Circulation

ORIGINAL RESEARCH ARTICLE

Frequency, Penetrance, and Variable Expressivity of Dilated Cardiomyopathy–Associated Putative Pathogenic Gene Variants in UK Biobank Participants

Ravi A. Shah, MB, BChir*; Babken Asatryan, MD, PhD*; Ghaith Sharaf Dabbagh, MD; Nay Aung, MD; Mohammed Y. Khanji, MBBCh, PhD; Luis R. Lopes, MD, PhD; Stefan van Duijvenboden, PhD; Anthony Holmes; Daniele Muser, MD; Andrew P. Landstrom, MD, PhD; Aaron Mark Lee, BSc(Hons), PhD, MBBS; Pankaj Arora, MD; Christopher Semsarian, MBBS, PhD, MPH; Virend K. Somers, MD, PhD; Anjali T. Owens, MD; Patricia B. Munroe, MSc, PhD; Steffen E. Petersen, MSc, MPH, MD, DPhil; C. Anwar A. Chahal, MBChB, PhD

BACKGROUND: There is a paucity of data regarding the phenotype of dilated cardiomyopathy (DCM) gene variants in the general population. We aimed to determine the frequency and penetrance of DCM-associated putative pathogenic gene variants in a general adult population, with a focus on the expression of clinical and subclinical phenotype, including structural, functional, and arrhythmic disease features.

METHODS: UK Biobank participants who had undergone whole exome sequencing, ECG, and cardiovascular magnetic resonance imaging were selected for study. Three variant-calling strategies (1 primary and 2 secondary) were used to identify participants with putative pathogenic variants in 44 DCM genes. The observed phenotype was graded DCM (clinical or cardiovascular magnetic resonance diagnosis); early DCM features, including arrhythmia or conduction disease, isolated ventricular dilation, and hypokinetic nondilated cardiomyopathy; or phenotype-negative.

RESULTS: Among 18 665 individuals included in the study, 1463 (7.8%) possessed ≥1 putative pathogenic variant in 44 DCM genes by the main variant calling strategy. A clinical diagnosis of DCM was present in 0.34% and early DCM features in 5.7% of individuals with putative pathogenic variants. ECG and cardiovascular magnetic resonance analysis revealed evidence of subclinical DCM in an additional 1.6% and early DCM features in an additional 15.9% of individuals with putative pathogenic variants. Arrhythmias or conduction disease (15.2%) were the most common early DCM features, followed by hypokinetic nondilated cardiomyopathy (4%). The combined clinical/subclinical penetrance was ≤30% with all 3 variant filtering strategies. Clinical DCM was slightly more prevalent among participants with putative pathogenic variants in definitive/strong evidence genes as compared with those with variants in moderate/limited evidence genes.

CONCLUSIONS: In the UK Biobank, ≈1 of 6 of adults with putative pathogenic variants in DCM genes exhibited early DCM features potentially associated with DCM genotype, most commonly manifesting with arrhythmias in the absence of substantial ventricular dilation or dysfunction.

Key Words: arrhythmias, cardiac ■ cardiomyopathies ■ death, sudden, cardiac ■ genetic testing ■ genetics ■ penetrance

Correspondence to: C. Anwar A. Chahal, MBChB, PhD, Center for Inherited Cardiovascular Diseases, WellSpan Health, 140 North Pointe Boulevard, Lancaster, PA 17601. Email anwar.chahal@pennmedicine.upenn.edu

*R.A. Shah and B. Asatryan contributed equally.

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/CIRCULATIONAHA.121.058143.

For Sources of Funding and Disclosures, see page XXX.

Circulation is available at www.ahajournals.org/journal/circ

Clinical Perspective

What Is New?

- Among individuals with putative pathogenic dilated cardiomyopathy (DCM) gene variants, DCM detected by ECG and cardiovascular magnetic resonance and early DCM features (ie, subclinical) were nearly 4 times more common than clinically manifest DCM or early features (23.7% versus 6.1%).
- More than 90% of participants with a putative pathogenic variant in DCM-associated genes did not have a history of DCM.
- Clinical DCM was slightly more prevalent among participants with putative pathogenic variants in definitive or strong evidence genes (13.9% for clinical and subclinical) as compared with those with variants in moderate or limited evidence genes, but there was no significant difference in combined clinical and subclinical phenotype by cluster.
- The overall clinical/subclinical penetrance of DCM-associated single putative pathogenic variants was highly variable between genes, ranging from 0 to 66.7%.

What Are the Clinical Implications?

- Arrhythmias and cardiac conduction disease are the most common early manifestations of putative pathogenic variants implicated in DCM, mostly occurring before the development of structural or functional abnormalities.
- A genotype-first screening approach for DCM using a large genetic panel is not suitable in the general population because of incomplete understanding of DCM genetic architecture and reduced penetrance of DCM-associated putative pathogenic variants.

Nonstandard Abbreviations and Acronyms

CCD	cardiac conduction disease
ClinGen	Clinical Genome Resource
CMR	cardiovascular magnetic resonance
DCM	dilated cardiomyopathy
FAF	filtering allele frequency
ICD-10	International Classification of Diseases, 10th revision
LV	left ventricular
WES	whole exome sequencing

Dilated cardiomyopathy (DCM) is a genetic heart disease that frequently leads to end-stage heart failure, characterized by progressive left ventricular (LV) or biventricular dilation and impaired contraction that is not explained exclusively by abnormal loading conditions (hypertension or valvular heart disease) or coronary artery disease.[1] Patients with DCM often present in adulthood and are prone to life-threatening ventricular arrhythmias, with 30% dying suddenly.[2] With the advent of next-generation sequencing technologies, there has been a dramatic increase in the number of genes tested and variants identified in patients with DCM.[3] To date, >250 genes from 10 gene ontologies have been reported in association with DCM, of which only 19 were recently found to have moderate, strong, or definitive evidence for causality in monogenic DCM by the Clinical Genome Resource (ClinGen) DCM Gene Curation Expert Panel.[4] It is estimated that a pathogenic or likely pathogenic variant can be identified in ≈20% to 35% of patients with DCM.[5,6]

The increasing availability, falling costs, and widespread use of genetic testing (including direct-to-consumer testing) offer an opportunity to use a genome-first method for diagnosis.[7] However, routine genetic screening is not justified because of the unknown frequency of putative pathogenic DCM gene variants in the general population, as well as uncertainties with incomplete penetrance and variable expressivity and challenges in variant calling.[8,9] These factors complicate the applicability and clinical implications of a given gene variant. Understanding the frequency and penetrance of DCM-associated gene variants in the general population is critical to patient and family counseling and clinical decision-making in those with incidental findings. However, the prevalence and penetrance of DCM-associated pathogenic variants in the general population remain insufficiently investigated.

Using the UK Biobank, we aimed to determine the frequency of putative pathogenic variants in the ClinGen DCM Gene Curation Expert Panel–asserted genes[4]; determine clinical DCM penetrance on the basis of electronic health records; identify patients with subclinical DCM or DCM features using advanced, quantitative 12-lead ECG and cardiovascular magnetic resonance (CMR) imaging data; and assess the effect of putative pathogenic variants in DCM-associated genes on patient outcomes. This study provides a large-scale genotype–phenotype correlation for DCM genes in the middle-aged to older adult population and in a subset of participants with clinically diagnosed DCM, with a focus on the expression of clinical and subclinical phenotype, and considering structural and arrhythmic features of DCM.

METHODS

Study Population

The UK Biobank study is a prospective study of 502 493 UK residents between 40 and 69 years of age at enrollment who were recruited at 22 assessment centers across the United Kingdom.[10] Participants attended a center visit undergoing deep phenotyping, including anthropometric measurements, extensive health and lifestyle questionnaires,

and biological samples. This provided information on baseline characteristics and self-reported medical conditions. Additional links to primary care records and external hospital data records provided data from hospital admissions in the form of International Classification of Diseases, 10th revision (ICD-10) diagnostic codes and OPCS-4 operation codes. The survival status was updated until January 2018, generating long-term follow-up data. A subset of participants in the UK Biobank have undergone a selection of whole exome sequencing (WES), CMR, and 12-lead ECG recordings; this subset comprised the cohort of this study. The UK Biobank received approval from the North West Multi-Center Research Ethics Committee.

Gene-First Approach to Identify the Study Population

Among UK Biobank participants, 200 000 underwent WES as previously described.[11] For this study, we used a panel of 44 genes recently asserted to be implicated in DCM by the ClinGen DCM Gene Curation Expert Panel.[4] This panel includes 11 genes with definite evidence (*BAG3*, *DES*, *FLNC*, *LMNA*, *MYH7*, *PLN*, *RBM20*, *SCN5A*, *TNNC1*, *TNNT2*, *TTN*), 1 with strong evidence (*DSP*), 7 with moderate evidence (*ACTC1*, *ACTN2*, *JPH2*, *NEXN*, *TNNI3*, *TPM1*, *VCL*), and 25 with limited evidence for causality in monogenic DCM (*ABCC9*, *ANKRD1*, *CSRP3*, *CTF1*, *DSG2*, *DTNA*, *EYA4*, *GATAD1*, *ILK*, *LAMA4*, *LDB3*, *MYBPC3*, *MYH6*, *MYL2*, *MYPN*, *NEBL*, *NKX2-5*, *OBSCN*, *PLEKHM2*, *PRDM16*, *PSEN2*, *SGCD*, *TBX20*, *TCAP*, *TNNI3K*).[4] We used 3 variant filtering strategies (1 primary and 2 secondary) to classify variants. For all strategies, we restricted the analysis to only high-quality (read depth ≥10, call quality ≥20, and genotype quality ≥20) and rare variants (minor allele frequency ≤0.001 in both gnomAD[12] and the UK Biobank exome dataset). A separate analysis was performed for the American College of Medical Genetics and Genomics clinically actionable DCM genes (*TNNT2*, *LMNA*, *FLNC*, and *TTN*).[13]

In the first filtering strategy (missense predicted loss-of-function variant filtering allele frequency [FAF]; main strategy), we used ANNOVAR[14] annotations and REVEL scores (a method for predicting deleterious missense variants[15]) to determine a set of putative pathogenic variants (as used elsewhere[16,17]). Variants with ANNOVAR annotations of frameshift insertions/deletions, gain/loss of stop codon, or disruption of canonical splice site dinucleotides were classified as predicted loss-of-function. Missense variants were determined as predicted pathogenic if the annotated REVEL score was ≥0.65.[16] For *TTN*, only radical variants (ie, nonsense, frameshift, and splice-site variants) were considered. We applied a FAF, removing all variants with a FAF of 8.4×10^{-5} or greater in gnomAD or UK Biobank[18] to produce our final set of variants. Because of the population prevalence of DCM, variants that occur more frequently than this are unlikely to be causative variants under a monogenic Mendelian model. This frequency threshold for DCM and other inherited cardiac conditions has been defined previously.[18]

Two secondary variant filtering strategies were performed (InterVar FAF and InterVar FAF ClinVar). Criteria used for these variant filtering strategies are provided in the Expanded Methods in the Supplemental Material.

Quality Control of Variant Filtering Strategy on the Basis of the Clinical DCM Population

Before applying our genetic testing approach to the study population, we performed a quality control analysis of the filtering and variant calling strategies on individuals with the clinical diagnosis of DCM WES (see Results).

ECG Analysis

All individuals who underwent CMR also underwent 12-lead ECG recording. Ten electrodes were placed in standard position, recorded at a frequency of 500 Hz for 10 seconds (Cardiosoft v6.51 GE), and stored in XML file format. These files were downloaded and reprocessed using GE MUSE v9.0 SP4, Marquette 12 SL.[19] Unusable ECG tracings were confirmed manually and removed. Of those remaining, 100 were randomly selected and underwent manual review by a board-certified cardiologist masked to the clinical diagnoses, CMR, or genetic status. These were then classified into bradyarrhythmias and tachyarrhythmias and conduction system disease using established criteria (see Table S1 for details).[20]

CMR Analysis

The UK Biobank CMR protocol has been described previously.[21] In brief, all CMR scans were acquired on a wide-bore 1.5 Tesla scanner (MAGNETOM Aera, Syngo Platform VD13A; Siemens Healthcare). The practical and ethical considerations posed by the large-scale and observational nature of the UK Biobank preclude the use of contrast or stress agents. The protocol includes bright blood anatomic assessment (sagittal, coronal, and axial), balanced and steady-state free precession sequences, left and right ventricular steady-state free precession cine images (long and short axes), myocardial tagging (3 short-axis slices), native T1 mapping, aortic flow, and imaging of the thoracic aorta. Typical measures were as follows: repetition time/echo time 2.6/1.1 ms, flip angle 80°, GRAPPA factor 2, voxel size $1.8\times1.8\times8$ mm^3 (6 mm for long axis). The actual temporal resolution of 32 ms was interpolated to 50 phases per cardiac cycle (≈20 ms). Analysis was performed using Circle CVI postprocessing software (version 5.1.1; Circle Cardiovascular Imaging Inc.).[22] Further details on phenotyping are given in the Appendix in the Supplemental Material.

Penetrance Analysis

We defined penetrant disease on the basis of the DCM clinical spectrum as laid out in the 2016 European Society of Cardiology position statement on DCM.[23] The spectrum includes DCM (LV dilation and hypokinesia), hypokinetic non-dilated cardiomyopathy (hypokinesia without LV dilatation), isolated LV dilation (LV dilation without hypokinesia), or arrhythmia or conduction disturbances.[23]

Phenotypic definitions were on the basis of a combination of clinical diagnosis (self-reported conditions and ICD-10 codes), procedures (self-reported and OPCS-4 codes), 12-lead ECGs, and CMR imaging (where available). A full list of phenotype definitions is shown in Table S1 and is adapted from definitions used elsewhere.[24–26] The observed phenotype was graded to clinically diagnosed DCM; early DCM features, including arrhythmia or cardiac conduction disease

(CCD), isolated ventricular dilation, and subclinical DCM; or phenotype-negative. Clinical DCM was defined by ICD-10 code I42.0; subclinical DCM was defined by fulfillment of the CMR criteria for DCM in the absence of a clinical history of DCM. In the classification of phenotype, the prioritization of phenotype categories was as follows: clinical DCM > subclinical DCM > hypokinetic nondilated cardiomyopathy > isolated ventricular dilatation > arrhythmia or CCD. For example, in the presence of ICD-10 code I42.0, the patient was considered to have clinically diagnosed DCM regardless of other history features and ECG or CMR features. The diagnosis of hypokinetic nondilated cardiomyopathy, isolated ventricular dilation, and subclinical DCM derived from analysis of CMR data. The phenotype category "arrhythmia or CCD" was defined as atrial fibrillation/flutter, bradyarrhythmia, CCD, preexcitation syndrome, or ventricular arrhythmia.

The penetrance and outcome analyses were stratified on the basis of gene-evidence clusters as defined by the ClinGen DCM Gene Curation Expert Panel.[4] Genes were clustered in the following categories: definitive/strong, moderate, or limited evidence.

Analysis of Genetic Yield in the Clinical DCM Population (Quality Control)

Patients with clinical DCM were identified from the UK Biobank population using ICD-10 code I42.0. Patients without clinically significant coronary artery disease were included; those with myocardial infarction or revascularization[26] were excluded. Genetic yield was determined for ClinGen DCM Gene Curation Expert Panel–asserted DCM-associated genes and classified according to evidence category,[4] using the same filtering strategies as described previously.

Statistical Analysis

Statistical analysis was performed with R statistical computing and graphics software, version 3.6.1,[27] using *tidyverse*[28] and *tableone*[29] packages. Continuous, normally distributed data are summarized as mean (SD) and nonnormally distributed data as median (interquartile range). Continuous data were compared using a 2-sample *t* test and categorical data using a χ^2 test to test for differences between genotype-positive and genotype-negative individuals. Details regarding outcome analysis are provided in the Expanded Methods in the Supplemental Material.

RESULTS

Quality Control of Variant Filtering Strategy on the Basis of Clinical DCM Population

Among 502 462 UK Biobank participants, there were 1415 (0.28%) individuals with the known clinical diagnosis of nonischemic DCM (30.2% female, mean age 59.8±7 years at enrollment). Table S2 shows the demographic characteristics of these patients. Among patients with DCM, 340 (24%) individuals underwent WES. Screening of genes ascertained to have at least limited evidence for causality in monogenic DCM revealed putative pathogenic variants in 55 (16%) patients (Figure S1). In accordance with previous observations,[30–32] truncating variants in the *TTN* gene were the most common (n = 17; 31% of genotype-positive DCM cases, 5% of all genotyped DCM cases), followed by *DSP* variants >1 putative pathogenic variants (n = 5 for each; 9.1% of genotype-positive DCM cases, 1.5% of all genotyped DCM cases). These observations validate our primary variant filtering strategy as one in line with that applied in clinical practice. Genetic test results in the clinical DCM subset using secondary variant filtering strategies are summarized in Figure S1.

Study Population

Out of 502 462 participants in the UK Biobank (54.4% female), 200 619 had undergone WES; 42 078 had 12-lead ECG, 39 616 had CMR. Given the staged approach to participant accrual, 18 665 participants had WES, 12-lead ECG, and CMR, forming the study population (52.7% female; 96.8% White; average age 55 years at recruitment and 64.4 years at last follow-up; Figure 1). Arrhythmia or cardiac conduction disease was present in 2729 (14.6%), isolated ventricular dilation in 522 (2.8%), hypokinetic nondilated cardiomyopathy in 645 (3.5%), and clinical/subclinical DCM in 189 (1%) participants (Table 1).

Prevalence of DCM-Associated Putative Pathogenic Variants in the UK Biobank

Among 18 665 individuals, 1463 (7.8%) were found to host at least 1 putative pathogenic variant in DCM-associated genes using the primary variant filtering strategy (Figure 2). Putative pathogenic variants were found in all 44 screened genes, and most frequently affected *OBSCN* (n = 153 [10.5% of all DCM genotype-positives]), *MYH6* (n = 149 [10.2%]), *SCN5A* (n = 140 [9.6%]), *MYH7* (n = 122 [8.3%]), *FLNC* (n = 121 [8.3%]), *MYBPC3* (n = 46 [3.1%]), and *TTN* genes (n = 44 [3%]). There were 30 individuals with *LMNA* variants (2%). Sixty-five individuals (4.4%) carried 2 or more putative pathogenic variants in the same or different genes. The prevalence of putative pathogenic variants according to secondary filtering strategies is provided in Figure S2.

Clinical Disease Penetrance of DCM-Associated Putative Pathogenic Variants in the UK Biobank

Among 1463 putative pathogenic variant carriers, 5 individuals had a clinical diagnosis of DCM; 14 additional individuals diagnosed with DCM did not host any putative pathogenic variants (Table 2). Those with putative pathogenic variants more frequently had heart failure (2.1 versus 1.2%; $P = 0.01$), but the risk of developing heart failure was not different between groups (hazard ratio, 1.46 [CI, 0.96–2.24]). Patients with putative pathogenic

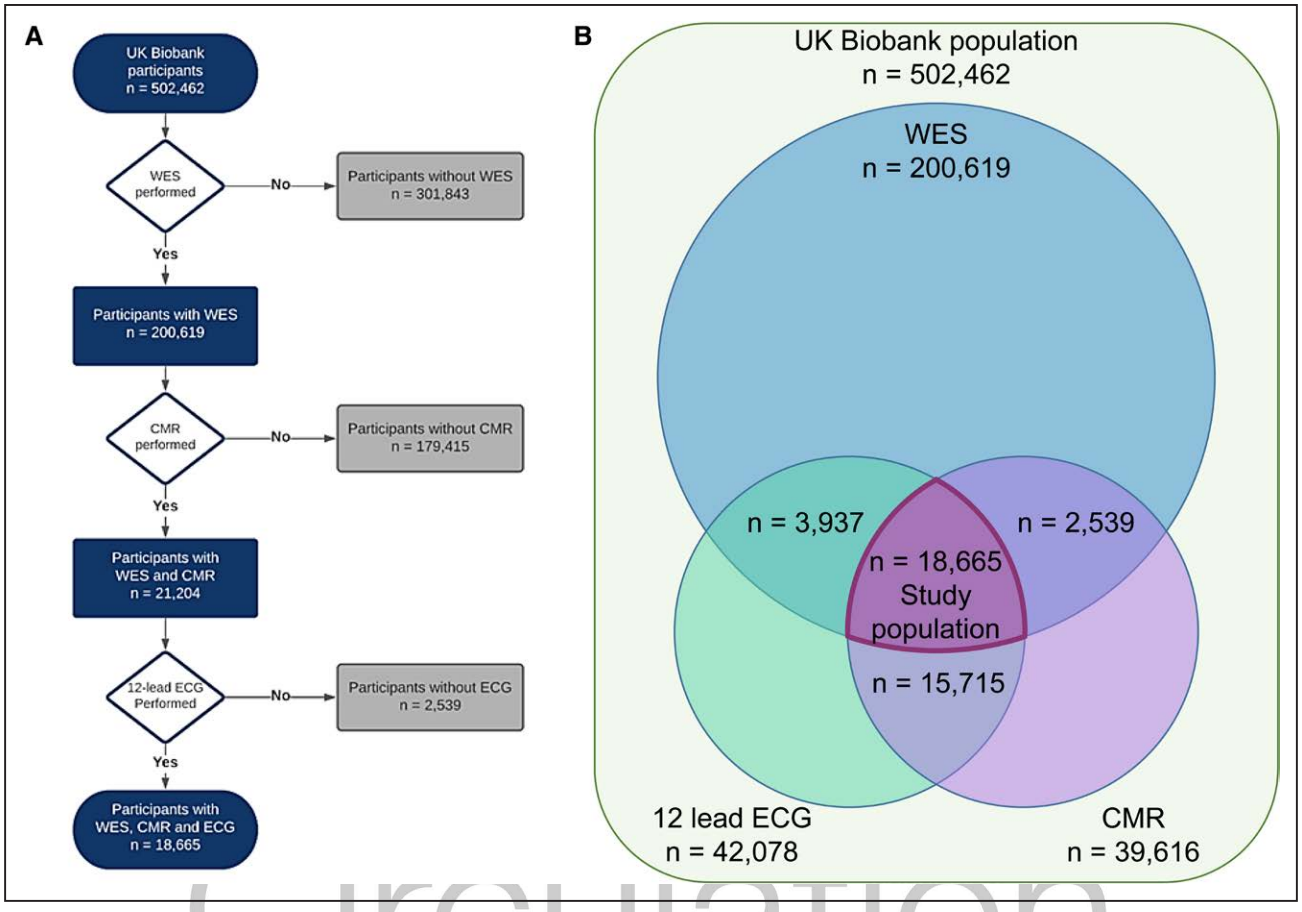

Figure 1. Study population selection criteria.
A, Flowchart demonstrating the sequential inclusion/exclusion criteria for the study population. **B**, Venn diagram showing the number of participants within the whole UK Biobank population with whole exome sequencing (WES), 12-lead ECG, and cardiac magnetic resonance (CMR) imaging.

variants did not show any difference in LV ejection fraction, LV end systolic volume, LV end diastolic volume, age at recruitment, death at follow-up, or age at death in comparison with those without (Table 3). A comparison of demographic and clinical characteristics of participants with and without putative pathogenic variants according to secondary variant filtering strategies is provided in Table S3.

Subclinical DCM and Early DCM Features in Individuals With Putative Pathogenic Variants

When assessed on the basis of the CMR data, 24 (1.6%) additional individuals with putative pathogenic variants met the diagnostic criteria for DCM (subclinical DCM). There were no differences in the frequency of early DCM features between genotype-positive and genotype-negative groups.

Combined Clinical and Subclinical Penetrance for DCM and Early DCM-Associated Features

Overall, 346 (23.7%) individuals with putative pathogenic variants had DCM or showed early phenotypic features that may in part be attributed to DCM, most frequently arrhythmia or cardiac conduction disease (n = 223 [64%]). The most common phenotypes within this category were first-degree heart block (n = 89), QRS duration of >110 ms (n = 81), atrial fibrillation/flutter (n = 67), and complete right bundle-branch block (n = 53). The overall penetrance of putative pathogenic variants combined for subclinical/clinical DCM and early DCM features varied between 0% and 66.7% in those with single putative pathogenic gene variants and between 0% and 100% in those with 2 or more putative pathogenic variants. Overall, individuals with putative pathogenic variants more frequently developed a clinical or subclinical DCM phenotype, as compared with those without putative pathogenic variants (2% versus 1%; $P = 0.00073$; Table 3). Individuals with 2 or more putative pathogenic variants did not demonstrate significantly different penetrance compared with those with a single variant ($P = 0.873$).

Penetrance Analysis on the Basis of Gene-Evidence Category

A gene-evidence cluster-based analysis revealed slightly higher frequency of clinical DCM in participants with putative pathogenic variants in definitive/strong evidence

Table 1. Demographic and Clinical Characteristics of the Study Population and of Phenotypic Subgroups

Characteristics	Overall	Phenotype-negative	Arrhythmia or cardiac conduction defect	Isolated ventricular dilation	Hypokinetic nondilated cardiomyopathy	Dilated cardiomyopathy*
No. of participants	18665	14580	2729	522	645	189
Female	9844 (52.7)	8345 (57.2)	938 (34.4)	281 (53.8)	195 (30.2)	85 (45.0)
Age at recruitment, y	54.96 (7.50)	54.33 (7.44)	57.33 (7.31)	55.45 (6.86)	57.62 (7.33)	58.70 (6.98)
Age at last follow-up, y	64.38 (7.47)	63.78 (7.41)	66.72 (7.32)	64.82 (6.81)	66.96 (7.25)	67.99 (6.96)
Ethnic background						
White	18057 (96.8)	14075 (96.6)	2650 (97.2)	513 (98.3)	634 (98.3)	185 (97.9)
Asian or Asian British	202 (1.1)	168 (1.2)	30 (1.1)	1 (0.2)	3 (0.5)	0 (0.0)
Black, Black British, Caribbean, or African	113 (0.6)	92 (0.6)	14 (0.5)	2 (0.4)	2 (0.3)	3 (1.6)
Other ethnic group	90 (0.5)	71 (0.5)	14 (0.5)	1 (0.2)	4 (0.6)	0 (0.0)
Mixed or multiple ethnic groups	87 (0.5)	74 (0.5)	7 (0.3)	4 (0.8)	2 (0.3)	0 (0.0)
Chinese	55 (0.3)	49 (0.3)	5 (0.2)	0 (0.0)	0 (0.0)	1 (0.5)
Not specified	55 (0.3)	47 (0.3)	7 (0.3)	1 (0.2)	0 (0.0)	0 (0.0)
Systolic BP automated reading, mm Hg	136.75 (18.66)	135.84 (18.37)	140.70 (19.33)	140.12 (19.75)	136.80 (18.27)	140.23 (20.10)
Diastolic BP automated reading, mm Hg	81.25 (10.37)	81.07 (10.27)	82.36 (10.63)	79.26 (10.59)	82.34 (10.76)	81.20 (11.24)
Body mass index, kg/m^2	26.51 (4.16)	26.43 (4.13)	26.89 (4.21)	25.90 (4.20)	27.06 (4.30)	27.09 (4.39)
Body surface area, m^2	1.86 (0.21)	1.85 (0.20)	1.92 (0.21)	1.83 (0.19)	1.93 (0.22)	1.88 (0.21)
Creatinine, μmol/L	72.12 (14.09)	71.34 (13.88)	75.68 (14.67)	70.44 (13.25)	76.46 (14.30)	71.92 (13.00)
Age at death, y	70.73 (6.71)	69.34 (6.94)	72.32 (5.66)	72.12 (6.71)	74.36 (6.57)	73.60 (4.56)
Coronary artery disease	846 (4.5)	429 (2.9)	281 (10.3)	32 (6.1)	72 (11.2)	32 (16.9)
Heart failure	240 (1.3)	60 (0.4)	96 (3.5)	20 (3.8)	40 (6.2)	24 (12.7)
Hypertension	6028 (32.3)	4324 (29.7)	1192 (43.7)	168 (32.2)	254 (39.4)	90 (47.6)
Stroke	330 (1.8)	213 (1.5)	88 (3.2)	13 (2.5)	13 (2.0)	3 (1.6)
Valvular heart disease	527 (2.8)	229 (1.6)	197 (7.2)	43 (8.2)	45 (7.0)	13 (6.9)
Atrial fibrillation/flutter	832 (4.5)	0 (0.0)	646 (23.7)	41 (7.9)	124 (19.2)	21 (11.1)
Bradyarrhythmia	352 (1.9)	0 (0.0)	273 (10.0)	26 (5.0)	31 (4.8)	22 (11.6)
Cardiac conduction defect	2448 (13.1)	0 (0.0)	2141 (78.5)	125 (23.9)	123 (19.1)	59 (31.2)
Sick sinus syndrome	19 (0.1)	0 (0.0)	13 (0.5)	5 (1.0)	1 (0.2)	0 (0.0)
Cardiac implantable electronic device†	148 (0.8)	0 (0.0)	120 (4.4)	16 (3.1)	8 (1.2)	4 (2.1)
Preexcitation syndrome	35 (0.2)	0 (0.0)	32 (1.2)	1 (0.2)	2 (0.3)	0 (0.0)
Ventricular arrhythmia	69 (0.4)	0 (0.0)	56 (2.1)	4 (0.8)	6 (0.9)	3 (1.6)

Values are n (%) or mean (SD). BP indicates blood pressure.
*Dilated cardiomyopathy was determined by magnetic resonance imaging data only.
†These devices include single- and dual-chamber permanent pacemaker, implantable cardioverter defibrillator, and cardiac resynchronization therapy.

genes as compared with those with variants in moderate/limited evidence genes. However, combined clinical and subclinical phenotype was not statistically different between groups stratified on the basis of gene-evidence category (Table 4).

Prevalence of DCM-Associated Putative Pathogenic Gene Variants Using Primary and Secondary Variant Calling Strategies

One or more putative pathogenic variants in DCM-associated genes were identified, in 1463 (7.8%) with the missense predicted loss-of-function FAF variant calling strategy, in 154 (0.8%) using the InterVar FAF strategy, and in 212 (1.1%) using the InterVar FAF ClinVar strategy (Table S4). The rate of diagnosis of clinical DCM varied between 0.3 and 1.4%. Early clinical features of DCM (5.7% to 7.6%) were present in <8% of individuals carrying putative pathogenic variants; an additional 16.89% to 17.6% had a subclinical phenotype on ECG or CMR using different variant calling strategies. With a combined clinical/subclinical prevalence of 12.3% to 15.6%, arrhythmias or conduction disease were the most common early DCM features, followed by hypokinetic nondilated cardiomyopathy, observed in 4.0% to 7.8%. The combined clinical/subclinical penetrance was ≤30% for all variant filtering strategies (Table 2 and Table S4).

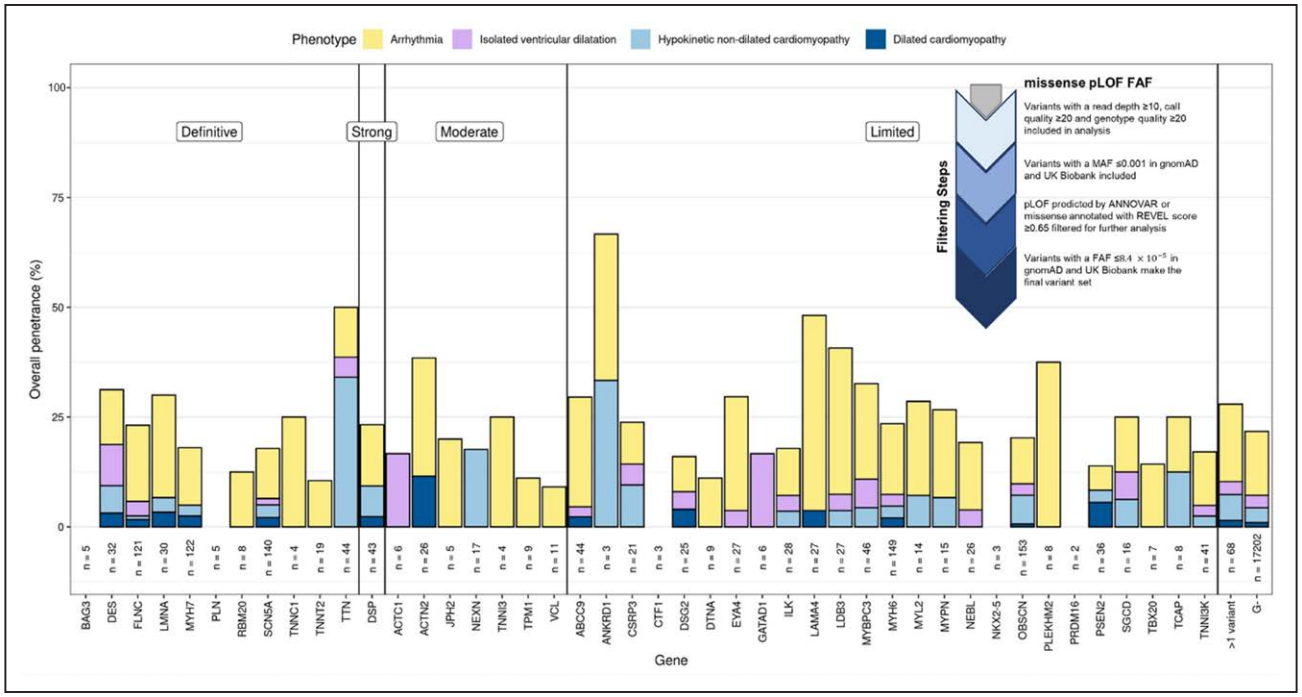

Figure 2. Clinical and subclinical penetrance of putative pathogenic variants in dilated cardiomyopathy–associated genes in middle-aged and older adults.

For each dilated cardiomyopathy (DCM)–associated gene, the height of the bar indicates the percentage of pathogenic/likely pathogenic variant carriers, by missense predicted loss-of-function (pLOF) filtering allele frequency (FAF) filtering strategy, with the specified phenotypes. Total number of participants with a pathogenic or likely pathogenic variant for each DCM-associated gene is indicated below the bar. The phenotype prevalence in those without a pathogenic or likely pathogenic mutation is shown on the far right (labeled "G–"). The phenotype prevalence in those with >1 putative pathogenic variant is shown in the second from right column. Genes are categorized according to the strength of evidence determined by the ClinGen DCM gene curation expert panel and ordered alphabetically within each category. MAF indicates mean allele frequency.

Gene-Based Analysis of Penetrance and Clinical Phenotype

Putative pathogenic variants were found in all 44 screened genes by the missense predicted loss-of-function FAF strategy and in a smaller number of genes when using the more strict secondary strategies. The gene-specific penetrance ranged from 0 to 66.7% (Figure 2). A subanalysis for *TNNT2*, *LMNA*, *FLNC*, and *TTN*, which are listed in the American College of Medical Genetics and Genomics clinically actionable gene list, revealed DCM or early DCM features in 10.5%, 26.7%, 22.3%, and 45.4% of cases, respectively, indicating higher penetrance than in the overall putative pathogenic population. Of the 44 individuals with *TTN* truncating variants, 31 (70.4%) were found in the A band (Table S5). Penetrance was 45.2% in A band versus 61.5% in non–A band *TTN* variant carriers. Participants with *TTN* variants in the A band did not show a statistically significant difference in phenotype, but there was a trend toward a lower penetrance of early DCM features (32% versus 61.5%; $P = 0.08$). All 4 clinical diagnoses of DCM within the *TTN* group were found among the participants with *TTN* variants in the A band. Individuals with A band variants had higher left ventricular end-systolic volumes (78.4 versus 62.9 mL; $P = 0.036$) and a trend toward increased end diastolic volumes (148.1 versus 130.0 mL; $P = 0.073$). A comparison of individuals with putative pathogenic variants in arrhythmogenic cardiomyopathy genes (*BAG3*, *DES*, *FLNC*, *LMNA*, *PLN*, *RBM20*, *SCN5A*, *DSP*, *DSG2*, *LDB3*, *NKX2-5*), as defined by the 2019 Heart Rhythm Society consensus statement,[33] compared with other DCM genes did not reveal any significant differences in arrhythmia or CCD phenotype (Table S6). Gene-based analysis of penetrance for additional filtering strategies is provided in Figure S2 and Tables S7 through S9.

Effect of DCM-Associated Putative Pathogenic Variants on Clinical Outcomes

To assess the effect of DCM-associated putative pathogenic variants on the clinical outcome, an event-free survival analysis was performed for genotype-positive versus genotype-negative patients for each of the 3 variant filtering strategies. Event-free survival was defined as survival without developing heart failure, stroke, or arrhythmia; requiring a cardiac implantable electronic device; or death. There was no statistical difference in survival between individuals with and without putative pathogenic variants (hazard ratio, 1.06 [95% CI, 0.87–1.29] for the primary strategy; Figure S4). Additional analysis regarding prevalence and incidence of

Table 2. Clinical and Subclinical Penetrance of DCM and DCM-Associated Clinical Features in Participants With Putative Pathogenic DCM Variants

Phenotype and filter	Overall	Phenotype-negative	Phenotype-positive			
			Early DCM features			
			Arrhythmia or cardiac conduction disease	Isolated ventricular dilatation	Hypokinetic nondilated cardiomyopathy	DCM
Clinical						
Total	18665	17545 (94)	1101 (5.9)	NA	NA	19 (0.1)
Missense pLOF FAF	1463 (7.8)	1374 (93.92)	84 (5.74)	NA	NA	5 (0.34)
Subclinical (ECG + CMR)						
Total	18665	15056 (80.66)	2253 (12.07)	522 (2.8)	645 (3.46)	189 (1.01)
Missense pLOF FAF	1463 (7.8)	1147 (78.4)	194 (13.26)	35 (2.39)	63 (4.31)	24 (1.64)
Combined						
Total	18665	14578 (78.1)	2725 (14.6)	522 (2.8)	635 (3.4)	205 (1.1)
Missense pLOF FAF	1463 (7.8)	1117 (76.35)	223 (15.24)	35 (2.39)	59 (4.03)	29 (1.98)

Values are n (%). The number and proportion for each phenotype are shown for those with a pathogenic or likely pathogenic variant by the missense pLOF FAF filtering strategy and in the total study population for comparison. The values in the "Overall" column represent the percentage of the total cohort with a putative pathogenic variant (G+) as determined by the relevant filtering strategy. The percentages in the remainder of the table represent the proportion of G+ individuals with the indicated phenotype. CMR indicates cardiac magnetic resonance; DCM, dilated cardiomyopathy; ECG, electrocardiography; FAF, filtering allele frequency; NA, not applicable; and pLOF, predicted loss-of-function variant.

DCM and potentially DCM-associated clinical features in genotype-positive and genotype-negative groups are provided in Tables S10 and S11.

DISCUSSION

Using the UK Biobank WES data, we analyzed the prevalence and penetrance of DCM-associated gene variants in a cohort of >18 000 individuals with 12-lead ECG and CMR. This is the first study providing insights into the clinical and subclinical penetrance of DCM-associated gene variants in a large population-scale dataset and has several important findings with direct clinical implications. First, the UK Biobank population of mainly middle-aged adults has a prevalence of nonischemic DCM of 1:355 (0.28%). Second, the variant filtering strategy used for those with a clinical diagnosis of DCM provided a yield of 16%. Third, using the same strategy and a genotype-first approach identified 1463 (7.8%) individuals with putative pathogenic DCM gene variants. Fourth, clinical/subclinical disease penetrance was highly variable, ranging from 0 to 66.7% between genes. Fifth, among individuals with putative pathogenic DCM variants, subclinical DCM and early DCM features, detected by 12-lead ECG or CMR, were 5 times more common than clinically manifest DCM (21.6% versus 3.8%; $P < 0.00001$; Figure 3). Last, participants with putative pathogenic variants in definitive/strong DCM genes appeared to have a slightly higher rate of clinical DCM than those with variants in lower evidence genes, but combined clinical and subclinical phenotype was not statistically different between groups stratified on the basis of gene-evidence categories.

Our analysis indicates a DCM prevalence of 1:355 in the UK Biobank population, with >3.3:1 male predominance, in line with previous population-based epidemiologic studies of nonischemic DCM.[34,35] Studies reporting a prevalence of 1 in 250 have perhaps had Black participants as well (in whom the rates of DCM are known to be high) or might have included other causes of a morphologic DCM phenotype. Among individuals with putative pathogenic variants in the screened DCM-associated genes, only 0.3% had a clinically diagnosed DCM, but an additional 1.6% (5-fold increase) were found to meet the CMR criteria for DCM, indicating that most DCM cases in the general adult population go unnoticed for many years. The 12-lead ECG and CMR screening in those with putative pathogenic variants identified 16.2% of individuals with subclinical early features of DCM, such as cardiac arrhythmias, isolated ventricular dilation, and hypokinetic nondilated DCM. The population with early DCM features was 11-fold larger than those with clinical and subclinical DCM combined (21.7% versus 2%), demonstrating the wide variability of phenotypic expression in individuals with putative pathogenic variants in DCM genes. Arrhythmias and CCD were the most common early manifestation of putative pathogenic variants implicated in DCM, mostly occurring before the development of structural or functional abnormalities. These findings are in line with observations in the large and well-phenotyped Geisinger database, showing very low penetrance of arrhythmogenic right ventricular cardiomyopathy-associated *PKP2* variants in the middle-aged population.[36]

An unsolved challenge in clinical genetics is that the principles applied to variant calling are probabilistic. Whereas the 2015 American College of Medical Genetics and Genomics guidelines integrate a robust data scheme to classify gene variants,[37] applying these

Table 3. Demographic and Clinical Characteristics of the Overall Study Population, Putative Pathogenic Variant Carriers (G+), and Those Without any Putative Pathogenic Variants in DCM-Associated Genes (G–) for the Primary Variant Filtering Strategy

Characteristics	Overall	Missense pLOF FAF		
		G–	G+	*P*
No. of participants	18665	17202	1463	
Female	9844 (52.7)	9076 (52.8)	768 (52.5)	0.866
Age at recruitment, y	54.96 (7.50)	54.94 (7.49)	55.16 (7.56)	0.283
Age at last follow-up, y	64.38 (7.47)	64.37 (7.46)	64.53 (7.58)	0.436
Ethnic background				
White	18057 (96.8)	16686 (97.0)	1371 (93.7)	
Asian or Asian British	202 (1.1)	163 (0.9)	39 (2.7)	
Black, Black British, Caribbean, or African	113 (0.6)	103 (0.6)	10 (0.7)	
Other ethnic group	90 (0.5)	74 (0.4)	16 (1.1)	
Mixed or multiple ethnic groups	87 (0.5)	79 (0.5)	8 (0.5)	
Not specified	55 (0.3)	50 (0.3)	5 (0.3)	
Chinese	55 (0.3)	41 (0.2)	14 (1.0)	
Systolic BP automated reading, mm Hg	136.75 (18.66)	136.75 (18.67)	136.67 (18.48)	0.866
Diastolic BP automated reading, mm Hg	81.25 (10.37)	81.27 (10.39)	81.09 (10.14)	0.544
Body mass index, kg/m^2	26.51 (4.16)	26.51 (4.17)	26.53 (4.07)	0.873
Body surface area, m^2	1.86 (0.21)	1.86 (0.21)	1.86 (0.21)	0.865
Creatinine, μmol/L	72.12 (14.09)	72.11 (14.06)	72.31 (14.45)	0.601
Heart failure	240 (1.3)	210 (1.2)	30 (2.1)	0.01
Early DCM features				0.14
Arrhythmia or CCD	2729 (14.6)	2505 (14.6)	224(15.3)	
Isolated ventricular dilation	522 (2.8)	487 (2.8)	35 (2.4)	
Hypokinetic nondilated cardiomyopathy	645 (3.5)	582 (3.4)	63 (4.3)	
DCM overall	205 (1.1)	176 (1.0)	29 (2.0)	0.00073
Subclinical	189 (1.0)	165 (1.0)	24 (1.6)	
Clinical	19 (0.1)	14 (0.1)	5 (0.3)	
LVESV, mL	63.46 (63.75)	63.50 (65.77)	62.92 (31.60)	0.737
LVEDV, mL	140.72 (136.83)	140.91 (141.97)	138.57 (43.22)	0.53
LVEF	55.67 (6.62)	55.69 (6.60)	55.39 (6.81)	0.094
Hypertension	6028 (32.3)	5574 (32.4)	454 (31.0)	0.295
Stroke	330 (1.8)	301 (1.7)	29 (2.0)	0.586
Valvular heart disease	527 (2.8)	487 (2.8)	40 (2.7)	0.894
Atrial fibrillation or flutter	832 (4.5)	765 (4.4)	67 (4.6)	0.865
Bradyarrhythmia	352 (1.9)	323 (1.9)	29 (2.0)	0.856
Cardiac conduction defect	2448 (13.1)	2238 (13.0)	210 (14.4)	0.155
First-degree heart block	971 (5.2)	882 (5.1)	89 (6.1)	0.129
Left anterior fascicular block	222 (1.2)	209 (1.2)	13 (0.9)	0.327
Left posterior fascicular block	1 (0.0)	1 (0.0)	0 (0.0)	1
Unspecified fascicular block	4 (0.0)	4 (0.0)	0 (0.0)	1
Left bundle-branch block	103 (0.6)	95 (0.6)	8 (0.5)	1
Right bundle-branch block	684 (3.7)	631 (3.7)	53 (3.6)	0.987
Bifascicular block	50 (0.3)	48 (0.3)	2 (0.1)	0.455
First degree and bifascicular block	4 (0.0)	4 (0.0)	0 (0.0)	1
Nonspecific intraventricular block	107 (0.6)	96 (0.6)	11 (0.8)	0.446

(*Continued*)

Table 3. Continued

Characteristics	Overall	Missense pLOF FAF		
		G−	G+	*P*
QRS >110 ms	911 (4.9)	830 (4.8)	81 (5.5)	0.25
Sick sinus syndrome	19 (0.1)	17 (0.1)	2 (0.1)	0.993
Cardiac implantable electronic devices*	148 (0.8)	133 (0.8)	15 (1.0)	0.373
Preexcitation syndrome	35 (0.2)	32 (0.2)	3 (0.2)	1
Ventricular arrhythmia	69 (0.4)	62 (0.4)	7 (0.5)	0.624
eGFR, mL/min/1.73 m^2	87.32 (16.24)	87.34 (16.36)	87.10 (14.82)	0.609
CKD ≥3	396 (2.2)	365 (2.2)	31 (2.2)	1

Values are n (%) or mean (SD). BP indicates blood pressure; CCD, cardiac conduction disease; CKD, chronic kidney disease; DCM, dilated cardiomyopathy; eGFR, estimated glomerular filtration rate; FAF, filtering allele frequency; LVEDV, left ventricular end diastolic volume; LVESV, left ventricular end systolic volume; LVEF, left ventricular ejection fraction; and pLOF, predicted loss-of-function variant.

*These devices include single- and dual-chamber permanent pacemaker, implantable cardioverter defibrillator, and cardiac resynchronization therapy.

criteria to biobank studies is challenging because expert–curator input is often necessary to assess variant pathogenicity and gene–disease association. Here, we applied 3 special variant calling models developed for this study to all genes, except for *TTN*, where we only considered radical variants as putative pathogenic variants. The 2 secondary filtering strategies we used can be considered more stringent than the primary strategy; these yielded a lower frequency of variants in the studied population with a similar low penetrance, indicating that the true DCM penetrance is very low regardless of variant calling strategy. On the other hand, many genes, such as *SCN5A*, *DSP*, *MYH7*, *DES*, and others, show broad pleiotropy and variability of phenotypic expression, making the definition of positive phenotype in large datasets complicated. To extend the breadth of recognizable phenotypes, we included early features of DCM. However, it is possible that the individuals with putative pathogenic variants in DCM genes, who showed no DCM features, manifest other cardiomyopathy phenotypes, or in the case of *SCN5A*, primary arrhythmia syndromes.[38]

Of the genes included in the ClinGen DCM panel used in this study to define a DCM-relevant gene panel, *TNNT2*, *LMNA*, *FLNC*, and *TTN* are among the 73 genes identified as medically actionable by the American College of Medical Genetics and Genomics,[13] for which clinical management guidelines have been established. In the UK Biobank subcohort, individuals with putative pathogenic variants in the *TNNT2*, *LMNA*, *FLNC*, and *TTN* genes showed signs of DCM or early DCM features in 10.5%, 26.7%, 22.3%, and 45.4% of cases, respectively. This is a significantly higher penetrance than that seen in the overall DCM gene panel.

Although individuals with putative pathogenic variants in DCM genes had a markedly higher observed frequency of clinical/subclinical DCM and twofold higher frequency of heart failure, our analysis reflects low penetrance of DCM-associated putative pathogenic variants in middle-aged to older individuals, indicating that most of these individuals are unlikely to develop disease. The presence of a putative pathogenic variant in the DCM-associated genes was not associated with a higher risk of heart failure or worse outcome, similar to the findings of Carruth et al[36] for pathogenic alleles in the *PKP2* gene. It should be noted that monogenic disease expressivity and penetrance, particularly in DCM, is dependent on genomic context, as indicated by family history or the effect of polygenic risk,[39,40] as well as the environment (eg, alcohol and pregnancy in *TTN* truncation variant–mediated DCM, inflammation in *DSP*-mediated cardiomyopathy).[41–44] Future studies investigating population penetrance of DCM should consider disease modifiers and approach to DCM as a multifactorial trait highly influenced by environmental and genetic modifiers.

A population-based genotype-first screening strategy must fulfill certain criteria to be cost-effective.[45] First, the condition should be a sufficient health problem, which DCM is. Second, the natural history should be understood. In DCM, this is partly understood, but as multimodality imaging and physician awareness improve, we have learned that the natural history can be different for this genetically heterogeneous group. Third, there should be a recognizable latent or early symptomatic stage, which for heart failure, isolated ventricular dilation, and hypokinetic nondilated cardiomyopathy is applicable to DCM. However, for those in whom the sentinel event is sudden death, this is not fulfilled. Fourth, there should be a suitable test or examination, which is fulfilled by echocardiography and CMR. Fifth, screening should be acceptable to the population, and, in general, screening for cardiovascular diseases is. Sixth and seventh, there should be an agreed policy on whom to treat and expected treatment, both of which are well satisfied regarding established heart failure syndrome and arrhythmia, but not agreed on asymptomatic mild or subclinical disease. Eighth, facilities for diagnosis and treatment should be available, which has been achieved throughout most of the developed

Table 4. Clinical and Subclinical Phenotype in Participants With Putative Pathogenic Variants in Dilated Cardiomyopathy Genes According to the Predicted Loss-of-Function Variant Missense Filtering Allele Frequency Classification, Stratified on the Basis of Gene–Evidence Class

Characteristics	Definitive/ strong	Limited	Moderate	*P* value	
				Comparing definitive/ strong versus moderate versus limited	Comparing definitive/ strong versus combined moderate/limited
Participants	608	770	85		
Clinical phenotype				0.133	0.029
Phenotype-negative	568 (93.4)	726 (94.3)	80 (94.1)		
Arrhythmia or CCD	35 (5.8)	44 (5.7)	5 (5.9)		
Dilated cardiomyopathy	5 (0.8)	0 (0.0)	0 (0.0)		
Subclinical phenotype (ECG + CMR)				0.542	0.488
Phenotype-negative	480 (78.9)	601 (78.1)	66 (77.6)		
Arrhythmia or CCD	74 (12.2)	109 (14.2)	11 (12.9)		
Isolated ventricular dilation	12 (2.0)	22 (2.9)	1 (1.2)		
Hypokinetic nondilated cardiomyopathy	31 (5.1)	28 (3.6)	4 (4.7)		
Dilated cardiomyopathy	11 (1.8)	10 (1.3)	3 (3.5)		
Combined phenotype				0.449	0.337
Phenotype-negative	468 (77.0)	584 (75.8)	65 (76.5)		
Arrhythmia or CCD	85 (14.0)	126 (16.4)	12 (14.1)		
Isolated ventricular dilatation	12 (2.0)	22 (2.9)	1 (1.2)		
Hypokinetic nondilated cardiomyopathy	27 (4.4)	28 (3.6)	4 (4.7)		
Dilated cardiomyopathy	16 (2.6)	10 (1.3)	3 (3.5)		

Values are n (%). For the purpose of analysis, patients with >1 putative pathogenic variant in genes classified in different evidence categories were considered in the higher evidence category. CCD indicates cardiac conduction disease; CMR, cardiac magnetic resonance; and ECG, electrocardiography.

world, although disparities persist and remain lacking in developing nations.[46] Ninth, the cost of case finding, diagnosis, and treatment should be economically balanced, which with clinical diagnosis is justified, but for genetics, is difficult to justify given upfront costs and the low signal-to-noise ratio. This may be improved with a targeted panel of very high-risk genes and higher penetrance. In our study, only *TTN* truncation variant would meet this criteria, whereas genes considered as high risk in consensus statements and guidelines, such as *DSP*, *LMNA*, *RBM20*, *SCN5A*, *FLNC*, and *PLN*,[33] had a low frequency of putative pathogenic variants and low penetrance. On the basis of current data, a genotype-first screening strategy in DCM would be difficult to justify owing to high cost and low penetrance. However, this should be interpreted in the context of this dataset being a relatively healthy and older population, where high-risk individuals may have experienced sudden death or not enrolled. Last, case finding should be a continuous process, and applying this to clinical screening for presentations of DCM may be better achieved through phenotype screening tools such as artificial intelligence–enabled ECG.[47] Years later, this method may be a more cost-effective approach than a genetic screening tool and continuous clinical phenotyping, which are prohibitively expensive.

Current standard of care in families with DCM is both clinical and genetic cascade screening in all first-degree relatives when a putative pathogenic gene variant is identified in the proband. Disease penetrance in genotype-positive family members of probands with DCM is significantly higher than the penetrance observed in our study examining a population-based cohort of individuals not selected because of cardiac disease.

Strengths and Limitations

Our study has a number of strengths. First, we used a robust methodology with 1 primary and 2 secondary variant calling strategies, which all confirmed the low yield of clinical and subclinical DCM in putative pathogenic variant carriers. Second, in order to include early DCM features as defined by Pinto et al.,[23] as well as subclinical DCM, we analyzed electronic health records, 12-lead ECG, and CMR data to enable deeper phenotyping. This is the first study to report detailed high-throughput computer interpretation, with manual validation, of 12-lead ECGs from the UK Biobank population. Third, our study included a cohort of >18 000 individuals who underwent CMR according to a standardized protocol.

The study results should be viewed in light of several limitations. First, some DCM genes show substantial pleiotropy with the development of distinct phenotypes. We did not consider phenotypic features other than DCM or early DCM features. This may have resulted in

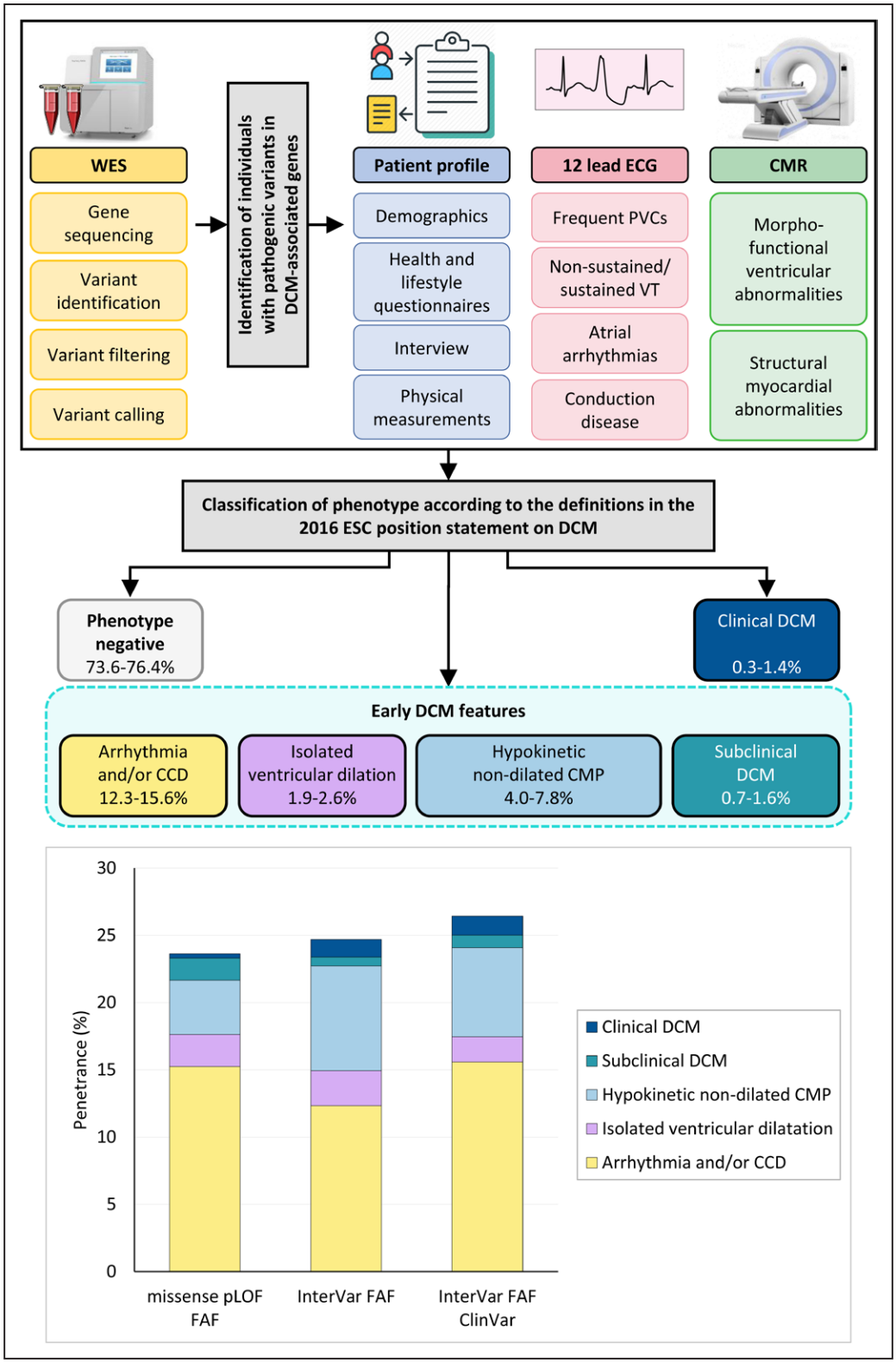

Figure 3. Central illustration of the study, demonstrating the methodology used for phenotype ascertainment, classification, and clinical/subclinical phenotype in participants with putative pathogenic variants in dilated cardiomyopathy genes according to the 3 variant filtering strategies applied.

CCD indicates cardiac conduction disease; CMP, cardiomyopathy; CMR, cardiac magnetic resonance; DCM, dilated cardiomyopathy; ESC, European Society of Cardiology; FAF, filtering allele frequency; pLOF, predicted loss-of-function variant; PVC, premature ventricular contractions; VT, ventricular tachycardia; and WES, whole exome sequencing.

underestimation of the clinical penetrance of these genes when it relates to other phenotypes. Second, the UK Biobank population may reflect volunteer bias and survivor bias with a sample of healthier individuals than the general UK population,[48] and thus may show lower frequency of putative pathogenic variants and lower penetrance. Third, the ethnicity is mainly White British individuals, making generalizability to other ethnicities challenging, particularly in a disease with known differences in genetic pathogenesis in different races.[49,50] Fourth, although 200 000 WES data are available, only 18 000 had both usable 12-lead ECG and CMR available; with the staged approach, and effects of the COVID-19 pandemic, the goals of 500 000 WES and 100 000 with 12-lead ECG and CMR are unlikely to be achieved on time. A larger sample of 100 000 and with longer follow-up may show higher penetrance, given that the age at onset of DCM is variable. In addition, the individuals invited for 12-lead ECG and CMR were determined by proximity to the testing centers; this may have introduced a zip code bias. The UK Biobank does not include recruits >40 years of age, which may underestimate prevalence and exclude patients with high-risk DCM who die of sudden death events. Detection of late gadolinium enhancement is not part of UK Biobank CMR protocol; right ventricular phenotypic expression was not evaluated for this study. Family pedigree information for cardiomyopathy and sudden death was not collected and return of results is not permitted in the UK Biobank. In clinical genomic and precision medicine, pedigrees, cosegregation analysis, familial risk, and heritability are best practice and not possible with UK Biobank.

Conclusions

More than 90% of middle-aged and older adults with putative pathogenic variants in DCM-associated genes did not have a history of DCM or of early DCM features. Nearly one-sixth of putative pathogenic variant carriers exhibited subclinical features on ECG or CMR, most commonly arrhythmias in the absence of substantial ventricular dysfunction. Given the difficulties in variant pathogenicity adjudication, low disease penetrance, and uncertainties in clinical actionability, applying a gene-first approach to DCM for clinical and investigative decision-making might currently be challenging for a broad gene panel, but might be useful for clinically actionable genes that show a relatively higher penetrance.

ARTICLE INFORMATION

Received November 24, 2021; accepted May 12, 2022.

Affiliations

Imperial College Healthcare NHS Trust, London, United Kingdom (R.A.S.). Department of Cardiology, Inselspital, Bern University Hospital, University of Bern, Switzerland (B.A.). Center for Inherited Cardiovascular Diseases, WellSpan Health, Lancaster, PA (G.S.D., C.A.A.C.). University of Michigan, Division of Cardiovascular Medicine, Ann Arbor (G.S.D.). Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London, United Kingdom (N.A., M.Y.K., L.R.L., A.M.L., S.E.P., C.A.A.C.). NIHR Barts Biomedical Research Centre, William Harvey Research Institute, Queen Mary University of London, United Kingdom (N.A., M.Y.K., S.v.D., A.M.L., P.B.M., S.E.P.). Newham University Hospital, Barts Health NHS Trust, London, United Kingdom (M.Y.K.). Centre for Heart Muscle Disease, Institute of Cardiovascular Science, University College London, United Kingdom (L.R.L.). GE Healthcare, Chicago, IL (A.H.). Cardiac Electrophysiology, Cardiovascular Division, Hospital of the University of Pennsylvania, Philadelphia (D.M., C.A.A.C.). Departments of Pediatrics, Division of Cardiology, and Cell Biology, Duke University School of Medicine, Durham, NC (A.P.L.). Division of Cardiovascular Disease, University of Alabama at Birmingham (P.A.). Agnes Ginges Centre for Molecular Cardiology at Centenary Institute (C.S.) and Sydney Medical School Faculty of Medicine and Health (C.S.), The University of Sydney, New South Wales, Australia. Department of Cardiology, Royal Prince Alfred Hospital, Sydney, New South Wales, Australia (C.S.). Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN (V.K.S., C.A.A.C.). Center for Inherited Cardiovascular Disease, Cardiovascular Division, University of Pennsylvania Perelman School of Medicine, Philadelphia (A.T.O.).

Acknowledgments

This study was conducted using the UK Biobank resource under access application 48286.

The authors thank all the participants and staff involved with planning, collection, and analysis, including core laboratory analysis of the cardiovascular magnetic resonance imaging data.

Sources of Funding

Dr Asatryan is supported by a Research Grant from the Swiss Heart Foundation. Dr Landstrom is supported by the American Academy of Pediatrics, Children's Cardiomyopathy Foundation, and the Dermer Foundation. Dr Lee reports funding from UKRI London Medical Imaging & Artificial Intelligence Center for Value Based Healthcare and a SmartHeart EPSRC program grant (grant EP/P001009/1). Dr Lopes is funded by a Medical Research Council UK Clinical Academic Research Partnership award (grant MR/T005181/1). Dr Semsarian is the recipient of a National Health and Medical Research Council Practitioner Fellowship grant (grant 1154992). Dr Petersen acknowledges the British Heart Foundation for funding the manual analysis to create a cardiovascular magnetic resonance imaging reference standard for the UK Biobank imaging resource in 5000 cardiovascular magnetic resonance scans (www.bhf.org.uk; PG/14/89/31194). Dr Aung recognizes the National Institute for Health Research Integrated Academic Training program, which supports his Academic Clinical Lectureship post. Dr Petersen and P.B. Munroe acknowledge support from the National Institute for Health Research Biomedical Research Center at Barts. Dr Petersen acknowledges support from the SmartHeart UK Engineering and Physical Sciences Research Council program grant (www.nihr.ac.uk; EP/P001009/1). This project was enabled through access to the Medical Research Council eMedLab Medical Bioinformatics infrastructure, supported by the Medical Research Council (www.mrc.ac.uk; MR/L016311/1). The funders provided support in the form of salaries for authors as detailed above, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosures

Dr Lopes serves on the advisory board of Bristol Myers Squibb. Dr Owens provided consulting to BMS and Cytokinetics. Dr Petersen provided consultancy to and is shareholder of Circle Cardiovascular Imaging Inc, Calgary, Alberta, Canada. The other authors report that they have no relevant conflicts of interest to disclose.

Supplemental Material

Expanded Methods

Appendix

Figures S1–S4

Tables S1–S12

REFERENCES

1. Schultheiss HP, Fairweather D, Caforio ALP, Escher F, Hershberger RE, Lipshultz SE, Liu PP, Matsumori A, Mazzanti A, McMurray J, et al. Dilated cardiomyopathy. *Nat Rev Dis Primers*. 2019;5:32. doi: 10.1038/s41572-019-0084-1

2. Patel V, Asatryan B, Siripanthong B, Munroe PB, Tiku-Owens A, Lopes LR, Khanji MY, Protonotarios A, Santangeli P, Muser D, et al. State of the art review on genetics and precision medicine in arrhythmogenic cardiomyopathy. *Int J Mol Sci*. 2020;21:E6615. doi: 10.3390/ijms21186615
3. McNally EM, Mestroni L. Dilated cardiomyopathy: genetic determinants and mechanisms. *Circ Res*. 2017;121:731–748. doi: 10.1161/CIRCRESAHA.116.309396
4. Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, Celeghin R, Edwards M, Fan J, Ingles J, et al. Evidence-based assessment of genes in dilated cardiomyopathy. *Circulation*. 2021;144:7–19. doi: 10.1161/CIRCULATIONAHA.120.053033
5. Asatryan B, Medeiros-Domingo A. Translating emerging molecular genetic insights into clinical practice in inherited cardiomyopathies. *J Mol Med (Berl)*. 2018;96:993–1024. doi: 10.1007/s00109-018-1685-y
6. Hershberger RE, Cowan J, Jordan E, Kinnamon DD. The complex and diverse genetic architecture of dilated cardiomyopathy. *Circ Res*. 2021;128:1514–1532. doi: 10.1161/CIRCRESAHA.121.318157
7. Ahmad F, McNally EM, Ackerman MJ, Baty LC, Day SM, Kullo IJ, Madueme PC, Maron MS, Martinez MW, Salberg L, et al. Establishment of specialized clinical cardiovascular genetics programs: recognizing the need and meeting standards: a scientific statement from the American Heart Association. *Circ Genom Precis Med*. 2019;12:e000054. doi: 10.1161/HCG.0000000000000054
8. Marian AJ, Asatryan B, Wehrens XHT. Genetic basis and molecular biology of cardiac arrhythmias in cardiomyopathies. *Cardiovasc Res*. 2020;116:1600–1619. doi: 10.1093/cvr/cvaa116
9. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol*. 2013;10:531–547. doi: 10.1038/nrcardio.2013.105
10. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12:e1001779. doi: 10.1371/journal.pmed.1001779
11. Szustakowski JD, Balasubramanian S, Sasson A, Khalid S, Bronson PG, Kvikstad E, Wong E, Liu D, Davis JW, Haefliger C, et al. Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. *Nat Genet*. 2021;53:942–948. doi: 10.1038/s41588-021-00885-0
12. 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
13. Miller DT, Lee K, Chung WK, Gordon AS, Herman GE, Klein TE, Stewart DR, Amendola LM, Adelman K, Bale SJ, et al; ACMG Secondary Findings Working Group. ACMG SF v3.0 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23:1381–1390. doi: 10.1038/s41436-021-01172-3
14. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164. doi: 10.1093/nar/gkq603
15. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*. 2016;99:877–885. doi: 10.1016/j.ajhg.2016.08.016
16. Park J, Levin MG, Haggerty CM, Hartzel DN, Judy R, Kember RL, Reza N, Ritchie MD, Owens AT, Damrauer SM, et al; Regeneron Genetics Center. A genome-first approach to aggregating rare genetic variants in LMNA for association with electronic health record phenotypes. *Genet Med*. 2020;22:102–111. doi: 10.1038/s41436-019-0625-8
17. Park J, Lucas AM, Zhang X, Chaudhary K, Cho JH, Nadkarni G, Dobbyn A, Chittoor G, Josyula NS, Katz N, et al. Exome-wide evaluation of rare coding variants using electronic health records identifies new gene-phenotype associations. *Nat Med*. 2021;27:66–72. doi: 10.1038/s41591-020-1133-8
18. Whiffin N, Walsh R, Govind R, Edwards M, Ahmad M, Zhang X, Tayal U, Buchan R, Midwinter W, Wilk AE, et al. CardioClassifier: disease- and gene-specific computational decision support for clinical genome interpretation. *Genet Med*. 2018;20:1246–1254. doi: 10.1038/gim.2017.258
19. General Electric. Marquette™ 12SL™ ECG Analysis Program Physician's Guide 2036070-006, Revision C. 2012; General Electric.
20. Mason JW, Hancock EW, Gettes LS, Bailey JJ, Childers R, Deal BJ, Josephson M, Kligfield P, Kors JA, Macfarlane P, et al; American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; American College of Cardiology Foundation; Heart Rhythm Society. Recommendations for the standardization and interpretation of the electrocardiogram: part II: electrocardiography diagnostic statement list: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology, the American College of Cardiology Foundation, and the Heart Rhythm Society endorsed by the International Society for Computerized Electrocardiology. *J Am Coll Cardiol*. 2007;49:1128–1135. doi: 10.1016/j.jacc.2007.01.025
21. Petersen SE, Matthews PM, Francis JM, Robson MD, Zemrak F, Boubertakh R, Young AA, Hudson S, Weale P, Garratt S, et al. UK Biobank's cardiovascular magnetic resonance protocol. *J Cardiovasc Magn Reson*. 2016;18:8. doi: 10.1186/s12968-016-0227-4
22. Suinesiaputra A, Sanghvi MM, Aung N, Paiva JM, Zemrak F, Fung K, Lukaschuk E, Lee AM, Carapella V, Kim YJ, et al. Fully-automated left ventricular mass and volume MRI analysis in the UK Biobank population cohort: evaluation of initial results. *Int J Cardiovasc Imaging*. 2018;34:281–291. doi: 10.1007/s10554-017-1225-9
23. Pinto YM, Elliott PM, Arbustini E, Adler Y, Anastasakis A, Böhm M, Duboc D, Gimeno J, de Groote P, Imazio M, et al. Proposal for a revised definition of dilated cardiomyopathy, hypokinetic non-dilated cardiomyopathy, and its implications for clinical practice: a position statement of the ESC working group on myocardial and pericardial diseases. *Eur Heart J*. 2016;37:1850–1858. doi: 10.1093/eurheartj/ehv727
24. Aragam KG, Chaffin M, Levinson RT, McDermott G, Choi SH, Shoemaker MB, Haas ME, Weng LC, Lindsay ME, Smith JG, et al. Phenotypic refinement of heart failure in a national biobank facilitates genetic discovery. *Circulation*. 2018;139:489–501. doi: 10.1161/CIRCULATIONAHA.118.035774
25. Khurshid S, Choi SH, Weng LC, Wang EY, Trinquart L, Benjamin EJ, Ellinor PT, Lubitz SA. Frequency of cardiac rhythm abnormalities in a half million adults. *Circ Arrhythm Electrophysiol*. 2018;11:e006273. doi: 10.1161/CIRCEP.118.006273
26. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018;50:1219–1224. doi: 10.1038/s41588-018-0183-z
27. R: A language and environment for statistical computing [computer program], version 3.6.1. R Foundation for Statistical Computing; 2019.
28. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. Welcome to the tidyverse. *J Open Source Softw*. 2019;4:1686.
29. tableone: Create 'Table 1' to Describe Baseline Characteristics with or without Propensity Score Weights, version R, package version 0.12.0 [computer program]. 2020.
30. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, Conner L, DePalma SR, McDonough B, Sparks E, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med*. 2012;366:619–628. doi: 10.1056/NEJMoa1110186
31. Tabish AM, Azzimato V, Alexiadis A, Buyandelger B, Knöll R. Genetic epidemiology of titin-truncating variants in the etiology of dilated cardiomyopathy. *Biophys Rev*. 2017;9:207–223. doi: 10.1007/s12551-017-0265-7
32. Esslinger U, Garnier S, Korniat A, Proust C, Kararigas G, Müller-Nurasyid M, Empana JP, Morley MP, Perret C, Stark K, et al. Exome-wide association study reveals novel susceptibility genes to sporadic dilated cardiomyopathy. *PLoS One*. 2017;12:e0172995. doi: 10.1371/journal.pone.0172995
33. Towbin JA, McKenna WJ, Abrams DJ, Ackerman MJ, Calkins H, Darrieux FCC, Daubert JP, de Chillou C, DePasquale EC, Desai MY, et al. 2019 HRS expert consensus statement on evaluation, risk stratification, and management of arrhythmogenic cardiomyopathy. *Heart Rhythm*. 2019;16:e301–e372. doi: 10.1016/j.hrthm.2019.05.007
34. Codd MB, Sugrue DD, Gersh BJ, Melton LJ III. Epidemiology of idiopathic dilated and hypertrophic cardiomyopathy: a population-based study in Olmsted County, Minnesota, 1975-1984. *Circulation*. 1989;80:564–572. doi: 10.1161/01.cir.80.3.564
35. Fairweather D, Cooper LT Jr, Blauwet LA. Sex and gender differences in myocarditis and dilated cardiomyopathy. *Curr Probl Cardiol*. 2013;38:7–46. doi: 10.1016/j.cpcardiol.2012.07.003
36. Carruth ED, Young W, Beer D, James CA, Calkins H, Jing L, Raghunath S, Hartzel DN, Leader JB, Kirchner HL, et al. Prevalence and electronic health record-based phenotype of loss-of-function genetic variants in arrhythmogenic right ventricular cardiomyopathy-associated genes. *Circ Genom Precis Med*. 2019;12:e002579. doi: 10.1161/CIRCGEN.119.002579
37. 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

38. Asatryan B. Cardiac sodium channel dysfunction and dilated cardiomyopathy: a contemporary reappraisal of pathophysiological concepts. *J Clin Med*. 2019;8:E1029. doi: 10.3390/jcm8071029
39. Pirruccello JP, Bick A, Wang M, Chaffin M, Friedman S, Yao J, Guo X, Venkatesh BA, Taylor KD, Post WS, et al. Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nat Commun*. 2020;11:2254. doi: 10.1038/s41467-020-15823-7
40. Haggerty CM, Damrauer SM, Levin MG, Birtwell D, Carey DJ, Golden AM, Hartzel DN, Hu Y, Judy R, Kelly MA, et al. Genomics-first evaluation of heart disease associated with titin-truncating variants. *Circulation*. 2019;140:42–54. doi: 10.1161/CIRCULATIONAHA.119.039573
41. Ware JS, Amor-Salamanca A, Tayal U, Govind R, Serrano I, Salazar-Mendiguchía J, García-Pinilla JM, Pascual-Figal DA, Nuñez J, Guzzo-Merello G, et al. Genetic etiology for alcohol-induced cardiac toxicity. *J Am Coll Cardiol*. 2018;71:2293–2302. doi: 10.1016/j.jacc.2018.03.462
42. Ware JS, Li J, Mazaika E, Yasso CM, DeSouza T, Cappola TP, Tsai EJ, Hilfiker-Kleiner D, Kamiya CA, Mazzarotto F, et al; IMAC-2 and IPAC Investigators. Shared genetic predisposition in peripartum and dilated cardiomyopathies. *N Engl J Med*. 2016;374:233–241. doi: 10.1056/NEJMoa1505517
43. Parikh VN. Promise and peril of population genomics for the development of genome-first approaches in Mendelian cardiovascular disease. *Circ Genom Precis Med*. 2021;14:e002964. doi: 10.1161/CIRCGEN.120.002964
44. Asatryan B, Asimaki A, Landstrom AP, Khanji MY, Odening KE, Cooper LT, Marchlinski FE, Gelzer AR, Semsarian C, Reichlin T, et al. Inflammation and immune response in arrhythmogenic cardiomyopathy: state-of-the-art review. *Circulation*. 2021;144:1646–1655. doi: 10.1161/CIRCULATIONAHA.121.055890
45. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. World Health Organization: 1968.
46. Albert MA, Carnethon MR, Watson KE. Disparities in cardiovascular medicine. *Circulation*. 2021;143:2319–2320. doi: 10.1161/CIRCULATIONAHA.121.055565
47. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, Pellikka PA, Enriquez-Sarano M, Noseworthy PA, Munger TM, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. 2019;25:70–74. doi: 10.1038/s41591-018-0240-2
48. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. *Am J Epidemiol*. 2017;186:1026–1034. doi: 10.1093/aje/kwx246
49. Coughlin SS, Gottdiener JS, Baughman KL, Wasserman A, Marx ES, Tefft MC, Gersh BJ. Black-white differences in mortality in idiopathic dilated cardiomyopathy: the Washington, DC, dilated cardiomyopathy study. *J Natl Med Assoc*. 1994;86:583–591.
50. Myers VD, Gerhard GS, McNamara DM, Tomar D, Madesh M, Kaniper S, Ramsey FV, Fisher SG, Ingersoll RG, Kasch-Semenza L, et al. Association of variants in BAG3 with cardiomyopathy outcomes in African American individuals. *JAMA Cardiol*. 2018;3:929–938. doi: 10.1001/jamacardio.2018.2541