-burden, and in whom the risk of stroke and the subsequent benefit of prophylactic anticoagulation is under investigation (16, 17).

In the prevention of SCD, specific medical treatment of the underlying predisposing disease and lifestyle adjustment have been shown to be effective in improving prognosis. (10-12) To this end, it is necessary to identify individuals with diseases that may predispose to SCD and initiate appropriate disease-specific treatment. Furthermore, adequate SCD risk stratification among individuals affected by SCD predisposing diseases is essential to identify those who would remain at high risk despite appropriate medical treatment and would benefit from further invasive treatment such as ICD placement. ICDs have demonstrated to be useful in the prevention of SCD, avoiding this devastating complication in sometimes very young individuals with no comorbidities other than those predisposing to SCD. On the other hand, it is important to avoid implanting ICDs in patients who do not need them, since the implantation of these devices is associated with a significant rate of complications and adverse effects. (18)

ICD implantation for secondary prevention of SCD, i.e., in those that have already experienced a life-threating VA, is generally well supported in most cardiovascular entities. Among individuals without previous history of VAs (primary prevention), different strategies are recommended depending on the underlying substrate that may predispose to SCD. Although there is no evidence of the benefit of mass screening programs in the asymptomatic general population for preventing SCD (19, 20), as SCD can still be the first manifestation of a cardiac disease, it is critical to be able to detect the relatively small, high-risk subgroups. This is the case in those affected with certain heritable diseases which may be at risk of SCD regardless of their symptomatology. When these individuals are diagnosed, generally by family screening assessment, adequate identification of those at elevated risk of SCD remains a challenge. Hence arrhythmic risk stratification schemes and risk predictive models, such as the hypertrophic cardiomyopathy (HCM) (21) and arrhythmogenic right ventricular cardiomyopathy (ARVC) (22) risk calculators, have been developed to help to better identify those that will benefit from ICD placement.

2. Pathophysiological mechanisms underlying cardiovascular diseases. Concept of heritability. Rare variants and common variants.

The heritability of CVDs is well demonstrated across the spectrum of the different conditions and, as such, family history for a specific CVD is a major risk factor for an individual to develop that disease. Heritability is defined as the proportion of phenotypical variance explained by the genetic substrate in a population. Heritable diseases, those with a genetic substrate and therefore transmissible from parent to offspring, have traditionally been classified as either *Mendelian* (monogenic) or *complex* (polygenic), signifying that an individual's disease occurrence results from a single gene perturbation or multiple gene perturbations, respectively.

2. 1 Mendelian cardiac diseases ("monogenic" inheritance)

Mendelian inheritance refers to diseases (or traits) that segregate in a family following a specific pattern, such as autosomal dominant or recessive. A simple understanding of Mendelian diseases is that they are *monogenic*, meaning that they are the consequence of a single gene genetic variant within a family. Heritable cardiac disorders, specifically

channelopathies and cardiomyopathies, leading known causes of SCD among young individuals, have been classically considered to have a Mendelian inheritance. These include ARVC, HCM, dilated cardiomyopathy (DCM), Brugada syndrome (BrS) and long QT syndrome (LQTS).

Since the discovery of the first gene related to a familial heart disease, the *MYH7* (Myosin Heavy Chain 7) gene for HCM (23), our understanding of the genetic architecture of heart disorders has markedly improved. The evolution from the techniques that are able to sequence a small amount of genes (Sanger) to techniques able to sequence a wide selection of target genes (Next Generation Sequencing, NGS) to the sequencing of the whole exome and the whole genome has allowed the identification of a greater number of genes and genetic variants that may contribute to heritable cardiac diseases, although not all of them have yet demonstrated a *definitive* association. That is, variants in hundreds of different genes have been reported in association with cardiac channelopathies or cardiomyopathies, of which ~45 have moderate, strong, or definitive evidence for CVD causality (24-29).

The discovery of causal genes and the implementation of genotype testing in clinical practice have clearly contributed to the diagnosis and management of patients carrying heritable cardiac diseases and their families. However, in some cases, sequencing techniques reveal genetic variants in genes lacking a conclusive clinical significance, i.e., variants of uncertain significance, thereby rendering uncertain the clinical relevance of the presence of the variant in non-affected carriers.

It is also important to note that variants with definite evidence of disease causality, i.e., pathogenic, are not determinant for disease onset. That is, variants can be pathogenic but still show “incomplete penetrance.” In this sense, the attribution of heritable cardiac diseases to a single site genetic variant is an oversimplification of a more complex phenomenon. In individuals carrying genetic variants associated with familial heart disease, the proportion of carriers that will develop the disorder, i.e., the disease penetrance, has been reported to be variable, sometimes because these individuals may carry variants in distinct genes, although penetrance may be variable even in individuals carrying distinct variants in the same gene (30).

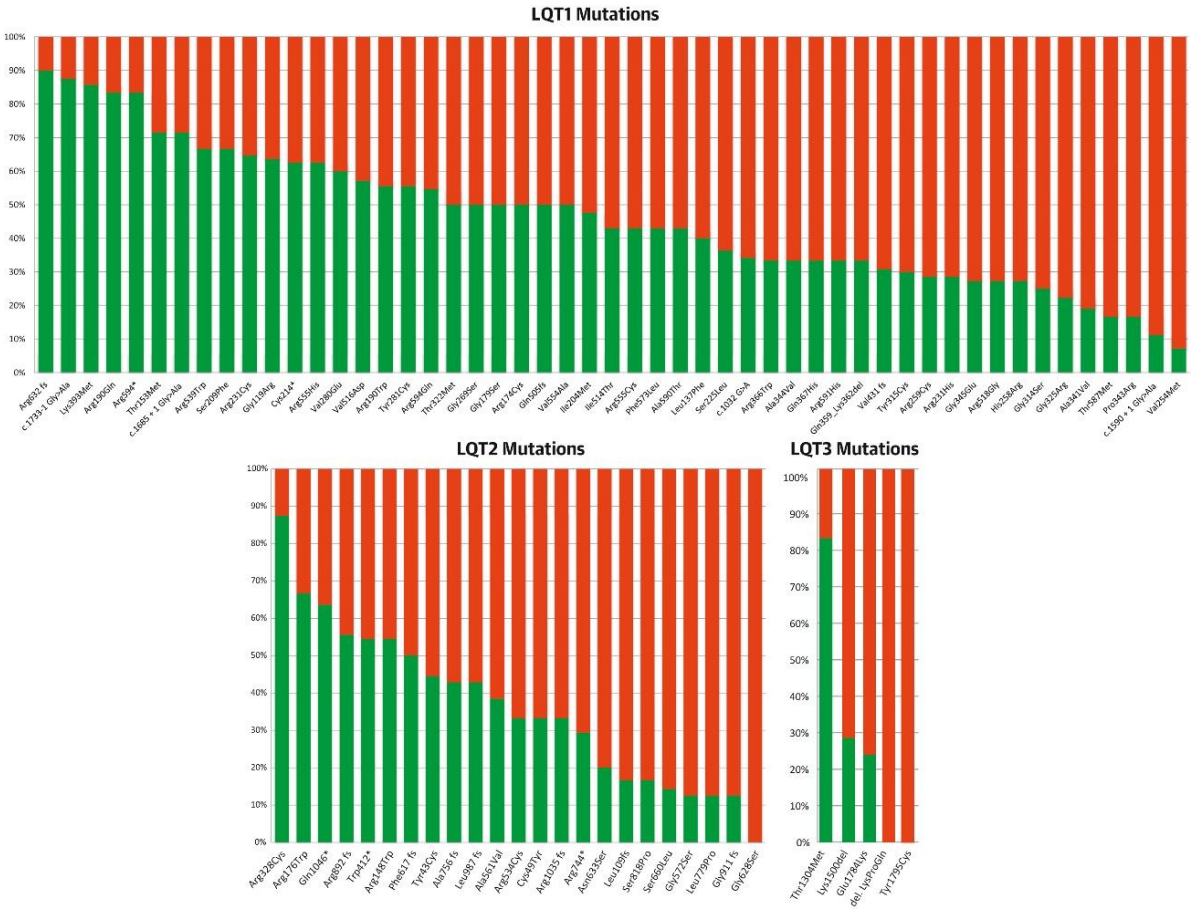


Figure 1. Penetrance of LQTS Syndrome Phenotype According to Genotype. In this figure from Mazzanti *et al.* (30) the penetrance of LQTS was calculated for each variant for which the study population included ≥6 carriers. LQTS 1 variants refer to *KCNQ1* (Potassium voltage-gated channel subfamily Q member 1) gene, LQTS 2 variants to *KCNH2* (Potassium voltage-gated channel subfamily H member 2) gene and LQTS 3 variants to *SCN5A* (Sodium voltage-gated channel alpha subunit 5) gene. Depicted in red is the percentage of patients with overt LQTS, whereas the proportion of patients with concealed LQTS is illustrated in green. This figure nicely displays how patients carrying variants in the same gene can have highly variable penetrance and how even in individuals with the same mutation, the variant may or may not be expressed, implying that other intervening variables come into play to determine the penetrance of the disease. Abbreviations: LQTS, Long QT syndrome.

Besides penetrance, the spectrum of disease manifestations can be highly variable among individuals with overt disease - variable expression-, even among individuals carrying the same pathogenic variant, or even among individuals from the same family, who carry the same variant and may share other genetic and environmental factors, but not all of them.

Although the genotype-phenotype relationship has the aforementioned limitations, it is important to highlight that the implementation of genetic testing in clinical practice is useful for detecting which relatives require clinical follow-up in families with genetic diseases in which the causal genetic variant is detected, and thus avoids follow-up in first-degree relatives who are not carriers of the causal variant, who in the absence of genetic testing or with a negative genetic test in the index case, would require medical follow-up.

2.1.1 Role of genotype testing in prognostic stratification

Beyond the usefulness of genetic testing in family screening, insight into the causal genetic variants in affected individuals can provide valuable guidance in understanding the disease as well as its clinical behavior and therapeutic considerations.

In the case of LQTS, for example, variants in genes encoding for distinct ion channels subunits in the cellular membrane of the cardiac myocyte have been reported. In this entity, variants in different genes imply a triggering of VAs by different stimuli. For example, patients carrying pathogenic variants in *KCNQ1* (LQTS 1) present arrhythmic events with adrenergic triggers, whereas those with pathogenic variants in *SCN5A* (LQTS 3) generally present with arrhythmic events during rest. (31) Likewise, event-free survival is lower in patients with pathogenic variants in sodium channels compared to those with pathogenic variants in the *KCNQ1* and *KCNH2* genes (LQTS 1 and 2, respectively). (31, 32) In the same way, patients carrying multiple mutations, as in those with Jervell and Lange-Nielsen syndrome, are known to have distinct phenotype and prognosis. In this sub-type of disease, the presence of biallelic pathogenic variants in either *KCNQ1* or *KCNE1* (potassium voltage-gated channel subfamily E regulatory subunit 1) is associated with sensorineural deafness, longer QT intervals and worse prognosis. (33)

In the setting of another heritable cardiac disease, DCM, variable expression is also manifest. For example, a patient with DCM may have different degrees of dilatation and dysfunction, of the left or both ventricles, dysfunction and dilatation can range from mild to severe, also dysfunction may be present without dilatation, and these patients may associate cardiac arrhythmias, or may require treatment for advanced HF such as ventricular assist devices or cardiac transplantation. The study of rare genetic variants in

patients with DCM has showed that patients with gene elusive disease have better prognosis than those carrying causative genetic variants; and that, among carriers, those patients carrying pathogenic variants in the *TTN* (titin) gene have a better survival than those with pathogenic variants in the *LMNA* (lamin A/C) gene or in desmosomal genes. (34) Other series have shown that patients with DCM carrying pathogenic variants in *LMNA* (35), *RBM20* (RNA binding motif protein 20) (36) or in *FLNC* (filamin C) (37) show greater predisposition to VAs.

Among other particular forms of presentation in DCM, individuals carrying pathogenic variants in the *DSP* (desmoplakin) gene present a form of arrhythmogenic cardiomyopathy with predominant involvement of the LV, which presents with LV dysfunction, fibrosis in the LV myocardial wall, and episodes of acute myocardial injury, which may overlap myocarditis presentation (chest pain, elevated troponins and normal coronary angiography) (38).

Beyond the affected gene, site-specific variants in a particular gene may confer characteristic prognosis. Following the examples in DCM, it has been suggested that, individuals carrying pathogenic variants in the *RBM20* hot-spot P-R-S-R-S-P between p.Pro633 and p.Pro638, which contains crucial amino acids for the protein structure and function, may show a worse prognosis (36).

The first article of this thesis describes a particular presentation of DCM in patients carrying a specific variant in the *BAG3* (BCL-2 associated athanogene 3) gene, which is associated with high risk of progressive HF. (39)

2.2 Complex cardiac diseases (polygenic inheritance)

The underlying genetic structure of most common adult-onset diseases remains incompletely characterized. The binary classification of CVDs in monogenic (or Mendelian) or polygenic partly stems from the technology and research designs that were primarily focused on detecting rare high-risk genetic factors (monogenic) through family-based linkage analysis versus common low-risk genetic factors (polygenic) via genome-wide association studies (GWAS). However, this naming may still be useful in clinical practice to guide genetic testing and family screening approaches.

GWAS classically investigate the association of common genetic variants, also called single nucleotide polymorphisms (SNPs) with binary traits/diseases or quantitative traits using regression methods. For each of the millions of common variants, i.e., allele frequency ≥0.5/1%, a *P*-value of association is reported, together with an effect size and its standard error.

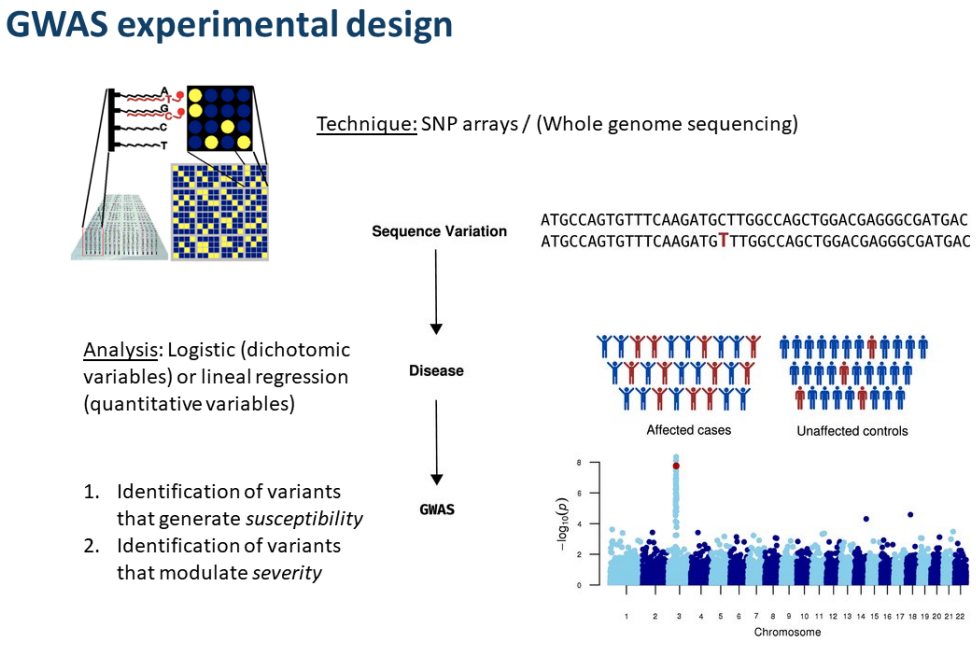


Figure 2. GWAS experimental design. Adapted from Palsson *et al.* (40). GWAS compare allele or genotype frequencies between individuals with a specific disease (cases) and those without the disease (controls). GWAS focus on single nucleotide polymorphisms (SNPs), which are sites in the DNA where individuals may have different nucleotides, e.g., T (for thymine) or C (for cytosine). The goal is to identify genetic variations associated with the disease. In the figure, a T variant (red) is enriched in cases compared to controls. Genotypes are determined using a chemical genotyping method, classically SNP arrays and more recently whole genome sequencing. In case-control studies, SNP genotyping is performed in both cases and matched controls. GWAS are also used for quantitative traits (e.g., cholesterol levels), and in that case linear regression is used to model the relationship between SNP genotypes and the quantitative trait. GWAS results are typically displayed in Manhattan plots (at the bottom right of the figure), which show the significance levels on a -log10 scale against the chromosomal positions of sequence variants. Association signals appear as peaks, indicating regions with significant genetic variations linked to the trait of interest. In general, genome-wide association signals are considered significant from a *P*-value < 5 x 10^{-8}.

In summary, GWAS provide a powerful study design for identifying genetic variations associated with diseases or traits, and they involve extensive genotyping and statistical analysis to detect these associations across the entire genome.

Recent large-scale genetic investigations and advanced analytical methods can offer insights into the most likely genetic architectures of complex traits. (41) Current evidence suggests that in reality, the genetic makeup of common adult-onset diseases likely exists on a spectrum, ranging from common low-risk to rare high-risk genetic variants that may cumulatively contribute to an individual's overall predisposition to a certain condition.

Thus, heritability of complex diseases arises from the contribution of common genetic variants with small additive effects (with effect sizes typically categorized as small: 1.0–1.5, moderate: >1.5, and intermediate: >3.0), combined with a relatively smaller contribution from rare and low-frequency variants with moderate or high effect size in genes known to be associated with familial forms of the disease. (42)

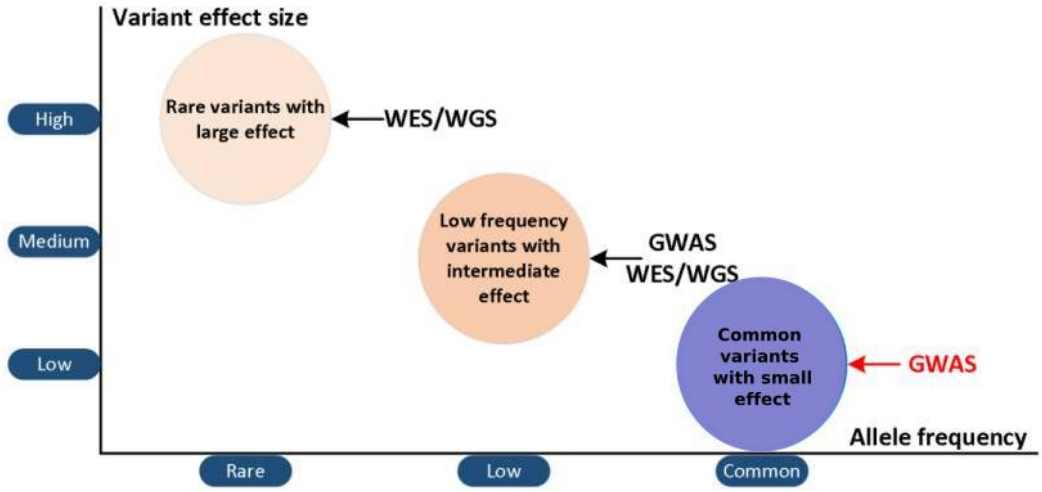


Figure 3. Genetic variants from high (rare) to low effect (common). Adapted from Bai *et al. (43)* WES, whole exome sequencing; WGS, whole genome sequencing. In summary, both statistical modeling predictions and empirical findings from comprehensive genomic studies reinforce the idea that the genetic architecture of many common adult-onset diseases consists of a spectrum of contribution from highly penetrant rare variants (linked to familial form presentations) to the amalgamation of common variants with small effects distributed throughout the genome, along with a smaller contribution from low-frequency variants with

moderate effects in genes associated with familial disease (which lead to forms of disease with less familial aggregation).

Variants with a modest effect, but more common, and variants with a moderate or large effect, but less common, can coexist, yielding the genetic architecture of the disease which may interact with additional environmental factors (Figure 4). This genetic background may vary between individuals of different ancestries, due to the founder effect and/or the variants that are co-heritable (depending on their location and the rate of genetic recombination). (44)

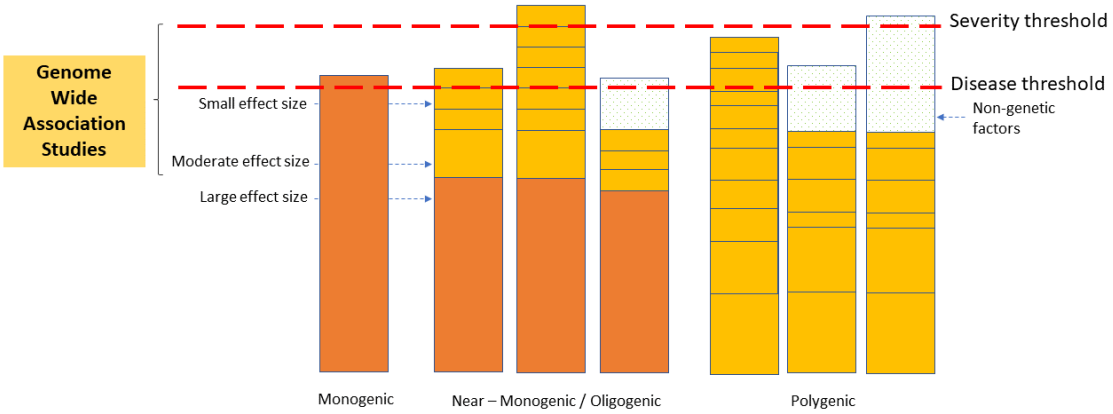


Figure 4. Genetic substrate and disease threshold. Adapted from Cerrone *et al.* (44)

GWAS have identified several genetic variants in association with complex cardiac diseases, which in turn have revealed relevant genes involved in cardiovascular pathophysiology. As an example, in AF, variants clustering on chromosome 4q25, mostly mapping to *PITX2* (paired like homeodomain 2) gene, have demonstrated to increase risk for AF among carriers. (43) This is the case of the allele rs6817105-C, which provides increased risk for early onset AF with an odds ratio (OR) of 2.43 [95% confidence interval (CI): 2.12-2.78] in individuals with Asian ancestry (45) and increases risk of AF at any age with an OR of 1.64 (95% CI: 1.55-1.73) in European individuals. (46) As another example, in CAD variants mapping *LPA* gene (lipoprotein A), on chromosome 6, have shown to be associated with increased risk of CAD [rs140570886-C, OR of 1.92 (95% CI: 1.68-2.20)] for CAD. (47) Genetic discovery of loci by GWAS has uncovered and/or reinforced relevant pathophysiological pathways implicated in CVD, although due to the small effect

of these SNPs, individually they do not confer a strong predisposition to the disease with which they are associated. In contrast, the cumulative effects of all common variants associated with a given risk/disease by polygenic scores (PGS) have been demonstrated to be linked to CVD risk (48).

2.2.1 Polygenic scores methodology

PGS aggregate the additive effect of each SNP on a specific phenotype by summing the effect of each allele multiplied by the dose of risk alleles (0, 1 or 2) present for each independent locus in an individual. This generates a score for each individual that separates individuals into different relative risk categories for a particular disease or event.

Traditionally, methods for constructing scores included only SNPs with significant *P*-value ($P<5x10^{-8}$). However, we know that genome-wide significant variants only explain a part of the phenotypical variance of complex diseases while a higher sub-group of variants might explain a higher proportion of this variance. For example, no single common variant explains more than 0.1% of schizophrenia risk; genome-wide significant variants explain only a 3%, whereas ~10,000 SNPs together explain 18% of the variance between schizophrenia cases and controls. (49, 50) Accordingly, the PGS methodology has evolved towards the inclusion of more variants, beyond the genome-wide significant threshold, in the calculation of the scores. Since the sum of the variants implies an independent additive genetic architecture, PGS methods account for the co-inheritance of certain variants, i.e., linkage disequilibrium (LD). SNPs and their effects can be included employing its raw effect or adjusting it according to Bayesian models and specific hyperparameters. Clumping and Thresholding methods select independent SNP (and their raw effects) which reach a certain *P*-value and LD r^2 thresholds. Bayesian methods, such as LDpred, its newer version LDpred2 (51), and PRS-CS (52), employ the fraction of causal variants (which is in correlation with the grade of polygenicity of the trait), and the heritability of the phenotype in LDpred and LDpred2, to adjust SNP effects according to a prior distribution specification. These values are usually pre-specified by the investigator and then the PGS that best predicts the trait in a validation cohort is selected to be applied in the population of interest in an independent testing dataset. The best

hyperparameters can also be learned directly from the summary statistics, so PGS methodologies also implement "auto" methods that allow the calculation of the score without the need of a validation dataset as in the case of LDpred2 and PRS-CS.

That is, PGS methodology modeling can undergo recalibration of the SNP effects which may help to address biases in effect sizes, which are often inflated in the initial discovery cohort. It could also entail accounting for multiple linked variants within each disease-associated locus, re-evaluating effect sizes for a specific sub-phenotype of interest or adjusting for ethnic or demographic factors that could impact the generalizability of the models. (42)

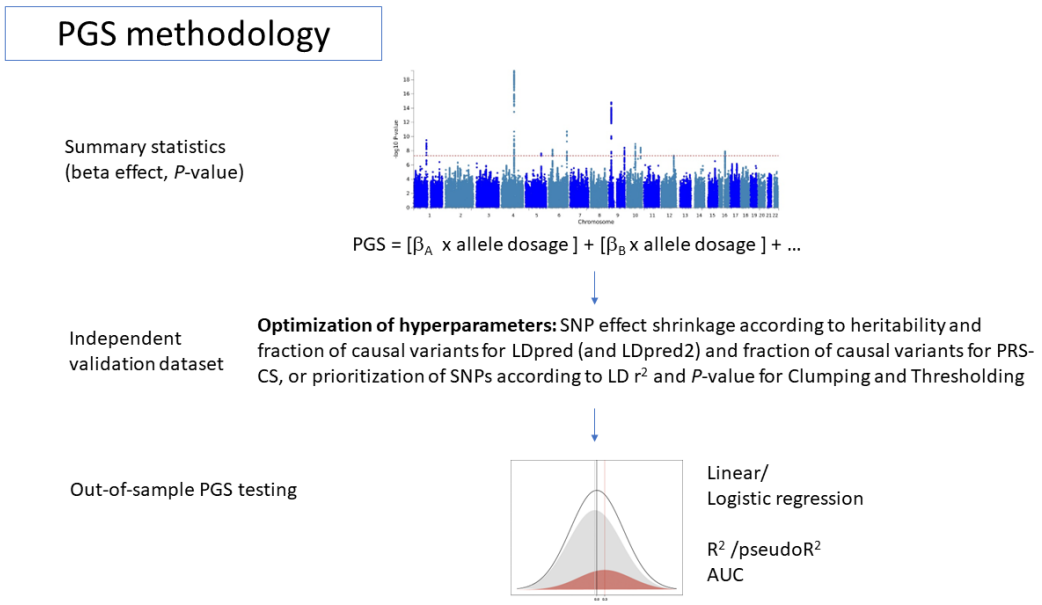


Figure 5. Polygenic scores methodology scheme. AUC, area under the curve; β, beta; PGS, polygenic score.

The units of the GWAS effect sizes determine the units of the PGS. In binary traits, i.e., case/control phenotype, PGS values are computed in relation to a hypothetical individual with the non-effect allele at every SNP, and, thus, they provide a relative estimate of risk rather than an absolute estimate.

While PGS are not yet integrated in routine clinical care, they have proven to be useful in the identification of individuals at risk of *complex* diseases, equivalent to risk attributable to "monogenic genetic variants" (48). Khera *et al.* demonstrated that a PGS in coronary heart disease, derived from a GWAS including 60,801 cases and 123,504

controls, identified 20 times more individuals with a comparable or greater risk than that found for familial hypercholesterolemia rare genetic variants. (48, 53)

2.2.2 Mendelian diseases are not only monogenic

Beyond the usefulness of PGS in complex diseases, recent data have demonstrated a significant contribution of common genetic variants also in the so-called *Mendelian* diseases. (54-56) In this more comprehensive paradigm the presence of a "pathogenic variant" is not always necessary or sufficient for phenotypic manifestation, but the disease may manifest under near monogenic, oligogenic or polygenic models (Figure 4). (57, 58)

In heritable cardiac diseases, the understanding of the contribution of polymorphisms may help identifying molecular subtypes of disease and the variability in disease expression, which could prompt more accurate risk stratification. In this setting PGS can be a valuable predictor of disease prognosis in individuals carrying high-risk rare genetic variants, either alone, or in combination with clinical predictors.

As an example, among carriers of pathogenic genetic variants for LQTS, modulatory variants may contribute to disease expression, such as is demonstrated with the *KCNE1* variant p.Asp85Asn (rs1805128), which increases QT duration by 7.42ms per minor allele (59). Likewise, multiple polymorphisms with a modest effect can jointly modify disease expression as demonstrated in a study by Turkowski *et al.,* in which a PGS including 61 SNPs previously associated with QT duration was significantly associated with increased QT duration in *KCNQ1* (LQTS 1), *KCNH2* (LQTS 2) and *SCN5A* (LQTS 3) variant carriers with no differences in the PGS effect according to the affected gene, and no prognostic impact of the PGS (no association with the occurrence of symptoms or events, *n*=423 patients, mean age at the end of the study 35.6±16.9). (60) In another study only including *KCNH2* pathogenic variant carriers*,* an analysis of the cumulative effect of 6 SNPs located in the *NOS1AP* gene by means of a PGS uncovered a strong linear relationship between *NOS1AP*-PGS and the QTc-interval (P=4.2×10^{-7}). Furthermore, individuals in the lowest quartile of the *NOS1AP*-PGS had a lower relative risk of cardiac events compared with those in the other quartiles combined (P=0.039) (61).

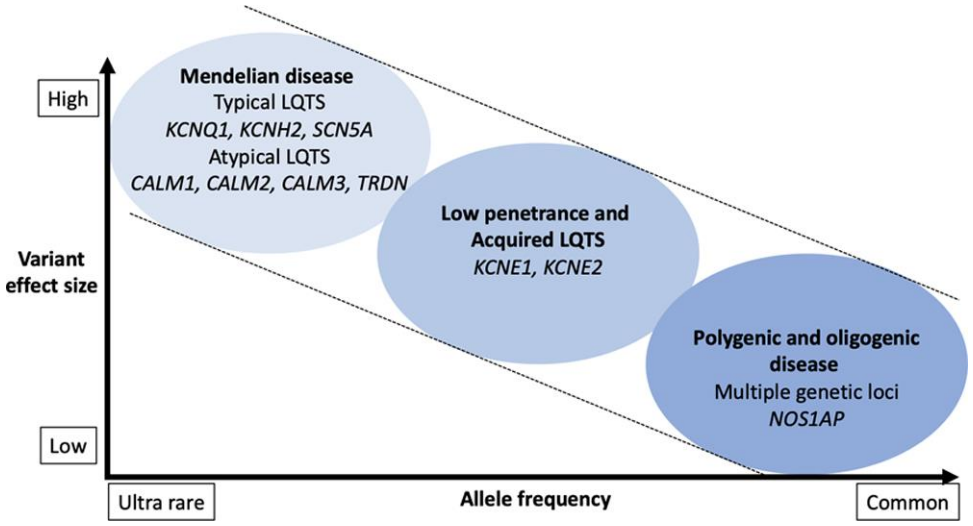


Figure 6. Genetic variants from high (rare) to low effect (common) in LQTS. Image from Ingles *et al*.(62)

To add flexibility to the genetic makeup of CVDs, we could reframe the previous vision of the *KCNE1* variant p.Asp85Asn as a modifying variant to a low-penetrance low-frequency variant that could, in turn, be modulated by other genetic or nongenetic factors such as QT-prolonging drugs. (44, 62)

Similarly, in a different disease such as HCM, a recently published investigation has provided a PGS (PGS_{HCM}) that is associated with maximal LV wall thickness (maxLVWT). In this study, each standard deviation (s.d.) of the PGS_{HCM} increase corresponded to a clinically relevant 1.4 mm absolute increase in maxLVWT. In this case, the PGS_{HCM} was also associated with adverse clinical events (a composite of septal reduction therapy, cardiac transplantation, sustained VA, SCD, appropriate defibrillator therapy or AF/atrial flutter), where each s.d. increase in PGS_{HCM} was associated with a 28% relative risk increase in adverse clinical events [hazard ratio 1.28 (95% CI: 1.06–1.54), $P=9\times10^{-3}$]. (55) Similarly, another study (63) demonstrated that a 1 s.d. unit increase in another PGS in HCM patients conferred a 0.71 ± 0.35 mm increase in maxLVWT ($P=0.048$) in carriers of *MYBPC3* truncating variants ($n=232$) and a 0.73 ± 0.36 mm increase ($P=0.037$) in carriers of *MYH7* missense variants ($n=186$). The most frequently observed pathogenic variant in HCM, *MYBPC3* p·Arg502Trp, was associated with a larger PGS effect size: 1.61 ± 0.80 mm increase per 1 s.d. unit increase in the PGS ($n=48$). These findings support the fact that the expression of the disease depends on monogenic and polygenic interaction, with

different contributions of the cumulative effect of common variants which may be gene and even variant specific.

Interestingly, in patients with clinically well-defined disease but not carrying rare genetic variants, greater influence of the cumulative effect of common polymorphisms (polygenicity) has been observed for LQTS(54), BrS (64) and HCM. (63) This underlying polygenic substrate may be captured by PGS and help to clarify susceptibility to disease among relatives and identify those who will benefit from clinical follow-up. Currently, clinical follow-up is recommended for almost lifelong in first-degree relatives of heritable cardiac disease patients, thus, implementation of PGS would allow to restrict it only to those with high scores. However, whereas gene-elusive cases can be due to polygenic inheritance, some might be linked to rare variants that have not been captured by clinical genetic testing including deep intronic variants affecting splicing, variants in regulatory regions, and variants in yet to be identified genes. (65)

PGS have emerged as a useful tool to assess the susceptibility to certain pathologies, allowing through a single measurement and with the information present from birth the screening for multiple conditions at once. Considering this framework in relation to the genetics of CVDs, PGS analyses can be used to predict disease-status or to point out in which individuals there is a greater pre-test probability of diagnosis, to ascertain diagnosis, to assess prognosis and to improve accuracy of clinical predictive models, and to contribute in disease-profiling in order to assess therapeutic response.

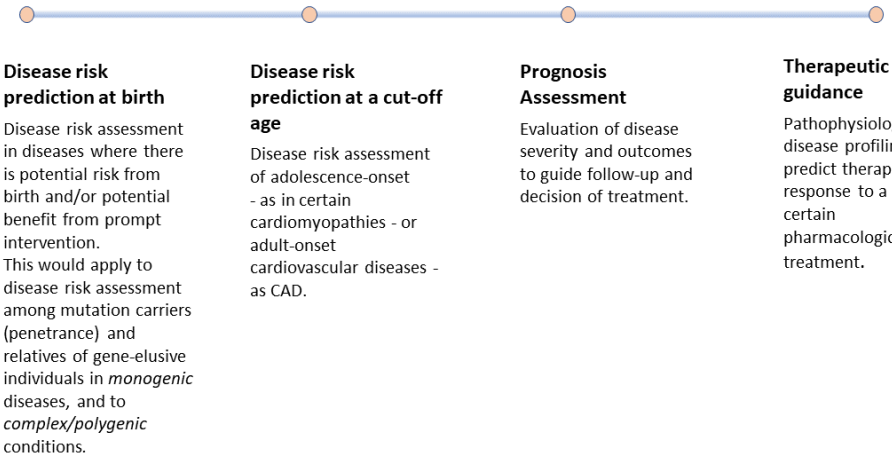


Figure 7. Potential applications of PGS at different points in the natural course of the disease and the diagnostic-therapeutic process. CAD, coronary artery disease.

On the process of evaluating the usefulness of PGS, it is important to consider how they are calculated and reported and how their clinical utility is assessed. Once the PGS has been calculated typically a regression is performed in the target sample, with the PGS as a predictor of the target trait and inclusion of covariates as appropriate. (66) The association between PGS and outcome can be measured with standard association or goodness-of-fit metrics, such as the *P*-value derived in testing a null hypothesis of no association, phenotypic variance explained (R^2), effect size estimate (beta or OR) per unit of PGS or between specific strata (e.g., high- versus low-risk individuals), or with measures of discrimination in disease prediction, such as the area under the receiver operator curve (AUC), the net reclassification index (NRI) and the integrated discrimination improvement.

There are many options on how PGS can be reported, including just indicating whether an individual has a score at the extreme high tail of the distribution (i.e., a binary report of identified/not identified as at high risk), or report of a continuous variable such as the PGS percentile, and report relative or absolute risk being the later the preferred option.

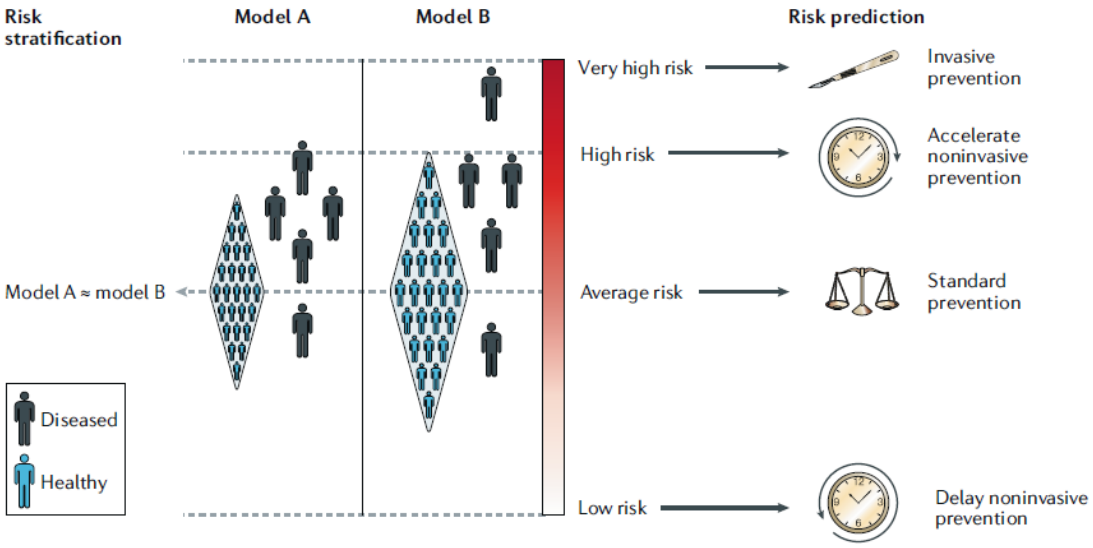


Figure 8. Contrasting risk stratification versus risk prediction. This figure from Torkamani *et al. (42)* illustrates the distinction between the usefulness of risk data for only ordering individuals according to their risk or informing both, risk strata and absolute risk of disease. In the middle of the diagram, two hypothetical models (A and B) are presented, with the vertical positioning of human figures representing the likelihood of an individual having or developing a

disease. When we compare these two models, the relative proportion of individuals with the disease (depicted in black) to those without it (depicted in blue) is approximately the same. In other words, the vertical alignment of the black figures relative to the blue ones is similar in both models, resulting in receiver operating characteristic curves with roughly equivalent area under the curve values. However, a significant difference exists in the absolute vertical placement of all figures, especially the highest-risk individual depicted in black, between the two models. This discrepancy leads to distinct conclusions regarding the utility of each model, as illustrated on the right side of the diagram. Specific thresholds for identifying individuals with a high or low risk of disease would vary depending on the specific disease and the balance between the risk and benefit of intervention. It is important to note that in this hypothetical example, altering disease probabilities for individuals at the extreme ends of the distribution does not significantly impact area under the curve measurements. However, it does result in varying implications for the utility of risk estimates.

2.2.3 Derivation data for polygenic scores.

2.2.3.1 Genetic substrate overlap across traits: Concept of endophenotype and pleiotropy.

PGS can be directly derived from a GWAS corresponding to the trait/disease under study or from an endophenotype of that trait/disease. An endophenotype is a heritable, quantifiable trait or characteristic that is intermediate in nature and believed to be related to a specific genetic component of a complex and multifactorial condition. These traits are often used in genetic research to help to identify the underlying genetic factors that contribute to the development of a particular condition. Quantitative endophenotypes are intrinsically more potent traits for genetic association studies compared to binary disease status, and they are frequently more accessible to be assessed.

In this sense, risk factors of certain diseases can be employed in their prediction. For example, a variant that increases low-density lipoprotein (LDL) cholesterol may also have corresponding downstream effects on coronary heart disease risk due to the causal relationship between these two traits. That variant exhibits "vertical" pleiotropy.

Vertical pleiotropy has been conceptualized and measured by explicit genetic methods such as Mendelian randomization. In contrast, a genetic cause that directly influences multiple traits, without one trait being mediated by another, would show "horizontal" pleiotropy. (67)

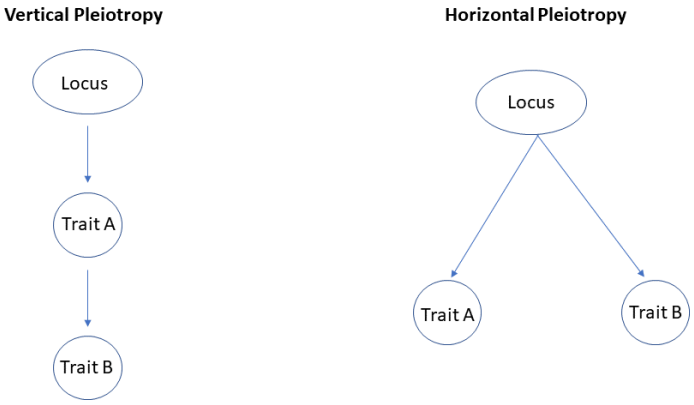


Figure 9. Depiction of vertical versus horizontal pleiotropy.

Following the endophenotype concept, the susceptibility to a certain disease can be a consequence of a heightened expression of a pathophysiological pathway or be the result of the cumulative effect of concurrent mechanisms, where the disease is an extreme form of the trait, i.e., the LQTS that can be perceived as the maximum expression of the mechanisms that lead to longer QT duration. This approach understands cardiovascular traits and diseases as a continuous spectrum underlying shared mechanisms; SNPs predicting an intermediate trait - endophenotype - are thus used to predict the disease - corresponding to an extreme form of the trait.

These shared pathways across different traits and diseases may be differentially involved as recently demonstrated by Tadros *et al*. (55) In this work, it is shown how previously identified lead risk alleles (rs10927875-C and rs2234962-T) predisposing to DCM (68) were found to be protective for HCM. Both loci were also found to be associated with LV ejection fraction (LVEF) in the general population (69), for whom DCM risk alleles decrease LVEF (contractility), while HCM risk alleles increase LVEF. This data indicates that genetic loci underlying variability of LV function in the general population may be differentially involved in susceptibility to HCM and DCM.

2.2.3.2 Genetic substrate overlap across traits: How to leverage shared genetic structure to improve GWAS estimates.

A useful tool to capture overlapping mechanisms across different traits/diseases is genetic correlation. This genetic overlap can be in turn used to leverage SNP to trait GWAS associations. Multi-trait analysis of GWAS (MTAG) is able to combine the effect of genetically correlated traits to improve trait-specific effect estimations. MTAG uses an inverse variance weighted meta-analysis, which employs the variance of the effect estimators of each study as a weight to combine the traits and adjust results according to the genetic correlation matrix to output trait-specific association statistics. The methodology employed, allows the inclusion of GWAS summary statistics from overlapping samples as an input in the analysis. (70)

GWAS have been successful in identifying thousands of associations between common genetic variation and human complex traits and diseases. However, a very large proportion of the phenotypic variance explained by significant SNPs remains unexplained for the vast majority of human traits and diseases. (71)

One of the main limitations of GWAS is that since the effects of each predictor element (SNP) on the outcome or phenotype of interest is small, large populations with thousands or even millions of individuals are needed to be able to detect significant associations and precision in the measurement of the effect. In this sense, MTAG, which enhances SNP to trait associations, has demonstrated to be useful in the discovery of new genomic loci and new pathophysiological pathways and in establishing more accurate genomic prediction. (70)

In the second article of this thesis, we employ MTAG enhanced SNP-phenotype associations for the discovery of new loci and relevant pathophysiologic pathways involved in three common cardiac diseases such as AF, CAD, and HF; and to test whether MTAG derived PGS increase predictive ability and discrimination in common CVDs.

3. From risk/prognostic factors to clinical models

3.1 Predictive models in common cardiovascular diseases

Prevention of CVD requires early identification of individuals at risk in order to apply appropriate primary prevention measures, such as dietary, lifestyle or pharmacological treatment. Over the past two decades, numerous prediction models have been developed, which combine multiple predictors each with a given weight within the score, the calculation of which places each individual in a particular risk stratum, implying a specific event probability. A large number of these models have focused on the prediction of CAD.

Estimation of risk of major cardiovascular events, including recurrences, serves as a valuable tool for patient education. These estimations empower patients to better understand their personal prognosis, leading to increased health motivation and improved adherence to prescribed therapies. Moreover, individual risk assessment plays a crucial role in guiding medical decision-making regarding preventive treatments.

In particular, individuals deemed to be at high risk stand to gain greater benefits from preventive treatments, such as cholesterol-lowering and blood pressure-lowering interventions. (72) High-risk patients experience more substantial absolute risk reduction, resulting in a lower number needed to treat (NNT) for any form of preventive treatment. (73)

The scores recommended in the 2021 clinical practice guidelines on cardiovascular prevention of the European Society of Cardiology (ESC) (74) are compiled in the U-Prevent platform interactive website. U-Prevent encompasses risk calculators for all categories of patients: the SCORE2 for European apparently healthy individuals (75), and a competing risk-adjusted elderly-specific score for people over 70 years, the SCORE-OP (76) , the ASCVD for North-American apparently healthy individuals (77), the ADVANCE risk score for patients with type 2 Diabetes Mellitus (78) and the SMART risk score for patients with previous vascular disease (79).

Furthermore, U-Prevent introduces innovative lifetime risk algorithms specific to each of these patient categories. These include the LIFE-CVD model for apparently healthy individuals aged 45–80 years (80), the DIAL-model designed for diabetes patients aged 30–85 years (81), and the SMART-REACH model tailored to vascular patients aged 45–80 years (82).

In addition, lifetime models provide lifetime treatment effect estimators for common preventive interventions such as blood pressure lowering, lipid lowering, and platelet aggregation inhibition which may facilitate shared decision-making regarding the initiation or discontinuation of cardiovascular preventive therapies.

According to the 2021 ESC guidelines on cardiovascular disease prevention (83), lifetime calculators can be particularly beneficial as educational aids for individuals under 50 years of age. However, it is anticipated that these calculators can also enhance patients' understanding of their cardiovascular prognosis and increase motivation for therapy in individuals over 50 years old.

Certain patient groups can derive significant benefits from lifetime estimations of cardiovascular prognosis and treatment effects. Younger patients might have a seemingly low 10-year risk due to their age, even if they have significant risk factors such as hypertension or hyperlipidemia. However, this low 10-year risk might obscure a substantially high lifetime risk. Conversely, older patients may have a naturally elevated 10-year risk, but the potential benefit of preventive treatment could be limited by their shorter life expectancy.

As an example, a 40-year-old male, with total cholesterol 7.0 mmol/L (272 mg/dl), LDL 4.4 mmol/L (171 mg/dL) and HDL 1.8 mmol/L (69 mg/dL) and no other comorbidities has a 1.6% 10-year risk SCORE2 (low-risk). In this setting, the treatment goal for LDL cholesterol would be < 100mg/dl, thus, initiation of treatment would be indicated. This same individual has a LIVE-CVD score of 31.3% which would be reduced to 12.2% with aggressive statin treatment (e.g., atorvastatin 80mg daily). When we check the CVD-free life years in the lifetime score model, the median life expectancy without vascular disease is 87 years old which increases to 89 years old with the treatment. The use of statins in this individual would then increase life expectancy free of events by two years,

but at a very advanced age in which maybe some other comorbidities may have appeared. Moreover, the use of statins, specially when using high doses and when started from young age, may predispose to diabetes mellitus (84, 85) and other well known side effects as myopathy. Therefore, even with the availability of all the abovementioned models, the benefit of treatment is not always so obvious.

In addition to age-related source of bias there is also an ancestry-related bias in risk estimation. Using data from the United Kingdom, it is recommended to apply specific correction factors when assessing CVD risk with risk calculators in individuals with African and Asian ancestry. (86)

Prediction models currently recommended in the clinical practice guidelines do not include PGS. PGS could contribute prognostic stratification, mainly because they can capture metabolic profiles with age-dependent penetrance that have not yet manifested and thus be especially valuable in the prediction of lifetime risk in the young. For example, in the case of the same patient discussed previously, in a scenario in which PGS would have proven their clinical utility, a high PGS would support the therapeutic intervention, e.g., treatment with statins, and a low PGS could support the decision of not initiating statins to avoid unnecessary side effects.

Regardless of the promising predictive ability of PGS in CAD (48), to investigate whether they capture pathophysiological underpinnings independently of risk factors included in the current models and to estimate the added value of PGS, we need to interrogate their usefulness in relation to already validated prediction models as well as to assess their implementation in relation to the current standard of care.

The clinical application of PGS may be particularly interesting if we consider that one of the main limitations of current predictive models is that they provide a snapshot of the risk of the individual at a given time, requiring iterative risk reassessment. PGS have the advantage that they require a single measurement and are likely to capture the potential risk of multiple disease risk factors in combination even before they manifest themselves. In patients with intermediate or high PGS this information can help us to decide on a closer clinical follow-up and to institute therapeutic measures when appropriate.

In other cardiovascular entities such as AF, several predictive measures have been suggested for anticipating the onset of AF, yet none have yet been extensively adopted in clinical settings. The CHARGE-AF (87) consortium published a 5-year predictive model derived from three cohorts from the United States which integrates predictors such as age, ancestry, height, weight, systolic and diastolic blood pressure, current smoking status, use of antihypertensive pharmacological treatment, diabetes mellitus, history of myocardial infraction and HF. Recent research denotes that the CHARGE-AF score provides good discrimination to predict incident AF. (88) The implementation of the CHARGE-AF score within clinical practice and the impending utilization of mobile health technologies and smartwatches to detect AF could likely yield improved prediction and more effective cardioembolic stroke prevention among individuals at high risk. (88) However, as mentioned before, risk stratification and benefit of anticoagulation in subclinical AF is under current investigation. AF and stroke PGS, in combination with clinical models and health technologies, can help to predict individuals at higher risk of AF-related cardioembolic events, allowing the definition of a target population in which screening measures have a higher yield, i.e., identify proarrhythmogenic and prothrombotic factors in one shot.

Despite all the expectations raised, PGS still have many limitations. Of those, it is particularly worrying that available data to derive PGS mostly comes from European ancestry populations, thus limiting transferability and guaranteed effectiveness in other populations, which may increase healthcare disparities. (89) In addition, PGS are limited to additive effects of genomic variants on the phenotype of interest and do not account for other type of genetic and DNA-environment interactions that may trigger or silence the underlying genetics.

3.2 Predictive models in heritable cardiomyopathies

Risk prediction models have demonstrated to be also useful among cardiac diseases classically considered as monogenic or Mendelian. One of the most feared complications in patients with these entities is SCD. As mentioned above, there are lifestyle modification and/or pharmacological measures that have been shown to be effective in the prevention of SCD. When these measures are insufficient, ICD implantation may be

necessary in high-risk individuals. In the field of cardiac cardiomyopathies, there is consensus that there is a high risk of SCD in patients who have already had a severe arrhythmic episode, and that in this case ICD implantation is beneficial in preventing SCD. The challenge appears when stratifying risk in individuals who have never had an event but are at high potential risk of it. When the prevention of SCD involves the implantation of an ICD, it is important to bear in mind that it is an effective treatment but associated with a high rate of complications, so that the harm derived from implantation in patients who are not at real risk can on some occasions outweigh that inherent to the baseline condition if we do not stratify accurately. Thus, we need stratifying tools with high sensitivity and high specificity to adequately protect at-risk individuals while avoiding unnecessary ICD implantation.

3.2.1 Arrhythmic risk stratification in dilated cardiomyopathy and non-dilated left ventricular cardiomyopathy

In recent years, multiple efforts have been made to combine the weighted value of risk factors through clinical prediction models. In the case of cardiomyopathies, and specifically in DCM and non dilated LV cardiomyopathy (NDLVC), there is growing consensus that LVEF alone does not adequately stratify all genetic and molecular subtypes of disease. DCM is defined by LV dilatation, usually accompanied by dysfunction, and NDLVC is defined by LV dysfunction (LVEF < 50%) without dilatation, potentially with the presence of abnormalities in late gadolinium enhancement in the cardiac magnetic resonance (LGE-CMR) imaging sequences. The treatment of these entities is mainly based on pharmacological agents that improve myocardial remodeling and prevent and ameliorate HF symptoms, and on the treatment and prevention of atrial and VAs, including appropriate risk stratification of SCD. Traditionally, individuals with DCM/NDLVC and LVEF < 35% were those considered at high risk of SCD. Current evidence suggests a pivotal role of LGE-CMR in the reclassification of DCM and NDLVC SCD risk groups (90, 91), with the identification of patients at risk of VA even with LVEF > 35% and high-risk LGE-CMR pattern, and at not so high risk even with LVEF < 35% in the absence of high-risk LGE-CMR pattern (91). Future efforts that combine the LGE-CMR pattern with the presence of pathogenic variants in genes associated with high arrhythmic risk will certainly be useful in risk stratification of DCM. In the case of high-risk genetic subtypes

[*LMNA*, *EMD* (Emerin), *TMEM43* (transmembrane protein 43), *DSP*, *RBM20*, *PLN* (phospholamban), *FLNC*-truncating variants] there is evidence of increased arrhythmic risk even with LVEF > 35%, although genotype alone is not sufficient to predict arrhythmic prognosis, and risk factors in addition to LVEF such as sex or the presence of non-sustained ventricular tachycardia (NSVT) or ventricular extrasystole count may play a role. In this regard, risk models such as the LMNA-risk-VTA (92) have been developed. This model aims to stratify the risk of SCD by integrating clinical variables as well as information on the mechanism of the deleterious variants in the *LMNA* gene. A similar approach, the *PLN* risk calculator, aims to stratify arrhythmic risk among carriers of the p.Arg14del variant in *PLN* (93).

3.2.2 Arrhythmic risk stratification in hypertrophic cardiomyopathy

In another cardiomyopathy, HCM, similar efforts have been undertaken with the development of the HCM Risk-SCD (21), a tool that has been assessed in multiple geographically diverse cohorts (94-99) -European, American, and one Asian - and in a meta-analysis of some of these published cohorts. (100)

HCM is defined by an increase in LV thickness (≥ 15mm, or ≥ 13mm in relatives or carriers of pathogenic variants), with or without right ventricle (RV) hypertrophy, which is not solely explained by abnormal loading conditions. Mutations in eleven genes encoding sarcomere-associated proteins have been demonstrated to be causative for HCM. (28) *MYH7* and *MYBPC3*, encoding β-myosin heavy chain and myosin binding protein C3, respectively, are the two most common genes involved, together accounting for about 50% of the HCM families. In approximately 40% of HCM patients the causal genes remain to be identified. (101) The treatment of HCM is focused on treating the possible diastolic dysfunction derived from the increase in myocardial LV thickness, to the reduction of the dynamic obstruction of the outflow tract of the LV (if present), and to treat atrial arrhythmias and VAs that develop in a myocardium with myocardial disarray, and with eventual fibrosis.

SCD is a devastating complication of HCM, which can sometimes occur in asymptomatic individuals or those with few symptoms. Although SCD occurs at an annual rate of less than 1%, (102) there is a subgroup of patients at increased risk. Because HCM is a disease

with a heterogeneous presentation, identifying the subgroup of patients at risk for life-threatening arrhythmias is crucial. Primary prevention HCM-SCD risk stratification algorithms and tools by established risk factors are depicted in Table 1.

	AHA/ACC 2011 (103)	**HCM SCD-risk score** (21)	**ESC 2014** (104)	**AHA/ACC 2020** (105)	**ESC 2023** (106)
Age	X	✓ (cont.)	✓[1]	Considered (see note)[2]	✓[1]
NSVT	Class IIb[3] (Class IIa if additional RM[4])	✓	✓[1]	Class IIb[3]	✓[1]
LV hypertrophy	Class IIa (≥30mm)	✓[5] (cont.)	✓[1,5]	Class IIa (≥30mm)	✓[1,5]
Abnormal blood pressure response to exercise	Class IIb (Class IIa if additional RM[4])	X	IIb[6]	X	X
Family history of SCD	Class IIa	✓	✓[1] (if multiple and at young age, IIb)[6]	Class IIa	✓[1]
Unexplained syncope	Class IIa	✓	✓[1]	Class IIa	✓[1]
LV outflow tract gradient	Considered as RM	✓ (cont.)	✓[1]	X	✓[1]
Left atrial size	X	✓ (cont.)	✓[1]	X	✓[1]
LVEF	X	X	X	IIa (≤ 50%)	IIb (< 50%)
LGE-CMR	Considered as RM	X	X	IIb (≥ 15%)	IIb (≥ 15%)
LV apical aneurysm	Considered as RM	X	X	IIa	X

Table 1. SCD-risk factors included in the risk stratification algorithms and tools for primary prevention of SCD and indication of implantable cardiac defibrillator in adult patients with HCM. Levels of recommendation: class IIa, implantable cardiac defibrillator is reasonable/can be useful; class IIb, implantable cardiac defibrillator (ICD) may or might be reasonable/may or might be considered.

Abbreviations: AHA/ACC, American Heart Association/American College of Cardiology; Cont., continuous; ESC, European Society of Cardiology; LGE-CMR, late gadolinium enhancement on cardiac magnetic resonance; LV, left ventricle; LVEF, left ventricle ejection fraction; NSVT, non-sustained ventricular tachycardia; RM, risk modifiers.

[1]As per HCM-SCD risk score estimated risk: ≥ 4 and < 6%, class IIb; ≥ 6 %class IIa.

[2]In the guidelines, it is stated that due to the very low SCD rate observed in individuals >60 years old with HCM, the risk stratification algorithm is most applicable to individuals below this age.

[3]According to AHA/ACC 2011 guidelines NSVT are more relevant for individuals < 30 years old. According to AHA/ACC 2020 the previous applies in addition to greater value of NSVT as a predictor if frequent, longer, and faster runs.

[4]Risk modifiers: LGE-CMR, LV apical aneurysm and genotype. LV outflow tract obstruction is considered an increasing risk modifier in patients with NSVT but renders uncertain utility of ICD in patients with HCM with an abnormal blood pressure response.

[5]Caution applies if hypertrophy ≥ 35mm as the HCM-SCD risk score does not perform well in this category.

[6]If HCM SCD-risk score < 4% and abnormal blood pressure or If history of SCD in multiple relatives and at young age decisions on ICDs must be made on an individual basis.

Of note, genotype was considered a risk modifier in 2011 but current evidence suggests that there is no conclusive information confirming that specific sarcomeric variants independently contribute to the prediction of SCD. (106)

There is agreement that there is a correlation between the sum of risk factors and the incidence of SCD but individual risk factors in isolation have poor positive predictive value for SCD and the simple summation model has a relatively poor predictive performance (103) and leads to the implantation of ICD in low-risk HCM individuals (107).

In light of the overestimation of risk reported for the 2011 American Heart Association/American College of Cardiology (AHA/ACC) guidelines (positive predictive value 9.8% at 1 year and 10.5% at 5 years; negative predictive value 94.1% at 1 year and 94.6% at 5 years) resulting in unnecessary prophylactic ICD implantation in a substantial number of patients (107) new strategies have been suggested.

Appropriately identifying patients that will benefit from and ICD is mandatory as these devices are associated to a significant rate of complications, especially among young individuals as is the case for HCM patients at highest risk for SCD. (18)

To overcome these limitations, O'Mahony *et al*. developed a new score based on a multicenter, retrospective cohort study that included 3,675 individuals in which risk factors independently associated with SCD in at least one multivariate study analysis were evaluated. The final Cox proportional hazards model was used to generate an online calculator [http://www.doc2do.com/hcm/webHCM.html] designed to provide individualized 5-year risk estimates for SCD. See variables included in Table 1. In this score, unlike previous recommendations, age, myocardial thickness, left atrial size, and LV outflow tract gradient are treated as continuous variables (accounting for a continuum of risk) instead of binary. (100) Internal validation showed calibration slope of 0.91 (95% CI: 0.74-1.08), C-index of 0.70 (95% CI: 0.68-0.72), D-statistic of 1.07 (95% CI: 0.81-1.32), depicting a number of ICD needed to implant to avoid 1 SCD of 16 when using an estimated risk threshold ≥ 4%.

The 2014 and 2023 ESC Guidelines use the model to generate consensus-based recommendations for ICD implantation (104, 106) where ICD placement is recommended in individual patients with an estimated 5-year risk of SCD ≥ 4 or ≥ 6% (level of recommendation class IIb and IIa, respectively). The defined cut-off for recommending ICD implantation is an arbitrary value that has been assumed by balancing medical and socioeconomical criteria.

Concerns about the sensitivity of the HCM-SCD risk calculator as compared with the 2011 AHA/ACC guidelines were raised in a study of 1,629 patients previously risk-stratified according to the AHA/ACC guidelines. The authors found that the calculator had adequate specificity, but poor sensitivity compared to the AHA/ACC recommendations, remarking that only 40% of individuals that had events at any time during the follow-up had a score ≥ 4-6%. (99) However, in a global validation study published in 2018 (*n*=3,703), (99) 64% of patients with events within 5 years of baseline evaluation had a HCM SCD-risk score ≥ 4-6% in the complete case analysis, and SCD was observed in 1.4% of individuals in the low risk group (estimated 5 year risk < 4%) and in

8.9% of individuals in the high risk group (estimated 5 year risk ≥ 6%). This validation study revealed a calibration slope of 1.02 (95% CI: 0.93-1.12), C-index of 0.70 (95% CI: 0.68-0.72), and D-statistic of 1.17 (95% CI: 1.05-1.29). The NNT of ICD required to save one live was 13. So even if the algorithm depicted good performance, still 12 unnecessary ICDs were needed to protect 1 patient.

Vriesendorp *et al.* showed a C-index of the HCM SCD-risk score of 0.69 (95% CI: 0.57-0.82), which performed significantly better than the conventional risk factor model based on the 2011 guidelines [C-index of 0.60 (95% CI: 0.50-0.70)] in a cohort of 706 patients with HCM. (94) Other external validation studies have confirmed good discrimination with C-index ranging from 0.69 to 0.92, and the pooled prevalence in a metanalysis of multiple published cohorts was in agreement with the estimated risk per category when using the calculator [1.0% (95% CI: 0.52-1.61) in the low-risk group, 2.43 (95% CI: 1.23-3.92) in the intermediate-risk group and 8.39 (95% CI: 6.68-10.25) in the high-risk group]. (100)

The latest North American guidelines, the 2020 AHA/ACC Guideline for the diagnosis and treatment of patients with HCM (105), still suggest ICD decisions based on the presence of individual risk criteria (see Table 1). The HCM SCD-risk score can be considered in the assessment of patients meeting one ore more risk factors to help in the information and decision-making discussion, although not specific threshold is recommended in order to allow a personalised balanced decision.

One of the limitations attributed to the HCM SCD-risk score is that the recommended cut-off threshold leaves individuals at risk undetected. It is postulated that this is because there are relevant risk factors not included in the score, such as LV apical aneurysm, LGE-CMR, and LV systolic dysfunction, and their impact on 5-year risk estimates is uncertain.

The 2023 ESC guidelines for the management of cardiomyopathies, suggest the use of the score similarly to what was stated in 2014, and adds the presence of LGE-CMR (≥15%) and LVEF < 50% as a class IIb recommendation for SCD primary prevention. (106)

As we can see, despite several studies addressing SCD risk stratification in HCM, there is no prospective validation of the recommended strategies and there is still no global

agreement on the best approach. Advances in the quantification of LGE-CMR, as well as the identification of other risk factors in these patients, could aid in more accurate risk stratification in the future.

Following this effort made in the arrhythmic risk stratification of HCM, similar approaches have been made in other cardiomyopathies.

3.2.3 Arrhythmic risk stratification in arrhythmogenic right ventricle cardiomyopathy

ARVC is a heritable cardiomyopathy characterized by replacement of myocardial tissue with fibrofatty infiltrate, which may predispose to VA and SCD. It has a prevalence of 1:1,000-5,000 inhabitants (108-110) and is one of the main causes of SCD in young people and athletes. (111) ARVC is caused by genetic variants in desmosome proteins. The reported rate of detection of causal variants in patients who meet diagnostic criteria is around 50-60%. (112) The most frequently affected gene is *PKP2* (plakophilin 2) (10-45%), followed by *DSP* (10-15%), *DSG2* (desmoglein 2) (7-10%) and *DSC2* (desmocollin 2) (2%). (113)

In the classic form of the disease the RV is the main affected cardiac structure. As the disease progresses, the LV may be involved in up to 70% of cases. The replacement of myocardial tissue by fibroadipose tissue results in the presence of myocardial wall deformities in the form of segmental alterations of contractility, sacculations and aneurysms. Depending on the extension of the fibrofatty replacement, there is also global dilatation of the ventricular cavity and reduction of global ventricular contractility. The entity frequently presents as T-wave inversion (TWI) in V1 to V4 leads on the ECG, VAs from the RV (with left bundle branch block morphology) and alterations in the RV on imaging tests. The clinical diagnosis of this entity is based on the 2010 ARVC Task Force criteria. (114)

Genotypic-phenotypic correlation studies and the increasing use of contrast-enhancement cardiac magnetic resonance (CMR) have led to the identification of fibro-fatty replacement of the myocardium as a key phenotypic feature of the disease that affects the myocardium of both ventricles. On some occasions, LV involvement may even exceed the severity of RV involvement. Initially, this led to the more inclusive term

of arrhythmogenic cardiomyopathy, and recently it has been suggested to classify this left-dominant cardiomyopathy as NDLVC (106), together with other entities that may course with LV dysfunction. In the genetic substrate of these patients, we may likely identify deleterious variants in the *DSP* gene. (38) Clinical features of patients with primary LV involvement include negative T waves in the inferolateral leads, low voltages in the limb leads, arrhythmias with right bundle branch block morphology (from the LV), LGE-CMR with subepicardial (less often midmyocardial) distribution, either solely or accompanied by segmental or global LV contractility abnormalities. (115, 116)

As the disease progresses, systolic dysfunction may lead to HF, which in advanced cases may require cardiac transplantation. Fibroadipose tissue may slow electrical conduction and favor the appearance of scar-related macroreentries. Likewise, desmosome altered function influences myocyte adhesion, excitability and coupling that can also favor VAs. (113) VAs range from frequent extrasystoles to ventricular tachycardia (VT), which may degenerate into ventricular fibrillation (VF) and SCD. ARVC has been described as responsible for up to 20% of SCD in young people in certain regions of Italy (117). SCD may be the first manifestation of the disease.

The primary objective in managing ARVC upon diagnosis is to prevent SCD, being ICDs the only proven effective treatment. The arrhythmic-event rate among ARVC patients has been reported from 1.0%/year in a cohort of predominantly asymptomatic genetic variant carriers to 30.1%/year in a cohort with definite ARVC (according to the 2010 Task Force criteria) (114) and severe phenotype, being the average proportion of arrhythmic events in studies with patients with definite ARVC 10.6%/year (range 3.0–30.1%/year). (118)

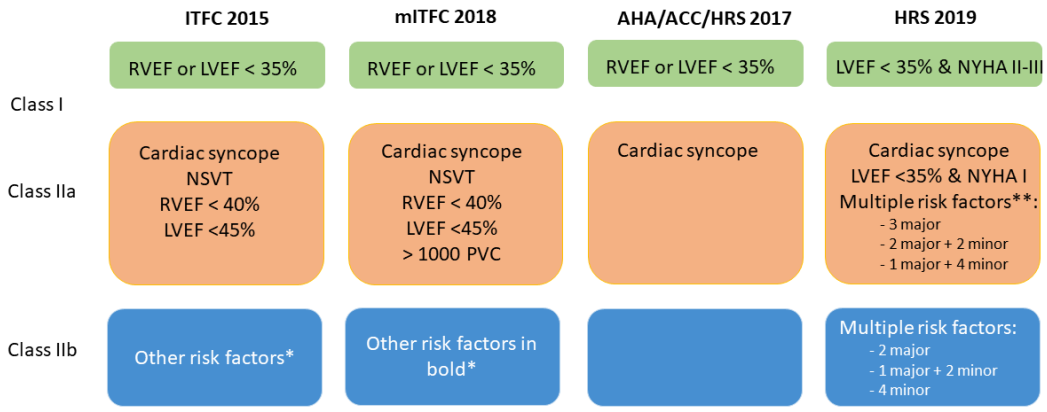
As exposed before, ICD themselves can occasion many complications, thus it is of paramount importance to adequately identify those patients that may benefit from them. Similar to other CVDs, consensus exists regarding the benefit of secondary prevention ICD implantation in ARVC patients with a history of sustained VA or resuscitated sudden cardiac arrest. However, there is still work to be done in the risk stratification of SCD in patients that never had a VA episode.

Multiple studies have attempted to adequately assess the clinical variables associated with sustained VAs in ARVC, identifying demographic, clinical, genetic and behavioral factors linked to arrhythmic risk, although the incidence of the disease (and the available cohorts of limited size) allow limited statistical power to assess the predictive value of the risk factors involved in this disease. (118)

The first comprehensive attempt to establish consensus in the assessment of arrhythmic risk and recommendation of ICD implantation in the context of ARVC occurred with the 2015 International Task Force consensus statement (ITFC). (119) This statement provided recommendations based on the presence of specific risk factors, which were used to categorize the level of risk for each patient. This effort was followed by a study involving 365 participants which suggested adjustment of the International Task Force's recommendations in order to improve differentiation between risk levels (mITFC). (120)

Additional approaches were provided by the 2017 AHA/ACC/Heart Rhythm Society (HRS) guideline for managing patients with VAs and preventing SCD (121), as well as in the 2019 HRS expert consensus statement on evaluation, risk stratification, and management of ARVC. (122).

An analysis of these four algorithms in a cohort of 617 showed that ITFC and mITFC provide higher sensitivity than AHA/ACC and HRS algorithms (94.0–97.8% vs. 76.7–83.5%), but lower specificity (15.9–32.0% vs. 42.7%-60.1%) in the prediction of any VA when considering class I and class IIa recommendations for ICD in the abovementioned algorithms (Figure 10). Despite worst performance of accuracy parameters for AHA and HRS algorithms, those were the ones with best discrimination ability (distinction of events from non-events) with a C-index of 0.72 for both algorithms, which was significantly higher than those for the ITFC and mITFC (0.63 and 0.61, respectively, $P < 0.001$) due to higher specificity and significantly less unnecessary ICD placements for the AHA and HRS algorithms. (123)



*LVEF < 50%, RVEF < 45% on angiography, biventricular dysfunction, RV or right atrial dilatation, heart failure, young age, **male sex**, compound or digenic heterozygosity, **proband status**, **inducible VT/VF**, extent of electroanatomic scar on RV endocardial voltage mapping, fragmented electrograms on RV endocardial voltage mapping, T-wave inversion in inferior leads, **inverted T waves in ≥3 precordial leads**, QRS fragmentation, precordial QRS amplitude ratio.

**Major: NSVT, inducibility to VT at EPS, LVEF < 49%.
Minor: Male sex, > 1000 PVCs/24h, RV dysfunction as per major 2010 TFC, proband status, multiple desmosomal variants.
If both NSVT and PVC criteria are present, only NSVT can be used.

Figure 10. Schematic overview of the four ICD placement algorithms for risk stratification in primary prevention of SCD in ARVC patients. Adapted from Bosman *et al.*(123) AHA/ACC, American Heart Association/ American College of Cardiology; HRS, Heart Rhythm Society; ITFC, international task force consensus statement; LVEF, left ventricular ejection fraction; mITFC, modified international task force consensus statement; NSVT, non-sustained ventricular tachycardia; PVC, premature ventricular complex; RV, right ventricle; RVEF, right ventricular ejection fraction; VF, ventricular fibrillation; VT, ventricular tachycardia.

In this work, the net clinical benefit analysis comparing the previous approaches revealed that the ITFC and mITFC provide the greatest net benefit when using a threshold range from 5 to 25%, while AHA and HRS have better performance for thresholds of risk above 25%.

These results showed that all four algorithms are able to stratify arrhythmic risk in ARVC patients but with limited accuracy, trading higher sensitivity (better protection rates) for lower specificity (ICD placements in patients not experiencing events) for the ITFC and mITFC and vice versa for the AHA and HRS ones, lower protection rates but decrease in unnecessary ICD placements.

To help to address those limitations, a retrospective ARVC cohort including patients from five comprehensive registries from North America and Europe was compiled. The primary aim of this effort was to develop an individualized multivariable predictive model for the first occurrence of sustained VAs in ARVC patients using easily accessible clinical variables and to compare the new model with the existing consensus-based algorithms for ICD placement. (22)

In that study, 528 patients with a definite diagnosis of ARVC according to the 2010 Task force criteria and no history of sustained VAs/SCD at baseline, aged 38.2 ± 15.5 years, 44.7% male, were retrospectively included. During a follow-up period of 4.83 (interquartile range 2.44–9.33), 146 (27.7%) experienced sustained VA, defined as SCD, aborted SCD, sustained VT, or appropriate ICD therapy. The predictive model estimated VA risk at 1 and 5 years by using Cox regression. Candidate predictors were predefined, drawing from clinical expertise and contemporary literature on the stratification of arrhythmic risk in ARVC (118, 120, 124), encompassing the following variables at diagnosis: age, sex, cardiac syncope in the prior 6 months, NSVT, number of premature ventricular complexes in 24 h, number of leads with TWI, and right and left ventricular ejection fractions. Among those pre-specified predictors all except LVEF were retained in the final model. Internal validation after bootstrapping re-sampling showed an adequate distinction of patients with and without events through an optimism-corrected C-index of 0.77 (95% CI: 0.73–0.81) and minimal over-optimism [calibration slope of 0.93 (95% CI: 0.92–0.95)]. By decision curve analysis, the clinical benefit of the model was superior to the ITFC consensus-based ICD placement algorithm with a 20.3% reduction of ICD placements with the same proportion of protected patients ($P < 0.001$).

The next step after the above-mentioned effort was to externally validate these results in an independent cohort of ARVC individuals. Thus, in the third section of this compilation of articles PhD thesis we present a study which aimed to externally validate the published ARVC arrhythmic risk calculator in a distinct, adequately powered, and geographically diverse cohort including patients from six countries across North America and Europe and assess the performance of the predictive model in comparison with other published guidelines and expert consensus recommendations for ICD use in ARVC patients in this independent cohort.

In this PhD thesis we focus on clinical prediction and prognostic stratification of cardiovascular diseases, either through prognostic information on rare genetic variants, or through clinical prediction models based on polygenic scores, which integrate the cumulative effect of common genetic variants across the whole genome in relation to a specific phenotype, or through clinical prediction models that integrate demographic and clinical markers.

Hypothesis

In the first article, we hypothesize that the carriers of the rare variant *BAG3* p.Gln88* present a form of dilated cardiomyopathy with a distinct clinical expression from those caused by other aetiologies.

In the second article, we hypothesize that multi-trait analysis of genome-wide association studies provides enhanced single nucleotide polymorphism to disease association effects allowing discovery of novel loci in genomic regions of common variation and that these enhanced single nucleotide polymorphism-effects will help to better predict disease-status in common cardiovascular diseases: atrial fibrillation, coronary artery disease and heart failure.

In the third article, we hypothesize that the arrhythmogenic right ventricle cardiomyopathy score published in 2019, which includes weighted information on age, biological sex, history of syncope in the past 6 months, history of non-sustained ventricular tachycardia, premature ventricular complexes count, sum of anterior and inferior leads with T-wave inversion and right ventricular ejection fraction, and was internally validated in the development cohort with a C-index of 0.77 [95% confidence interval: 0.73-0.81] is still valid in an independent external cohort of patients with arrhythmogenic right ventricle cardiomyopathy.

Objectives

Article 1:

- To describe the prognosis of individuals carrying the *BAG3* p.Gln88*variant, a rare pathogenic variant in a gene associated with dilated cardiomyopathy.

Article 2:

- To analyze the ability of multi-trait analysis of genome-wide association studies to empower genomic locus discovery in atrial fibrillation, coronary artery disease and heart failure by leveraging relevant genetic correlations of these traits with other cardiovascular traits.
- To assess the added value of multi-trait analysis of genome-wide association studies to improve polygenic scores-based disease prediction and discrimination in common cardiovascular diseases: atrial fibrillation, coronary artery disease and heart failure.

Article 3:

- To externally validate the arrhythmogenic right ventricle cardiomyopathy risk calculator for predicting ventricular arrhythmias in patients with arrhythmogenic right ventricle cardiomyopathy.
- To compare the performance of the arrhythmogenic right ventricle cardiomyopathy risk calculator with current risk assessment for ventricular arrhythmias recommended by published guidelines and expert consensus.

Material, methods, and results

PDFS of article 1, 2 and 3.

Article 1

Circulation: Heart Failure

IMAGES AND CASE REPORTS IN HEART FAILURE

BAG3 Genetic Cardiomyopathy May Overlap Fulminant Myocarditis Clinical Findings

Paloma Jordà, MD; Daniel Martínez, MD, PhD; Marta Farrero, MD, PhD; María Ángeles Marcos, MD, PhD; Elena Sandoval, MD; María Ángeles Castel, MD, PhD; Daniel Pereda, MD, PhD; Eduard Quintana, MD, PhD; Elena Arbelo, MD, PhD; Aurora Sánchez, MC, PhD; Òscar Campuzano, PhD; Coloma Tirón de Llano, MD; Ramon Brugada, MD, PhD; Josep Brugada, MD, PhD; Manuel Castellá, MD, PhD; Félix Pérez-Villa, MD, PhD; Ana García-Álvarez, MD, PhD

We present a clinical report on severe cardiomyopathy caused by a rare variant of the *BAG3* gene among individuals with European ancestry that debuted as acute heart failure (HF) with clinical features of acute myocarditis but coinciding with suspicious familial history.

The first case occurred in 2004. A 15-year-old girl (IV.4) was admitted for dyspnea, preceded by 15 days of abdominal pain and vomiting, that rapidly progressed to cardiogenic shock. ECG demonstrated sinus tachycardia, low voltages and flattened T-waves in limb leads, late R-wave transition, J-point elevation in anterior leads, and isodiphasic T-waves in lateral leads (Figure S1, 1); and in the echocardiography severe biventricular dysfunction and left ventricle (LV) dilatation was observed. She required a biventricular assist device and emergent heart transplantation. Diagnostic suspicion was fulminant myocarditis as the cardiac biopsy showed myocardial neutrophilic infiltration (borderline diagnosis of myocarditis according to Dallas criteria; Figure [A and B]). The explanted heart depicted fibrin-hemorrhagic pericarditis, thickened LV wall, and biventricular dilatation.

In 2005, her mother's cousin (III.1) was admitted at the age of 38 due to a 45-day onset of progressive HF preceded by fever and myalgias. Echocardiography depicted severe LV impairment and dilatation. He was transplanted 2 days later in an Interagency Registry for Mechanically Assisted Circulatory Support 2 situation. Clinical course suggested myocarditis, but no inflammation was observed in the explanted heart.

Nine years later, in 2016, the sister of the index case (IV.5) was admitted due to cardiogenic shock at 17 years old after a 15-day course of rhinorrhea and dry coughing. ECG demonstrated sinus tachycardia and nonspecific repolarization abnormalities resembling those of her sister (Figure S1, 2). Echocardiography showed dilated LV and severe biventricular systolic dysfunction. A biventricular assist device assist device was implanted (Interagency Registry for Mechanically Assisted Circulatory Support 2), and she underwent heart transplantation 12 days later. Cardiac biopsy did not demonstrate inflammatory infiltrates but diffuse distortion of myocardial fibers with interfibrillary edema (Figure [C through G]).

The third case in the same family raised the suspicion of a familial disease and genetic testing (next-generation sequencing panel including 50 genes associated with inherited cardiomyopathies) was performed revealing a *BAG3* p.Gln88* pathogenic variant (American College of Medical Genetics and Genomics classification[1]), causing premature truncation of the protein. The index case (IV.4), her mother (III.4), grandmother (II.3), and her mother's uncle and cousin (II.1 and III.1) carried the same rare pathogenic variant (Figure [H]). To our knowledge, this report is the first clinical description in literature of individuals carrying this variant. The 2 sisters (IV.4 and IV.5) also had a novel rare variant -p.(Pro52Arg)- in *BAG3*, absent in their mother, and classified as a variant of uncertain significance.[1] This variant is absent in large

Key Words: cardiomyopathies ■ dyspnea ■ genetics ■ heart failure ■ inflammation ■ myocarditis ■ shock, cardiogenic

Correspondence to: Paloma Jordà, MD, or Ana García-Álvarez, MD, PhD, Cardiology Department, Institut Clínic Cardiovascular, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Hospital Clínic de Barcelona, Universitat de Barcelona, C/Villarroel 170, 6 Esc. 3, 08036 Barcelona, Spain. Email pjordab@clinic.cat or anagarci@clinic.cat

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/CIRCHEARTFAILURE.121.008443.
For Sources of Funding and Disclosures, see page 293.

Circulation: Heart Failure is available at www.ahajournals.org/journal/circheartfailure

Nonstandard Abbreviations and Acronyms

HF	heart failure
LV	left ventricle

international population databases[2] and is predicted to be deleterious by in silico data.

Clinical evaluation of the proband's mother (III.4) at 45 years old demonstrated dilated cardiomyopathy with isolated LV involvement. Since she was at New York Heart Association functional class II and had left bundle branch block on ECG (Figure S1, 3), cardiac resynchronization therapy device was implanted. Proband's grandmother (II.3) and great-uncle (II.1) had no dilated cardiomyopathy at the age of 67 and 76, respectively, although her grandmother's ECG showed nonspecific negative T-waves in lateral leads (Figure S1, 4) along with moderate septal hypertrophy on echocardiography.

Of note, we have found the same *BAG3* p.Gln88* rare variant in an unrelated patient from our cohort of

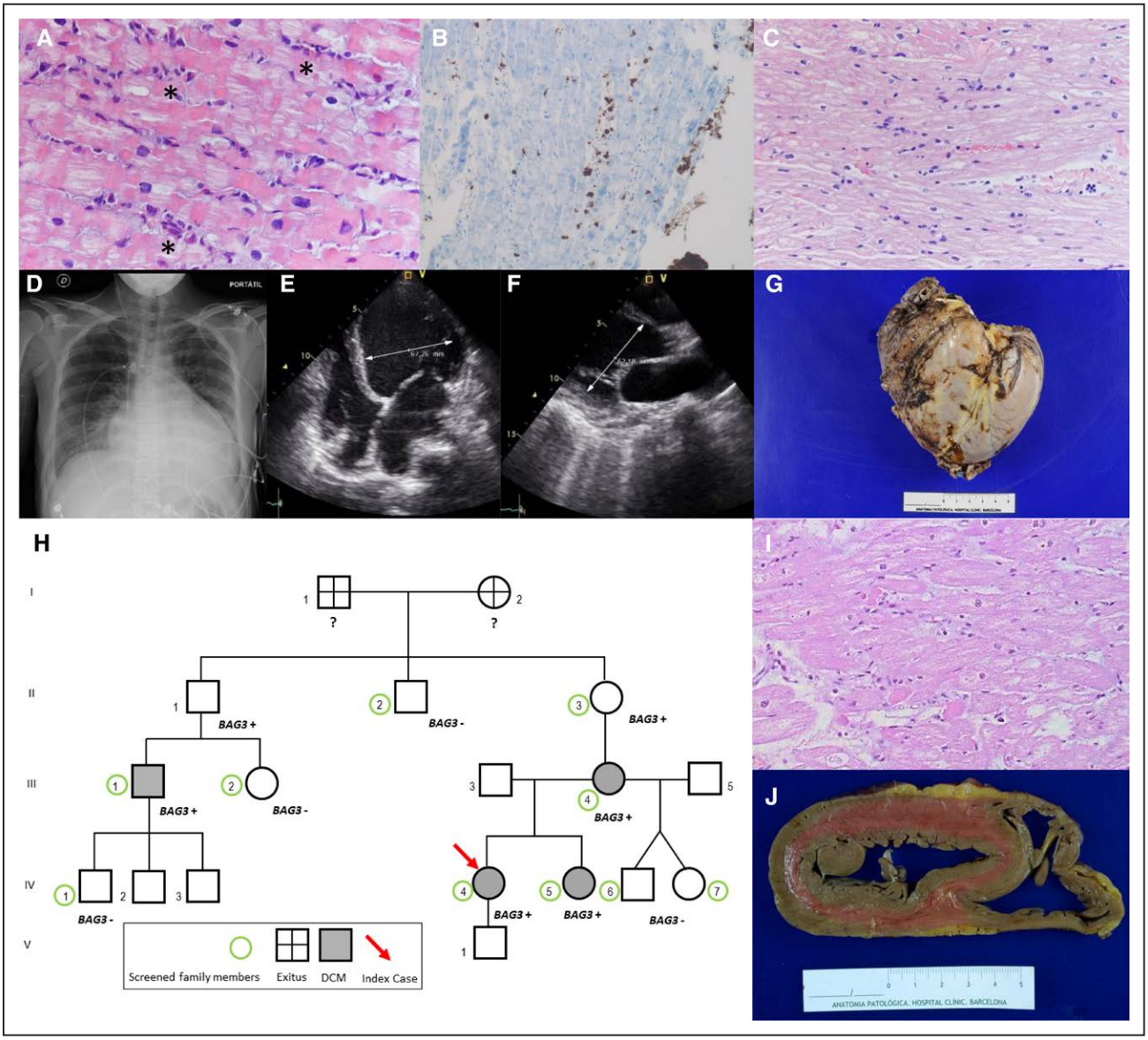


Figure. Family tree of the related cases and iconography of the radiologic, echocardiographic, and histopathologic findings. **A** and **B** correspond to case IV.4. **A** depicts hematoxylin-eosin staining of the cardiac biopsy specimen which shows foci of neutrophilic perivascular infiltration (marked with black asterisks) and occasional eosinophils but no myocyte degeneration. **B** shows CD15 immunohistochemical staining for neutrophils in brown. **C** through **G** correspond to case IV.5. **C** depicts vacuolated cardiomyocytes, diffuse distortion of myocardial fibers, and interfibrillary edema but no inflammation infiltrates nor necrosis in the cardiac biopsy tissue. Thoracic x-ray (**D**) and echocardiography (**E** and **F**) at admission show dilated left ventricle (LV) and pulmonary congestion. Biventricular dilatation of the explanted heart is shown in **G**. **H** depicts cosegregation of the *BAG3* p.88* pathogenic mutation in the related individuals. **I** and **J** describe the findings in the unrelated patient with the same *BAG3* mutation. Histopathology of the explanted heart depicted mild lymphocytic infiltration and focal myocyte necrosis (**I**) in a thickened and dilated LV (**J**). DCM indicates dilated cardiomyopathy.

transplanted patients. Reviewing medical history, the story repeated. At 30 years old, he had a 10-day course of fever and progressive dyspnea presumed to be severe pneumonia. Echocardiography showed dilated and severely impaired LV. He was supported with intraaortic balloon pump and left ventricle assist device due to cardiogenic shock and underwent emergent heart transplantation subsequently. The diagnostic suspicion was acute myocarditis based on lymphocytic infiltration and focal myocyte necrosis in the explanted heart (Figure [I and J]). No other cases in this family have been found to date. Genotyping was performed in the course of a protocol investigating the genetic background of myocarditis.

No patient was positive for IgM titers of cardiotropic viruses. Polymerase chain reaction viral analysis on the heart has been retrospectively performed, following current recommendations,[3] in all transplanted patients except in case III.1 (no sample remained). A low viral load of parvovirus B19 was found in the index case and the unrelated patient. The absence of other cardiotropic viruses and the low viral load suggest that its presence lacks etiopathogenic relevance, as the role of parvovirus B19 in endomyocardial biopsies still remains unclear.[4] Patient characteristics at diagnosis or at the time of family screening are presented in the Table.

More than 100 genes have been reported associated with inherited dilated cardiomyopathy; 33 of these genes carry sufficient evidence to be classified as definitive disease-causing genes, including *BAG3*.[5] In the largest cohort of *BAG3* carriers described so far (n=129), 15.5% of them developed severe HF defined as the composite outcome of heart transplantation, left ventricle assist device or death due to HF.[6] Our work describes other interesting findings. Severe cardiomyopathy with acute HF onset appeared at very young ages, specifically in 2 individuals (15 and 17 years old). Early onset in these patients may be partially explained by compound heterozygous rare variants in *BAG3*. In addition, 4 out of 7 individuals carrying the same *BAG3* p.Gln88* mutation (one of them unrelated) were transplanted due to cardiogenic shock (57%) suggesting an aggressive disease course. Our findings regarding this specific variant do not replicate male-associated aggressive course previously reported.[4] Of note, acute and severe presentation of this genetic cardiomyopathy in our patients associates clinical features (viral infection-like symptoms and acute onset) and pathological findings (myocardial inflammatory infiltrates and distortion of myocardial fibers) that

Table. Patient Characteristics at the Time of Diagnosis or Family Screening

Case	Mutations	Age	Viral symptoms	HF symptoms	Echocardiography/CMR	Suspected myocarditis[3]/Dallas criteria/ cardiac viral PCR	Biomarkers	VAD or HTx
Index case (IV.4)	*BAG3* p.Gln88*, p.(Pro52Arg)	15	Gastrointestinal (also attributable to low cardiac output)	HF rapidly progressing to CS (15 d)	Severe LV dilatation and impairment. Moderated RV dilatation with severe RV impairment.	Yes/Yes, borderline/Parvovirus B19: 92 copies per mL	↑ CRP (×4), ↑ CK and TnI (×1.2)	IABP, BiVAD, and HTx
Sister (IV.5)	*BAG3* p.Gln88*, p.(Pro52Arg)	17	Respiratory (rhinorrhea and cough)	HF rapidly progressing to CS (15 d)	Dilated and impaired LV (LVEF 15%, LVEDd 66 mm). Nondilated but severe hypokinetic RV.	Yes/No, but diffuse distortion of myocardial fibers with interfibrillary edema/Negative	Normal values of CRP, TnI and CK	BiVAD and HTx
Mother (III.4)	*BAG3* p.Gln88*	45	No	NYHA functional class II	Dilated and impaired LV (LVEDd 67 mm). CMR: Indexed LVED 185 mL/m^2, LVEF 21%. Normal RV. No edema nor LGE.	No/ – / –	Normal values of TnI and CK	No
Grandmother (II.3)	*BAG3* p.Gln88*	67	No	No	Hypertensive cardiomyopathy (septal thickness 14 mm) with nondilated LV and normal LVEF.	No/ – / –	–	No
Great-uncle (II.1)	*BAG3* p.Gln88*	76	No	No	Normal echocardiography.	No/ – / –	–	No
Cousin (III.1)	*BAG3* p.Gln88*	38	Unspecific (fever and myalgias)	HF progressing to CS (45 d)	LV severely impaired and dilated and RV dysfunction.	Yes/No/ –	↑↑ CRP (×10), normal CK and TnI	IABP, BiVAD, and HTx
UR	*BAG3* p.Gln88*	30	Fever	HF progressing to CS (10 d)	LVEF 29%, LVEDd 57 mm. RV with mild dilatation and hypokinesia. CMR without edema nor LGE.	Yes/Yes/Parvovirus B19: 9 copies per ml	↑ CRP (×4), normal CK and TnI	IABP, LVAD, and HTx

Description of the clinical, genetic, biochemical, and imaging characteristics of patients with acute presentation (IV.4, IV.5, III.1, and UR) at the time of disease onset or at the time of family screening in other related individuals. BiVAD indicates biventricular assist device; CK, creatine kinase; CMR, cardiac magnetic resonance; CRP, C-reactive protein; CS, cardiogenic shock; HF, heart failure; HTx, heart transplantation; IABP, intraaortic balloon pump; LGE, late gadolinium enhancement; LV, left ventricle; LVAD, left ventricle assist device; LVEDd, left ventricle end-diastolic diameter; LVEDv, left ventricle end-diastolic volume; LVEF, left ventricle ejection fraction; NYHA, New York Heart Association; PCR, polymerase chain reaction; RV, right ventricle; TnI, troponin I; UR, unrelated patient; and VAD, ventricular assist device.

may overlap with myocarditis typical findings, albeit other mechanisms may be present.

Further research and characterization of myocardial infiltrates and immune activation (including immunohistochemical study of the heart and serum cardiac autoantibodies), as well as the presence of viral genome in the myocardium and peripheral blood,[3] will help clarify the contribution of immune-mediated mechanisms to fulminant presentations of genetic cardiomyopathies. Furthermore, assessment of disease progression according to site-specific mutations in future cohorts may aid prognosis stratification.

ARTICLE INFORMATION

Affiliations

Cardiology Department, Institut Clínic Cardiovascular (P.J., M.F., M.A.C., E.A., J.B., F.P.-V., A.G.-Á.) and Cardiovascular Surgery Department, Institut Clínic Cardiovascular (E.S., D.P., E.Q., M.C.), Department of Genetics, Institut de Bioquímica i Genètica Molecular (A.S.), Microbiology Department (M.Á.M.), and Pathology Department (D.M.), Hospital Clínic de Barcelona, Universitat de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Spain (P.J., D.M., M.F., E.S., M.A.C., D.P., E.Q., E.A., A.S., J.B., M.C., F.P.-V., A.G.-Á). Universitat de Girona, Spain (Ò.C., R.B.). Cardiovascular Genetics Center, Institut d'Investigació Biomèdica de Girona Dr. Josep Trueta, Spain (Ò.C., C.T.d.L., R.B.). Cardiology Department, Hospital Universitari Dr. Josep Trueta, Girona, Spain (C.T.d.L., R.B.). Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares, Madrid, Spain (M.A.C., E.A., Ò.C., R.B., J.B.).

Acknowledgments

The authors thank the anonymous reviewers for their contributions to this manuscript. Coauthor Félix Pérez-Villa, MD, PhD, died July 28, 2021. The authors want to acknowledge the achievements of this remarkable colleague and mentor; his legacy will be remembered forever.

Disclosures

None.

Supplemental Material

Figure S1

REFERENCES

1. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al.; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–424. doi: 10.1038/gim.2015.30
2. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, et al.; Genome Aggregation Database Consortium. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581:434–443. doi: 10.1038/s41586-020-2308-7
3. Caforio AL, Pankuweit S, Arbustini E, Basso C, Gimeno-Blanes J, Felix SB, Fu M, Heliö T, Heymans S, Jahns R, et al.; European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. Current state of knowledge on aetiology, diagnosis, management, and therapy of myocarditis: a position statement of the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J*. 2013;34:2636–48, 2648a. doi: 10.1093/eurheartj/eht210
4. Verdonschot J, Hazebroek M, Merken J, Debing Y, Dennert R, Brunner-La Rocca HP, Heymans S. Relevance of cardiac parvovirus B19 in myocarditis and dilated cardiomyopathy: review of the literature. *Eur J Heart Fail*. 2016;18:1430–1441. doi: 10.1002/ejhf.665
5. McNally EM, Mestroni L. Dilated cardiomyopathy: genetic determinants and mechanisms. *Circ Res*. 2017;121:731–748. doi: 10.1161/CIRCRESAHA.116.309396
6. Domínguez F, Cuenca S, Bilińska Z, Toro R, Villard E, Barriales-Villa R, Ochoa JP, Asselbergs F, Sammani A, Franaszczyk M, et al.; European Genetic Cardiomyopathies Initiative Investigators. Dilated cardiomyopathy due to BLC2-Associated Athanogene 3 (BAG3) Mutations. *J Am Coll Cardiol*. 2018;72:2471–2481. doi: 10.1016/j.jacc.2018.08.2181

Article 2

Title: **Multi-trait analysis of genome-wide association studies for the discovery of novel genetic loci and genomic prediction in common cardiovascular diseases**

Co-authors: Paloma Jordà[1], Yiwei Lai[1], Amélie Jeuken[1], Louis-Philippe Lemieux Perreault[1], Elisabeth Goulet[1], Najim Lahrouchi[2], CHARGE Inflammation Working Group[3], Anna Nozza[4], Michael W Tanck[2], Peter Guerra[1], Julia Cadrin-Tourigny[1], Simon de Denus[5], Connie R Bezzina[2,6], Guillaume Lettre[1], David Busseuil[1], Marie-Pierre Dubé[1], Jean-Claude Tardif[1], Rafik Tadros[1]

[1]Montreal Heart Institute, Faculty of Medicine, Université de Montréal, Montreal, Québec, Canada.

[2]Department of Experimental Cardiology, Heart Center, Amsterdam Cardiovascular Sciences, University of Amsterdam, Amsterdam UMC, Amsterdam, the Netherlands.

[3]Helsinki Collegium for Advanced Studies, University of Helsinki, Helsinki 00014, Finland; Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki 00014, Finland; Folkhälsan Research Centre, Helsinki 00250, Finland.

[4]Montreal Health Innovations Coordinating Center, Montreal, Québec, Canada

[5]Montreal Heart Institute, Faculty of Pharmacy, Université de Montréal, Montreal, Québec, Canada.

[6]European Reference Network for Rare and Low Prevalence Complex Diseases of the Heart (ERN GUARD-HEART).

Journal: Nature Cardiovascular Research

Abstract

Common genetic variation detected by genome-wide association studies (GWAS) partially explains variability in cardiovascular conditions, which may have overlapping mechanisms. We aimed to leverage multi-trait analysis of GWAS (MTAG) drawing power from genetically correlated traits to improve genetic loci discovery and genomic prediction in 3 common cardiovascular diseases: atrial fibrillation and flutter (AF), coronary artery disease (CAD), and heart failure (HF). MTAG allowed the identification of 19 novel disease-specific genomic loci for AF, 131 for CAD and 141 for HF. Polygenic scores (PGS) analyses in 15,177 Canadian individuals yielded similar results when comparing single-trait versus multi-trait PGS for AF and CAD, although it revealed improved prediction and classification of disease-status in CAD for females ($\Delta r^2_{CAD\text{-}F}$ 1.735% (95%CI: 0.609-2.856); $NRI_{CAD\text{-}F}$ 0.208, P<0.001). HF-PGSs models depicted poor prediction and calibration, impeding discrimination assessment. These findings support MTAG utility for enhancing loci discovery in common cardiac diseases and potentially for improving CAD-PGS performance in females.

Note: Until the article is published, Tables 1 and 2 (large Excel tables), as well as Supplementary Tables and Supplementary Note, will be available on MedRxiv.

INTRODUCTION

Genome-wide association studies (GWAS) have mapped numerous single nucleotide polymorphisms (SNP) in loci implicated in molecular mechanisms altered in common complex diseases, thus providing guidance in the investigation of novel drug targets.(1)

Atrial fibrillation (AF), coronary artery disease (CAD), and heart failure (HF) are common heart diseases, with a prevalence of 8%(2), 2.5%(3) and 2.5%(4), respectively. The genetic substrate of these diseases is partly explained by genomic loci harboring common genetic variants in the general population; this substrate may overlap within cardiovascular diseases and traits, which can be quantified by genetic correlation.

Polygenic scores (PGS) intend to account for the genetic susceptibility due to common genetic variants, i.e., SNPs, and may be useful for the susceptibility to disease assessment in order to set up preventive strategies among high-risk individuals.

Multi-trait analysis of genome-wide association studies (MTAG) boosts statistical power by joint analysis of genetically correlated traits, fitting the results to the genetic correlation matrix and allowing the inclusion of overlapping samples among the included GWAS. MTAG has demonstrated to enhance loci discovery and increase the variance explained by PGS.(5)

The objectives of this study were 1) to empower genomic locus discovery in AF, CAD and HF by leveraging genetic correlations with relevant cardiovascular traits using MTAG, and 2) to assess the added value of MTAG to improve PGS-based disease prediction and discrimination in common cardiovascular diseases.

RESULTS

Heritability and genetic correlation

Cardiovascular traits and diseases with publicly available GWAS summary statistics were compiled, prioritizing those with the largest size and power, encompassing 56 traits/diseases including common cardiovascular diseases such as AF(3, 6), CAD(7) and HF(4) as well as rare heritable heart diseases and cardiometabolic traits. The complete list of phenotypes and included studies appears in Supplementary Table 1. The study workflow is shown in Fig. 1.

We computed h^2_{SNP} using Linkage Disequilibrium Score (LDSC) regression (8) for all the included cardiovascular phenotypes, which ranged from 0.05 to 0.27 (Supplementary Table 2). For binary traits (diseases), h^2_{SNP} on a liability scale ranged from 0.03 (for HF) to 0.17 (for type 2 diabetes mellitus, T2DM) (Supplementary Table 3).

We then computed pairwise genetic correlations (r_g) for 1,770 trait and/or disease pairs using LDSC regression. A total of 443 unique pairwise r_g were significant at a *P*-value<0.05 (corrected for false discovery rate, Extended Fig. 1 and Supplementary Table 2), from which 325 yielded absolute r_g above 0.20. AF, CAD and HF showed 12, 21 and 23 significant r_g, respectively, depicted in Fig. 2 and Supplementary Tables 4-6. These findings and their contextualization in the literature are described in the Supplementary Note. Notably, in this work we describe novel genetic correlations of myocardial trabeculation, measured as a fractal dimension of the endocardium trabeculae (9) with cardiac electrical parameters, such as cardiac repolarization in inferior and aVR leads (r_g -0.35 and 0.28,

respectively) and AF (r_g 0.15 to 0.21) for the mid-ventricular and apical layers; and exercise heart rate (HR) and HR variability (r_g -0.31 to -0.33) for basal layers.

Multi-trait analysis of GWAS (MTAG)

In the quest to unravel novel genetic loci implicated in AF, CAD, and HF, we performed MTAG for each of those common heart diseases (5). MTAG leverages genetically correlated traits to increase power for locus discovery. For each of the 3 MTAG analyses (AF, CAD and HF), we included 3 cardiovascular traits with the highest genetic correlation with the disease of interest. Restricting the number of included traits to 3 allows for estimation of the upper bound of the false discovery rate (maxFDR; see Online Methods).

For each of the 3 diseases, correlated traits were clustered in 3 groups based on pairwise r_g (Fig.2). Within each of the 3 clusters, we then selected the trait with the highest r_g with the disease of interest (i.e., AF, CAD or HF). The following traits were included in the MTAG. For the AF MTAG using Nielsen *et al.* GWAS(6), we included body mass index (BMI)(10), HF(4) and another AF GWAS by Roselli *et al*. (3), which have pairwise r_g with AF Nielsen *et al.* GWAS of 0.21, 0.56 and 1, respectively. For the CAD MTAG employing the Nelson *et al.* study (7), we included HF(4), low-density lipoprotein (LDL) (11) and systolic blood pressure (SBP) (12), which have pairwise r_g with CAD Nelson *et al.* GWAS of 0.72, 0.22 and 0.34, respectively. For the HF MTAG using the GWAS by Shah *et al*.(4), we included AF(6), CAD(7) and dilated cardiomyopathy (DCM) (13), which showed pairwise r_g of 0.56, 0.72 and 0.56 with HF. We estimated maxFDR to be 0.2%, 1.6% and 0.4% for the AF, CAD and HF analysis respectively (Supplementary Table 7).

MTAG significantly increased the association test statistic, i.e., the mean χ^2, which was equivalent to an increase in the effective sample size for a conventional GWAS (N_{eff}) from 228,221 to 296,021 for AF, from 224,727 to 310,368 for CAD, and from 180,076 to 358,418 for HF (Supplementary Table 7).

We identified 19 *novel* loci at genome-wide significance ($P<5x10^{-8}$) in AF, 89 in CAD, and 52 in HF (Table 1, Supplementary Tables 8-10). Novel loci were defined as not identified by the original published single-trait GWAS summary statistics, i.e., not in Nielsen *et al*., Roselli *et al*., nor in meta-analysis reported in Roselli *et al*. for AF, not in Nelson *et al*. for CAD, and not reported in Shah *et al.* in the case of HF. The Miami plots comparing single-trait and multi-trait results are depicted in Fig. 3. Supplementary Tables 8-10 describe the Z-score - beta estimate divided by the standard error - for each lead SNP in novel loci in both the MTAG and all 4 traits included in the analysis, the latter also depicted in Extended Data Fig. 2. To assess the robustness of the findings, we looked up the novel loci reported in the present MTAG in other GWAS publications. We observed that 32%, 71% and 62% of the *novel* MTAG loci have been associated with AF, CAD and HF, respectively, at either 1) a genome wide significant level in partially overlapping or independent cohorts, or 2) at $P<5x10^{-3}$ in an independent cohort (FinnGen), as detailed in Supplementary Tables 8-10. Two recently published GWAS in CAD (Aragam *et al*.(14)) and HF (Levin *et al*.(15)) were employed for the identification of *additional novel loci* by MTAG, leading to the identification of 42 and 89 loci not identified by the conventional GWAS (Supplementary Tables 25-26). Of those, 62% and 4.5% of CAD Aragam *et al.* and HF Levin *et al.* loci were "validated" following the above described approach. Interestingly 19 and 37 overlapped with the CAD Nelson *et al.* and HF Shah *et al.* MTAGs, and an *additional* 23 and 52 respectively were further identified (Table 1, Supplementary Figs. 1 and 2).

MAGMA gene-set functional analysis and tissue expression from MTAG summary statistics are depicted in Supplementary Tables 14-22, 29-34 and Supplementary Fig. 3. The gene-set analyses

identified cardiac depolarization, repolarization, and muscle contraction in AF; macromolecular complex remodeling, heart development, collagen, triglyceride, chylomicron and LDL metabolism in CAD; and muscle structure, tissue development and RNA transcription initiation in HF.

Genes mapping to *novel* MTAG loci (using Open Targets variant to gene algorithm) are listed in Table 1 and described in Supplementary Tables 8-10, 25-26 and the more relevant ones due to mechanistic insight or support by other studies are described briefly in the Discussion and in the Supplementary Note. GTEx Tissue expression for all prioritized genes is depicted in Supplementary Fig. 4.

Genomic risk prediction

We next sought to assess whether PGS derived from MTAG outperform those derived from original single-trait studies in discriminating individuals with known (prevalent) AF, CAD and HF. Specifically, we derived multiple PGS from single-trait and multi-trait (MTAG) summary statistics (PGSsingleT and PGSmultiT, respectively) using LDpred2(16). LDpred2 is a genome-wide PGS derivation method which adjusts GWAS summary statistics SNP effect estimates based on their probability of being causal and linkage disequilibrium (LD) information from a reference panel. For its calculation the following hyperparameters are employed: fraction of causal variants (ρ, the Greek letter "rho", not the *P*-value), heritability estimates (h^2) and a sparse option which fits some effects to 0. It has an infinitesimal method that considers all variants as causal (ρ=1); an auto method, which learns the hyperparameters from the summary statistics and therefore does not require a validation cohort; and a grid method which employs up to 168 combinations of h^2, ρ, and sparse option, from which the combination with the best Wald-Z is selected from the validation dataset.(16) For a comprehensive analysis we also derived PGS using LDpred1 and traditional LD-based clumping and *P*-value thresholding (C-T) (17) (see full description of the approach in Online Methods and Supplementary Note, results available in Supplementary Tables 38-40).

PGS performance was assessed in a cohort of 15,177 unrelated individuals enrolled in the Montreal Heart Institute Biobank (MHIBB). Specifically, 2,500 MHIBB participants were used as a validation cohort to select the best PGSsingleT and best PGSmultiT; which were subsequently compared in the remaining 12,677 of the testing dataset. The basal characteristics of the MHIBB cohort are presented in Table 2.

Disease-specific PGSsingleT and PGSmultiT were significantly associated with AF, CAD and HF in the testing cohort, using both univariable logistic regression and multivariable logistic regression with correction for sex, age and the first 4 genotypic principal components, with the exception of HF PGSsingleT univariable logistic model (Supplementary Table 35). PGSmultiT yielded very similar area under the curve (AUC) and phenotype explained variance by the model (Nagelkerke R^2, referred to hereafter as R^2) to PGSsingleT in AF [ΔAUC=0.001, *P*=0.63; ΔR^2 0.053% (95%CI: -0.305, 0.421)] and CAD [ΔAUC=0.003, *P*=0.36; ΔR^2 0.103% (95%CI: -0.493, 0.687)]. In contrast, HF PGSmultiT model significantly increased the AUC, although AUCs were overall low, and did not significantly increase R^2 [ΔAUC=0.017, *P*=0.03; ΔR^2 0.072% (95%CI: -0.029, 0.224)] (Fig.4, Supplementary Table 35).

PGSsingleT and PGSmultiT were compared using the Net Reclassification Index (NRI), which is the difference of the proportion of individuals whose risk is correctly reclassified minus the proportion of individuals whose risk is incorrectly reclassified, and the Integrated Discrimination Improvement (IDI), a measurement of the improvement of the risk estimate difference between affected and non-affected individuals. NRI and IDI supported the ΔAUC results. PGSmultiT did not provide significant improvement compared to PGSsingleT for AF and CAD (NRI_{AF} 0.002, *P*=0.91, IDI_{AF} 0.000 *P*=0.64; NRI_{CAD} 0.005, *P*=0.76, IDI_{CAD} 0.001, *P*=0.52). Regarding HF, PGS models were poorly calibrated for the

univariable PGS models, thus the NRI should be interpreted with great caution (NRI_{HF} 0.057, *P*=0.025; IDI_{HF} 0.0004, *P*=0.017) (18) (Extended Data Fig. 3, Supplementary Table 35). In HF, PGS performed differently in the validation compared to the testing dataset, as the grid model performed better in the validation cohort, while the infinitesimal model performed better in the larger testing cohort (Supplementary Fig. 5, Supplementary Table 40). If we compare the results of LDpred2-infinitesimal PGSs in the testing dataset, the PGSmultiT models obtained better calibration and significant reclassification (Extended Data Fig. 3, Supplementary Table 35). The lack of model calibration and the different characteristics of the validation and testing cohorts in HF may be due to its etiological diversity, which is covered in more detail in the discussion. PGS performance of all LDpred2 methods in the testing dataset is depicted in Supplementary Fig. 6.

CAD PGSsingleT and PGSmultiT were derived from a meta-analysis of CAD GWAS(7), which includes samples genotyped on the Illumina Infinium Exome beadchip from the MHIBB. The overlap between MHIBB and the derived PGS concerned ~1.5% of the SNPs in ~6% of cases and 2% of controls in the CAD GWAS meta-analysis. We performed sensitivity analyses which confirmed significant association of CAD PGS when removing the overlapping samples from the MHIBB, and when deriving PGS from an older CAD GWAS which did not include MHIBB samples(19), and supported the findings in the main analysis. Results of these sensitivity analyses are presented in the Supplementary Note and Supplementary Table 35.

When analyzing MTAG results per sex stratum we observed comparable results to the overall cohort for AF and HF. However, in CAD we observed significantly increased explained variance and discrimination ability in females for the PGSmultiT models compared to PGSsingleT [ΔAUC 0.019 (*P*=0.004); ΔR^2 1.725% (95%CI: 0.609, 2.856)], resulting in significantly improved reclassification: $NRI_{CAD\text{-}F}$ 0.208 (*P*<0.001) and $IDI_{CAD\text{-}F}$ 0.011 (*P*<0.001). PGSmultiT did not enhance the model results in males (ΔAUC 0.001 (*P*=0.68); ΔR^2 -0.100% (95%CI: -0.879, 0.676)], with an $NRI_{CAD\text{-}M}$ -0.012 (*P*=0.59) and $IDI_{CAD\text{-}M}$ 0.000 *(P*=0.74) (Fig.2, Supplementary Table 36). Of interest, PGSmultiT provided consistently higher OR values for the females in top PGS percentiles (1%, 5%, 10% 20%) versus the remaining cohort of females when compared to the PGSsingleT predictions (Supplementary Table 37, Extended Fig. 4).

DISCUSSION

The findings of our work can be summarized as follows: 1) Cardiovascular traits and common diseases have shared genetic architectures as demonstrated by the 443 significant pairwise genetic correlations found among the 56 included traits; 2) By leveraging such genetic correlations, multi-trait analysis of GWAS using MTAG empowers locus discovery in common cardiovascular diseases, identifying 19, 131, and 141 novel loci for AF, CAD, and HF, respectively; 3) Genomic risk prediction through genome-wide PGS provides similar results for single-trait and multi-trait derived scores in AF and CAD, and inconclusive results in HF; 4) Multi-trait derived PGS shows improved disease explained variance and discriminative ability for CAD in females, but not in males.

In our genetic correlation analyses, we reproduced prior findings from genetic and epidemiological studies and the new correlations fall within the expected. See detailed explanation in the Supplementary Note. Among the novel correlations in AF, we highlight the correlation of this electrical disease with trabeculation in LV mid-ventricular and apical segments, which in turn correlate with ECG repolarization in inferior and aVR leads. These genetic correlations are probably attributable to the development of the conduction system, bundle branches and the peripheral ventricular conduction

system, including Purkinje fibers, from the ventricular trabeculae.(20) Moreover, the Purkinje-muscle junction plays a role in the heterogeneity of transmural repolarization. (21)

The overlapping genetic structure of the included cardiometabolic traits allowed us to select the most relevant ones to enhance trait-specific effect estimates for AF, CAD and HF through MTAG. MTAG methodology has been previously used mostly in neurologic and psychiatric diseases (5, 22-34), in hypertrophic cardiomyopathy (HCM) (13, 35), and recently in HF(15). The hierarchical clustering approach our group uses for selecting the traits to be included in MTAG may be useful to other researchers in the setting of numerous multiple relevant associations with the trait of interest. By selecting traits from different clusters, we include traits with less overlapping genetic architecture in the MTAG analysis. The joint MTAG analysis of 4 traits, including those with the highest r_g with the trait of interest from each cluster, allows the calculation of the maxFDR, which controls the conformity of the included traits with the MTAG assumptions.

Recently, Levin *et al.* conducted an MTAG analysis on HF using their HF GWAS (N_{eff}=428,755) together with LV ejection fraction (LVEF, r_g -0.29, P=4.51x10^{-11}), LV end-systolic volume (r_g 0.36, P=3.73x10^{-16}) and LV end-diastolic volume (r_g 0.33 P=2.98x10^{-13}). In the MTAG conducted by Levin *et al*., 23 loci not identified by the conventional GWAS were found using MTAG. The increase in χ^2 and maxFDR were not reported. In contrast, when we performed an MTAG analysis employing the summary statistics of Levin *et al.* we found 89 additional loci to the GWAS. The main difference between the approaches is that while Levin *et al*. included highly inter-correlated traits in the MTAG analysis (LV volumetric traits), with relatively small sample size (N=36,041) and moderate r_g with HF, the clustering approach we employed led to the inclusion of more diverse traits which also displayed high genetic correlation with HF, namely AF(r_g 0.56, P=3.10x10^{-61}), CAD (r_g 0.72, P=7.38x10^{-80}) and DCM (r_g 0.56, P=1.81x10^{-8}).

In our study, MTAG uncovered 19, 131, and 141 loci in AF, CAD, and HF that had not been identified by single-trait summary statistics, of which a large proportion are reported for the first time, i.e., not reported in other GWAS contained in the GWAS Catalog, UK biobank (UKB), and FinnGen databases. In the absence of independent replication, caution is warranted since it is possible that some of those novel loci are false positive. MTAG allows for an estimation of the upper bound of the false discovery rate. In our 3 analyses, maxFDR was <5% suggesting that risk for false positive results is low. To further assess the robustness of our implementation of MTAG, we queried the GWAS catalog and PheWAS in FinnGen and UKB for the 291 novel loci; 134 (46%) were either reported at P<5x10^{-8} in other published GWAS (which may include the used datasets) or had P<0.005 in an independent case-control GWAS (mostly FinnGen). This further supports the robustness of MTAG to uncover novel loci.

Among genes mapping to *novel* loci, we identify pathophysiological pathways congruent with the mechanisms of each disease and with relevant potential therapeutic implications. Among the new AF loci, rs2209073 on chromosome 17 mapped to the *SRR* gene, responsible for the synthesis of D-serine, a co-agonist of NMDA receptors whose activation has shown a role in reduced HR variability and enhanced AF inducibility in animal models. (36) This locus has been previously associated with QRS in another GWAS. (37) In CAD, the *NPR3* gene mapping rs1173727 on chromosome 5 codes for a natriuretic peptide receptor whose inhibition provides cardioprotective role for CAD according to Mendelian randomization studies(38). In HF, genes mapped to novel loci on chromosome 5 (*TBX5*, lead SNP rs10507248) and 12 (*NKX2-5,* lead SNP rs2277923), are known genes in congenital heart disease involving septal defects and conduction abnormalities (39, 40). In addition, MTAG loci in HF mapped numerous genes involved in cytoskeleton structure, which has a known role in cardiomyopathy and emerged as a therapeutic target for reducing stiffness in HF (41, 42), such as *CORO6* and *BCAS3*

(mapped to rs11655587 and rs4968535, respectively, on chromosome 17) and *SWAP70* (mapped to rs360153 on chromosome 11). See detailed description in the Supplementary Note.

To further evaluate the utility of MTAG in common cardiovascular diseases, we assessed the performance of the multi-trait derived PGSs in disease prediction. To overcome statistical power challenges in genomic prediction we employed MTAG under the hypothesis that MTAG-derived PGS would improve SNP effect estimates and PGS-based disease prediction. To our knowledge, this is the first work employing MTAG-based genomic prediction in common cardiovascular diseases.

Our PGSs were derived using LDpred2 in a cohort of 15,177 patients from the MHIBB, a hospital-based cohort including individuals predominantly of French-Canadian ancestry (97.7%). The MHIBB was previously used to successfully validate other CAD PGS (43). In this work, we replicate previous results of the PGS tested in the MHIBB and those tested in other cohorts such as the UKB (16, 44) (Supplementary Table 39). The MHIBB is a cardiology hospital cohort in which there is obvious enrichment of cardiovascular diseases. The discrimination parameters employed in this work (AUC, continuous NRI and IDI) are not affected by disease prevalence. (45)

In Turley *et al.*(5), MTAG developers obtained improvements in R^2 values of PGS models derived from MTAG versus single-trait GWAS. Such R^2 improvements were close to what would be expected based on the observed increase in the mean χ^2 statistics. In our study, we also obtained an increase in mean χ^2 in the MTAG summary statistics of the three diseases, but this was not reflected in improved R^2 with MTAG in the overall cohort results (Supplementary Table 35). The main differences with the Turley *et al*. study, apart from the nature of the diseases of interest, is that herein we calculate the PGS by LDpred2-grid, which prioritizes SNPs effect sizes based on a prior assumption of the genetic architecture of the disease according to a diverse spectrum of hyperparameters (ρ and h^2), rather than LDpred1-infinitesimal, which adjusts SNP effects employing nominal h^2 and assuming all variants are causal (ρ=1). From our work we deduce that the improvement of SNP-phenotype associations provided by MTAG in common cardiac diseases does not translate into improved genetic prediction when employing LDpred2-grid-derived PGS. We hypothesize this is because common cardiac diseases dispose of large sample-size GWAS and LDpred2-grid adequately identifies the genetic architecture of the disease and weights SNPs according to its probability of being causal, altogether rendering relevant SNPs effect estimates for disease prediction adequately captured and re-estimated. In the current work LDpred2-grid method outperformed LDpred2-infinitesimal, supporting that prioritization of relevant variants effects instead of all improves disease prediction. LDpred2 methodology is different from classical methods such as C-T, that employ *unadjusted* SNP effects prioritized according to *P*-value. In the latter scenario, MTAG enhanced summary statistics, with a larger number of significant loci ($P<5x10^{-8}$), outperform conventional GWAS in predicting common cardiovascular diseases (CAD and HF) when employing low *P*-value thresholds (Supplementary Fig. 9).

Interestingly, even if MTAG-derived PGSs did not enhance PGSsingleT models, they did not impair genomic prediction either, thereby MTAG did not add noise to the single-trait summary statistics, i.e., MTAG did not bias the estimate of the effect toward correlated traits away from the real effect on the trait of interest.

HF analyses yielded low h^2_{SNP} and accordingly HF models showed low R^2. PGS in HF showed poor calibration and different performance in the validation and testing datasets, i.e., the randomly created cohorts for validation and testing of PGSs had different clinical characteristics (Table 2) and different performance of the PGS hyperparameters (Supplementary Tables 38-40). PGS including a higher fraction of causal variants provided the best predictions when assessing the testing dataset (see

further details in Supplementary Note). This points to greater polygenicity or etiological diversity of the disease which, in our opinion, is consistent with the fact that HF is a syndrome with several underlying etiopathologies, and potentially opposing mechanisms (e.g., hypercontractile in HCM vs. hypocontractile in DCM)(13). Similarly, Shah *et al*. stated that HF GWAS revealed lower statistical power and lower number genetic associations compared to other studies with similar N_{eff}, suggesting the heritability of HF may be more attributable to underlying disease subtypes than to the common final pathway (4). Despite low h^2_{SNP}, HF had many significant genetic correlations. In particular, HF depicted high genetic overlap with CAD *(*r_g 0.72) and several overlapping correlations with CAD, consistent with the high CAD prevalence in the cohorts included in the HF GWAS by Shah *et al.* This may raise concerns as to whether the HF correlations obtained were specific to HF or mediated by CAD, but it is reassuring that traits with high genetic overlap with HF such as DCM (defined as *non-ischemic*, r_g 0.60, P=1.81x10^{-8}) were not correlated to CAD (r_g 0.10, P=0.20), as expected. In addition, PGSmultiT grid and infinitesimal models depicted similar and better performance, respectively, than the PGSsingleT ones. These last two findings suggest that MTAG does not introduce a bias toward CAD-related mechanisms independent of HF. Likewise, the enhanced performance of the infinitesimal model in HF suggests that MTAG can be useful in etiologically diverse diseases.

The stratified analysis by sex yielded similar results for AF and HF as those observed in the overall cohort. This was not the case for CAD where there was no improvement by PGSmultiT in males but PGSmultiT in females significantly outperformed PGSsingleT. In females, PGSmultiT obtained a significant increase in prediction and discrimination parameters, with increased explained variance by the model, an improved classification as represented by the NRI and significantly better IDI showing better differentiation of prediction between affected and non-affected. This suggests that MTAG better captures disease mechanisms in this underrepresented group in CAD. We provide further explanation in the Supplementary Note. Even after the enhancement, CAD explained variance by PGSmultiT in females was lower than those in males, similarly to what was previously reported. (46)

One of the limitations of MTAG is that only those SNPs represented in all included traits are retained in the analysis, so some significant SNPs for the trait of interest may be excluded and limit novel loci discovery and predictive power of the final summary statistics, because even if the SNP to trait association statistic (i.e., χ^2) are improved, the number of represented SNPs is dropped. Similarly, for the use of LDpred2 it is recommended to restrict SNPs to the subset of HapMap3 SNPs (N=1,120,696), which may also compromise the final predictive capacity due to the possible omission of relevant SNPs. Another important limitation of our work is that in order to calculate the LDSC genetic correlation through summary statistics, only predominantly European populations were included, therefore our results may not be transferable to other cohorts with different ancestry. Finally, it is important to note that MTAG assumes that all SNPs share the same variance-covariance matrix of effect sizes across the included traits in the analysis, which might lead to deviations in the actual estimates. The maxFDR calculation estimates the maximum probability that the above assumption is violated, which was low in all our analyses. In addition, we confirmed a high proportion of the MTAG loci. However, we still recommend interpreting the new loci with great caution, specifically for those that were nonsignificant in the single-trait summary statistics at a nominal *P*-value threshold.

Altogether, our research provides evidence that the integration of shared genetic underpinnings among CV traits and diseases using MTAG joint analysis is an interesting and relevant tool for the discovery of novel trait-specific loci and unravelling the knowledge underlying their pathophysiology in AF, CAD, and HF. Multi-trait PGSs yield similar discrimination and prediction in CAD and AF, and non-conclusive results for HF, although they point to better performance (when infinitesimal PGS

methods are considered). Furthermore, MTAG enabled improvement in discriminative and predictive capabilities for CAD in females.

METHODS

Compilation and quality control on study summary-level data

Compilation of publicly available summary statistics of CV phenotypes published to date, prioritizing those with the largest size and power, resulted in the inclusion of 56 CV traits/diseases and 62 summary statistics (Supplementary Table 1).

We included studies comprising European ancestry individuals (or mostly European individuals) to be able to compute LDSC pairwise genetic correlation.

The statistics were generated as described in the original GWAS studies (see Supplementary Table 1 for references). They were processed according to the standard quality control procedures (47) and transformed into the input format required for the software tools to be used: LDSC regression (8) and MTAG (5).

All summary statistics were converted to the same genome assembly: GRCh37. When SNPs identification was missing, they were annotated using ANNOVAR and avsnp150 left normalized SNP database(48).

During the quality control we removed insertions, deletions and monomorphic SNPs (frequency equal to 0 or 1) as per MTAG software requirements.

LDSC genomic correlation

LDSC r_g of the CV included traits and diseases was conducted to assess their genetic overlap and as a guidance for the selection of the traits that we incorporated in the joint analysis using MTAG. We followed the procedure recommended in the LDSC score tutorial (https://github.com/bulik/ldsc/wiki/LD-Score-Estimation-Tutorial), using the munge step to restrict it to well imputed HapMap3 SNPs. The munge_sumstats.py script, which is used by both LDSC and MTAG, filters SNPs with MAF ≤ 0.01, out-of-bounds *P*-values [0-1] and also the values smaller than the smallest possible float value in python ($P<2.23 \times 10^{-308}$), strand-ambiguous SNPs, duplicated rs numbers and it also restricts SNP sample size variation by computing the 90^{th} percentile of the SNP sample-size distribution and dropping SNPs with a sample size below 66% of this value for LDSC, and 75% of this value for MTAG. For the LDSC analyses we employed the LD scores reference panel for European population from the Broad Institute 'eur_w_ld_chr', available from https://data.broadinstitute.org/alkesgroup/LDSCORE and computed using 1000 Genomes Project European individuals. (49)

Among 1770 non repeated comparisons (marked in the Unique Pairs column in Supplementary Table 2), LDSC r_g *P*-value < 0.05 after correcting for multiple comparisons (P_{FDR}, using Benjamini & Hochberg false discovery rate method) were considered significant (Supplementary Table 2). To select the traits significantly correlated with AF, CAD and HF, the correction was performed individually for each disease of interest (Supplementary Tables 4-6).

LDSC regression outputs h^2_{SNP} on an observed scale for binary traits, i.e., using the prevalence of the condition in the sample. h^2_{SNP} on a liability scale was computed for these traits and reported in the Supplementary Table 3, along with the population prevalence used for its calculation.

MTAG analysis

We performed MTAG analysis for AF, CAD and HF.

In addition to the filters discussed in the previous section, MTAG restricts the SNPs to those common to all the summary statistics in the analysis.

For binary traits input summary statistics, the effective sample size (N_{eff}) was calculated using the formula N_{eff} = 4 / (1/Number of cases + 1/Number of controls) as recommend in the MTAG publication. (5) The GWAS equivalent sample size after MTAG was obtained from the MTAG log file, which calculates it per SNP as $N_{eff\,MTAG} = N_{eff\,GWAS} * (\text{mean } \chi^2_{\ MTAG} - 1)/(\text{mean } \chi^2_{GWAS} - 1)$.

To ensure the reliability of our results, we calculated the maxFDR, which calculates the type I error in the dataset for the worst-case scenario: SNPs that are truly null for one trait but non-null for another trait. The maxFDR increases with lower genetic correlation and wider difference in the summary statistics mean χ^2 between the included traits. χ^2 scales with sample size and h^2.

To be able to calculate MTAG maxFDR, we conducted MTAG using 4 traits per each analysis (otherwise its calculation is computationally not feasible).

To select the traits to be included in each MTAG analysis we conducted a cluster analysis for the traits significantly correlated with each of the diseases of interest: AF, CAD and HF. Clustering was performed for traits with at least moderate LDSC r_g (≥0.20) that is statistically significant (P_{FDR}≤0.01). The trait with the highest genetic correlation from each cluster was selected, which allowed to include traits with high overlapping genetic structure with the trait of interest but not between them.

Hierarchical clustering was performed using the absolute value of the r_g, the Euclidean distance and the complete method (Fig. 2).

For the AF MTAG we used Nielsen *et al.* summary statistics as the main trait, instead of Roselli *et al.* ones, as the first one had higher summary statistics mean $\chi^{2.}$ (Supplementary Table 7).

For the HF MTAG using Shah *et al*.(4), we initially included AF, CAD, and SBP(12) (Supplementary Fig. 10) but it yielded a maxFDR of 40.9%. This was presumed to be due to imbalanced statistical power between HF (mean χ^2 = 1.14) and SBP (mean χ^2 = 2.84). As described in the MTAG method paper, imbalance of statistical power between traits is a well-recognized cause of type 1 error inflation. We then excluded blood pressure traits from the HF MTAG analysis. New clustering was performed leading to the selection of DCM as the fourth trait for the HF MTAG, yielding a maxFDR of 0.4% (Supplementary Table 7).

For the new MTAG analysis using new summary statistics published in CAD (Aragam *et al.*) and HF (Levin *et al.*), we kept the same features we used for the main analysis, using updated summary statistics if available (Supplementary Table 23) and did not combine 1) Nelson *et al.* with Aragam *et al.* nor 2) Shah *et al.* with Levin *et al.* (similarly to what we did for AF), as these cohorts were already included in the Aragam *et al.* and Levin *et al.* GWAS, respectively.

Annotation and functional analysis

The annotation and functional mapping of the new generated summary statistics was conducted using FUMA v1.5.3 (https://fuma.ctglab.nl/)(50) and Open Targets v22.10 (https://genetics.opentargets.org/)(51, 52).

Genomic risk loci were defined in FUMA. We clumped independent SNPs at a LD r^2 >0.1 and merged LD blocks with distance ≤ 250kb. This led to a number of loci slightly different to the ones reported in the original summary statistics (Supplementary Tables 8-10, 25-26). MTAG *novel* loci were defined as those loci not identified as significant in the original GWAS summary statistics of the trait of interest (at a significance level $P \leq 5x10^{-8}$) and without overlapping boundaries with them. For the AF results, we reviewed the new loci with respect to Nielsen *et al.*(6) and Roselli *et al.(3)* and the meta-analysis of non-overlapping samples of both reported in Roselli et al.

Prioritization of causal genes was conducted using Open Targets, an open approach to systematically prioritize causal variants and genes and which integrates information on molecular phenotype quantitative trait loci experiments, chromatin interaction experiments, *in silico* functional predictions and the distance between the variant and each gene's canonical transcription start site. The genes with the highest variant to gene score for each lead variant are reported in Table 1 and in Supplementary Tables 8-10, 25-26, and the whole list of associated genes is reported in Supplementary Tables 11-13, 27-28.

Summary statistics functional analysis was conducted using MAGMA gene-set and tissue expression functions in FUMA.

The association of the novel loci with the traits of interest (AF, CAD and HF) in other studies was explored through FUMA and OpenTargets. FUMA retrieves GWAS associations from the GWAS Catalog for all the variants included in a locus (defined using the parameters previously specified). The GWAS Catalog contains variants reported to be associated with any phenotype at a significance level of $P \leq 1x10^{-5}$. In our case, we filtered for $P \leq 5x10^{-8}$. OpenTargets retrieves GWAS associations of pre-specified variants, in our case the list of the novel MTAG lead variants, with the trait of interest in the UKB, FinnGen and/or GWAS Catalog repositories. In the query, OpenTargets uses the specified variants and those that are in high LD with them (LD r^2 >0.7) or in the same credible set, i.e., set of independent variants in a locus (after conditional analysis) likely to be causal at a 95% confidence interval. In addition, Open Targets also provides PheWAS associations ($P \leq 5x10^{-3}$) for the pre-specified list of variants, in our case, i.e., the list of the novel MTAG lead variants we input.

Polygenic scores

Derivation data

We used the enhanced summary statistics from the MTAG analyses to construct PGS (PGSmultiT) and compared them to PGS obtained from single-trait summary statistics (PGSsingleT) to assess whether MTAG improved prediction and discrimination ability. The PGSsingleT were derived from the previously published publicly available GWAS summary statistics on AF (Nielsen et al.(6)), CAD (Nelson et al.(7)), and HF (Shah et al.(4)), and the PGSmultiT from the summary statistics that we generated through MTAG for each disease. As recommended by LDpred2 developers the SNPs used were restricted to those available in HapMap3. Quality control of the summary statistics before the score calculation included the removal of indels, SNPs with a MAF ≤ 0.01, imputation info ≤ 0.7 when available, SNPs with no rsID or position available and duplicate and ambiguous SNPs. Prior to the calculation of the PGSs, the summary statistics were converted to GRCh38 assembly using the *liftover* function available at the UCSC Genome Browser(53), to be consistent with the genetic build of the MHIBB, the target data.

PGS calculation in the target data

PGS were calculated in a population of 15,177 unrelated individuals with complete array genotype and detailed phenotypic data. Individuals were recruited from the André and France Desmarais Hospital Cohort of the MHIBB. All participants provided informed consent to be included in the Biobank. The MHIBB protocol was approved by Montreal Heart Institute's Scientific and Ethics Committees. This study was approved by these committees and by the Cohort Management Committee. Complete information on quality control and the imputation of the genotypes is described in the Supplementary Note. The validation cohort was composed of 2,500 individuals, which was used to select the best performing hyperparameters, and 12,677 were included as the testing cohort. The baseline characteristics of the MHIBB cohort are described in Table 2. Baseline data analyses were conducted in R (version 4.0.2). Continuous variables were presented as mean ± standard deviation or median [interquartile range (IQR)], and these were compared using the t-test for independent samples or the Mann-Whitney U test. Categorical variables were reported as frequencies (%) and the Fisher's exact test was used for comparison.

The scores were derived using LDpred2 in the main analysis, as well as a comprehensive analysis including LDpred1 and C-T methods. Further details on the hyperparameters employed for each method are available in the Supplementary Note.

Using the LDpred2 score with the best performance in the validation dataset based on the Wald-Z from the logistic regression model, the performance of the PGSs was assessed in the 12,677 MHIBB in the testing dataset. Logistic regression models with the disease as the outcome and the disease-specific candidate PGS as a predictor, and another model including sex, age at last follow-up and the first four genotypic principal components as covariates, were compared for PGSsingleT vs PGSmultiT.

PGS logistic regression models Nagelkerke R^2 and AUC were assessed through the rms R package (version 6.2-0), and confidence intervals were created for 10,000 bootstrapping samples using boot (version 1.3-2), and AUCs compared using pROC package (version 1.18.2) and the roc.test() function which employs the Delong method(54). NRI and IDI were assessed using the R package Hmisc (version 4.7-0). For calibration plots, 1,000 bootstrapping samples were used and the continuous calibration curve was smoothed using the lowess method(55). All analyses were conducted in R (version 4.0.2).

The definitions of the disease or outcome used for the PGS analyses were as follows:

For AF we coded as prevalent AF all patients coded as having atrial fibrillation or atrial flutter in the MHIBB, same diagnosis that were included in Nielsen *et al*.(6) Controls were selected as not having atrial fibrillation or atrial flutter.

CAD was defined as the occurrence at any time of fatal or nonfatal myocardial infarction (MI), percutaneous transluminal coronary angioplasty (PTCA) or coronary artery bypass grafting (CABG) and controls were selected to be free of MI, PTCA or CABG. Our definition was similar to the one used by Nelson *et al*.(7), but this one also included patients with unstable angina and chronic ischemic heart disease as cases for the summary statistics we employed (SOFT).

HF was defined as HF diagnosed by a cardiologist based on clinical notes, hospitalization or death report. In addition, individuals with a left ventricle ejection fraction ≤45% were also coded under the diagnosis of HF in the MHIBB. Controls were selected to be free of HF. Our definition was in line with the one used by Shah *et al*.(4), in which the diagnosis of the 26 cohorts was based on the combination of the following parameters: Physician diagnosis or adjudication in 18 cohorts, ICD codes (discharge ±

cause of death) in 12 cohorts, imaging in 15 cohorts, natriuretic peptides in 1 cohort and HF treatment in 9 cohorts. In the search for a more precise definition of HF, the PGS was also tested using two additional definitions:

a. Cases were those with HF who also had an NT-proBNP ≥ 125 pg/ml when available, excluding those patients without NT-proBNP and a diagnosis of HF from the analysis, and keeping as controls those without HF.

b. Cases were defined as having severe HF which included patients who received a heart transplant, those which have carried ventricular assist device, or those which had experienced HF NYHA functional class III-IV.

These results are available in Supplementary Table 40.

DATA AVAILABILITY

Upon publication of the study following peer-review, summary statistics generated through the use of MTAG will be made available on the GWAS catalog, and polygenic scores derived from single-trait and multi-trait summary statistics will be made available on the PGS catalog using both the GRCh37 and GRCh38 builds.

ACKNOWLEDGMENTS

We would like to thank all the patients who volunteered to be included in the Montreal Heart Institute Biobank cohort. As well, we would like to acknowledge Vincent Bourgeault from the Montreal Heart Institute Hospital for his support during this project.

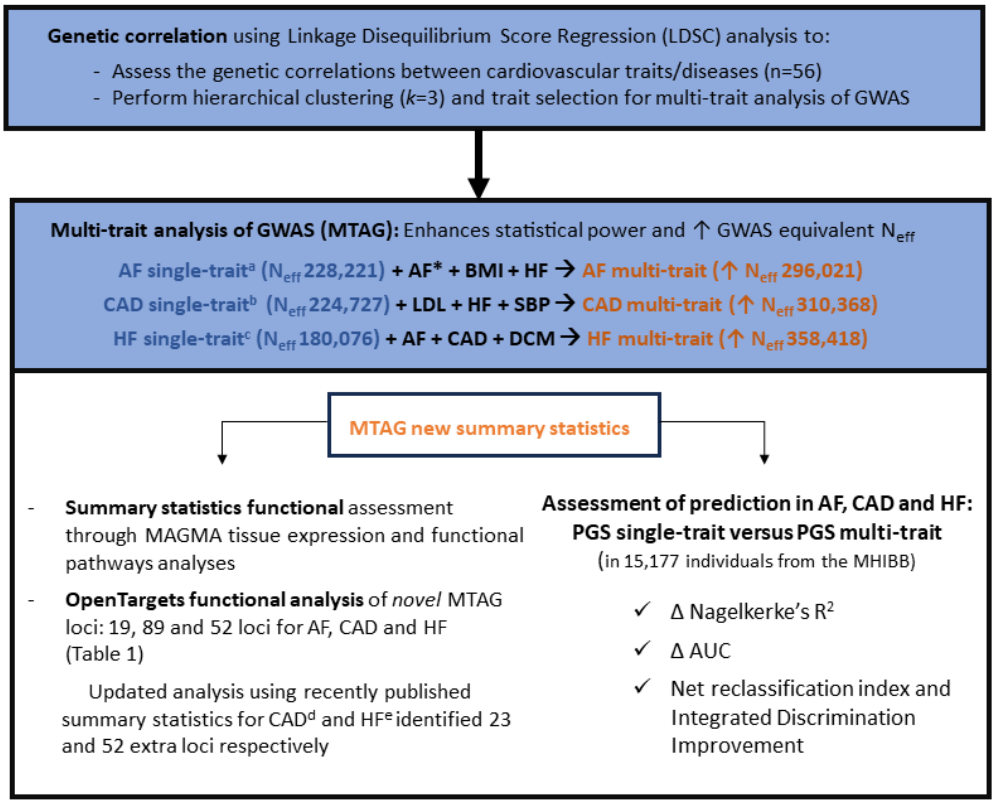
REFERENCES

1. Behan FM, Iorio F, Picco G, Goncalves E, Beaver CM, Migliardi G, et al. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. Nature. 2019;568(7753):511-6.
2. Yang C, Starnecker F, Pang S, Chen Z, Guldener U, Li L, et al. Polygenic risk for coronary artery disease in the Scottish and English population. BMC Cardiovasc Disord. 2021;21(1):586.
3. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, et al. Multi-ethnic genome-wide association study for atrial fibrillation. Nat Genet. 2018;50(9):1225-33.
4. Shah S, Henry A, Roselli C, Lin H, Sveinbjornsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nat Commun. 2020;11(1):163.
5. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat Genet. 2018;50(2):229-37.
6. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nat Genet. 2018;50(9):1234-9.
7. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet. 2017;49(9):1385-91.
8. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015;47(11):1236-41.
9. Meyer HV, Dawes TJW, Serrani M, Bai W, Tokarczuk P, Cai J, et al. Genetic and functional insights into the fractal structure of the heart. Nature. 2020;584(7822):589-94.
10. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet. 2019;28(1):166-74.
11. Surakka I, Horikoshi M, Magi R, Sarin AP, Mahajan A, Lagou V, et al. The impact of low-frequency and rare variants on lipid levels. Nat Genet. 2015;47(6):589-97.
12. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet. 2018;50(10):1412-25.
13. Tadros R, Francis C, Xu X, Vermeer AMC, Harper AR, Huurman R, et al. Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. Nat Genet. 2021;53(2):128-34.
14. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Atri DS, et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. Nat Genet. 2022;54(12):1803-15.
15. Levin MG, Tsao NL, Singhal P, Liu C, Vy HMT, Paranjpe I, et al. Genome-wide association and multi-trait analyses characterize the common genetic architecture of heart failure. Nat Commun. 2022;13(1):6914.
16. Prive F, Arbel J, Vilhjalmsson BJ. LDpred2: better, faster, stronger. Bioinformatics. 2020.
17. Vilhjalmsson BJ, Yang J, Finucane HK, Gusev A, Lindstrom S, Ripke S, et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. Am J Hum Genet. 2015;97(4):576-92.
18. Thomas LE, O'Brien EC, Piccini JP, D'Agostino RB, Pencina MJ. Application of net reclassification index to non-nested and point-based risk prediction models: a review. Eur Heart J. 2019;40(23):1880-7.
19. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015;47(10):1121-30.

20. van Weerd JH, Christoffels VM. The formation and function of the cardiac conduction system. Development. 2016;143(2):197-210.
21. Walton RD, Martinez ME, Bishop MJ, Hocini M, Haissaguerre M, Plank G, et al. Influence of the Purkinje-muscle junction on transmural repolarization heterogeneity. Cardiovasc Res. 2014;103(4):629-40.
22. Deak JD, Zhou H, Galimberti M, Levey DF, Wendt FR, Sanchez-Roige S, et al. Genome-wide association study in individuals of European and African ancestry and multi-trait analysis of opioid use disorder identifies 19 independent genome-wide significant risk loci. Mol Psychiatry. 2022.
23. Seligowski AV, Misganaw B, Duffy LA, Ressler KJ, Guffanti G. Leveraging Large-Scale Genetics of PTSD and Cardiovascular Disease to Demonstrate Robust Shared Risk and Improve Risk Prediction Accuracy. Am J Psychiatry. 2022:appiajp21111113.
24. Carcel-Marquez J, Muino E, Gallego-Fabrega C, Cullell N, Lledos M, Llucia-Carol L, et al. A Polygenic Risk Score Based on a Cardioembolic Stroke Multitrait Analysis Improves a Clinical Prediction Model for This Stroke Subtype. Front Cardiovasc Med. 2022;9:940696.
25. Guo P, Gong W, Li Y, Liu L, Yan R, Wang Y, et al. Pinpointing novel risk loci for Lewy body dementia and the shared genetic etiology with Alzheimer's disease and Parkinson's disease: a large-scale multi-trait association analysis. BMC Med. 2022;20(1):214.
26. Wang H, Yi Z, Shi T. Novel loci and potential mechanisms of major depressive disorder, bipolar disorder, and schizophrenia. Sci China Life Sci. 2022;65(1):167-83.
27. Meyers JL, Zhang J, Chorlian DB, Pandey AK, Kamarajan C, Wang JC, et al. A genome-wide association study of interhemispheric theta EEG coherence: implications for neural connectivity and alcohol use behavior. Mol Psychiatry. 2021;26(9):5040-52.
28. Liu K, Zhu L, Yu M, Liang X, Zhang J, Tan Y, et al. A Combined Analysis of Genetically Correlated Traits Identifies Genes and Brain Regions for Insomnia. Can J Psychiatry. 2020;65(12):874-84.
29. Liu L, Wen Y, Ning Y, Li P, Cheng B, Cheng S, et al. A trans-ethnic two-stage polygenetic scoring analysis detects genetic correlation between osteoporosis and schizophrenia. Clin Transl Med. 2020;9(1):21.
30. Yang XL, Zhang SY, Zhang H, Wei XT, Feng GJ, Pei YF, et al. Three Novel Loci for Infant Head Circumference Identified by a Joint Association Analysis. Front Genet. 2019;10:947.
31. Chung J, Marini S, Pera J, Norrving B, Jimenez-Conde J, Roquer J, et al. Genome-wide association study of cerebral small vessel disease reveals established and novel loci. Brain. 2019;142(10):3176-89.
32. Pei YF, Liu L, Liu TL, Yang XL, Zhang H, Wei XT, et al. Joint Association Analysis Identified 18 New Loci for Bone Mineral Density. J Bone Miner Res. 2019;34(6):1086-94.
33. Chung W, Chen J, Turman C, Lindstrom S, Zhu Z, Loh PR, et al. Efficient cross-trait penalized regression increases prediction accuracy in large cohorts using secondary phenotypes. Nat Commun. 2019;10(1):569.
34. Lam M, Trampush JW, Yu J, Knowles E, Davies G, Liewald DC, et al. Large-Scale Cognitive GWAS Meta-Analysis Reveals Tissue-Specific Neural Expression and Potential Nootropic Drug Targets. Cell Rep. 2017;21(9):2597-613.
35. Tadros R, Zheng SL, Grace C, Jorda P, Francis C, Jurgens SJ, et al. Large scale genome-wide association analyses identify novel genetic loci and mechanisms in hypertrophic cardiomyopathy. medRxiv. 2023.
36. Shi S, Liu T, Wang D, Zhang Y, Liang J, Yang B, et al. Activation of N-methyl-d-aspartate receptors reduces heart rate variability and facilitates atrial fibrillation in rats. Europace. 2017;19(7):1237-43.
37. Norland K, Sveinbjornsson G, Thorolfsdottir RB, Davidsson OB, Tragante V, Rajamani S, et al. Sequence variants with large effects on cardiac electrophysiology and disease. Nat Commun. 2019;10(1):4803.
38. Cronje HT, Karhunen V, Hovingh GK, Coppieters K, Lagerstedt JO, Nyberg M, et al. Genetic evidence implicating natriuretic peptide receptor-3 in cardiovascular disease risk: a Mendelian randomization study. BMC Med. 2023;21(1):158.

39. Xu JH, Gu JY, Guo YH, Zhang H, Qiu XB, Li RG, et al. Prevalence and Spectrum of NKX2-5 Mutations Associated With Sporadic Adult-Onset Dilated Cardiomyopathy. Int Heart J. 2017;58(4):521-9.
40. Steimle JD, Moskowitz IP. TBX5: A Key Regulator of Heart Development. Curr Top Dev Biol. 2017;122:195-221.
41. Caporizzo MA, Prosser BL. The microtubule cytoskeleton in cardiac mechanics and heart failure. Nat Rev Cardiol. 2022;19(6):364-78.
42. Ehler E. Actin-associated proteins and cardiomyopathy-the 'unknown' beyond troponin and tropomyosin. Biophys Rev. 2018;10(4):1121-8.
43. Wunnemann F, Sin Lo K, Langford-Avelar A, Busseuil D, Dube MP, Tardif JC, et al. Validation of Genome-Wide Polygenic Risk Scores for Coronary Artery Disease in French Canadians. Circ Genom Precis Med. 2019;12(6):e002481.
44. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018;50(9):1219-24.
45. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med. 2011;30(1):11-21.
46. Nikpay M, Stewart AFR, McPherson R. Partitioning the heritability of coronary artery disease highlights the importance of immune-mediated processes and epigenetic sites associated with transcriptional activity. Cardiovasc Res. 2017;113(8):973-83.
47. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, et al. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. Int J Methods Psychiatr Res. 2018;27(2):e1608.
48. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
49. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.
50. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826.
51. Ghoussaini M, Mountjoy E, Carmona M, Peat G, Schmidt EM, Hercules A, et al. Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. Nucleic Acids Res. 2021;49(D1):D1311-D20.
52. Mountjoy E, Schmidt EM, Carmona M, Schwartzentruber J, Peat G, Miranda A, et al. An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. Nat Genet. 2021;53(11):1527-33.
53. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002;12(6):996-1006.
54. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.
55. Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. Stat Med. 2014;33(3):517-35.

Fig. 1: Flowchart of the study design

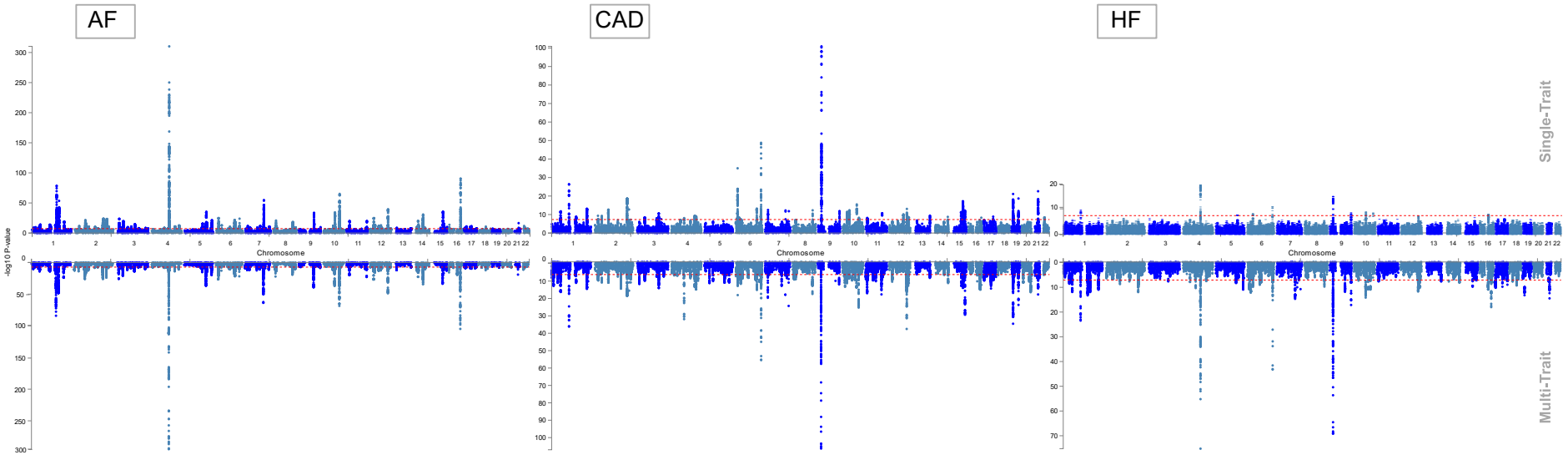


Description of the design of the study. Abbreviations: AF, atrial fibrillation and atrial flutter; AUC, area under the curve; BMI, body mass index; CAD, coronary artery disease; DCM, dilated cardiomyopathy; HF, heart failure; LDL, low-density lipoprotein; MHIBB, Montreal Heart Institute biobank; N_{eff}, effective sample size; PGS, polygenic risk score; SBP, systolic blood pressure. [a]Nielsen *et al.(6)*, [b]Nelson *et al.(7)*, [c]Shah *et al.(4)*, [d]Aragam *et al.(14)*, [e]Levin *et al.(15)* *Another AF study different from Nielsen *et al.* was included.

Fig.2: Linkage Desequilibrium Score Regression Genetic Correlations and Hierarchichal Clustering for Atrial Fibrillation, Coronary Artery Disease and Heart Failure

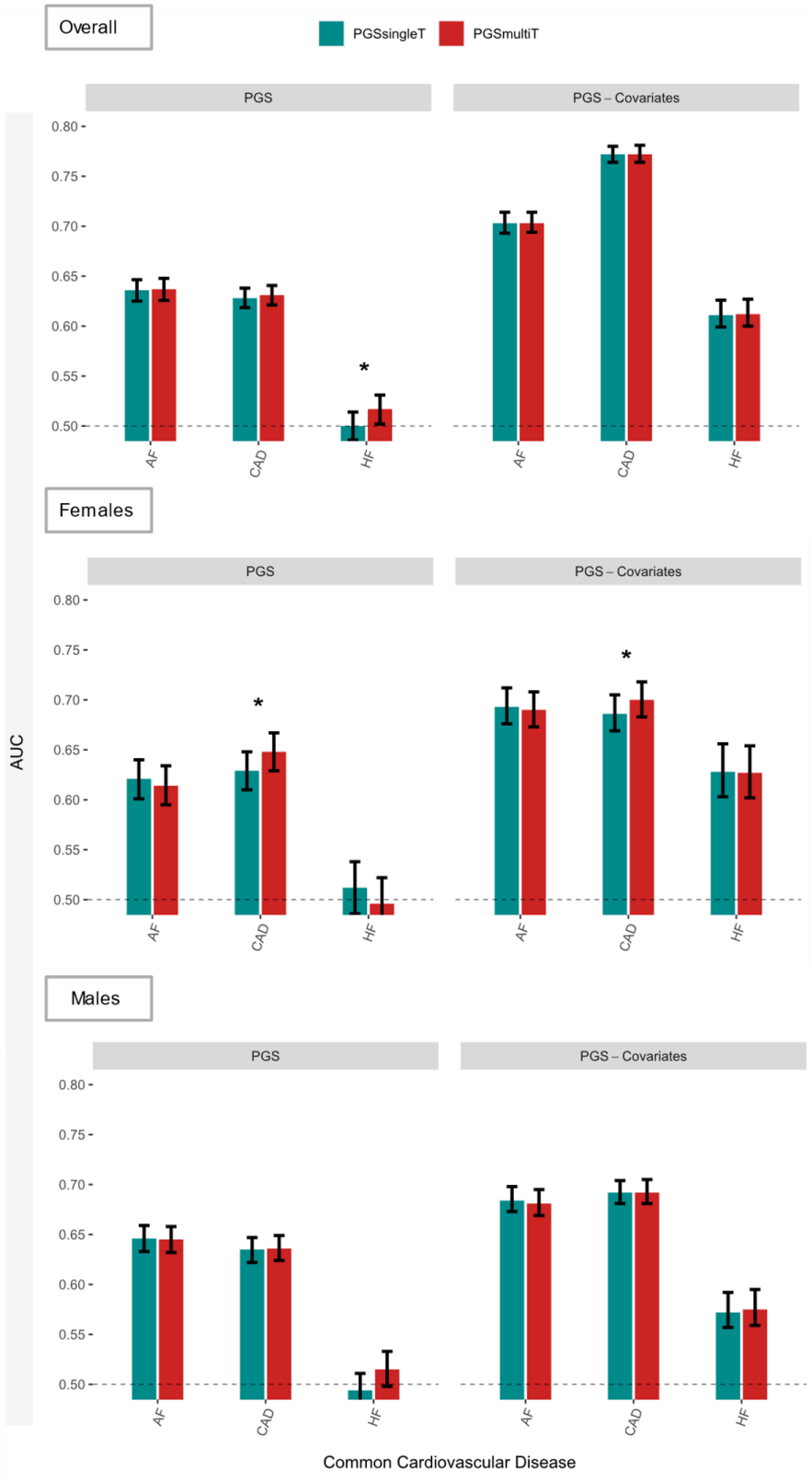
On the left, barplots depicting significant genetically correlated traits with AF, CAD, and HF (P_{FDR}≤0.05; r_g and 95% confidence intervals are represented). All r_g values appear in Supplementary Tables 4-6. On the right, hierarchical clustering of the traits with r_g>0.20 and P_{FDR}≤0.01 with AF, CAD and HF. The Euclidean distance and the complete method were used to obtain three clusters grouping the traits significantly correlated with AF, CAD, and HF. For each of the three clusters, the trait with the highest genetic correlation with the trait of interest was selected for MTAG analysis (marked with an asterisk). P_{FDR}, *P*-value corrected by false discovery rate; r_g, genetic correlation. Other abbreviations as per Supplementary Table 1.

Fig.3: Miami plots of single-trait and multi-trait summary statistics



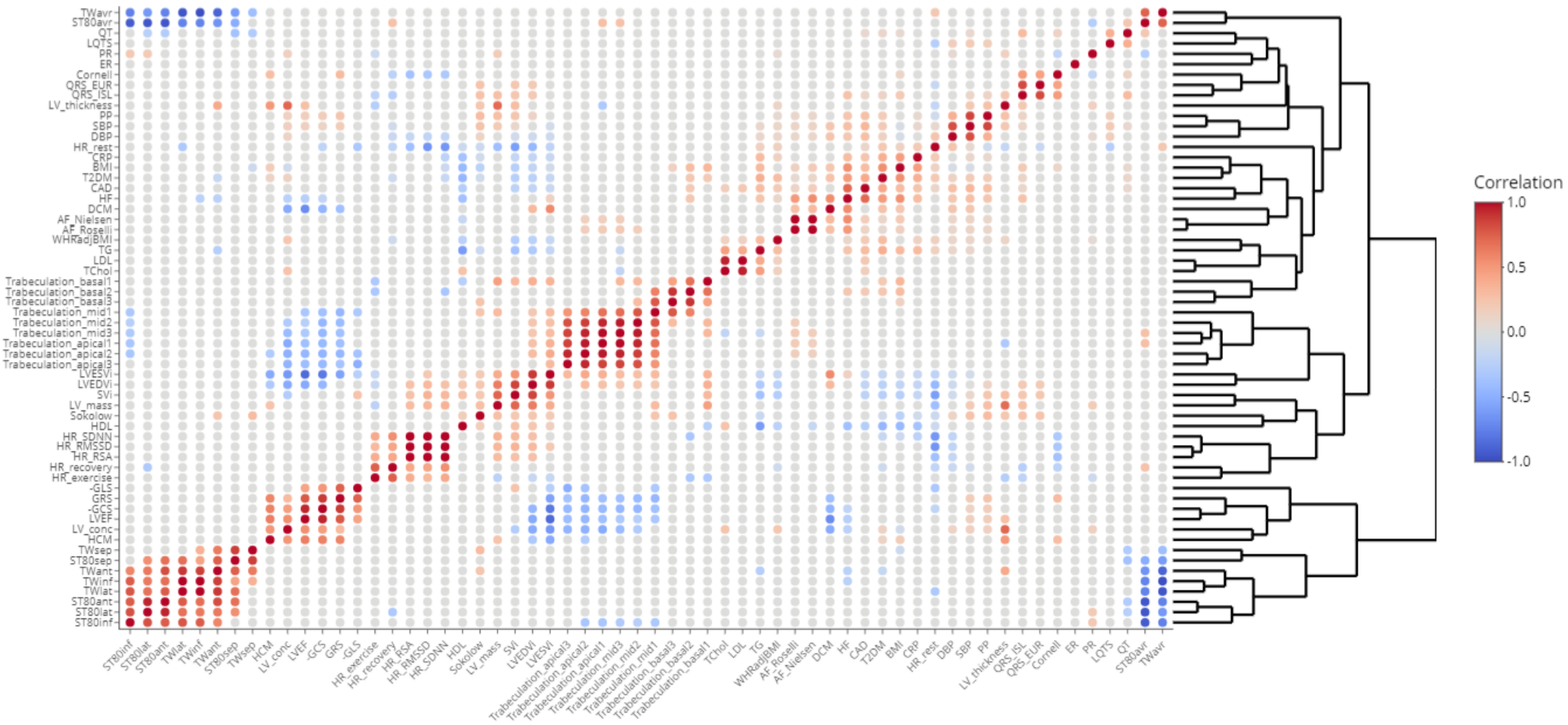
Miami plots from the summary statistics of the original studies, Nielsen *et al.* for atrial fibrillation (AF), Nelson *et al.* for coronary artery disease (CAD) and Shah *et al.* for heart failure (HF), are depicted at the top, and from the MTAG results at the bottom. MTAG increased the number of significant identified loci from 119 to 138 in AF, from 50 to 135 for CAD, and from 11 to 61 in HF. This led to the identification of 19 novel loci not reported in the original summary statistics for AF, 89 for CAD and 52 for HF. Red dashed line shows the genome-wide significance threshold (P=5x10^{-8}).

Fig.4: Single-trait versus Multi-trait Polygenic Scores: comparison of the discriminative ability

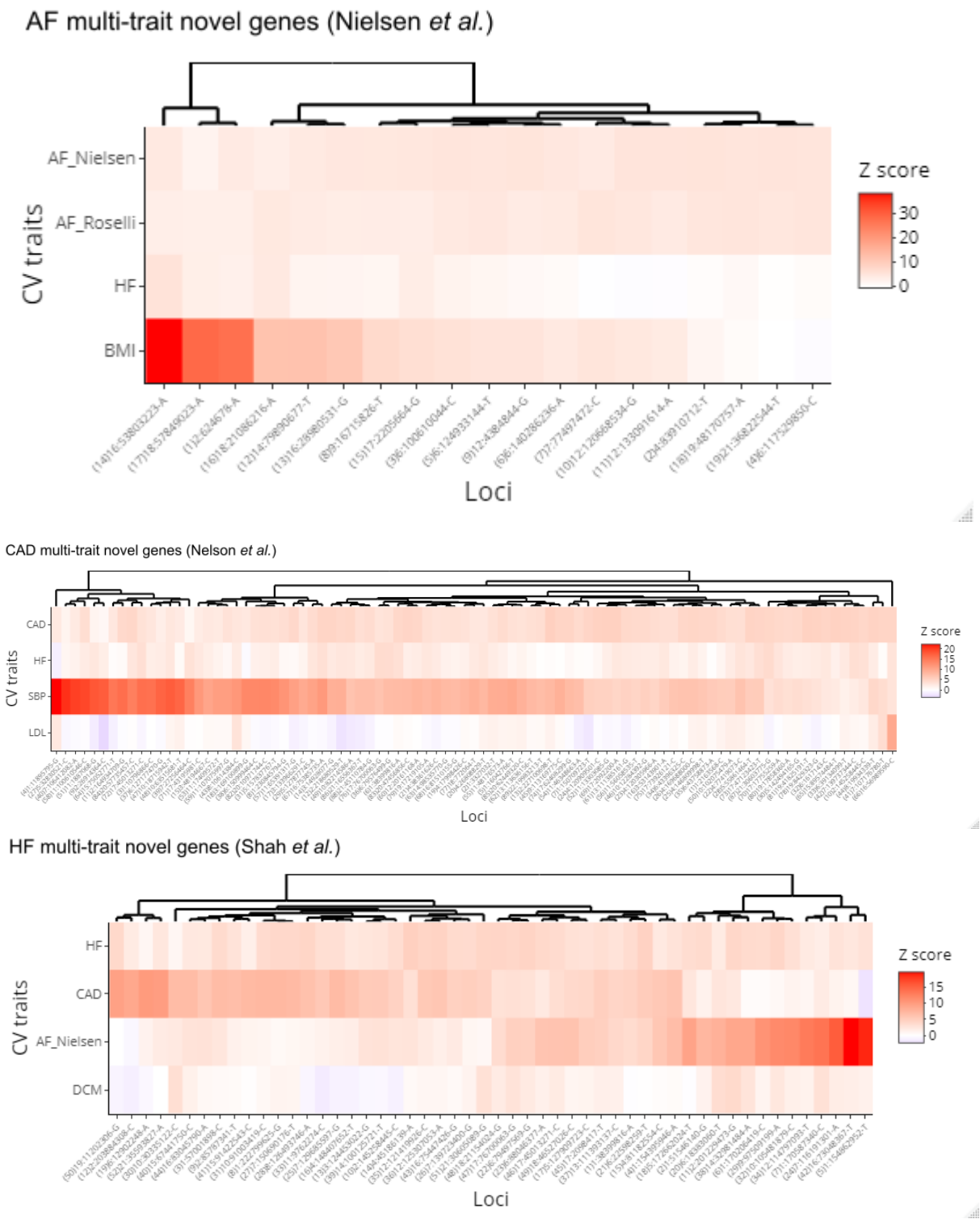


Area under the curve (AUC) from logistic regression model with the disease as an outcome (AF, CAD and HF) and the disease-specific candidate polygenic score as a predictor variable. The models including the single-trait summary statistics (PGSsingleT) are represented in blue and the multi-trait ones (PGSmultiT) in red. Scores developed by LDpred2, using the UK Biobank linkage disequilibrium reference panel (N= 362,320 individuals). The validation (N=2,500) and testing datasets (N=12,677) consisted in unrelated individuals from the Montreal Heart Institute Biobank cohort. AUC shown are those from the testing dataset. For the female cohort, 997 individuals were included in the validation dataset and 5,152 in the testing dataset. For the male cohort, 1,503 individuals in the validation dataset and 7,525 in the testing dataset. Representation of the AUCs corresponding to the best LDpred2 method, defined as the one with highest Wald Z in the validation dataset in each disease: LDpred2-grid. 95% confidence intervals are represented. Significant differences in the discrimination ability assessed using ΔAUC, Net Reclassification Index and Integrated Discrimination Improvement marked with an asterisk (*). The lack of model calibration in HF only-PGS models indicates caution in the interpretation of these results (Supplementary Table 35). Interestingly, PGS derived from the multi-trait summary statistics (PGSmultiT) improved discrimination ability for CAD in females (Supplementary Table 36).

Extended Data Fig.1: Genetic correlations

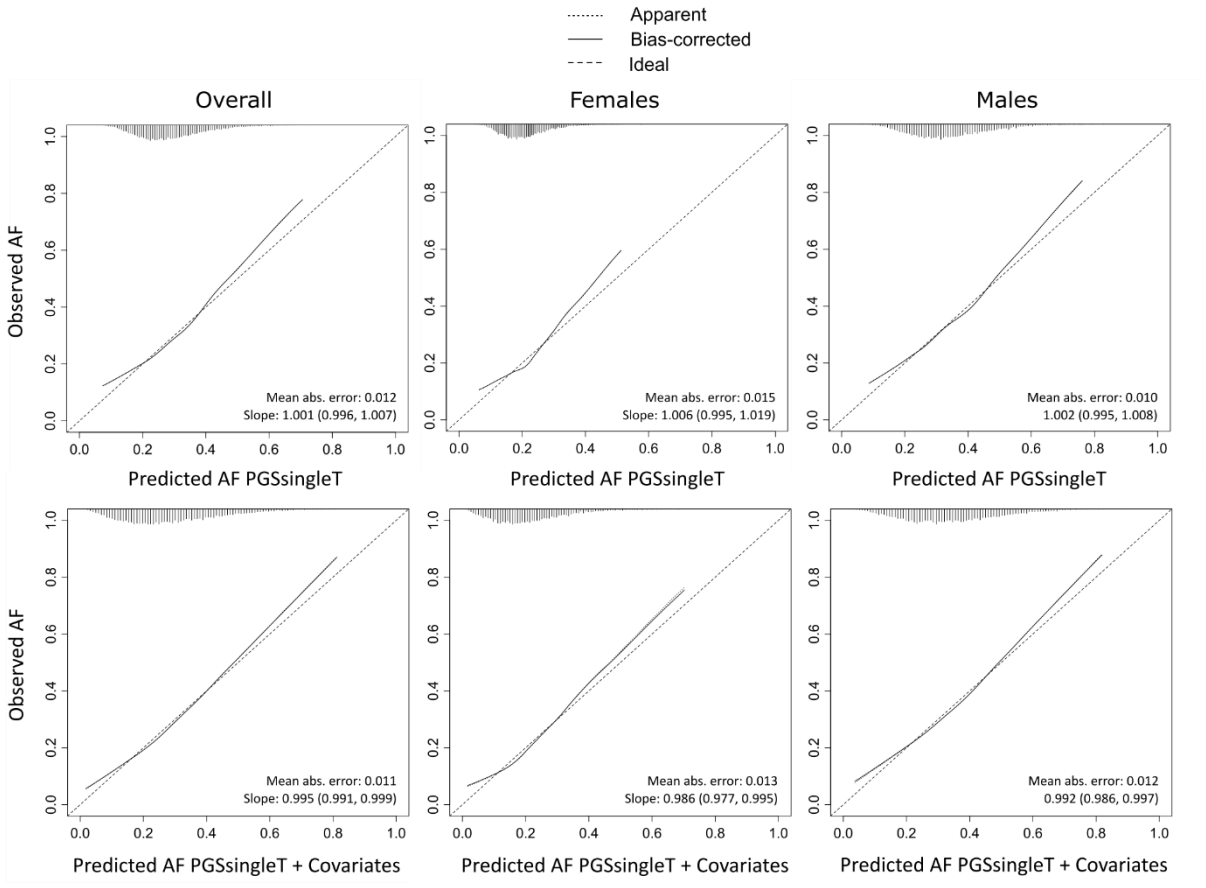


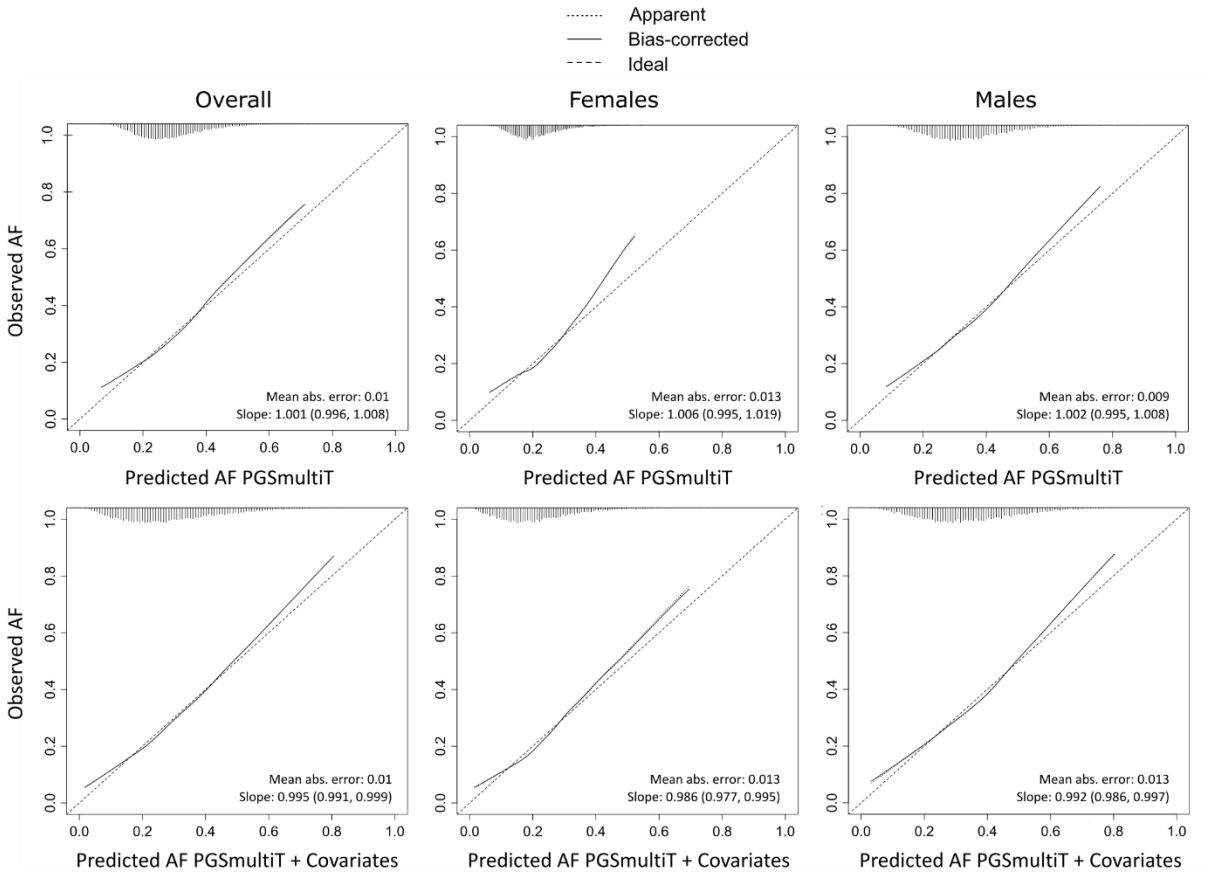
Extended Data Fig.1: Genetic correlations. Genetic correlations of the 56 traits and cardiovascular diseases with publicly available summary statistics (using Linkage Disequilibrium Score Regression (1), all values in supplementary table 2). Non-significant correlations (*P* value > 0.05 after correction for false discovery rate) are set to 0 and shaded in grey. Abbreviations specified in Supplementary Table 1.

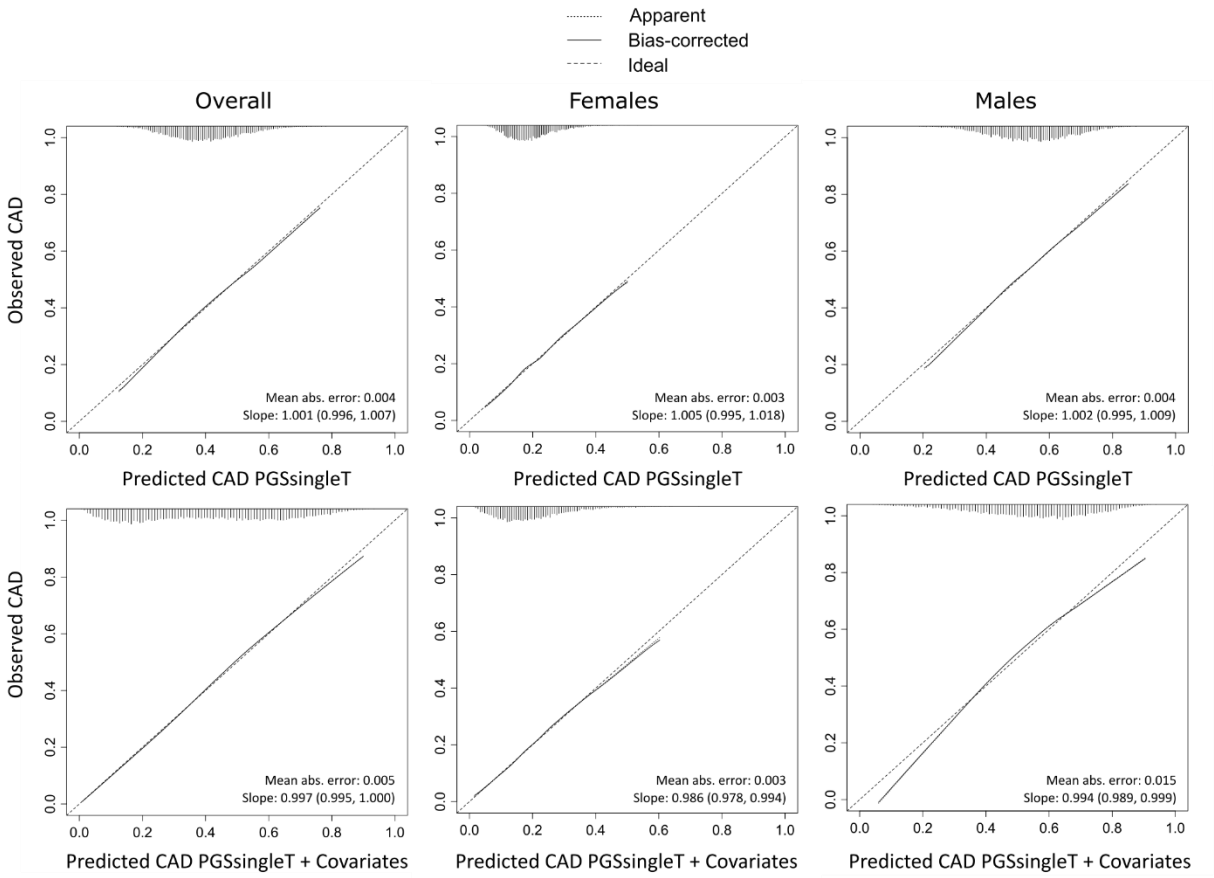


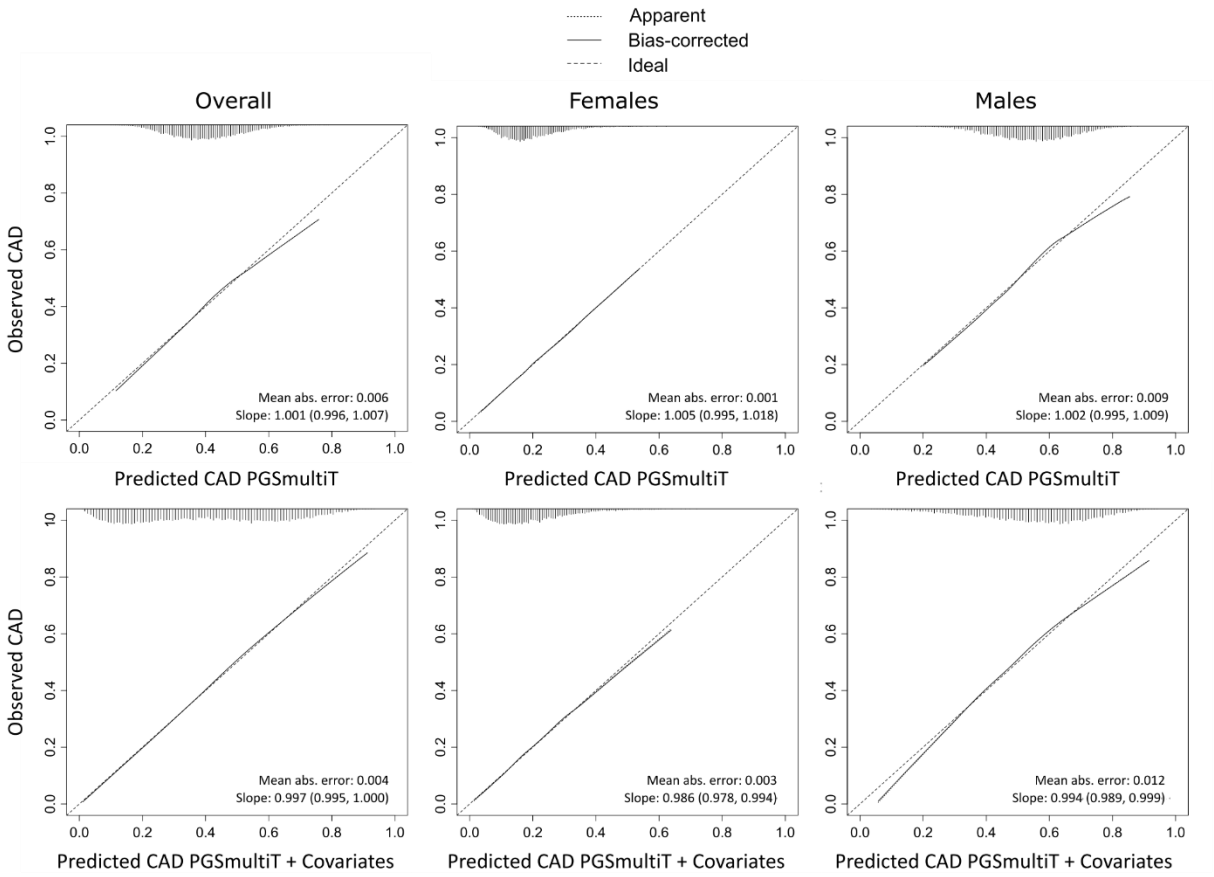
Extended Data Fig.2: Per SNP Z-score of the novel loci for each of the traits included in the MTAG analysis.

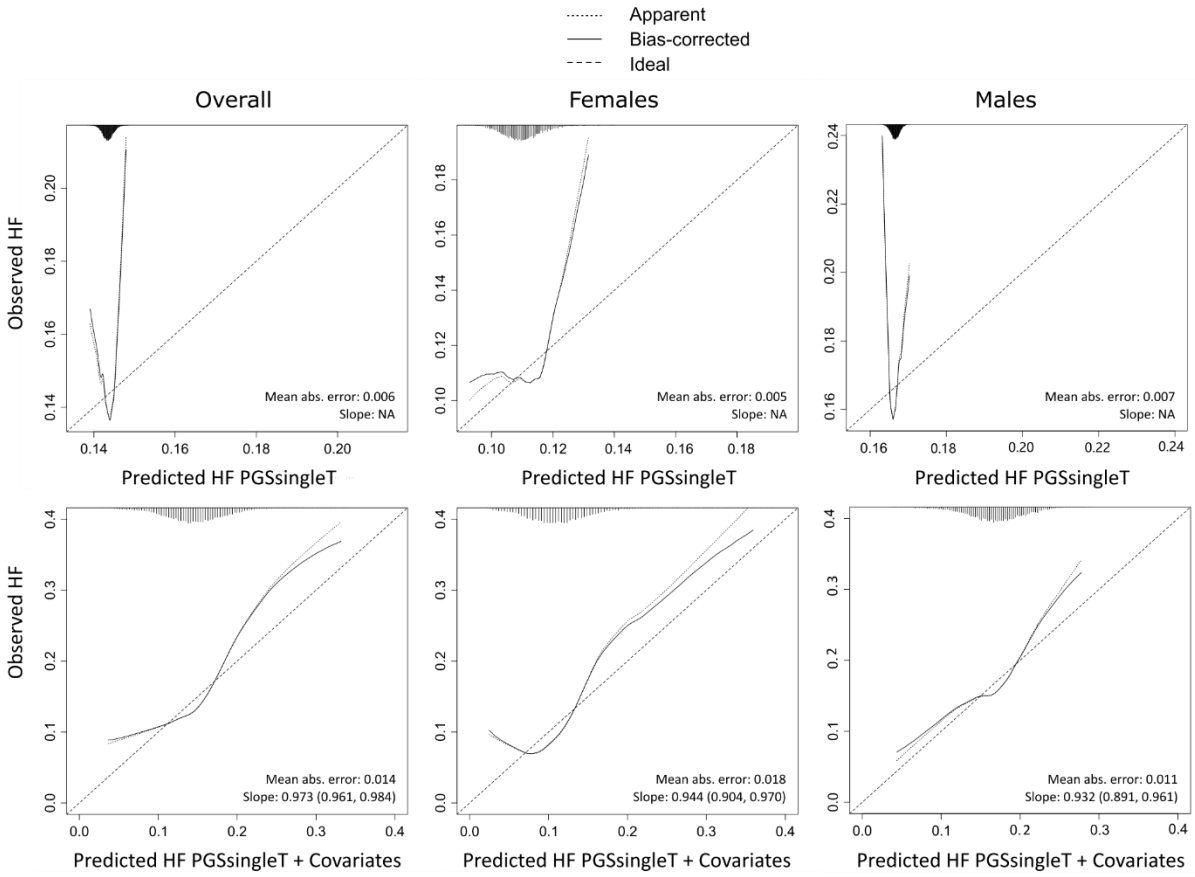
For each novel MTAG loci the lead SNP Z-score of the disease risk allele (beta divided standard error), from each of the traits included in the MTAG analysis is represented. SNP identifiers along the x-axis represent the locus number (Supplementary Tables 8-10) followed by the genomic position (in GRCh37) and the risk allele. Abbreviations as per Extended Data Fig. 1.

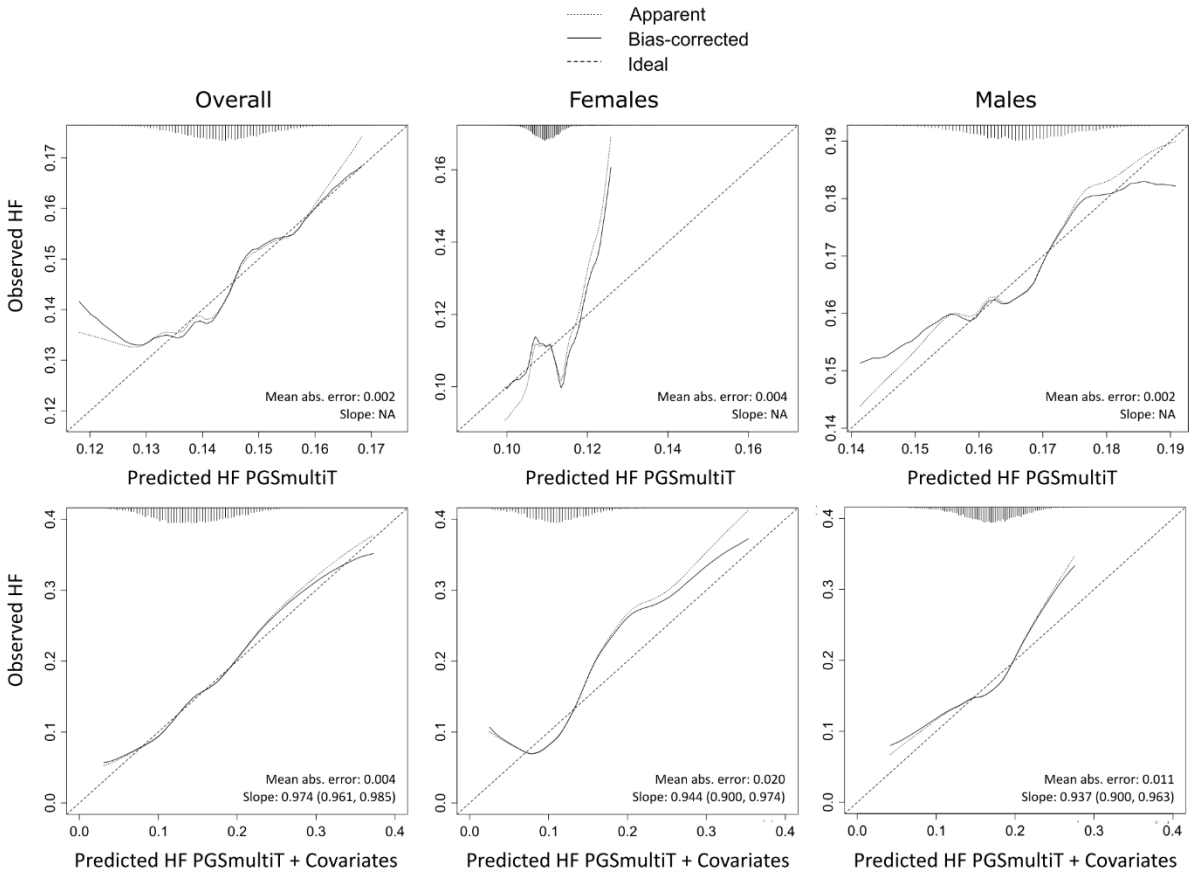

Apparent
Bias-corrected
Ideal
Overall
Females
Males
Observed AF
Mean abs. error: 0.012
Slope: 1.001 (0.996, 1.007)
Mean abs. error: 0.015
Slope: 1.006 (0.995, 1.019)
Mean abs. error: 0.010
1.002 (0.995, 1.008)
Predicted AF PGSsingleT
Mean abs. error: 0.011
Slope: 0.995 (0.991, 0.999)
Mean abs. error: 0.013
Slope: 0.986 (0.977, 0.995)
Mean abs. error: 0.012
0.992 (0.986, 0.997)
Predicted AF PGSsingleT + Covariates

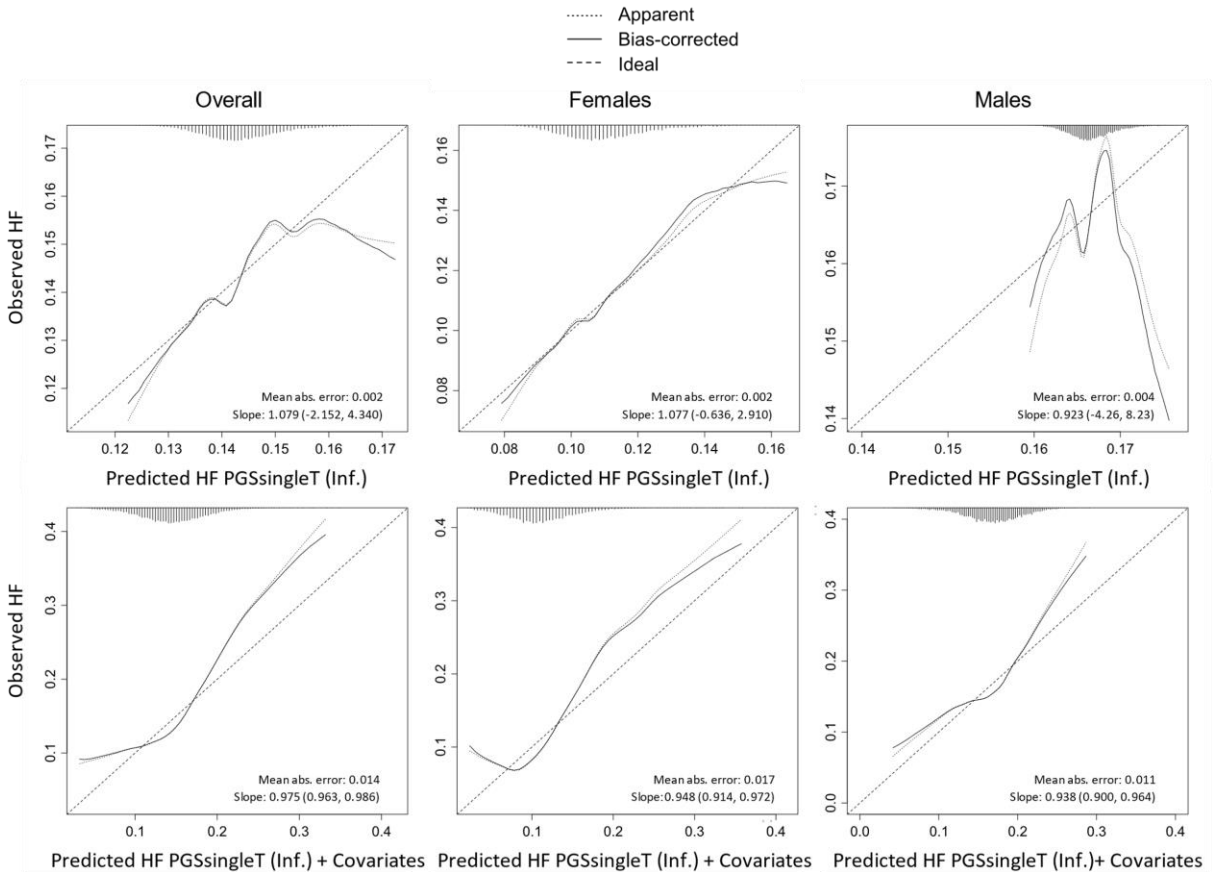

Apparent
Bias-corrected
Ideal
Overall
Females
Males
Observed AF
Mean abs. error: 0.01
Slope: 1.001 (0.996, 1.008)
Mean abs. error: 0.013
Slope: 1.006 (0.995, 1.019)
Mean abs. error: 0.009
Slope: 1.002 (0.995, 1.008)
Predicted AF PGSmultiT
Mean abs. error: 0.01
Slope: 0.995 (0.991, 0.999)
Mean abs. error: 0.013
Slope: 0.986 (0.977, 0.995)
Mean abs. error: 0.013
Slope: 0.992 (0.986, 0.997)
Predicted AF PGSmultiT + Covariates

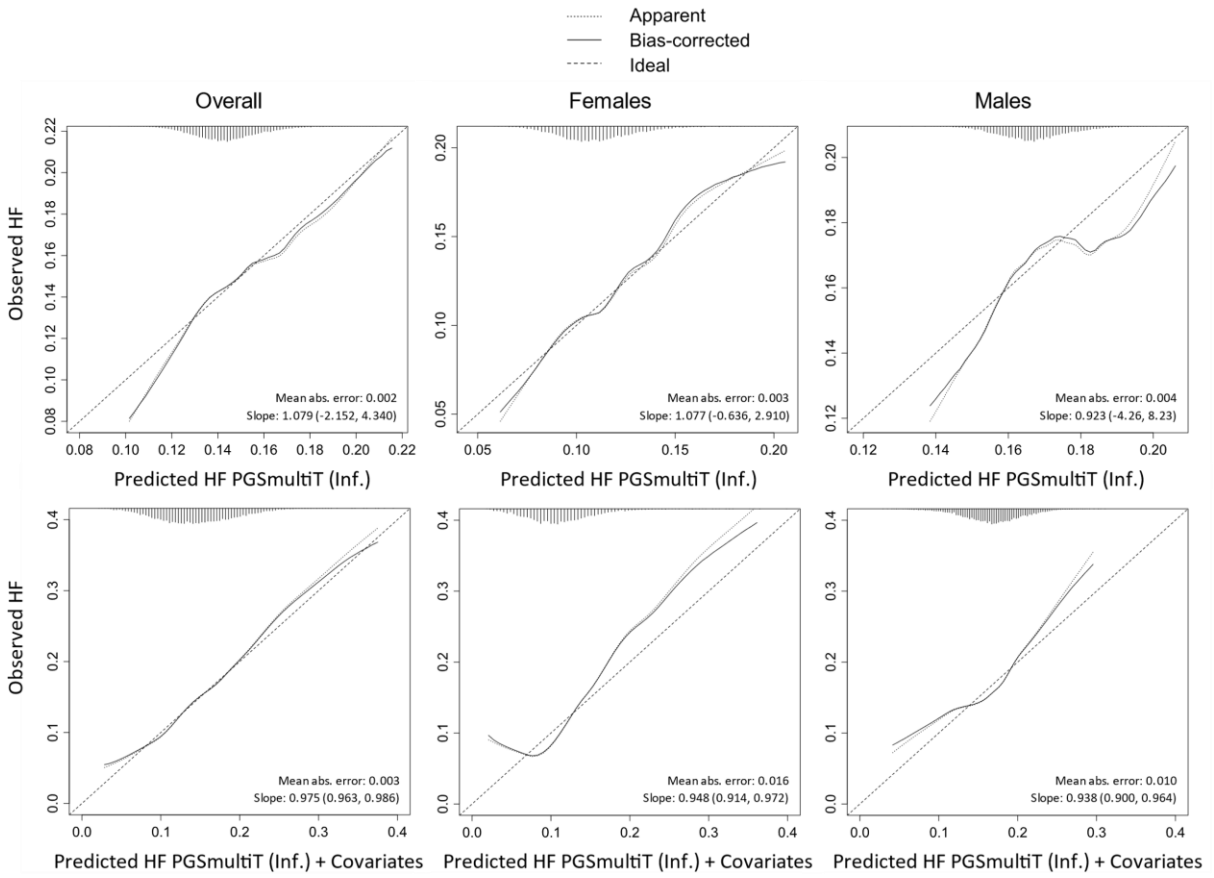

Apparent
Bias-corrected
Ideal
Overall
Females
Males
Observed CAD
Predicted CAD PGSsingleT
Mean abs. error: 0.004
Slope: 1.001 (0.996, 1.007)
Mean abs. error: 0.003
Slope: 1.005 (0.995, 1.018)
Mean abs. error: 0.004
Slope: 1.002 (0.995, 1.009)
Predicted CAD PGSsingleT + Covariates
Mean abs. error: 0.005
Slope: 0.997 (0.995, 1.000)
Mean abs. error: 0.003
Slope: 0.986 (0.978, 0.994)
Mean abs. error: 0.015
Slope: 0.994 (0.989, 0.999)


Apparent
Bias-corrected
Ideal
Overall
Females
Males
Observed CAD
Predicted CAD PGSmultiT
Mean abs. error: 0.006
Slope: 1.001 (0.996, 1.007)
Mean abs. error: 0.001
Slope: 1.005 (0.995, 1.018)
Mean abs. error: 0.009
Slope: 1.002 (0.995, 1.009)
Predicted CAD PGSmultiT + Covariates
Mean abs. error: 0.004
Slope: 0.997 (0.995, 1.000)
Mean abs. error: 0.003
Slope: 0.986 (0.978, 0.994)
Mean abs. error: 0.012
Slope: 0.994 (0.989, 0.999)

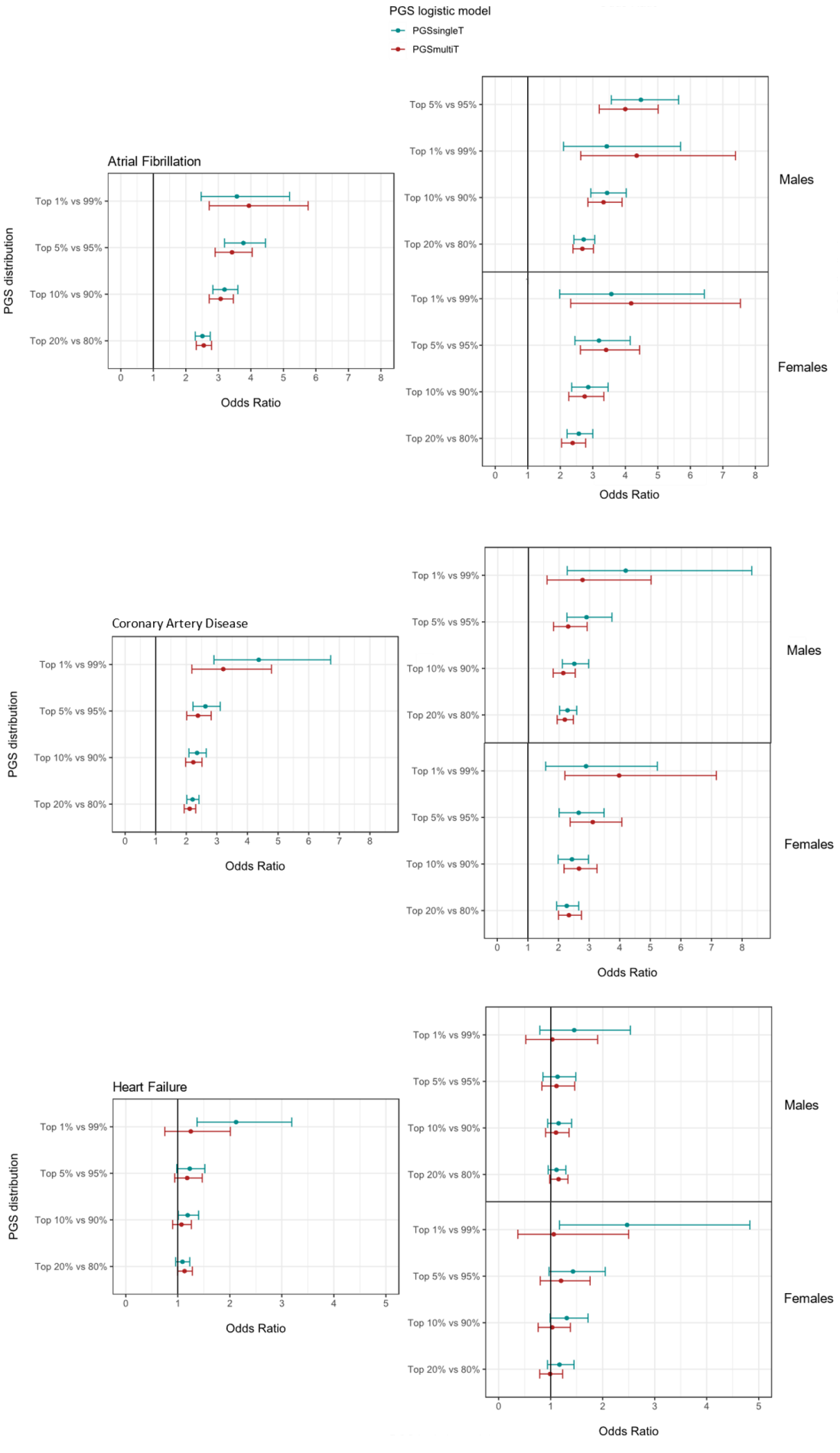

Apparent
Bias-corrected
Ideal
Overall
Females
Males
Observed HF
Predicted HF PGSsingleT
Mean abs. error: 0.006
Slope: NA
Mean abs. error: 0.005
Slope: NA
Mean abs. error: 0.007
Slope: NA
Predicted HF PGSsingleT + Covariates
Mean abs. error: 0.014
Slope: 0.973 (0.961, 0.984)
Mean abs. error: 0.018
Slope: 0.944 (0.904, 0.970)
Mean abs. error: 0.011
Slope: 0.932 (0.891, 0.961)


Apparent
Bias-corrected
Ideal
Overall
Females
Males
Observed HF
Predicted HF PGSmultiT
Mean abs. error: 0.002
Slope: NA
Mean abs. error: 0.004
Slope: NA
Mean abs. error: 0.002
Slope: NA
Predicted HF PGSmultiT + Covariates
Mean abs. error: 0.004
Slope: 0.974 (0.961, 0.985)
Mean abs. error: 0.020
Slope: 0.944 (0.900, 0.974)
Mean abs. error: 0.011
Slope: 0.937 (0.900, 0.963)


Apparent
Bias-corrected
Ideal
Overall
Females
Males
Observed HF
Mean abs. error: 0.002
Slope: 1.079 (-2.152, 4.340)
Mean abs. error: 0.002
Slope: 1.077 (-0.636, 2.910)
Mean abs. error: 0.004
Slope: 0.923 (-4.26, 8.23)
Predicted HF PGSsingleT (Inf.)
Mean abs. error: 0.014
Slope: 0.975 (0.963, 0.986)
Mean abs. error: 0.017
Slope:0.948 (0.914, 0.972)
Mean abs. error: 0.011
Slope: 0.938 (0.900, 0.964)
Predicted HF PGSsingleT (Inf.) + Covariates


Apparent
Bias-corrected
Ideal
Overall
Females
Males
Observed HF
Mean abs. error: 0.002
Slope: 1.079 (-2.152, 4.340)
Mean abs. error: 0.003
Slope: 1.077 (-0.636, 2.910)
Mean abs. error: 0.004
Slope: 0.923 (-4.26, 8.23)
Predicted HF PGSmultiT (Inf.)
Mean abs. error: 0.003
Slope: 0.975 (0.963, 0.986)
Mean abs. error: 0.016
Slope: 0.948 (0.914, 0.972)
Mean abs. error: 0.010
Slope: 0.938 (0.900, 0.964)
Predicted HF PGSmultiT (Inf.) + Covariates

Extended Data Fig. 3: Calibration plots showing agreement between observed (*y*-axis) and predicted probability of event (*x*-axis) in the testing dataset. The 45º degrees line represents perfect calibration. On the *x*-axis, the spike histogram represents the number of individuals with a predicted risk corresponding to the value on the x-axis. Continuous calibration curve smoothed using lowess method. Good calibration is observed for all models except HF.


PGS logistic model
PGSsingleT
PGSmultiT
Atrial Fibrillation
Coronary Artery Disease
Heart Failure
PGS distribution
Top 1% vs 99%
Top 5% vs 95%
Top 10% vs 90%
Top 20% vs 80%
Odds Ratio
Males
Females

Extended Data Fig. 4: Odds ratio across the different percentile groups for the overall cohort and stratified by sex.

Odds ratios and 95% confidence intervals were calculated by comparing predicted risk for those individuals at the top 1%, 5%, 10% and 20% PGS percentiles with the remaining individuals below the cutoff value for each percentile of risk. Predicted risk was calculated through a logistic regression model including the PGS as a predictor and each disease of interest, AF, CAD or HF, as the outcome. These analyses were conducted in the testing dataset for the overall cohort and per sex strata.

Article 3

ESC European Society of Cardiology

European Heart Journal (2022) **43**, 3041–3052
https://doi.org/10.1093/eurheartj/ehac289

CLINICAL RESEARCH
Arrhythmias

Arrhythmic risk prediction in arrhythmogenic right ventricular cardiomyopathy: external validation of the arrhythmogenic right ventricular cardiomyopathy risk calculator

Paloma Jordà [1,2,3†], Laurens P. Bosman[4†], Alessio Gasperetti[5], Andrea Mazzanti[6,7,8], Jean-Baptiste Gourraud[9], Brianna Davies[10], Tanja Charlotte Frederiksen[11,12], Zoraida Moreno Weidmann[13], Andrea Di Marco[14], Jason D. Roberts[15,16,17], Ciorsti MacIntyre[18], Colette Seifer[19], Antoine Delinière [20], Wael Alqarawi[21], Deni Kukavica [6,7,8], Damien Minois [9], Alessandro Trancuccio [6,7,8], Marine Arnaud [9], Mattia Targetti[22], Annamaria Martino[23], Giada Oliviero[23], Daniel C. Pipilas[24], Corrado Carbucicchio[25], Paolo Compagnucci[26], Antonio Dello Russo[26], Iacopo Olivotto[22], Leonardo Calò [23], Steven A. Lubitz[24], Michael J. Cutler[28], Philippe Chevalier [20], Elena Arbelo [2,3,29,30], Silvia Giuliana Priori [6,7,8], Jeffrey S. Healey[15,16], Hugh Calkins [5], Michela Casella [27], Henrik Kjærulf Jensen[11,12], Claudio Tondo[25,31], Rafik Tadros [1], Cynthia A. James[5], Andrew D. Krahn[10], and Julia Cadrin-Tourigny [1*]

[1]Cardiovascular Genetics Center, Montreal Heart Institute, Université de Montréal, Montréal, QC, Canada; [2]Arrhythmia Section, Cardiology Department, Hospital Clínic, Universitat de Barcelona, Barcelona, Spain; [3]Institut d'Investigació August Pi i Sunyer (IDIBAPS), Barcelona, Spain; [4]Division of Heart and Lungs, Department of Cardiology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; [5]Division of Cardiology, Department of Medicine, Johns Hopkins Hospital, Baltimore, MD, USA; [6]Department of Molecular Medicine, University of Pavia, Pavia, Italy; [7]Department of Molecular Cardiology, IRCCS Istituti Clinici Scientifici Maugeri, Pavia, Italy; [8]Department of Molecular Cardiology, Centro Nacional de Investigaciones Cardiovasculares, Madrid, Spain; [9]Department of Cardiology, Centre Hospitalier Universitaire Nantes, Nantes, France; [10]Centre for Cardiovascular Innovation, Division of Cardiology, University of British Columbia, Vancouver, BC, Canada; [11]Department of Cardiology, Aarhus University Hospital, Aarhus N, Denmark; [12]Department of Clinical Medicine, Health, Aarhus University, Aarhus N, Denmark; [13]Department of Cardiology, Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Universitat Autònoma de Barcelona, CIBERCV, Barcelona, Spain; [14]Arrhythmia Unit, Department of Cardiology, Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Barcelona, Spain; [15]Population Health Research Institute, McMaster University, Hamilton, ON, Canada; [16]Division of Cardiology, Hamilton Health Sciences, Hamilton, ON, Canada; [17]Section of Cardiac Electrophysiology, Division of Cardiology, Department of Medicine, Western University, London, ON, Canada; [18]Cardiac Electrophysiology Service, Quenn Elisabeth II Health Sciences Center, Dalhousie University, Halifax, NS, Canada; [19]St-Boniface Hospital, University of Manitoba, Winnipeg, MB, Canada; [20]National Reference Center for Inherited Arrhythmias of Lyon, Louis Pradel Cardiovascular Hospital, Hospices Civils de Lyon, Lyon, France; [21]Cardiac Electrophysiology Service, Ottawa Heart Institute, University of Ottawa, Ottawa, ON, Canada; [22]Cardiomyopathy Unit, Department of Cardiology, Careggi University Hospital, Florence, Italy; [23]Department of Cardiology, Policlinico Casilino, Rome, Italy; [24]Cardiovascular Research Center and Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, MA, USA; [25]Department of Clinical Electrophysiology and Cardiac Pacing Centro Cardiologico Monzino, IRCCSC, Milan, Italy; [26]Department of Biomedical Sciences and Public Health, Cardiology and Arrhythmology Clinic, University Hospital Umberto I-Salesi-Lancisi, Marche Polytechnic University, Ancona, Italy; [27]Department of Clinical, Special and Dental Sciences, Cardiology and Arrhythmology Clinic, University Hospital Umberto I-Salesi-Lancisi, MarchePolytechnic University, Ancona, Italy; [28]Intermountain Medical Center Heart Institute, Intermountain Medical Center, Murray, UT, USA; [29]Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Madrid, Spain; [30]European Reference Network for rare, low prevalence and complex diseases of the heart - ERN GUARD-Heart; and [31]Department of Biomedical, Surgical and Dentistry Sciences, University of Milan, Milan, Italy

Received 17 November 2021; revised 16 April 2022; accepted 18 May 2022; online publish-ahead-of-print 29 June 2022

See the editorial comment for this article 'Arrhythmogenic right ventricular cardiomyopathy: the never-ending quest for a risk calculator', by Estelle Gandjbakhch and Annina S. Vischer, https://doi.org/10.1093/eurheartj/ehac324.

* Corresponding author. Tel: +1 514 376 3330, Fax: +1 514 593 2496, Email: julia.cadrin-tourigny@umontreal.ca
† Share first authorship.

Abstract

Aims Arrhythmogenic right ventricular cardiomyopathy (ARVC) causes ventricular arrhythmias (VAs) and sudden cardiac death (SCD). In 2019, a risk prediction model that estimates the 5-year risk of incident VAs in ARVC was developed (ARVCrisk.com). This study aimed to externally validate this prediction model in a large international multicentre cohort and to compare its performance with the risk factor approach recommended for implantable cardioverter-defibrillator (ICD) use by published guidelines and expert consensus.

Methods and results In a retrospective cohort of 429 individuals from 29 centres in North America and Europe, 103 (24%) experienced sustained VA during a median follow-up of 5.02 (2.05–7.90) years following diagnosis of ARVC. External validation yielded good discrimination [C-index of 0.70 (95% confidence interval-CI 0.65–0.75)] and calibration slope of 1.01 (95% CI 0.99–1.03). Compared with the three published consensus-based decision algorithms for ICD use in ARVC (Heart Rhythm Society consensus on arrhythmogenic cardiomyopathy, International Task Force consensus statement on the treatment of ARVC, and American Heart Association guidelines for VA and SCD), the risk calculator performed better with a superior net clinical benefit below risk threshold of 35%.

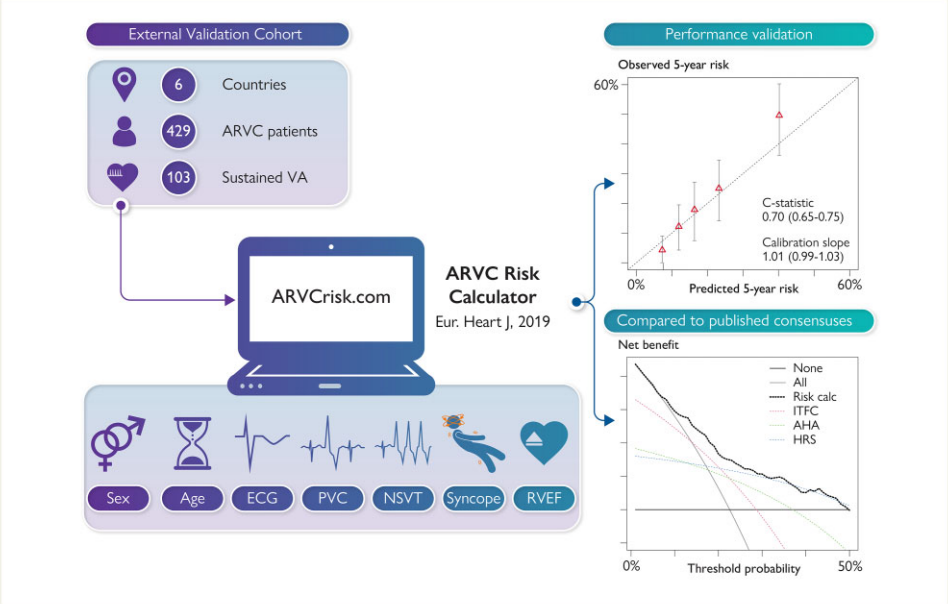
Conclusion Using a large independent cohort of patients, this study shows that the ARVC risk model provides good prognostic information and outperforms other published decision algorithms for ICD use. These findings support the use of the model to facilitate shared decision making regarding ICD implantation in the primary prevention of SCD in ARVC.

Structured Graphical Abstract

Key Question
Does the published arrhythmogenic right ventricular cardiomyopathy (ARVC) risk calculator (ARVCrisk.com) adequately predicting sustained ventricular arrhythmia (VA) in a distinct geographically diverse cohort?.

Key Finding
The ARVC risk calculator:
1) Predicted adequately with good discrimination and calibration.
2) Performed better than other consensus-based implantable cardioverter defibrillator (ICD) implantation algorithms.

Take Home Message
The ARVC risk calculator provides reliable information and can facilitate shared decision-making regarding ICD implantation in primary prevention ARVC.



Validation of the arrhythmogenic right ventricular cardiomyopathy (ARVC) risk calculator in a distinct cohort. AHA, American Heart Association; ECG, electrocardiogram; HRS, Heart Rhythm Society; ICD, implantable cardioverter-defibrillator; ITFC, International Task Force Criteria; NSVT, non-sustained ventricular tachycardia; PVC, premature ventricular complex; VA, ventricular arrhythmia.

Keywords Arrhythmogenic right ventricular cardiomyopathy • Implantable cardioverter-defibrillator • Sudden cardiac death • Ventricular arrhythmias • Risk stratification • Genetic cardiomyopathies

Introduction

Arrhythmogenic right ventricular cardiomyopathy (ARVC) is a significant cause of sustained ventricular arrhythmia (VA) and sudden cardiac death (SCD), especially in young individuals and athletes. Preventing this catastrophic outcome through the prophylactic use of implantable cardioverter-defibrillators (ICDs) is a cornerstone of the disease management. Given the significant drawbacks associated with ICDs in this young and active population, appropriate patient selection is essential.

Over the past 25 years, numerous studies have identified predictors of sustained VA and SCD in ARVC and consensus documents have integrated these in decision algorithms for ICD use.[1–3] Building on this knowledge, a risk prediction model for sustained VA and SCD in ARVC was recently developed in a multinational cohort ($n = 528$, designed as the derivation cohort) mostly including high volume referral centres for ARVC.[4] This prediction model provides individualized prediction of the risk of VA in patients with ARVC without a prior history of sustained VA. Since its online publication, the risk calculator's official site (http://www.ARVCrisk.com) has been used ~20 000 times illustrating its uptake in clinical practice.

The model has been internally and externally validated in small studies.[4–9] However, adequately powered external validation is still lacking,[10] yet is paramount to confirm the reproducibility, generalizability, and need to update the model in an independent population.

The aims of the present study are thus (i) to conduct external validation of the published risk calculator in a distinct, adequately powered, and geographically diverse cohort including patients from six countries across North America and Europe and (ii) to compare the performance of the risk prediction model with other published guidelines and expert consensus recommendations for ICD use. During the current validation study, our group detected an inaccuracy in the formula of the original ARVC risk calculator published in 2019. It was corrected both on the website (ARVCrisk.com) and in the published manuscript.[11] We base the present study on the corrected risk calculator.

Methods

Study design

We conducted an observational, retrospective, longitudinal cohort study in accordance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement.[12]

Study population

The study population was derived from 29 centres (see Supplementary material online, *Table S1*) in six European and North American countries. This current cohort will be designated as the 'validation cohort' while the cohort leading to the published model will be designated as the 'derivation cohort'. New patients from two centres participating in the original study (Montreal Heart Institute and Johns Hopkins Hospital) were included (52 patients; 12% of the cohort). No patients in the current cohort were included in the original ARVC derivation cohort. From each site, consistent with the derivation cohort, consecutive patients who (i) were diagnosed with definite ARVC as per 2010 Task Force Criteria (TFC),[13] (ii) were alive at presentation, and (iii) had not experienced spontaneous sustained VA or sudden cardiac arrest (SCA) at diagnosis were included. The study conforms to the Helsinki declaration and was approved by local ethics and/or institutional review boards. To maintain patient confidentiality, data, and study materials will not be made available to other researchers for purposes of replicating the results. A limited dataset may be made available upon request.

Data collection

Data were collected independently by each of the participating centres using uniform definitions. A complete list of variables and their definitions can be found in Supplementary material online, *Table S2*. Genetic variants were reviewed according to the American College of Medical Genetics and Genomics guidelines by cardiologists specialized in cardiovascular genetics (R.T. and J.C.T.).[14]

Missing data

Patients with >50% of predictors missing were excluded from the analysis. Missingness was assumed to be at random and imputed using multiple imputation by chained equations.[15] Missing quantitative values for right ventricular ejection fraction (RVEF) and left ventricular ejection fraction (LVEF) were imputed manually when only qualitative assessment was available as done previously[4] and detailed in Supplementary material online, *Table S2*. The multiple imputation model included all pre-specified predictors, proband status and genotype together with the outcome, and cumulative baseline hazard estimation.[16] A total of 25 imputed datasets were generated, and the final inference estimations were combined using Rubin's rules.[17]

Study outcomes

In accordance with the published ARVC risk prediction model which this study aims to validate, the primary outcome was the first sustained VA following the definite diagnosis as per the TFC. Sustained VA was defined as a composite of the occurrence of SCD, SCA, spontaneous sustained ventricular tachycardia (VT; lasting ≥30 s at ≥100 b.p.m. or with haemodynamic compromise requiring cardioversion), ventricular fibrillation/flutter, or appropriate ICD intervention. Heart transplantation, cardiovascular mortality, and all-cause mortality were also collected.

Predictor variables and risk calculator

The same candidate predictors as those selected in the published model based on prior literature were considered.[18–20] These include sex, age, recent cardiac syncope (here defined as transient loss of consciousness and postural tone with spontaneous recovery with a likely arrhythmic mechanism, within a year of diagnosis), non-sustained VT (NSVT: defined as hemodynamically stable VT at ≥100 b.p.m., for ≥3 beats <30 s), number of premature ventricular complexes (PVCs) on 24-h Holter monitoring, the extent of T-wave inversion (TWI) on anterior and inferior leads, and RVEF. Each predictor variable was determined at the time of

diagnosis, defined as 1 year before to 1 year after the date of diagnosis as per 2010 TFC and prior to the occurrence of the primary outcome.

The 5-year risk of sustained VA for an individual patient as per the published model is calculated using the following equation[4]:

$$P(\text{VA at 5 years}) = 1 - 0.8396^{\exp(\text{LP})}$$

where the linear predictor (LP) was calculated according to the equation:

$$\begin{aligned}\text{LP} = {} & 0.488 \times \text{male sex} - 0.022 \times \text{age} + 0.657 \\ & \times \text{history of recent cardiac syncope} + 0.811 \times \text{history of NSVT} \\ & + 0.170 \times \ln(\text{24-h PVC count}) + 0.113 \\ & \times \text{sum of anterior and inferior leads with TWI} - 0.025 \\ & \times \text{RVEF}.\end{aligned}$$

Of note, the baseline hazard for 5-year prediction (0.8396) has been corrected since the initial publication in 2019.[11]

Statistical analysis

Analyses were performed with RStudio version 1.3.1093 (Boston, MA, USA). Continuous variables were expressed as mean ± standard deviation or median [interquartile range (IQR)] and compared using either the independent sample *t*-test or the Mann–Whitney *U* test. Categorical variables were presented as frequencies (%) and compared using the Fisher's exact test. Follow-up duration was calculated as the time interval between the time of definite diagnosis according to TFC and the endpoint or censoring. Censoring was defined as death from any other cause, heart transplantation or the most recent follow-up visit at which the endpoint could be ascertained. Event-free survival probability was estimated using the Kaplan–Meier method and Cox proportional hazard regression analysis.

Model validation

The approach to external validation follows the method suggested by Royston and Altman for Cox prognostic models.[21] First, the overall discriminative performance of the model was measured using Harrell's C-statistic, and the model fit by calculating the calibration slope, the regression of the LP (i.e. the product of the variable part of the Cox model) in the current cohort (validation cohort). Graphical evaluation of calibration was performed by plotting the predicted risk against the observed risk of sustained VA, using grouped Kaplan–Meier estimates and the continuous hazard regression function. The choice of the number of groups presented was based on the balance between providing sufficient spread in group risk, while maintaining adequate group sizes for precision. For the complete cohort, five groups are presented while for subgroup analyses, four groups are presented.

Subsequently, a more in-depth analysis of the model fit was performed by a Cox's model including the same predictor variables in combination with the LP of the original model (as an offset variable) to evaluate potential differences in the regression coefficients of each individual predictor. The result indicating the validity of the model would be that if all coefficients β^* equalled 0, reflecting that all the variability in the validation sample is accounted for by the published model. In addition, the baseline survival function of the validation dataset was compared to that of the derivation dataset to see if the overall predictions need to be globally shifted upward or downward. Lastly, a new prediction model using the same predictor variables was fitted to the validation dataset and compared to the fit of the original model using the Akaike information criterion (AIC), with a difference of >2 defined as statistically significant. This allows testing whether a model specifically fitted to the validation dataset performs better than the original model in the validation dataset.

Subgroup analyses

We visually explored the performance of the model specifically in different populations of interest by comparing calibration plots for these subgroups. We stratified the cohort by geographic origin (Europe vs. North America), by proband status and by plakophilin 2 carrier status (*PKP2*; causal variant carrier vs. non-carrier). We did not report quantitative markers of performance such as the C-statistic as this study was not powered adequately for these subgroups.

To assess the impact of carrying an ICD on prediction accuracy, we also presented calibration plots based on ICD carrier status at baseline defined as ICD implantation prior to a year following diagnosis and first VA outcome, whichever came first.

Clinical utility

To assess the relative clinical utility of the risk prediction model, it was compared to three other published expert consensus algorithms for ICD implantation in ARVC: the 2015 International Task Force Consensus for the treatment of ARVC (ITFC),[18] the 2017 American Heart Association (AHA) guidelines for the management of VA and prevention of SCD,[2] and the 2019 Heart Rhythm Society (HRS) consensus on arrhythmogenic cardiomyopathy (excluding programmed ventricular stimulation)[22] through decision curve analysis. In a decision curve analysis,[23] the clinical benefit is assessed by the 'net benefit' representing the balance between useful (i.e. in patients with events) vs. useless (i.e. in patients without events) ICD placement at 5 years weighted according to the threshold used for ICD implantation. More specifically, the decision curve uses the following formula:

$$\begin{aligned}&\text{True positives/total sample size} \\ &\quad - \text{false positives/total sample size} \times (\text{pt}/1 - \text{pt})\end{aligned}$$

where 'pt' represents threshold probability, in the current case, threshold for ICD implantation. Therefore, the higher the threshold used, the greater the harm of useless ICD use (i.e. false positive) is valued. Higher values indicate greater benefit while a value of 0 indicates no benefit.

To present the consequence of setting different thresholds for ICD implantation, we evaluated and plotted the proportion of patients who would receive ICDs and the proportion of treated and missed events at each threshold. We compared these with the recommendations for ICD use by the three published consensus mentioned above [ITFC(1), AHA (2), HRS(3)].

Results

The study population included 429 definite ARVC patients without a history of sustained VA or SCA at the time of diagnosis aged 43.1 ± 15.8 years and slightly more than half ($n = 235$, 54.8%) were male. Probands accounted for two-thirds of the cohort ($n = 278$, 64.8%). Half ($n = 198$, 46.6%) of patients had a pathogenic or likely pathogenic variant in a gene with definite or moderate association with ARVC,[24] which represents 70% (198 patients) of the 282 patients for whom the complete genetic information was available. *PKP2* was the most common genotype, carried by 111 patients (26%) followed by DSP in 38 patients (9%). Compared to PKP2 patients, DSP patients were more likely to have a decrease in LVEF <

Table 1 Baseline clinical characteristics

	Overall ($n = 429$)	Non-sustained VA ($n = 326$)	Sustained VA ($n = 103$)	P-value
Demographics and genetics				
Age at diagnosis (years)	43.1 ± 15.8	44.1 ± 15.7	40.1 ± 16.0	0.025
Male sex	235 (54.8)	159 (48.8)	76 (73.8)	<0.001
Proband status	278 (64.8)	197 (60.4)	81 (78.6)	0.001
(Likely) pathogenic variants ($n = 282$)	198 (46.2)	150 (46.0)	48 (46.6)	0.480
Genotype				0.302
PKP2	111 (25.6)	84 (25.8)	27 (26.2)	
DSP	38 (8.9)	33 (10.1)	5 (4.9)	
DSG2	27 (6.3)	22 (6.7)	5 (4.9)	
DSC2	3 (0.7)	1 (0.3)	2 (1.9)	
JUP	0 (0.0)	0 (0.0)	0 (0.0)	
TMEM43	10 (2.3)	4 (1.2)	6 (5.8)	
PLN	3 (0.7)	3 (0.9)	0 (0.0)	
Multiple mutations	6 (1.4)	3 (0.9)	3 (2.9)	
Clinical history				
Recent cardiac syncope ($n = 424$)	37 (8.6)	16 (4.9)	21 (20.4)	<0.001
ECG/continuous ECG monitoring				
TWI in ≥3 precordial leads ($n = 409$)	250 (58.3)	187 (57.4)	63 (61.2)	0.295
TWI in ≥2 inferior leads ($n = 403$)	109 (25.4)	81 (24.8)	28 (27.2)	0.589
PVC count ($n = 324$)	1434 (439–3601)	1354 (400–3719)	1676 (602–3492)	0.160
NSVT ($n = 359$)	148 (34.5)	105 (32.2)	43 (41.7)	0.001
Imaging				
RVEF (%) ($n = 410$)	45 (36–53)	47 (38–53)	40 (35–48.5)	<0.001
LVEF (%) ($n = 404$)	57 (51–60)	57 (51–61)	57 (50–60)	0.049
Treatment at baseline				
ICD	175 (40.8)	113 (34.7)	62 (60.2)	<0.001
Anti-arrhythmic drugs ($n = 408$)				0.041
Amiodarone	23 (6.0)	16 (4.9)	10 (9.8)	
Sotalol	79 (18.4)	55 (16.9)	24 (23.3)	
Propafenone/flecainide	15 (3.5)	9 (2.8)	6 (5.8)	
β-blockers ($n = 407$)	206 (48.0)	156 (47.9)	50 (48.5)	0.50
Follow-up	5.02 (2.05–7.90)	4.48 (1.86–7.32)	6.12 (2.60–10.08)	0.002

Variables are expressed as frequency (%), mean ± standard deviation, or median (interquartile range). Total number of patients with available data for a given variable are mentioned in parenthesis for variables with missing data.
DSC2, desmocollin-2; *DSG2*, desmoglein-2; *DSP*, desmoplakin; ICD, implantable cardioverter-defibrillator; LVEF, left ventricular ejection fraction; NSVT, non-sustained ventricular tachycardia; *JUP*, junction plakoglobin; *PKP2*, plakophilin-2; *PLN*, phospholamban; PVC, premature ventricular complex; RVEF, right ventricular ejection fraction; *TMEM43*, transmembrane protein 43; TWI, T-wave inversion; VA, ventricular arrhythmia.

50% (44.7% vs. 6.4%) but less likely to have VA events at follow-up (13.2% vs. 24.5%). Baseline characteristics according to genotype are presented in Supplementary material online, *Table S3*. Other clinical and demographic characteristics are summarized in *Table 1*. Baseline characteristics by country of origin are presented in Supplementary material online, *Table S4*, and a comparison of the derivation and validation cohort populations is presented in Supplementary material online, *Table S5*.

Overall, 299 (70.0%) patients had complete data for the pre-specified predictors. Six of the eight predictors had missing data:

recent cardiac syncope ($n = 5$, 1.17%), NSVT (70 = 16.32%), PVC count ($n = 105$, 24.48%), extent of leads with TWI ($n = 26$, 6.06%) and RVEF ($n = 19$, 4.43%). From an initial cohort of 433 patients, four patients were excluded as >50% of their predictors were missing (four predictors or more).

Outcomes

During a median follow-up of [5.02 (2.05–7.90)] years, 103 patients (24%) experienced sustained VA events corresponding to an annual event rate of 4.98% [95% confidence interval (CI) 4.07–6.04]. *Figure 1* shows the cumulative survival free from first sustained VA.

Among patients who experienced sustained VA during follow-up, the most common events were ICD treated VAs, which represented 59.2% of events ($n = 61$), followed by sustained VT ($n = 32$, 31.1%), SCA ($n = 7$, 6.8%), and SCD ($n = 3$, 2.9%). In patients with sustained or ICD treated VT events, the median cycle length (available in 57/93 events) was 280 ms (IQR: 246–315) which corresponds to 214 b.p.m. (190–243).

At last follow-up, 9 (2.1%) patients had died, including 2 from non-cardiac causes, and 7 (1.6%) had undergone heart transplantation.

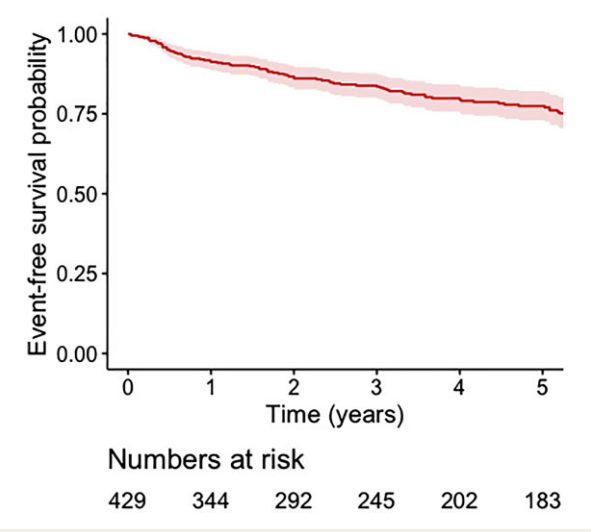


Figure 1 Survival free from sustained ventricular arrhythmia at follow-up. The cumulative event-free survival for any ventricular arrhythmia with 95% confidence intervals (shaded area) is plotted.

External validation

Model validation revealed a Harrell C-index of 0.70 (95% CI 0.65–0.75). The calibration slope was 1.01 (95% CI 0.99–1.03) showing no significant difference in discrimination. The calibration of the model is graphically presented in *Figure 2* demonstrating good overall agreement between predicted and observed shorter-term (1 year) and longer-term durations (5 year) with no significant over or under prediction across the complete risk spectrum. The distribution of patients according to their risk is presented in Supplementary material online, *Figure S1* and calibration plots for intermediate durations (1, 2, 3, and 5 years) in Supplementary material online, *Figure S2*.

Two different aspects of the model fit or potential misspecification were evaluated. First, the assessment of individual predictor coefficients (*Figure 3A*) all showed no significant diversion from the original model in this cohort. This finding means that none of the individual coefficient would benefit from being modified from their original values to improve prediction in this cohort.

Second, the baseline survival function (i.e. predictors-adjusted survival) was assessed through the comparison of the baseline survival

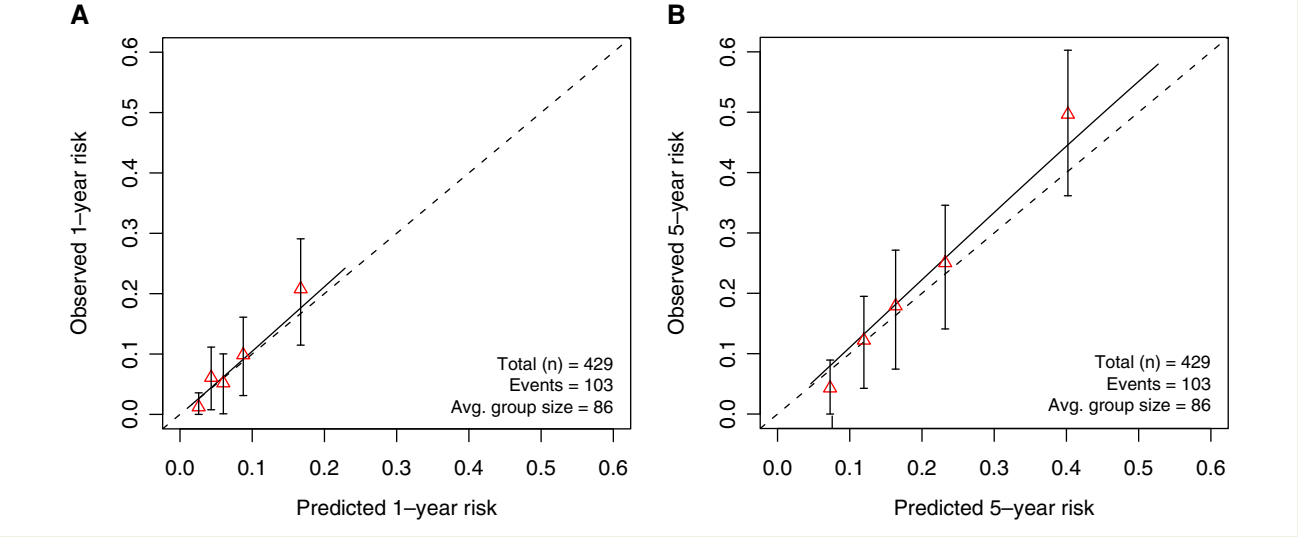


Figure 2 Calibration plots presenting the agreement between predicted (*x*-axis) and observed (*y*-axis) 1-year (Panel *A*) and 5-year (Panel *B*) risk of ventricular arrhythmia. Triangles represent binned Kaplan–Meier estimates with 95% confidence intervals for quintiles of predicted risk. The straight line is the continuous calibration hazard regression with the dotted line represents optimal calibration (i.e. perfect correspondence between predictions and observations across the risk spectrum). The calibration is shown to be acceptable across the risk spectrum with no significant under or over prediction in any risk category. VA, ventricular arrhythmia.

A

	Individual predictor coefficients	
Predictors	Coefficient	P value
Sex	0.305	0.188
Age	0.013	0.051
PVC count	-0.117	0.595
NSVT	-0.361	0.171
Leads with TWI	-0.062	0.156
Cardiac Syncope	0.321	0.225
RVEF	0.001	0.946

B

	Probability of survival	
Time	Derivation cohort	Validation cohort
1 year	0.938	0.938
2 years	0.901	0.899
3 years	0.880	0.880
4 years	0.870	0.850
5 years	0.840	0.831

C

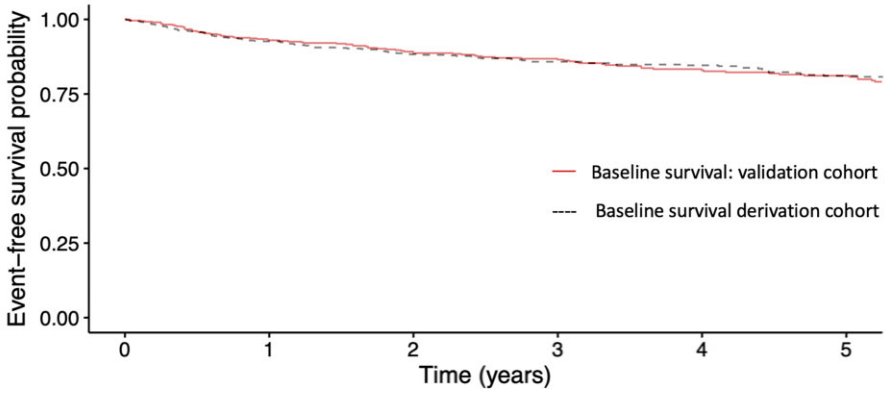


Figure 3 Assessment of the model fit. Assessment of the individual predictors (*A*) show an absence of diversion from the initial model as all coefficients are non-significantly different from 0. Compared survival probability of the derivation and validation cohorts (*B*) and baseline survival hazard (i.e. predictors-adjusted survival) presented as survival curves (*C*) both show similar expected survival. NSVT, non-sustained ventricular tachycardia; PVC, premature ventricular complex; TWI, T-wave inversion; RVEF, right ventricular ejection fraction.

probabilities (i.e. predictors-adjusted survival) in the derivation and the validation cohorts at different time points showing similar expected survival curves as shown numerically and visually in *Figure 3B* and *C*. These findings suggest that the survival function does not need to be modified to improve prediction in this cohort.

Finally, the potential need to update the model was assessed by comparing the fit of the published model with the derivation of a new model in the validation cohort. The AIC of the published model in the current cohort (1059.14) and of a model derived in this cohort (1060.93) were not significantly different (absolute difference in AIC of 1.79) indicating the absence of significant improvement in predictions when fitting a model to this population.

As a sensitivity analysis, we repeated the process in patients with complete data ($n = 299$) resulting in a similar C-statistic, calibration slope, baseline risk, and calibration plot (see Supplementary material online, *Figure S3*).

Clinical utility

We compared the performance of the risk calculator with published consensus-based decision algorithms for ICD use in ARVC. As illustrated in *Figure 4*, the risk calculator generally had a superior net clinical benefit when compared to the other published algorithms for ICD use. Its performance becomes similar to the HRS consensus above a risk of ~35%.

Finally, we graphically presented the impact of different threshold for ICD implantation on the proportion of ICD use and the protection rate and compared to the published decision algorithms (*Figure 5*). Higher thresholds result in less ICD use but less protection from VA. As an example, a threshold of 15% would results in implanting 59.4% of patients with ICDs while protecting 85.7% of patients with incident VA events.

Subgroups analyses

The performance in subgroups of interest was visually explored by calibration plots presented in Supplementary material online, *Figure S4*. This cohort was not sufficiently powered to provide definite answers in these subgroups.

Calibration appeared acceptable in patients from both Europe and North America, although this analysis had low precision in the North American population due to its smaller size.

The model performed well both in probands and family members with a possible trend toward overestimation in family members in the lower risk spectrum. The calibration was also visually acceptable both in *PKP2* carriers and non-carriers.

Calibration plots according to the presence of an ICD show an acceptable agreement between predictions and observations with a tendency towards overestimation in non-ICD carriers and underestimation in ICD carriers in the higher risk spectrum (see Supplementary material online, *Figure S5*).

Discussion

In this study, we validated the published ARVC risk calculator in an independent cohort of patients from 29 centres in 6 countries in North America and Europe. Since its publication in 2019, the risk calculator had a significant uptake in clinical practice. Ensuring its

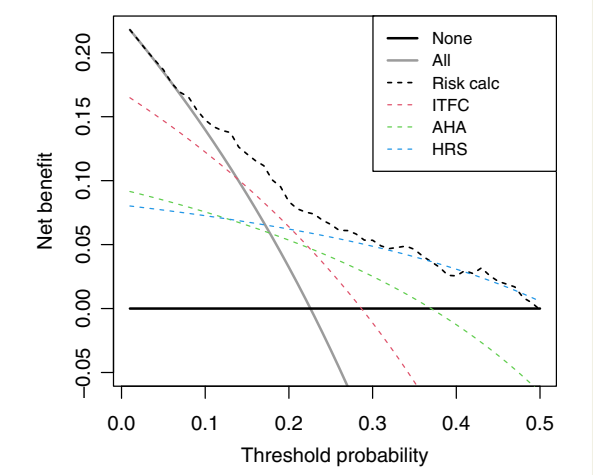


Figure 4 Decision curve analysis comparing the clinical utility of our model (dashed thick black line) with the 2015 International Task Force Consensus Statement algorithm for the treatment of arrhythmogenic right ventricular cardiomyopathy (dashed red line), the 2017 American Heart Association algorithm for the management of ventricular arrhythmia and prevention of sudden cardiac death (dashed green line) and the 2019 Heart Rhythm Society consensus on arrhythmogenic cardiomyopathy with exclusion of the Programmed ventricular stimulation (dashed blue line). The clinical utility of each treatment strategy is compared by plotting the net benefit (*y*-axis) for a range of possible implantable cardioverter-defibrillator placement thresholds based on the 5-year risk of ventricular arrhythmia (*x*-axis). Higher net benefit values indicate greater benefit while a value of 0 indicates no benefit. The published risk calculator depicted a better net benefit than the other published algorithms for implantable cardioverter-defibrillator implantation thresholds below a 35%. Above this threshold its performance was similar to the Heart Rhythm Society consensus algorithm. ICD, implantable cardioverter-defibrillator; ARVC, arrhythmogenic right ventricular cardiomyopathy, VA ventricular arrhythmia, SCD sudden cardiac death.

reproducibility and accuracy in an independent patient population is crucial to ensure both usefulness and safety.

The main findings are as follows:

(1) Demonstration that the model is accurate in its predictions with an adequate discrimination and calibration in a cohort with a sufficient sample size.[10,25] The performance of the risk calculator was indeed comparable to what was reported initially and its prediction accuracy in this cohort would not be improved by recalibration.[21]
(2) Demonstration that the risk calculator generally outperforms various risk factor approaches recommended in published consensus-based algorithms for ICD use in ARVC.

These findings thus support the clinical use of this risk prediction model as a valuable tool for sustained VA and SCD risk stratification in definite ARVC and, consequently, for guiding decisions about primary prevention ICD indications (*Structured Graphical abstract*).

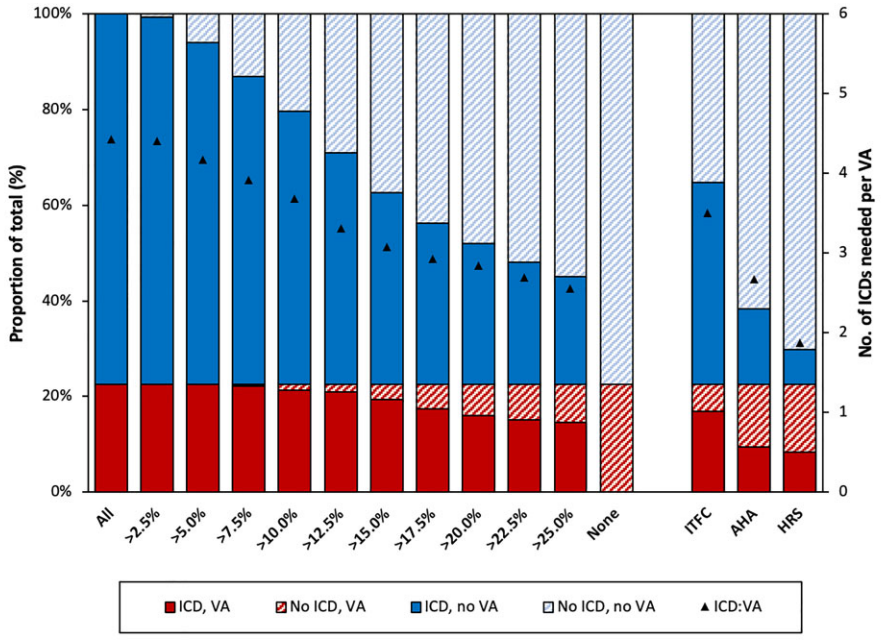
Comparison of the internal and external validation populations

While based on the same inclusion criteria (i.e. definite diagnosis of ARVC and no prior history of sustained VA at the time of diagnosis), the initial risk calculator included a high proportion of patients treated at highly specialized ARVC referral centres. Thus, a significant concern regarding this population is a possible selection bias due to the preferential referral of patients for adverse disease progression (i.e. recurrent VA referred for ablation and severe heart failure for advanced therapies). This could potentially hamper external validity. The present cohort derived from 29 different centres in 6 countries is thus likely to reflect a more diverse ARVC population. Expectedly, the annual event rate in this validation cohort (4.98%, 95% CI 4.07–6.04) was slightly lower, although non-significantly, than in the derivation population (5.6%, 95% CI 4.7–6.6) during a similar follow-up period [5.02 (2.05–7.90) years in the validation versus 4.83 (2.44–9.33) years in the derivation cohort]. This reflects the overall high risk of VA events in definite ARVC patients such as those included in this study which is consistent with prior literature and often preceding structural changes.[4,19,25–28]

Some differences between the two cohorts (shown in Supplementary material online, *Table S4*) might have limited the potential discrepancy in event rates, such as a higher proportion of probands (64.8% vs. 49.8%, $P < 0.001$) and males (54.8% vs. 44.7%, $P = 0.002$) in the current cohort. Conversely, patients in the present cohort were slightly older (43.1 vs. 38.2 years of age, $P < 0.001$), had less recent cardiac syncope and NSVT. The proportion of patients with decreased LVEF (<50%) was also higher in this cohort (17.7 vs. 12.7, $P = 0.002$). Although individuals in the current population were more likely to receive anti-arrhythmic drugs ($P < 0.001$) and β-blockers (48.0 vs. 37.9, $P = 0.001$), the proportion of ICD carriers at baseline was similar (41.1 vs. 40.8 $P = 0.98$). Finally, while still representing the predominant genotype, the proportion of patients with *PKP2* causal variants was lower than in the derivation cohort (39.4% vs. 51.1% of tested patients) factoring that the current cohort has a lower proportion of patients with known genetic information. This predominance of PKP2 genotype is consistent with prior literature including patients with definite ARVC diagnosis.[29] The proportion of patients with DSP causal variants was also higher (8.9% vs. 4.4%) than in the derivation cohort.

Model performance

The current validation cohort included 429 patients, of whom 103 had events. This met the minimally recommended sample size of 100 patients with and 100 patients without events to attain sufficient power for external validation.[30] The initial study and internal validation using bootstrapping yielded an optimism corrected C-statistic of 0.77 (95% CI 0.73–0.81) and a calibration slope of 0.93 (95% CI 0.92–0.95). In the current study, we obtained comparable results with a slightly lower C-statistic of 0.70 (95% CI 0.65–0.75) showing acceptable discrimination and a calibration slope of 1.01 (95% CI 0.99–1.03) demonstrating almost perfect agreement between predictions and observations for sustained VA. As illustrated in the calibration plot, this concordance between observations and predictions was consistent across the risk spectrum



	VA	ICD	>2.5%	>5.0%	>7.5%	>10.0%	>12.5%	>15.0%	>17.5%	>20.0%	>25.0%	ITFC	AHA	HRS
True negative	-	-	0.7%	6.0%	13.1%	20.4%	29.0%	37.3%	43.8%	48.0%	54.9%	35.3%	61.7%	70.1%
False positive	-	+	76.7%	71.4%	64.4%	57.0%	48.4%	40.1%	33.6%	29.4%	22.5%	42.1%	15.8%	7.3%
False negative	+	-	0.0%	0.1%	0.5%	1.3%	1.6%	3.2%	5.1%	6.6%	8.1%	5.7%	13.2%	14.3%
True positive	+	+	22.6%	22.5%	22.1%	21.3%	21.0%	19.4%	17.4%	16.0%	14.5%	16.9%	9.4%	8.3%
ICD:VA			4.4	4.2	3.9	3.7	3.3	3.1	2.9	2.8	2.6	3.5	2.7	1.9
Sensitivity			100.0%	99.7%	98.0%	94.3%	93.1%	85.7%	77.2%	71.0%	64.3%	74.8%	41.7%	36.9%
Specificity			0.9%	7.7%	16.9%	26.3%	37.4%	48.2%	56.6%	62.0%	70.9%	45.6%	79.6%	90.6%
ICD proportion			99.3%	93.9%	86.5%	78.3%	69.5%	59.4%	51.0%	45.5%	37.1%	59.0%	25.2%	15.6%
Protection rate			100.0%	99.7%	98.0%	94.3%	93.1%	85.7%	77.2%	71.0%	64.3%	74.8%	41.7%	36.9%

Figure 5 Impact of implantable cardioverter-defibrillator use threshold on clinical outcomes. The potential impact of different thresholds for implantable cardioverter-defibrillator use according to the model is presented on the left side and the proportion of patients who would get an implantable cardioverter-defibrillator according to the different consensus statements is presented on the right side. For each threshold (*x*-axis) the proportion of patients (*y*-axis) who have events (red) who do not have events (blue), who would receive an implantable cardioverter-defibrillator (solid colours) or not receive one (hashed colours) are presented. The triangles represent the number of implantable cardioverter-defibrillator needed per event prevented for each threshold (right-sided *y*-axis). The numerical values are presented in the table below. Implantable cardioverter-defibrillator:ventricular arrhythmia, ratio of implantable cardioverter-defibrillator placements required to protect one patient developing ventricular arrhythmia; other abbreviations as in figure 4.

(*Figure 2*). Calibration in subgroups based on geographical origin, pedigree position, and genotype did not reveal major discrepancies although the study was not adequately powered to arrive at definitive conclusions in these subgroups.

The results of the current study are consistent with five small studies which have addressed the external validation of the ARVC risk calculator since its publication. The risk calculator was shown to perform well in patients with a definite diagnosis of ARVC[5,6,8] and regardless of their exercise status.[7] The validation study by Baudinaud *et al.*,[6] on a cohort of 115 patients, only 15 with VA events, of whom only one had an ICD at baseline, reported a C-statistic of 0.84 (CI 0.74–0.93) while reporting an overestimation of the risk in lower risk patients.

Clinical utility

The model generally showed a superior net clinical benefit when compared to a risk factor approach as recommended in the three published consensus documents.[2,18,22] The model was similarly shown to outperform the ITFC and HRS consensus in two separate cohorts.[5,6] These studies, however, suggested highly different thresholds for ICD implantation (10% and 37%), assuming an equal weight to unprotected VA and unnecessary ICDs. We did not present such an analysis as we do not propose that these adverse events are equivalent and rather preferred the use of the weighted analysis along with the graphical presentation of the clinical implications of the different threshold. The question of the threshold for ICD implantation is a legitimate concern when using the risk calculator.

Establishing a single perfect threshold is a delicate undertaking as every cut-off point comes with a trade-off between unnecessary ICDs with their potential complications versus the potential for unprotected SCA. The relative weight of these opposing undesirable events varies significantly from one individual to another. In the individualized decision-making process; however, a few points should be considered when reflecting on the threshold for ICD use. First, when tempted to use a similar threshold as suggested by the guidelines for the hypertrophic cardiomyopathy (HCM) risk calculator (i.e. ≥6% within 5 years),[31,32] the breakdown of the type of events is relevant. In ARVC cohorts, including the current study and in the derivation cohort, most events were either ICD treated events or sustained VA, while most events in the cohort leading to the HCM risk calculator cohort were SCD or SCA.[33] Although most clinicians agree that sustained or ICD treated VAs represent significant events, supported by guidelines,[2,34] the exact number of treated VA events corresponding to a potential SCD is unknown in ARVC. Another important aspect to consider is that none of these studies are prospective evaluations of the role of ICDs in SCD prevention. Such an undertaking would not be feasible in contemporary high-risk ARVC populations. However, from such prior studies in the general cardiomyopathy population the one which established a benefit for primary prevention ICDs with the lowest annual risk of mortality, SCD-HeFT, had an annual risk of SCD of 3.5%.[35] Finally, the cost of ICDs is rarely a significant determinant nowadays in countries where ICDs can be considered in primary prevention.[36] Factoring the low number of ICDs needed to treat one VA event in ARVC, decreases in the cost of devices, the lifespan of modern ICDs reaching 10 years, and the potential number of quality-adjusted live years (QALY) saved in this young, usually otherwise healthy population (only five individuals had non-arrhythmic death during follow-up in this cohort), the common, although debated thresholds for a QALY between 50 000 and 100 000 USD[37] remains far of reach. Conversely, the rate of short- and long-term complications of ICDs remain significant in ARVC patients (annual rate of complications of 4.2% and of inappropriate shocks of 3.9%),[38] and although subcutaneous-ICDs have become an appealing alternative, there is no evidence of a lesser risk.[39,40]

Thus, in light of these different considerations, we do believe that the best use of the risk calculator is as a shared decision making tool balancing the opposing risks of SCD and ICD use. It appears reasonable that the predicted 5-year risk threshold for recommending an ICD would range from 5% to 25%, depending on the patient's values and preferences, and the clinician's judgement. We acknowledge that the threshold may change in the future with advances in non-invasive treatments and innovations in ICD technology which may lower risks associated with devices.

Future improvements in the model

While the model demonstrated a better performance compared to other published decision algorithms, it remains imperfect as illustrated by a C-statistic of 0.70. While it is unlikely that any risk stratification tool for SCD could predict the totality of these events, different elements could potentially improve prediction in the future. The addition of more refined parameters indicating left ventricular involvement, including late gadolinium enhancement were recently suggested.[9] Genotype may also improve SCD risk prediction as recently proposed for patients with phospholamban associated disease.[41] Finally, additional invasive parameters such as programmed ventricular stimulation[42,43] might add additional accuracy in intermediate-risk cases. Moreover, the model is based on prediction of risk from the time of diagnosis of ARVC; a time-updated model for repeated risk prediction may have practical clinical utility.

Limitations

In this study, the majority of sustained VA outcomes are ICD treated events. While this fact is not possible to overcome in most modern ARVC populations and while most would agree that these still represent significant events, they do not directly represent the underlying risk of SCD. However, this is a limitation shared with most of previous studies in this field, including most of those used to elaborate prior consensus-based risk stratification algorithms. While underpowered for events, calibration plots by ICD carrier status show acceptable correlation between predictions and observations. This reflects both that ICDs are implanted in patients believed to be at higher risk (selection bias), but also increase the detection of some arrhythmia that might have gone undetected otherwise (information bias) (see Supplementary material online, *Figure S4*). While family members are well represented in the derivation cohort (50.2%), they are less prevalent in the current cohort (35.2%), and contribute to a lower proportion of events (21.1%). The calibration plot in this specific subgroup, although underpowered, suggests possible overestimation in the lower risk patients which should be taken in consideration when using the model.

Missing data also represent a limitation of this retrospective cohort. Although a complete case analysis reassuringly demonstrates similar results with regard to performance, missing data could influence the relative benefit of the model over consensus-based methods.

Finally, this validation only applies to patients who were well represented in the derivation and validation cohorts. The model should thus not be used in patients who do not meet definite ARVC diagnosis as per 2010 TFC such as those with left dominant forms and in patients with rare malignant genotypes such as *TMEM43*-p.S358L, of which only 10 patients were included in this cohort.

Conclusion

In this external validation study, we demonstrated that the published ARVC risk prediction model not only provides accurate prognostic information in patients with ARVC without a prior history of sustained VA at diagnosis, but also performs generally better than other published decision algorithms. These findings support its clinical use as a valuable tool for risk stratification enabling consistent and effective shared decision making for ICD implantation.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Acknowledgements

We thank the ARVC patients and families who have made this work possible.

Funding

P.J. is supported by the Daniel Bravo Foundation grant and Spanish Society of Cardiology Magda Heras mobility grant, A.G. by the Wilton W. Webster Fellow of Heart Rhythm Society, The Johns Hopkins ARVD/C Programme by the Leonie-Wild Foundation, Leyla Erkan Family Fund for ARVD Research, The Hugh Calkins, Marvin H. Weiner, and Jacqueline J. Bernstein Cardiac Arrhythmia Center, Dr Francis P. Chiramonte Private Foundation, Dr Satish, Rupal, and Robin Shah ARVD Fund at Johns Hopkins, Bogle Foundation, Campanella Family, Patrick J. Harrison Family, Peter French Memorial Foundation, and Wilmerding Endowments, H.K.J. by grants from Novo Nordisk Foundation, Denmark (NNF18OC0031258 and NNF20OC0065151), S.A.L. by NIH grant 1R01HL139731 and American Heart Association 18SFRN34250007, R.T. by the Canada Research Chairs programme, J.C.T. by the Philippa and Marvin Carsley Cardiology Research Chair. S.G.P. receives support from Ricerca Corrente funding scheme of the Italian Ministry of Health and Italian Ministry of Research and University Dipartimenti di Eccellenza 2018 to 2022 grant to the Molecular Medicine Department (University of Pavia).

Conflicts of interest: C.M.: honoraria from Abbott, I.O.: grants, consulting fees or honoraria from BSM, Cytokinetics, Shire, Genzyme, Amicus, Menarini International, Boston Scientific, and Tenaya, S.A.L.: grants from BMS/Pfizer, Boehringer Ingelheim, fitbit, IBM, and consulting fees from BMS/Pfizer, Invitae, and Blackstone, J.S.H.: research grants from Boston Scientific, Abbott, and Medtronic and is on Scientific Advisory Board for Boston Scientific, H.K.J.: grant from Novo Nordisk foundation and honoraria from Abbott and Biosense Webster, H.C.: consultant for Medtronic Inc., Biosense Webster, Pfizer, and Abbott. He receives research support from Boston Scientific Corp. C.A.J. receives salary support from this grant and consulting fees from Pfizer, J. C.-T. consulting fees from BMS/Pfizer and Bayer.

References

1. Corrado D, Wichter T, Link MS, Hauer RN, Marchlinski FE, Anastasakis A *et al.* Treatment of arrhythmogenic right ventricular cardiomyopathy/dysplasia: an international task force consensus statement. *Circulation* 2015; **132**: 441–453.
2. Al-Khatib SM, Stevenson WG, Ackerman MJ, Bryant WJ, Callans DJ, Curtis AB *et al.* 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2018; **72**: e91–e220.
3. Towbin JA, McKenna WJ, Abrams DJ, Ackerman MJ, Calkins H, Darrieux FCC *et al.* 2019 HRS expert consensus statement on evaluation, risk stratification, and management of arrhythmogenic cardiomyopathy. *Heart Rhythm* 2019; **16**: e301–e372.
4. Cadrin-Tourigny J, Bosman LP, Nozza A,Wang W, Tadros R, Bhonsale A *et al.* A new prediction model for ventricular arrhythmias in arrhythmogenic right ventricular cardiomyopathy. *Eur Heart J* 2019; **40**: 1850–1858.
5. Aquaro GD, De Luca A, Cappelletto C, Raimondi F, Bianco F, Botto N, *et al.* Comparison of different prediction models for the indication of implanted cardioverter defibrillator in patients with arrhythmogenic right ventricular cardiomyopathy. *ESC Heart Fail* 2020; **7**: 4080–4088.
6. Baudinaud P, Laredo M, Badenco N, Rouanet S, Waintraub X, Duthoit G, *et al.* External validation of a risk prediction model for ventricular arrhythmias in arrhythmogenic right ventricular cardiomyopathy. *Can J Cardiol* 2021; **37**: 1263–1266.
7. Gasperetti A, Dello Russo A, Busana M, Dessanai M, Pizzamiglio F, Saguner AM, *et al.* Novel risk calculator performance in athletes with arrhythmogenic right ventricular cardiomyopathy. *Heart Rhythm* 2020; **17**: 1251–1259.
8. Casella M, Gasperetti A, Gaetano F, Busana M, Sommariva E, Catto V, *et al.* Long-term follow-up analysis of a highly characterized arrhythmogenic cardiomyopathy cohort with classical and non-classical phenotypes-a real-world assessment of a novel prediction model: does the subtype really matter. *Europace* 2020; **22**: 797–805.
9. Aquaro GD, De Luca A, Cappelletto C, Raimondi F, Bianco F, Botto N, *et al.* Prognostic value of magnetic resonance phenotype in patients with arrhythmogenic right ventricular cardiomyopathy. *J Am Coll Cardiol* 2020; **75**: 2753–2765.
10. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016;**35**: 214–226.
11. Corrigendum to: A new prediction model for ventricular arrhythmias in arrhythmogenic right ventricular cardiomyopathy. *Eur Heart J* 2022;**43**:2712.
12. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, *et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; **162**: W1–73.
13. Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, *et al.* Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. *Eur Heart J* 2010; **31**: 806–814.
14. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015; **17**: 405–424.
15. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999;**18**:681–694.
16. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;**30**:377–399.
17. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons, 1987.
18. Corrado D, Wichter T, Link MS, Hauer R, Marchlinski F, Anastasakis A, *et al.* Treatment of arrhythmogenic right ventricular cardiomyopathy/dysplasia: an international task force consensus statement. *Eur Heart J* 2015; **36**: 3227–3237.
19. Bosman LP, Sammani A, James CA, Cadrin-Tourigny J, Calkins H, van Tintelen JP, *et al.* Predicting arrhythmic risk in arrhythmogenic right ventricular cardiomyopathy: a systematic review and meta-analysis. *Heart Rhythm* 2018; **15**: 1097–1107.
20. Orgeron GM, Te Riele A, Tichnell C, Wang W, Murray B, Bhonsale A, *et al.* Performance of the 2015 International Task Force Consensus Statement risk stratification algorithm for implantable cardioverter-defibrillator placement in arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circ Arrhythm Electrophysiol* 2018; **11**: e005593.
21. Royston P, Altman DG. External validation of a cox prognostic model: principles and methods. *BMC Med Res Methodol* 2013; **13**: 33.
22. Towbin JA, McKenna WJ, Abrams DJ, Ackerman MJ, Calkins H, Darrieux FCC, *et al.* 2019 HRS expert consensus statement on evaluation, risk stratification, and management of arrhythmogenic cardiomyopathy. *Heart Rhythm* 2019; **16**: e301–e372.
23. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res* 2019; **3**: 18.
24. James CA, Jongbloed JDH, Hershberger RE, Morales A, Judge DP, Syrris P, *et al.* International evidence based reappraisal of genes associated with arrhythmogenic right ventricular cardiomyopathy using the clinical genome resource framework. *Circ Genom Precis Med* 2021; **14**: e003273.
25. Bhonsale A, James CA, Tichnell C, Murray B, Gagarin D, Philips B, *et al.* Incidence and predictors of implantable cardioverter-defibrillator therapy in patients with arrhythmogenic right ventricular dysplasia/cardiomyopathy undergoing implantable cardioverter-defibrillator implantation for primary prevention. *J Am Coll Cardiol* 2011; **58**: 1485–1496.
26. Mazzanti A, Ng K, Faragli A, Maragna R, Chiodaroli E, Orphanou N, *et al.* Arrhythmogenic right ventricular cardiomyopathy: clinical course and predictors of arrhythmic risk. *J Am Coll Cardiol* 2016; **68**: 2540–2550.
27. Battipaglia I, Scalone G, Macchione A, Pinnacchio G, Laurito M, Milo M, *et al.* Association of heart rate variability with arrhythmic events in patients with arrhythmogenic right ventricular cardiomyopathy/dysplasia. *Circ J* 2012; **76**: 618–623.
28. Santangeli P, Dello Russo A, Pieroni M, Casella M, Di Biase L, Burkhardt JD, *et al.* Fragmented and delayed electrograms within fibrofatty scar predict arrhythmic events in arrhythmogenic right ventricular cardiomyopathy: results from a prospective risk stratification study. *Heart Rhythm* 2012; **9**: 1200–1206.
29. Groeneweg JA, Bhonsale A, James CA, te Riele AS, Dooijes D, Tichnell C, *et al.* Clinical presentation, long-term follow-up, and outcomes of 1001 arrhythmogenic right ventricular dysplasia/cardiomyopathy patients and family members. *Circ Cardiovasc Genet* 2015; **8**: 437–446.
30. Vergouwe Y, Steyerberg EW, Eijkemans MJ *et al.* Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005; **58**: 475–483.
31. Authors/Task Force m, Elliott PM, Anastasakis A, Borger MA, Borggrefe M, Cecchi F, *et al.* 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy: the Task Force for the Diagnosis and Management of Hypertrophic

Cardiomyopathy of the European Society of Cardiology (ESC). *Eur Heart J* 2014; **35**: 2733–2779.
32. O'Mahony C, Jichi F, Pavlou M, Monserrat L, Anastasakis A, Rapezzi C, *et al.* A novel clinical risk prediction model for sudden cardiac death in hypertrophic cardiomyopathy (HCM risk-SCD). *Eur Heart J* 2014; **35**: 2010–2020.
33. Cadrin-Tourigny J, Tadros R. Predicting sudden cardiac death in genetic heart disease. *Can J Cardiol* 2022;**38**:479–490.
34. Priori SG, Blomstrom-Lundqvist C. European Society of Cardiology Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death summarized by co-chairs. *Eur Heart J* 2015: **36**(41):2757–2759.
35. Poole JE, Olshansky B, Mark DB, Anderson J, Johnson G, Hellkamp AS, *et al.* Long-term outcomes of implantable cardioverter-defibrillator therapy in the SCD-HeFT. *J Am Coll Cardiol* 2020; **76**: 405–415.
36. Smith T, Jordaens L, Theuns DA, van Dessel PF, Wilde AA, Hunink MG. The cost-effectiveness of primary prophylactic implantable defibrillator therapy in patients with ischaemic or non-ischaemic heart disease: a European analysis. *Eur Heart J* 2013; **34**: 211–219.
37. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness–the curious resilience of the $50,000-per-QALY threshold. *N Engl J Med* 2014;**371**:796–797.
38. Nordkamp LR O, Postema PG, Knops RE, van Dijk N, Limpens J, Wilde AA, *et al.* Implantable cardioverter-defibrillator harm in young patients with inherited arrhythmia syndromes: a systematic review and meta-analysis of inappropriate shocks and complications. *Heart Rhythm* 2016;**13**:443–454.
39. Knops RE, Olde Nordkamp LRA, Delnoy PHM, Boersma LVA, Kuschyk J, El-Chami MF, *et al.* Subcutaneous or transvenous defibrillator therapy. *N Engl J Med* 2020;**383**: 526–536.
40. Orgeron GM, Bhonsale A, Migliore F, James CA, Tichnell C, Murray B, *et al.* Subcutaneous implantable cardioverter-defibrillator in patients with arrhythmogenic right ventricular cardiomyopathy/dysplasia: a transatlantic experience. *J Am Heart Assoc* 2018; **7**: e008782.
41. Verstraelen TE, van Lint FHM, Bosman LP, de Brouwer R, Proost VM, Abeln BGS, *et al.* Prediction of ventricular arrhythmia in phospholamban p.Arg14del mutation carriers-reaching the frontiers of individual risk prediction. *Eur Heart J* 2021; **42**: 2842–2850.
42. McKenna WJ, Asaad NA, Jacoby DL. Prediction of ventricular arrhythmia and sudden death in arrhythmogenic right ventricular cardiomyopathy. *Eur Heart J* 2019; **40**:1859–1861.
43. Saguner AM, Medeiros-Domingo A, Schwyzer MA, On CJ, Haegeli LM, Wolber T, *et al.* Usefulness of inducible ventricular tachycardia to predict long-term adverse outcomes in arrhythmogenic right ventricular cardiomyopathy. *Am J Cardiol* 2013; **111**: 250–257.

Discussion

1. Main findings

The findings of this PhD thesis are: 1) Rare variants related to the pathogenesis of DCM, as such described in the first article of this PhD, *BAG3* p.Gln88*, can confer specific prognosis information and help in the management of the affected patients and their families. 2) The analysis of common genetic variants with post-GWAS tools such as the multi-trait analysis of GWAS (MTAG) improves SNP effect estimates from conventional GWAS. MTAG allowed the discovery of 19 novel disease-specific genomic loci for AF, 131 for CAD and 141 for HF. However, PGS derived from the MTAG results did not improve disease-status prediction and discrimination in these three common CVDs in the overall cohort, although it revealed improved prediction and classification of disease-status in CAD for females [ΔNagelkerke's $R^2_{CAD\text{-}F}$ 1.735% (95% CI: 0.609-2.856); $NRI_{CAD\text{-}F}$ 0.208 (95% CI: 0.139-0.277)]. 3) The ARVC risk calculator is a valid tool to predict the risk of VAs. This model had previously been described although in the development study, only internal validation using bootstrapping was used. In the third article of this PhD, we demonstrate the external validity of the previous results in an independent multicentric cohort stemming from 29 centers from Europe and North America and provide a comparison with the previously recommended algorithms for VA risk stratification in ARVC patients through decision curve analysis. The external validation showed adequate calibration and discrimination of the model [good agreement between predicted and observed risk in the calibration plots, C-index of 0.70 (95% CI: 0.65-0.75), calibration slope of 1.01 (95% CI: 0.99-1.03)] and a better performance of the ARVC risk score at an estimated risk of VAs below 35%.

In this thesis we offer a vision on different predictive variables, from genetic to clinical, which are valuable in the comprehension of the substrate and manifestations of CVDs and how predictive models considering these variables, individually and jointly integrated, are useful in the prognostic stratification of patients affected with these diseases.

2. Genotype-phenotype prognostic implications in dilated cardiomyopathy

In the first two articles of this PhD, we employ genetic profiling alone as a predictor of disease. In the first, we analyzed the value of a rare variant in the *BAG3* gene as responsible for a distinctive phenotype of DCM (definition of rare variant as MAF <1% in the general population; the variant we describe, *BAG3* p.Gln88*, has never been described in the general population). *BAG3* encodes a co-chaperone with an anti-apoptotic role, which has been associated with cardiomyopathy and myofibrillar myopathy. Pathogenic variants causing *BAG3*-cardiomyopathy have shown altered location of the protein in the nucleus and cytoplasm, disorganization of the myofibrillar structure and increased sensitivity to apoptosis of the cardiomyocyte. (125)

In our work we present a series of patients carrying the pathogenic variant *BAG3* p.Gln88* which causes a premature truncation of the encoded protein, where carriers exhibited a penetrance of 71%, and a progression to severe HF and cardiogenic shock of 47%. Acute HF onset occurred at very young ages, specifically in 2 individuals (15 and 17 years old). Another study had shown that rare truncating variants in the *BAG3* gene (*BAG* p.Arg309*) cause a form of DCM with disease onset at earlier ages than individuals carrying deleterious variants in other DCM-causing genes (P=0.037) although no differences in severe HF-related events prognosis (defined as heart transplantation, ventricular assist device or HF-related death at young age) related to *BAG3* carrier status were observed in the survival analysis (P=0.74). (126)

Interestingly, among the initial work supporting the deleterious role of common variants in the *BAG3* gene in the pathogenesis of DCM there is a GWAS that identified two significant loci associated with HF due to DCM. In that work, two SNPs were identified as independently associated with DCM: rs2234962 [c.T757C, p.Cys151Arg, OR 1.52 (95% CI: 1.22-1.89] and rs3858340 [c.1220C>T, p.Pro407Leu], both located in the *BAG3* gene sequence (10q25-26), only the first one reaching genome-wide significance. This discovery promoted the sequencing of the *BAG3* gene exons in DCM cases and controls, which led to the identification of 6 variants identified as deleterious, of which four caused an early truncation of the protein (p.Gln251Argfs*56, p.Arg396Glyfs*48, p.Ser385Glnfs*56, p.Arg309*) and the two others were missense variants (p.Glu455Lys,

p.Val468Met). Among the 18 carriers of these six variants, 72% were diagnosed with DCM. Among those affected, four needed a heart transplant and five relatives with no genotyping available but diagnosed with DCM died prematurely.(68)

Similarly, in the largest published series of individuals carrying pathogenic variants in the *BAG3* gene (*n*=129), a high penetrance of DCM (68.4%) and a high prevalence of severe HF were described, with 15.5% of patients presenting the combined endpoint of heart transplantation, LV assist device or death due to HF. The presence of truncating versus missense variant in the *BAG3* gene was not related to HF-related events (*P*=0.95), but male sex (*P*=0.001), and LV diameter together with LVEF at diagnosis showed association to HF-related events in the univariate analysis (*P*=0.01 and *P*<0.001, respectively). Interestingly, in this work they depict the immunohistochemistry of 3 patients with pathogenic truncating variants in the *BAG3* gene in which the protein displays a disorganized pattern in the Z-disc in the presence of myofibrillar disarray. (39)

Our work provides additional insight into the understanding of this form of the disease. Besides the abovementioned histopathological characterisation by Dominguez *et al.* (39), which is in line with previously described similar findings (125), we report the presence of acute myocardial damage (vacuolization of cardiomyocytes, distortion of myocardial fibers and interfibrillar edema and even focal myocyte necrosis), in some cases associated with inflammatory infiltrates as part of the pathogenesis (onset/progression) of this entity. Despite the presence of suggestive viral symptomatology in some individuals, only a low parvovirus B19 load was detected in the myocardial tissue polymerase chain reaction of two patients and IgM antibodies of cardiotropic virus were absent. These findings would support an apoptotic/immune-mediated component in genetically predisposed individuals in the pathogenesis of this disease and raises the question of whether therapy targeting these mechanisms could be useful in the treatment of *BAG3*-DCM. However, as this is a small series of few patients and without a control group, whether this presentation is typical of this variant, of variants in this same exon, of *BAG3* truncating variants or of deleterious variants in the *BAG3* gene in general, remains to be clarified by future work that includes characterization of the immunological and histopathological profile of *BAG3*-DCM.

3. Genetic loci discovery and genomic prediction in atrial fibrillation, coronary artery disease and heart failure

Beyond the genetic basis of cardiac diseases traditionally considered to be inherited, GWAS have contributed to the discovery of numerous genetic loci harboring common variation associated with complex cardiac diseases contributing to a deeper understanding of their genetic underpinnings. Since not all heritability of common heart diseases can be explained by significant SNPs, in the second article of this thesis we aimed to use MTAG to improve genome-wide SNP effect estimates and unmask relevant loci involved in the mechanisms of these entities.

In our study, MTAG allowed the discovery of novel loci associated with CVD. Overall, we report a total of 19, 131, and 141 loci in AF, CAD, and HF, uncovered by MTAG (and not identified by conventional GWAS), being a substantial proportion described for the first time. MTAG allows for the calculation of the maximum false discovery rate, which was <5% in all our MTAG analyses. A look-up of the novel associations in the GWAS Catalog, United Kingdom Biobank and FinnGen databases confirmed GWAS associations for 46% of the loci (either at a $P<5x10^{-8}$ in studies potentially including partially overlapping samples with those we employed or at a $P<0.005$ in the independent FinnGen cohort). Finally, MTAG methodology had been employed in other CVDs such as HCM, in which loci discovered by MTAG were completely replicated (55), altogether supporting reliability of our results.

The new loci we describe point to some interesting pathways related to the diseases to which they have been associated. Among those, the locus mapping serine racemase (*SRR*, lead SNP rs2209073, chromosome 17), has been reported for the first time in AF. SRR is an enzyme that produces D-serine, a co-agonist of NMDA (N-methyl-D-aspartate) receptors, which are involved in myriad functions of the central nervous system and have exhibited a role in reduced heart rate variability and enhanced AF inducibility in animal models. (127) In CAD, among the genes mapping new loci there is natriuretic peptide receptor 3 (*NPR3,* lead SNP rs1173727, chromosome 5), a receptor for natriuretic peptides that has been reported to have a protective role for coronary heart disease according to Mendelian randomization studies. (128) Likewise, among the genes mapped to new loci associated with HF there are numerous genes involved in the

cytoskeleton structure, an emerging target in the treatment of HF. (129) The loci reported in this work, together with their mapped genes and associated pathophysiological pathways, may be interesting departure sites for the investigation of new therapeutic targets in these diseases. However, in the absence of an independent replication of the current results, they need to be interpreted cautiously, specially for those loci that did not reach significance level at a nominal *P*-value ($P<0.05$) in the original GWAS.

The “gene-to-function-to-therapy" approach has proven successful in CVDs. To illustrate so, since the role of *PCSK9* (proprotein convertase subtilisin/kexin type 9) gene as a cause of autosomal dominant hypercholesterolemia was discovered in 2003 (130), several drugs have been developed that safely and effectively inhibit PCSK9. PCSK9 is a serine protease involved in the degradation of LDL receptors by preventing them from capturing LDL and clearing it from the circulation. Drugs specifically targeting this pathway, such as the anti-PCSK9 antibodies alirocumab and evolocumab, have been shown to reduce cholesterol levels by up to >50% and cardiovascular events by ~ 50%, in both individuals with familial hypercholesterolemia due to deleterious variants in *PCSK9* as well as in individuals with hypercholesterolemia not carrying variants in this gene. (131)

Another example of genetically guided therapy is the recently approved drug for HCM patients mavacamten. HCM has traditionally been considered an autosomal dominant disease. The first HCM-related gene, *MYH7*, which encodes the myosin heavy chain of the cardiac sarcomere (the contractile molecular machinery of the heart), was reported in 1989. (132) This group later demonstrated the increase in contractility as a mechanism underlying HCM in mice carrying missense mutations in *MYH7*, and that the compound MYK-461 (mavacamten), inhibited the development of HCM in mice carrying these variants through a reduction in ATPase activity. (133) Mutations in 11 genes have been associated with HCM (28), whose alterations lead to increased myocyte contractility and ATP utilization, resulting in a cascade of secondary molecular changes leading to the hallmark findings of HCM. (134) Subsequently, human clinical trials in obstructive HCM have demonstrated the efficacy of mavacamten through reduction of myocardial contractility and improvement of myocardial metabolism in patients with symptomatic

HCM, improving hemodynamic parameters and symptoms versus placebo (135-137), with a further positive impact on LV mass and wall thickness and left atrial volume (138), and being recommended as a second-line drug in the current therapeutic guidelines published this year 2023. (106)

These exemplify the impact of human genetic discoveries in improving patient care through a "gene-to-function-to-therapy" approach. Beyond mechanistical insight provided by rare variants, common variants associated with CVD may be helpful in pinpointing molecular therapeutic targets, as genetically linked treatment has demonstrated higher rate of success in treatment trials after internal review of pharmaceutical research and development companies. (139)

Regarding disease prediction, we observed that conventional GWAS-derived PGS were associated with disease in AF and CAD, but not in HF. MTAG-derived PGS were associated with all three diseases: AF, CAD, and HF. Multi-trait PGS yielded similar discrimination and prediction in CAD and AF, and non-conclusive results for HF, although they pointed to better performance. MTAG also revealed enhanced prediction and classification of disease-status in CAD for females, which might suggest that MTAG better captured disease mechanisms in this underrepresented group in CAD. The absence of improvement in the overall cohort was most likely driven by already well-powered GWAS with large sample size in these entities, and LDpred2 PGS method ability to prioritize and re-estimate SNP effects appropriately, altogether rendering relevant SNP effects in the prediction of the phenotype of interest adequately captured. Consequently, SNP-phenotype effects enhanced by MTAG, did not translate into improved prediction in the global cohort.

An important limitation of our work is that in order to calculate the LDSC genetic correlation through GWAS summary statistics, only predominantly European populations were included, therefore our results may not be transferable to other cohorts with different ancestry.

4. Arrhythmic risk stratification in arrhythmogenic right ventricle cardiomiopathy

In the third and last article of this PhD we extend our vision beyond predictive models integrating the whole genome to the analysis of predictive models that jointly integrate demographic and clinical variables in prognostic stratification.

This is the case of the model for the prediction of VAs in ARVC [https://arvcrisk.com/] (22). As previously mentioned, this model integrated variables that had previously been associated with the occurrence of VAs in ARVC and combined them through a Cox proportional-hazards model. VAs were defined as: SCD, sudden cardiac arrest, sustained VT persisting ≥30s at ≥100 beats per minute (bpm) or with haemodynamic compromise. In the derivation study (22), corresponding to a cohort of 528 patients with definite diagnosis of ARVC according to the 2010 Task Force criteria (114), the model was only validated internally using bootstrapping. In the third article of this thesis, we present the external validation of the ARVC risk calculator in an independent cohort of patients including 429 patients from 29 European and North American centers. In this cohort we show that: 1) The coefficients and survival curve of the model do not require a re-estimation; 2) the previously developed model has good discrimination [C-index 0.70 (95% CI: 0.65–0.75), calibration slope 1.01 (95% CI 0.99-1.03)] and is adequately calibrated in this new cohort, i.e., adequate agreement between predicted and observed risk, globally and in risk subgroups represented by quintiles in the calibration plots. 3) Finally, when comparing the ARVC risk calculator to the ITFC, mITFC, AHA and HRS algorithms using net clinical benefit, we depict that the risk calculator offers greater benefit for risk thresholds for ICD implantation below 35%, being equaled by the HRS algorithm from a risk of VA >35% and remaining superior to the other algorithms also in this risk range. Therefore, the ARVC risk score is proven to perform reliably.

The ARVC risk calculator has been validated in other cohorts in which it has also exhibited good discrimination [Uno's concordance index 0.75 (95% CI: 0.70–0.81) (140), C-index 0.84 (95% CI: 0.74-0.93)] (141) but also showing an overestimation of risk in some patient populations (140, 141). The risk calculator performs well, but caution applies in populations for which it has not been developed as showed in a cohort of ARVC patients meeting borderline diagnosis criteria (not definite), where the score only performed overall well and specifically well in individuals with lone RV involvement but

underestimated risk in those with biventricular or left dominant forms (142). Another cohort including patients with definite diagnosis of ARVC also showed that the risk calculator provides good prediction in patients with isolated RV involvement but underestimates the risk in patients with concomitant LV involvement, with an observed risk 28% higher (95% CI: 21% to 36%) than that calculated in these patients. The same study showed that LV involvement - defined as LVEF <50%, LV wall motion abnormalities, LV fatty infiltration or LGE-CMR with nonischemic pattern-, and a dominant LV presentation were associated with the presence of VA independently from the ARVC risk score. Interestingly, in this work there were no events in individuals with exclusive RV involvement and a 5-year ARVC risk < 15%. (143) Another cohort, assessed the performance of the score in the distinct genetic subgroups of ARVC, the global cohort depicted Uno's concordance index of 0.75, while the sub-group analysis in gene-positive, *PKP2*, *DSP*, other desmosomal and no variant carriers depicted Uno's concordance index ranging 0.67-0.88, being the best discrimination provided for *PKP2* carriers, followed by gene-positive, *DSP* carriers and gene-elusive patients. *PKP2* carriers showed Uno's concordance of 0.88, while *DSP* carriers of 0.77. In the same work they suggest that LV involvement might be relevant in arrhythmic stratification in the *DSP* subgroup.

In the development of the ARVC score, RVEF and LVEF were included as predictors, but LVEF was removed from the model after stepwise backward selection. (22) Given the abovementioned evidence, the inclusion of LV involvement could help to improve VA prediction in these patients, and especially in those carrying *DSP* deleterious variants which typically cause left-dominant biventricular arrhythmogenic cardiomyopathy. (140)

The model does not integrate genetic information as a predictive variable. Although complex genotype (double mutations or compound heterozygosis) is associated with worse arrhythmic prognosis (112, 144), their prevalence was very low in the derivation and validation cohorts (2-2.5%), and there is no evidence that mutations in specific genes confer worse prognosis, except for a better prognosis in carriers of variants in *DSC2* (145).

Among other predictors that have been suggested to enhance the ARVC risk calculator, the induction of sustained VAs in the electrophysiological study has shown to be associated with the combined outcome of VA independently from the ARVC risk calculator, depicting sensitivity of 76% and specificity of 68% (146), and increasing C-

index from 0.72 to 0.75 in comparison with the model alone (log-likelihood ratio for nested models, $P<0.001$). Since this study selected only patients with available electrophysiological study at the time of diagnosis, which was conducted following standard clinical indication, selection bias of higher risk patients resulting in underrepresentation of ARVC individuals with less advanced stage disease cannot be discarded.

As we can observe, the ARVC risk calculator is useful in the management of ARVC patients, being a valuable tool that provides integrated information of different arrhythmic predictors in a unified assessment, while accounting for their weight/relevance, the possible interaction among the included predictors and without losing the information on continuous variables, which altogether, is a great advance. Even so, it works well in the patients for whom it was designed, patients with definitive diagnosis of ARVC according to 2020 Task Force criteria, although in certain cohorts it shows a tendency to overestimate risk and it has been observed that it can potentially be improved by adding other clinical variables with independent predictive value (such as LV involvement), which will be especially relevant in specific patient subgroups such as those with predominant or concomitant LV involvement, including *DSP* variant carriers. (140)

One of the limitations of the ARVC calculator is that it is not clear which is the best threshold for recommending or discarding ICD implantation for SCD prevention. It has been postulated that this decision should take into account patient's preferences considering the calculated risk and the potential benefit and harms derived from an ICD implantation.

Another limitation of the current stratification algorithms as a whole is that they predict VA globally, including ICD therapies (generally programmed from 180bpm) and sustained VT with hemodynamic tolerance, in addition to rapid VA (>250bpm), VF and cardiac arrest.

In our external validation paper, 59% of the patients reached the endpoint of sustained VA due to ICD intervention. Of note, the proportion of ICD interventions that would lead to a SCD if these patients were not wearing and ICD is unknown. In addition, in this work

we show a tendency to overestimate risk in non-ICD carriers in the sub-group analysis, similarly to what has been shown in cohorts with 1% of ICD carriers. (141) Thus, caution needs to be applied when using tools predicting any VA, such do the ARVC risk calculator and some criteria in the current algorithms, for SCD risk stratification.

As mentioned in the introduction, the ITFC and mITFC algorithms have shown to be those that offer the greatest protection, but with a high rate of patients treated with ICDs unnecessarily. The AHA and HRS recommendations depict the highest global accuracy, as they are more specific, but leave a large proportion of patients unprotected (~7%) from a possibly lethal ARVC complication, SCD. (123) When comparing the ARVC risk calculator with other risk stratification algorithms, Cadrin-Tourigny *et al.*(22) showed that a cutoff point of > 18% provides a protection rate similar to the ITFC (~90%), but a much higher specificity and thus a lower number of unnecessary ICD implantations (with an absolute reduction of 20%) in spite of the limitation that these data are shown in the same cohort in which the score was derived. In the external validation cohort, a threshold of 17.5-20% (22) offers a protection rate ~75% (similar to ITFC), with a reduction of the unnecessary ICD of 8,5-12,7%. Aquaro *et al.* employed a cohort of 140 patients to determine the best cutoff point for the risk calculator to recommend ICD implantation, finding that the cutoff point of >10% (while using the maximally selected rank statistic) protects 95% of patients at the expense of ICD implantation in 81% of the cohort. (147) Interestingly, in a cohort with a 1% rate of ICD implantation (and thus with a more unbiased VA endpoint), it was shown that a risk calculator cutoff <10% has a negative predictive value of 100% for the combined endpoint of VA. (141)

In a comparison of the ITFC, mITFC, AHA and HRS algorithms based on risk factors, Bosman *et al.* showed that, when plotting the net benefit analysis for only rapid ventricular events -being rapid ventricular events defined as any sustained VA >250 bpm, SCD or sudden cardiac arrest-, to protect individuals at a 5-year observed risk >6% (as in HCM), the clinical algorithm for risk stratification that yielded the best risk assessment based on the decision curve analysis was the ITFC consensus. (123) The ARVC risk calculator was not included in this comparison, as the cohort employed overlapped with the cohort from which the ARVC score was derived. The recent ESC guidelines for the management of cardiomyopathies have included the same indications for primary

prevention of VA in ARVC as in the ITFC consensus, adding inducibility in the electrophysiological study as a class IIa recommendation, and endorsing the ARVC risk calculator but without a specific threshold value. (106) These indications will probably lead to implantation rates similar to those of the ITFC/mITFC (123) while the ARVC score can be a complementary tool considering its good discrimination and high negative predictive value. (141, 147)

In addition to the ARVC risk calculator, the same group developed a model for the prediction of rapid sustained VAs (>250bpm and SCD or arrest), which is probably a closer proxy for SCD than sustained VAs in general (>100bpm and SCD or arrest). (148) To date, it remains a pending issue to externally validate and analyze the clinical value of the rapid VA model, which is potentially very promising in approaching the prediction of the event intended to prevent, SCD.

Furthermore, it is important to note that these models have been derived and validated in European and North American cohorts, and therefore their validation in individuals of different ancestral origin and in healthcare systems with other socioeconomic backgrounds remains to be done.

5. The balance between benefits and harms in prediction models in cardiovascular diseases

Prediction of cardiovascular risk, with the intention of preventing the disease and its progression, is based on probabilistic (and nondeterministic) models. Therefore, it is essential to consider the impact of all interventions undertaken, from the prognostic assessment and transmission of results to the implementation of therapeutic measures, when necessary, and their potential nocebo effects. (149)

To facilitate this task, clinicians need adequate evidence and stratification tools which allow providing accurate prognostic information to jointly make decisions according to the patient's individual preferences after receiving the evidence-based information.

That is, for an individual such as the one presented in the introduction, with cholesterol values in the treatment range but without other comorbidities, in whom treatment with statins for 47 years (from 40 to 87) will prolong his median life free of vascular events from 87 to 89 years, the decision to treat or not to treat may be arguable. It is interesting

to note that, considering that the LIFE-CVD (80) model incorporates the effect of treatment on lifetime risk modification, it would also be interesting to incorporate the prediction of relevant treatment-related side effects, e.g. myalgia and/or diabetes mellitus, in the case of statin treatment. As another example, in an individual with ARVC, the utility of implantation of an ICD to prevent SCD will probably vary depending on whether what we are predicting, preventing, and treating is a VT with good hemodynamic tolerance, or a VT with poor hemodynamic tolerance or VF that can lead to SCD. Therefore, it is important to inform patients that the current consensus and models on ARVC, for example, are mostly based on studies that have evaluated arrhythmic ventricular risk globally. In addition, as we have highlighted many times, ICDs placements are associated to a significant rate of associated complications. To assess the global impact of the risk reduction and harms derived from the intervention, similarly to the therapeutic benefit is integrated into the LIFE-CVD model, it would be interesting to integrate so into other clinical models together with the probability of complications derived from treatment, in the case of ARVC for example, complications derived from the ICD implantation.

6. Final considerations and future directions

In this thesis we present distinct clinical tools useful in the prognostic stratification of patients with CVDs. From the exploration of the underpinnings of *BAG3*-DCM, to the use of predictive models using polygenic scores, to the implementation of predictive models integrating clinical variables.

In all those, the predictive capacity will depend both on the precision of the estimation of the effects on the variables included (which mainly depends on the representation of the predictors in the cohort and on the statistical power derived from the sample size), as well as on whether the individual fits the characteristics of the population from which these effects were derived and validated. Therefore, efforts need to be made to adequately represent the full spectrum of individuals affected by the diseases we intend to study in the cohorts employed in the derivation and validation of these models, including different forms of presentation of the disease and individuals with diverse ancestral and geographic backgrounds.

In the first article we describe a molecular subtype of DCM that presents with characteristic clinical presentation and histopathologic findings. Future cohorts that study in-depth the characteristics of these patients, including clinical presentation and molecular characterization, and contextualize their findings with those in other subtypes of DCM, will help in the characterization of DCM caused by *BAG3*, aiding in the prognostic stratification and management of these patients, and maybe allowing the identification of therapeutic targets for molecularly distinct forms of DCM.

In the second section of this thesis, we analyze the role of MTAG in genetic loci discovery and its incremental value in genomic prediction. While we do not demonstrate a benefit provided by MTAG in predicting disease status, we do reveal its role in the discovery of new loci that may point to interesting pathophysiological pathways, which may help in the understanding of CVDs and be useful as a starting point in the research and development of treatment targeted to those pathways. Notably, MTAG may be an interesting tool to help elucidate the underlying mechanisms in underrepresented subgroups in certain entities. In this work, we further evaluated the benefit of MTAG-derived PGS in predicting disease status, without MTAG improving the predictive or discriminative ability for PGS in AF, CAD, and HF, and for this reason, we did not delve into the analysis of the new PGS developed in combination with clinical models. However, to assess the added value of genomic prediction in common CVDs may certainly be of interest in future approaches.

The third section of this thesis covers the external validation of a clinical model for arrhythmic prognostic prediction in patients with ARVC, a strategy that has been used in other entities such as HCM. Several studies support the added prognostic value of LV involvement parameters to the ARVC risk score, especially in those carriers of *DSP* variants, which could help to improve the accuracy of the ARVC risk calculator strategy. In addition, there are current efforts attempting to evaluate the role of common genetic variants in the susceptibility and severity of ARVC, and if such is demonstrated relevant as in other cardiomyopathies as HCM (55, 150), the evaluation of their contribution in prognosis stratification would certainly be of interest.

As a final consideration, it is worth noting that in addition to research on instruments for CVD prediction and prognostic stratification, it is important to evaluate the impact (positive and negative) of their implementation in sometimes healthy and/or asymptomatic individuals, and to promote health education and a health system that accompanies and empowers individuals in making decisions related to their health.

Conclusions

I. The *BAG3* p.Gln88* variant identifies a high-penetrant form of dilated cardiomyopathy leading to acute heart failure onset at young ages and a specific histopathological pattern. This may serve to understand the basis of this molecular sub-type of disease in order to improve disease-specific treatment.

II. The implementation of multi-trait analysis of genome-wide association studies allows the discovery of disease-specific genomic loci for atrial fibrillation, coronary artery disease and heart failure, which are associated with relevant pathophysiological pathways in these diseases.

III. Multi-trait analysis of genome-wide association studies does not improve polygenic scores-based disease prediction and discrimination in common cardiovascular diseases: atrial fibrillation, coronary artery disease and heart failure.

IV. The arrhythmogenic right ventricle cardiomyopathy risk calculator external validation demonstrates good discrimination and calibration, together with a net clinical benefit superior to the other current treatment algorithms for predicting ventricular arrhythmias in patients with arrhythmogenic right ventricular cardiomyopathy.

V. Altogether, this PhD thesis evaluates the role of cardiovascular genetics in the understanding of the pathophysiological mechanisms underlying cardiovascular diseases, its diagnostic and prognostic value, as well as the role of disease prediction and prognostic stratification models in these entities.

References

1. Mizusawa Y, Wilde AA. Brugada syndrome. Circ Arrhythm Electrophysiol. 2012;5(3):606-16.
2. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart Disease and Stroke Statistics-2022 Update: A Report From the American Heart Association. Circulation. 2022;145(8):e153-e639.
3. Myerburg RJ, Junttila MJ. Sudden cardiac death caused by coronary heart disease. Circulation. 2012;125(8):1043-52.
4. Glinge C, Jabbari R, Risgaard B, Lynge TH, Engstrom T, Albert CM, et al. Symptoms Before Sudden Arrhythmic Death Syndrome: A Nationwide Study Among the Young in Denmark. J Cardiovasc Electrophysiol. 2015;26(7):761-7.
5. Marijon E, Uy-Evanado A, Dumas F, Karam N, Reinier K, Teodorescu C, et al. Warning Symptoms Are Associated With Survival From Sudden Cardiac Arrest. Ann Intern Med. 2016;164(1):23-9.
6. Fox CS, Evans JC, Larson MG, Kannel WB, Levy D. Temporal trends in coronary heart disease mortality and sudden cardiac death from 1950 to 1999: the Framingham Heart Study. Circulation. 2004;110(5):522-7.
7. Eckart RE, Shry EA, Burke AP, McNear JA, Appel DA, Castillo-Rojas LM, et al. Sudden death in young adults: an autopsy-based series of a population undergoing active surveillance. J Am Coll Cardiol. 2011;58(12):1254-61.
8. Risgaard B, Winkel BG, Jabbari R, Behr ER, Ingemann-Hansen O, Thomsen JL, et al. Burden of sudden cardiac death in persons aged 1 to 49 years: nationwide study in Denmark. Circ Arrhythm Electrophysiol. 2014;7(2):205-11.
9. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomstrom-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. Eur Heart J. 2021;42(5):373-498.
10. Zeppenfeld K, Tfelt-Hansen J, de Riva M, Winkel BG, Behr ER, Blom NA, et al. 2022 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. Eur Heart J. 2022;43(40):3997-4126.
11. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Bohm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J. 2021;42(36):3599-726.
12. Packer M. Major reduction in the risk of sudden cardiac death in patients with chronic heart failure with the use of drug and device combinations that favourably affect left ventricular structure. Eur J Heart Fail. 2019;21(7):823-6.
13. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. 1991;22(8):983-8.
14. Lin HJ, Wolf PA, Benjamin EJ, Belanger AJ, D'Agostino RB. Newly diagnosed atrial fibrillation and acute stroke. The Framingham Study. Stroke. 1995;26(9):1527-30.
15. Force USPST, Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, et al. Screening for Atrial Fibrillation: US Preventive Services Task Force Recommendation Statement. JAMA. 2022;327(4):360-7.
16. Lopes RD, Alings M, Connolly SJ, Beresh H, Granger CB, Mazuecos JB, et al. Rationale and design of the Apixaban for the Reduction of Thrombo-Embolism in Patients With Device-Detected Sub-Clinical Atrial Fibrillation (ARTESiA) trial. Am Heart J. 2017;189:137-45.

17. Benjamin EJ, Go AS, Desvigne-Nickens P, Anderson CD, Casadei B, Chen LY, et al. Research Priorities in Atrial Fibrillation Screening: A Report From a National Heart, Lung, and Blood Institute Virtual Workshop. Circulation. 2021;143(4):372-88.
18. Olde Nordkamp LR, Postema PG, Knops RE, van Dijk N, Limpens J, Wilde AA, et al. Implantable cardioverter-defibrillator harm in young patients with inherited arrhythmia syndromes: A systematic review and meta-analysis of inappropriate shocks and complications. Heart Rhythm. 2016;13(2):443-54.
19. Yoshinaga M, Kucho Y, Sarantuya J, Ninomiya Y, Horigome H, Ushinohama H, et al. Genetic characteristics of children and adolescents with long-QT syndrome diagnosed by school-based electrocardiographic screening programs. Circ Arrhythm Electrophysiol. 2014;7(1):107-12.
20. Jorgensen T, Jacobsen RK, Toft U, Aadahl M, Glumer C, Pisinger C. Effect of screening and lifestyle counselling on incidence of ischaemic heart disease in general population: Inter99 randomised trial. BMJ. 2014;348:g3617.
21. O'Mahony C, Jichi F, Pavlou M, Monserrat L, Anastasakis A, Rapezzi C, et al. A novel clinical risk prediction model for sudden cardiac death in hypertrophic cardiomyopathy (HCM risk-SCD). Eur Heart J. 2014;35(30):2010-20.
22. Cadrin-Tourigny J, Bosman LP, Nozza A, Wang W, Tadros R, Bhonsale A, et al. A new prediction model for ventricular arrhythmias in arrhythmogenic right ventricular cardiomyopathy. Eur Heart J. 2019;40(23):1850-8.
23. Geisterfer-Lowrance AA, Kass S, Tanigawa G, Vosberg HP, McKenna W, Seidman CE, et al. A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. Cell. 1990;62(5):999-1006.
24. James CA, Jongbloed JDH, Hershberger RE, Morales A, Judge DP, Syrris P, et al. International Evidence Based Reappraisal of Genes Associated With Arrhythmogenic Right Ventricular Cardiomyopathy Using the Clinical Genome Resource Framework. Circ Genom Precis Med. 2021;14(3):e003273.
25. Adler A, Novelli V, Amin AS, Abiusi E, Care M, Nannenberg EA, et al. An International, Multicentered, Evidence-Based Reappraisal of Genes Reported to Cause Congenital Long QT Syndrome. Circulation. 2020;141(6):418-28.
26. Hosseini SM, Kim R, Udupa S, Costain G, Jobling R, Liston E, et al. Reappraisal of Reported Genes for Sudden Arrhythmic Death: Evidence-Based Evaluation of Gene Validity for Brugada Syndrome. Circulation. 2018;138(12):1195-205.
27. Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, et al. Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. Circulation. 2021;144(1):7-19.
28. Ingles J, Goldstein J, Thaxton C, Caleshu C, Corty EW, Crowley SB, et al. Evaluating the Clinical Validity of Hypertrophic Cardiomyopathy Genes. Circ Genom Precis Med. 2019;12(2):e002460.
29. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24.
30. Mazzanti A, Maragna R, Vacanti G, Monteforte N, Bloise R, Marino M, et al. Interplay Between Genetic Substrate, QTc Duration, and Arrhythmia Risk in Patients With Long QT Syndrome. J Am Coll Cardiol. 2018;71(15):1663-71.
31. Schwartz PJ, Priori SG, Spazzolini C, Moss AJ, Vincent GM, Napolitano C, et al. Genotype-phenotype correlation in the long-QT syndrome: gene-specific triggers for life-threatening arrhythmias. Circulation. 2001;103(1):89-95.
32. Priori SG, Schwartz PJ, Napolitano C, Bloise R, Ronchetti E, Grillo M, et al. Risk stratification in the long-QT syndrome. N Engl J Med. 2003;348(19):1866-74.

33. Schwartz PJ, Spazzolini C, Crotti L, Bathen J, Amlie JP, Timothy K, et al. The Jervell and Lange-Nielsen syndrome: natural history, molecular basis, and clinical outcome. Circulation. 2006;113(6):783-90.
34. Gigli M, Merlo M, Graw SL, Barbati G, Rowland TJ, Slavov DB, et al. Genetic Risk of Arrhythmic Phenotypes in Patients With Dilated Cardiomyopathy. J Am Coll Cardiol. 2019;74(11):1480-90.
35. van Rijsingen IA, Arbustini E, Elliott PM, Mogensen J, Hermans-van Ast JF, van der Kooi AJ, et al. Risk factors for malignant ventricular arrhythmias in lamin a/c mutation carriers a European cohort study. J Am Coll Cardiol. 2012;59(5):493-500.
36. Jorda P, Toro R, Diez C, Salazar-Mendiguchia J, Fernandez-Falgueras A, Perez-Serra A, et al. Malignant Arrhythmogenic Role Associated with RBM20: A Comprehensive Interpretation Focused on a Personalized Approach. J Pers Med. 2021;11(2).
37. Ortiz-Genga MF, Cuenca S, Dal Ferro M, Zorio E, Salgado-Aranda R, Climent V, et al. Truncating FLNC Mutations Are Associated With High-Risk Dilated and Arrhythmogenic Cardiomyopathies. J Am Coll Cardiol. 2016;68(22):2440-51.
38. Smith ED, Lakdawala NK, Papoutsidakis N, Aubert G, Mazzanti A, McCanta AC, et al. Desmoplakin Cardiomyopathy, a Fibrotic and Inflammatory Form of Cardiomyopathy Distinct From Typical Dilated or Arrhythmogenic Right Ventricular Cardiomyopathy. Circulation. 2020;141(23):1872-84.
39. Dominguez F, Cuenca S, Bilinska Z, Toro R, Villard E, Barriales-Villa R, et al. Dilated Cardiomyopathy Due to BLC2-Associated Athanogene 3 (BAG3) Mutations. J Am Coll Cardiol. 2018;72(20):2471-81.
40. Palsson R, Indridason OS, Edvardsson VO, Oddsson A. Genetics of common complex kidney stone disease: insights from genome-wide association studies. Urolithiasis. 2019;47(1):11-21.
41. Timpson NJ, Greenwood CMT, Soranzo N, Lawson DJ, Richards JB. Genetic architecture: the shape of the genetic contribution to human traits and disease. Nat Rev Genet. 2018;19(2):110-24.
42. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet. 2018;19(9):581-90.
43. Bai J, Lu Y, Zhu Y, Wang H, Yin D, Zhang H, et al. Understanding PITX2-Dependent Atrial Fibrillation Mechanisms through Computational Models. Int J Mol Sci. 2021;22(14).
44. Cerrone M, Remme CA, Tadros R, Bezzina CR, Delmar M. Beyond the One Gene-One Disease Paradigm: Complex Genetics and Pleiotropy in Inheritable Cardiac Disorders. Circulation. 2019;140(7):595-610.
45. Lee JY, Kim TH, Yang PS, Lim HE, Choi EK, Shim J, et al. Korean atrial fibrillation network genome-wide association study for early-onset atrial fibrillation identifies novel susceptibility loci. Eur Heart J. 2017;38(34):2586-94.
46. Ellinor PT, Lunetta KL, Albert CM, Glazer NL, Ritchie MD, Smith AV, et al. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. Nat Genet. 2012;44(6):670-5.
47. Klarin D, Zhu QM, Emdin CA, Chaffin M, Horner S, McMillan BJ, et al. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. Nat Genet. 2017;49(9):1392-7.
48. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018;50(9):1219-24.
49. Martin AR, Daly MJ, Robinson EB, Hyman SE, Neale BM. Predicting Polygenic Risk of Psychiatric Disorders. Biol Psychiatry. 2019;86(2):97-109.
50. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511(7510):421-7.
51. Prive F, Arbel J, Vilhjalmsson BJ. LDpred2: better, faster, stronger. Bioinformatics. 2020.

52. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. 2019;10(1):1776.
53. Abul-Husn NS, Manickam K, Jones LK, Wright EA, Hartzel DN, Gonzaga-Jauregui C, et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. Science. 2016;354(6319).
54. Lahrouchi N, Tadros R, Crotti L, Mizusawa Y, Postema PG, Beekman L, et al. Transethnic Genome-Wide Association Study Provides Insights in the Genetic Architecture and Heritability of Long QT Syndrome. Circulation. 2020;142(4):324-38.
55. Tadros R, Francis C, Xu X, Vermeer AMC, Harper AR, Huurman R, et al. Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. Nat Genet. 2021;53(2):128-34.
56. Barc J, Tadros R, Glinge C, Chiang DY, Jouni M, Simonet F, et al. Author Correction: Genome-wide association analyses identify new Brugada syndrome risk loci and highlight a new mechanism of sodium channel regulation in disease susceptibility. Nat Genet. 2022;54(5):735.
57. Yap YG, Behr ER, Camm AJ. Drug-induced Brugada syndrome. Europace. 2009;11(8):989-94.
58. Lopez-Medina AI, Chahal CAA, Luzum JA. The genetics of drug-induced QT prolongation: evaluating the evidence for pharmacodynamic variants. Pharmacogenomics. 2022;23(9):543-57.
59. Arking DE, Pulit SL, Crotti L, van der Harst P, Munroe PB, Koopmann TT, et al. Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. Nat Genet. 2014;46(8):826-36.
60. Turkowski KL, Dotzler SM, Tester DJ, Giudicessi JR, Bos JM, Speziale AD, et al. Corrected QT Interval-Polygenic Risk Score and Its Contribution to Type 1, Type 2, and Type 3 Long-QT Syndrome in Probands and Genotype-Positive Family Members. Circ Genom Precis Med. 2020;13(4):e002922.
61. Kolder I, Tanck MWT, Postema PG, Barc J, Sinner MF, Zumhagen S, et al. Analysis for Genetic Modifiers of Disease Severity in Patients With Long-QT Syndrome Type 2. Circ Cardiovasc Genet. 2015;8(3):447-56.
62. Ingles J, Semsarian C. Time to Rethink the Genetic Architecture of Long QT Syndrome. Circulation. 2020;141(6):440-3.
63. Harper AR, Goel A, Grace C, Thomson KL, Petersen SE, Xu X, et al. Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. Nat Genet. 2021;53(2):135-42.
64. Jimmy Juang JM, Liu YB, Julius Chen CY, Yu QY, Chattopadhyay A, Lin LY, et al. Validation and Disease Risk Assessment of Previously Reported Genome-Wide Genetic Variants Associated With Brugada Syndrome: SADS-TW BrS Registry. Circ Genom Precis Med. 2020;13(4):e002797.
65. Bagnall RD, Ingles J, Dinger ME, Cowley MJ, Ross SB, Minoche AE, et al. Whole Genome Sequencing Improves Outcomes of Genetic Testing in Patients With Hypertrophic Cardiomyopathy. J Am Coll Cardiol. 2018;72(4):419-29.
66. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020;15(9):2759-72.
67. Jordan DM, Verbanck M, Do R. HOPS: a quantitative score reveals pervasive horizontal pleiotropy in human genetic variation is driven by extreme polygenicity of human traits and diseases. Genome Biol. 2019;20(1):222.
68. Villard E, Perret C, Gary F, Proust C, Dilanian G, Hengstenberg C, et al. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. Eur Heart J. 2011;32(9):1065-76.

69. Aung N, Vargas JD, Yang C, Cabrera CP, Warren HR, Fung K, et al. Genome-Wide Analysis of Left Ventricular Image-Derived Phenotypes Identifies Fourteen Loci Associated With Cardiac Morphogenesis and Heart Failure Development. Circulation. 2019;140(16):1318-30.
70. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat Genet. 2018;50(2):229-37.
71. Mayhew AJ, Meyre D. Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities. Curr Genomics. 2017;18(4):332-40.
72. Dorresteijn JA, Visseren FL, Ridker PM, Wassink AM, Paynter NP, Steyerberg EW, et al. Estimating treatment effects for individual patients based on the results of randomised clinical trials. BMJ. 2011;343:d5888.
73. van der Leeuw J, Ridker PM, van der Graaf Y, Visseren FL. Personalized cardiovascular disease prevention by applying individualized prediction of treatment effects. Eur Heart J. 2014;35(13):837-43.
74. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Back M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. Eur J Prev Cardiol. 2022;29(1):5-115.
75. group Sw, collaboration ESCCr. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur Heart J. 2021;42(25):2439-54.
76. group SOw, collaboration ESCCr. SCORE2-OP risk prediction algorithms: estimating incident cardiovascular event risk in older persons in four geographical risk regions. Eur Heart J. 2021;42(25):2455-67.
77. Goff DC, Jr., Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation. 2014;129(25 Suppl 2):S49-73.
78. Kengne AP, Patel A, Marre M, Travert F, Lievre M, Zoungas S, et al. Contemporary model for cardiovascular risk prediction in people with type 2 diabetes. Eur J Cardiovasc Prev Rehabil. 2011;18(3):393-8.
79. Dorresteijn JA, Visseren FL, Wassink AM, Gondrie MJ, Steyerberg EW, Ridker PM, et al. Development and validation of a prediction rule for recurrent vascular events based on a cohort study of patients with arterial disease: the SMART risk score. Heart. 2013;99(12):866-72.
80. Jaspers NEM, Blaha MJ, Matsushita K, van der Schouw YT, Wareham NJ, Khaw KT, et al. Prediction of individualized lifetime benefit from cholesterol lowering, blood pressure lowering, antithrombotic therapy, and smoking cessation in apparently healthy people. Eur Heart J. 2020;41(11):1190-9.
81. Berkelmans GFN, Gudbjornsdottir S, Visseren FLJ, Wild SH, Franzen S, Chalmers J, et al. Prediction of individual life-years gained without cardiovascular events from lipid, blood pressure, glucose, and aspirin treatment based on data of more than 500 000 patients with Type 2 diabetes mellitus. Eur Heart J. 2019;40(34):2899-906.
82. Kaasenbrood L, Bhatt DL, Dorresteijn JAN, Wilson PWF, D'Agostino RB, Sr., Massaro JM, et al. Estimated Life Expectancy Without Recurrent Cardiovascular Events in Patients With Vascular Disease: The SMART-REACH Model. J Am Heart Assoc. 2018;7(16):e009217.
83. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Back M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. Eur Heart J. 2021;42(34):3227-337.
84. Hoogwerf BJ. Statins may increase diabetes, but benefit still outweighs risk. Cleve Clin J Med. 2023;90(1):53-62.
85. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. Lancet. 2016;388(10059):2532-61.

86. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ. 2017;357:j2099.
87. Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB, et al. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. J Am Heart Assoc. 2013;2(2):e000102.
88. Goudis C, Daios S, Dimitriadis F, Liu T. CHARGE-AF: A Useful Score For Atrial Fibrillation Prediction? Curr Cardiol Rev. 2023;19(2):e010922208402.
89. 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(4):584-91.
90. Mirelis JG, Escobar-Lopez L, Ochoa JP, Espinosa MA, Villacorta E, Navarro M, et al. Combination of late gadolinium enhancement and genotype improves prediction of prognosis in non-ischaemic dilated cardiomyopathy. Eur J Heart Fail. 2022;24(7):1183-96.
91. Di Marco A, Brown PF, Bradley J, Nucifora G, Claver E, de Frutos F, et al. Improved Risk Stratification for Ventricular Arrhythmias and Sudden Death in Patients With Nonischemic Dilated Cardiomyopathy. J Am Coll Cardiol. 2021;77(23):2890-905.
92. Wahbi K, Ben Yaou R, Gandjbakhch E, Anselme F, Gossios T, Lakdawala NK, et al. Development and Validation of a New Risk Prediction Score for Life-Threatening Ventricular Tachyarrhythmias in Laminopathies. Circulation. 2019;140(4):293-302.
93. Verstraelen TE, van Lint FHM, Bosman LP, de Brouwer R, Proost VM, Abeln BGS, et al. Prediction of ventricular arrhythmia in phospholamban p.Arg14del mutation carriers-reaching the frontiers of individual risk prediction. Eur Heart J. 2021;42(29):2842-50.
94. Vriesendorp PA, Schinkel AF, Liebregts M, Theuns DA, van Cleemput J, Ten Cate FJ, et al. Validation of the 2014 European Society of Cardiology guidelines risk prediction model for the primary prevention of sudden cardiac death in hypertrophic cardiomyopathy. Circ Arrhythm Electrophysiol. 2015;8(4):829-35.
95. O'Mahony C, Jichi F, Ommen SR, Christiaans I, Arbustini E, Garcia-Pavia P, et al. International External Validation Study of the 2014 European Society of Cardiology Guidelines on Sudden Cardiac Death Prevention in Hypertrophic Cardiomyopathy (EVIDENCE-HCM). Circulation. 2018;137(10):1015-23.
96. Magri D, Limongelli G, Re F, Agostoni P, Zachara E, Correale M, et al. Cardiopulmonary exercise test and sudden cardiac death risk in hypertrophic cardiomyopathy. Heart. 2016;102(8):602-9.
97. Choi YJ, Kim HK, Lee SC, Park JB, Moon I, Park J, et al. Validation of the hypertrophic cardiomyopathy risk-sudden cardiac death calculator in Asians. Heart. 2019;105(24):1892-7.
98. Fernandez A, Quiroga A, Ochoa JP, Mysuta M, Casabe JH, Biagetti M, et al. Validation of the 2014 European Society of Cardiology Sudden Cardiac Death Risk Prediction Model in Hypertrophic Cardiomyopathy in a Reference Center in South America. Am J Cardiol. 2016;118(1):121-6.
99. Maron BJ, Casey SA, Chan RH, Garberich RF, Rowin EJ, Maron MS. Independent Assessment of the European Society of Cardiology Sudden Death Risk Model for Hypertrophic Cardiomyopathy. Am J Cardiol. 2015;116(5):757-64.
100. O'Mahony C, Akhtar MM, Anastasiou Z, Guttmann OP, Vriesendorp PA, Michels M, et al. Effectiveness of the 2014 European Society of Cardiology guideline on sudden cardiac death in hypertrophic cardiomyopathy: a systematic review and meta-analysis. Heart. 2019;105(8):623-31.
101. Marian AJ, Braunwald E. Hypertrophic Cardiomyopathy: Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. Circ Res. 2017;121(7):749-70.
102. Maron BJ, Olivotto I, Spirito P, Casey SA, Bellone P, Gohman TE, et al. Epidemiology of hypertrophic cardiomyopathy-related death: revisited in a large non-referral-based patient population. Circulation. 2000;102(8):858-64.

103. Gersh BJ, Maron BJ, Bonow RO, Dearani JA, Fifer MA, Link MS, et al. 2011 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation. 2011;124(24):2761-96.
104. Authors/Task Force m, Elliott PM, Anastasakis A, Borger MA, Borggrefe M, Cecchi F, et al. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy: the Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). Eur Heart J. 2014;35(39):2733-79.
105. Ommen SR, Mital S, Burke MA, Day SM, Deswal A, Elliott P, et al. 2020 AHA/ACC Guideline for the Diagnosis and Treatment of Patients With Hypertrophic Cardiomyopathy: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation. 2020;142(25):e558-e631.
106. Arbelo E, Protonotarios A, Gimeno JR, Arbustini E, Barriales-Villa R, Basso C, et al. 2023 ESC Guidelines for the management of cardiomyopathies. Eur Heart J. 2023.
107. O'Mahony C, Tome-Esteban M, Lambiase PD, Pantazis A, Dickie S, McKenna WJ, et al. A validation study of the 2003 American College of Cardiology/European Society of Cardiology and 2011 American College of Cardiology Foundation/American Heart Association risk stratification and treatment algorithms for sudden cardiac death in patients with hypertrophic cardiomyopathy. Heart. 2013;99(8):534-41.
108. Rampazzo A, Nava A, Danieli GA, Buja G, Daliento L, Fasoli G, et al. The gene for arrhythmogenic right ventricular cardiomyopathy maps to chromosome 14q23-q24. Hum Mol Genet. 1994;3(6):959-62.
109. Peters S, Trummel M, Meyners W. Prevalence of right ventricular dysplasia-cardiomyopathy in a non-referral hospital. Int J Cardiol. 2004;97(3):499-501.
110. Basso C, Corrado D, Marcus FI, Nava A, Thiene G. Arrhythmogenic right ventricular cardiomyopathy. Lancet. 2009;373(9671):1289-300.
111. Basso C, Corrado D, Thiene G. Arrhythmogenic right ventricular cardiomyopathy: what's in a name? From a congenital defect (dysplasia) to a genetically determined cardiomyopathy (dystrophy). Am J Cardiol. 2010;106(2):275-7.
112. Bhonsale A, Groeneweg JA, James CA, Dooijes D, Tichnell C, Jongbloed JD, et al. Impact of genotype on clinical course in arrhythmogenic right ventricular dysplasia/cardiomyopathy-associated mutation carriers. Eur Heart J. 2015;36(14):847-55.
113. Corrado D, Link MS, Calkins H. Arrhythmogenic Right Ventricular Cardiomyopathy. N Engl J Med. 2017;376(1):61-72.
114. Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, et al. Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. Eur Heart J. 2010;31(7):806-14.
115. Corrado D, van Tintelen PJ, McKenna WJ, Hauer RNW, Anastastakis A, Asimaki A, et al. Arrhythmogenic right ventricular cardiomyopathy: evaluation of the current diagnostic criteria and differential diagnosis. Eur Heart J. 2020;41(14):1414-29.
116. Corrado D, Perazzolo Marra M, Zorzi A, Beffagna G, Cipriani A, Lazzari M, et al. Diagnosis of arrhythmogenic cardiomyopathy: The Padua criteria. Int J Cardiol. 2020;319:106-14.
117. Thiene G, Nava A, Corrado D, Rossi L, Pennelli N. Right ventricular cardiomyopathy and sudden death in young people. N Engl J Med. 1988;318(3):129-33.
118. Bosman LP, Sammani A, James CA, Cadrin-Tourigny J, Calkins H, van Tintelen JP, et al. Predicting arrhythmic risk in arrhythmogenic right ventricular cardiomyopathy: A systematic review and meta-analysis. Heart Rhythm. 2018;15(7):1097-107.
119. Corrado D, Wichter T, Link MS, Hauer RN, Marchlinski FE, Anastasakis A, et al. Treatment of Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia: An International Task Force Consensus Statement. Circulation. 2015;132(5):441-53.

120. Orgeron GM, Te Riele A, Tichnell C, Wang W, Murray B, Bhonsale A, et al. Performance of the 2015 International Task Force Consensus Statement Risk Stratification Algorithm for Implantable Cardioverter-Defibrillator Placement in Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy. Circ Arrhythm Electrophysiol. 2018;11(2):e005593.
121. Al-Khatib SM, Stevenson WG, Ackerman MJ, Bryant WJ, Callans DJ, Curtis AB, et al. 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. Heart Rhythm. 2018;15(10):e73-e189.
122. Towbin JA, McKenna WJ, Abrams DJ, Ackerman MJ, Calkins H, Darrieux FCC, et al. 2019 HRS expert consensus statement on evaluation, risk stratification, and management of arrhythmogenic cardiomyopathy: Executive summary. Heart Rhythm. 2019;16(11):e373-e407.
123. Bosman LP, Nielsen Gerlach CL, Cadrin-Tourigny J, Orgeron G, Tichnell C, Murray B, et al. Comparing clinical performance of current implantable cardioverter-defibrillator implantation recommendations in arrhythmogenic right ventricular cardiomyopathy. Europace. 2022;24(2):296-305.
124. Corrado D, Wichter T, Link MS, Hauer R, Marchlinski F, Anastasakis A, et al. Treatment of arrhythmogenic right ventricular cardiomyopathy/dysplasia: an international task force consensus statement. Eur Heart J. 2015;36(46):3227-37.
125. Arimura T, Ishikawa T, Nunoda S, Kawai S, Kimura A. Dilated cardiomyopathy-associated BAG3 mutations impair Z-disc assembly and enhance sensitivity to apoptosis in cardiomyocytes. Hum Mutat. 2011;32(12):1481-91.
126. Chami N, Tadros R, Lemarbre F, Lo KS, Beaudoin M, Robb L, et al. Nonsense mutations in BAG3 are associated with early-onset dilated cardiomyopathy in French Canadians. Can J Cardiol. 2014;30(12):1655-61.
127. Shi S, Liu T, Wang D, Zhang Y, Liang J, Yang B, et al. Activation of N-methyl-d-aspartate receptors reduces heart rate variability and facilitates atrial fibrillation in rats. Europace. 2017;19(7):1237-43.
128. Cronje HT, Karhunen V, Hovingh GK, Coppieters K, Lagerstedt JO, Nyberg M, et al. Genetic evidence implicating natriuretic peptide receptor-3 in cardiovascular disease risk: a Mendelian randomization study. BMC Med. 2023;21(1):158.
129. Caporizzo MA, Prosser BL. The microtubule cytoskeleton in cardiac mechanics and heart failure. Nat Rev Cardiol. 2022;19(6):364-78.
130. Abifadel M, Varret M, Rabes JP, Allard D, Ouguerram K, Devillers M, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet. 2003;34(2):154-6.
131. Kaddoura R, Orabi B, Salam AM. PCSK9 Monoclonal Antibodies: An Overview. Heart Views. 2020;21(2):97-103.
132. Jarcho JA, McKenna W, Pare JA, Solomon SD, Holcombe RF, Dickie S, et al. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. N Engl J Med. 1989;321(20):1372-8.
133. Green EM, Wakimoto H, Anderson RL, Evanchik MJ, Gorham JM, Harrison BC, et al. A small-molecule inhibitor of sarcomere contractility suppresses hypertrophic cardiomyopathy in mice. Science. 2016;351(6273):617-21.
134. Marian AJ. Molecular Genetic Basis of Hypertrophic Cardiomyopathy. Circ Res. 2021;128(10):1533-53.
135. Olivotto I, Oreziak A, Barriales-Villa R, Abraham TP, Masri A, Garcia-Pavia P, et al. Mavacamten for treatment of symptomatic obstructive hypertrophic cardiomyopathy (EXPLORER-HCM): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2020;396(10253):759-69.
136. Desai MY, Owens A, Geske JB, Wolski K, Naidu SS, Smedira NG, et al. Myosin Inhibition in Patients With Obstructive Hypertrophic Cardiomyopathy Referred for Septal Reduction Therapy. J Am Coll Cardiol. 2022;80(2):95-108.

137. Desai MY, Owens A, Geske JB, Wolski K, Saberi S, Wang A, et al. Dose-Blinded Myosin Inhibition in Patients With Obstructive Hypertrophic Cardiomyopathy Referred for Septal Reduction Therapy: Outcomes Through 32 Weeks. Circulation. 2023;147(11):850-63.
138. Saberi S, Cardim N, Yamani M, Schulz-Menger J, Li W, Florea V, et al. Mavacamten Favorably Impacts Cardiac Structure in Obstructive Hypertrophic Cardiomyopathy: EXPLORER-HCM Cardiac Magnetic Resonance Substudy Analysis. Circulation. 2021;143(6):606-8.
139. Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. Nat Rev Drug Discov. 2014;13(6):419-31.
140. Protonotarios A, Bariani R, Cappelletto C, Pavlou M, Garcia-Garcia A, Cipriani A, et al. Importance of genotype for risk stratification in arrhythmogenic right ventricular cardiomyopathy using the 2019 ARVC risk calculator. Eur Heart J. 2022;43(32):3053-67.
141. Baudinaud P, Laredo M, Badenco N, Rouanet S, Waintraub X, Duthoit G, et al. External Validation of a Risk Prediction Model for Ventricular Arrhythmias in Arrhythmogenic Right Ventricular Cardiomyopathy. Can J Cardiol. 2021;37(8):1263-6.
142. Casella M, Gasperetti A, Gaetano F, Busana M, Sommariva E, Catto V, et al. Long-term follow-up analysis of a highly characterized arrhythmogenic cardiomyopathy cohort with classical and non-classical phenotypes-a real-world assessment of a novel prediction model: does the subtype really matter. Europace. 2020;22(5):797-805.
143. Aquaro GD, De Luca A, Cappelletto C, Raimondi F, Bianco F, Botto N, et al. Prognostic Value of Magnetic Resonance Phenotype in Patients With Arrhythmogenic Right Ventricular Cardiomyopathy. J Am Coll Cardiol. 2020;75(22):2753-65.
144. Rigato I, Bauce B, Rampazzo A, Zorzi A, Pilichou K, Mazzotti E, et al. Compound and digenic heterozygosity predicts lifetime arrhythmic outcome and sudden cardiac death in desmosomal gene-related arrhythmogenic right ventricular cardiomyopathy. Circ Cardiovasc Genet. 2013;6(6):533-42.
145. Protonotarios A, Anastasakis A, Panagiotakos DB, Antoniades L, Syrris P, Vouliotis A, et al. Arrhythmic risk assessment in genotyped families with arrhythmogenic right ventricular cardiomyopathy. Europace. 2016;18(4):610-6.
146. Gasperetti A, Carrick RT, Costa S, Compagnucci P, Bosman LP, Chivulescu M, et al. Programmed Ventricular Stimulation as an Additional Primary Prevention Risk Stratification Tool in Arrhythmogenic Right Ventricular Cardiomyopathy: A Multinational Study. Circulation. 2022;146(19):1434-43.
147. Aquaro GD, De Luca A, Cappelletto C, Raimondi F, Bianco F, Botto N, et al. Comparison of different prediction models for the indication of implanted cardioverter defibrillator in patients with arrhythmogenic right ventricular cardiomyopathy. ESC Heart Fail. 2020;7(6):4080-8.
148. Cadrin-Tourigny J, Bosman LP, Wang W, Tadros R, Bhonsale A, Bourfiss M, et al. Sudden Cardiac Death Prediction in Arrhythmogenic Right Ventricular Cardiomyopathy: A Multinational Collaboration. Circ Arrhythm Electrophysiol. 2021;14(1):e008509.
149. Turnwald BP, Goyer JP, Boles DZ, Silder A, Delp SL, Crum AJ. Learning one's genetic risk changes physiology independent of actual genetic risk. Nat Hum Behav. 2019;3(1):48-56.
150. Tadros R, Zheng SL, Grace C, Jorda P, Francis C, Jurgens SJ, et al. Large scale genome-wide association analyses identify novel genetic loci and mechanisms in hypertrophic cardiomyopathy. medRxiv. 2023.