

1 **Prevalence and disease expression of pathogenic and likely pathogenic variants associated**
2 **with inherited cardiomyopathies in the general population**

3

4 Mimount Bourfiss, MD^{1*}; Marion van Vugt, MSc^{1*}; Abdulrahman I. Alasiri, MSc¹; Bram Ruijsink, MD,
5 PhD^{1,2}; Jessica van Setten, MSc, PhD¹; Amand F. Schmidt, MSc, PhD^{1,3}; Dennis Dooijes, PhD⁴; Esther
6 Puyol-Antón, PhD²; Birgitta K. Velthuis, MD, PhD⁵; J. Peter van Tintelen, MD, PhD⁴; Anneline S.J.M. te
7 Riele, MD, PhD^{1,6}; Annette F. Baas, MD, PhD⁴; Folkert W. Asselbergs, MD, PhD^{1,3,7}.

8 *Shared first authorship

9 *Affiliations*

- 10 1. Department of Cardiology, University Medical Center Utrecht, Utrecht University, Utrecht,
11 the Netherlands
12 2. School of Biomedical Engineering and Imaging Sciences, King's College London, London,
13 United Kingdom
14 3. Institute of Cardiovascular Science, London, Faculty of Population Health Sciences,
15 University College London, United Kingdom
16 4. Department of Genetics, University Medical Center Utrecht, Utrecht, Utrecht University,
17 Utrecht, the Netherlands
18 5. Department of Radiology, University Medical Center Utrecht, Utrecht University, Utrecht,
19 the Netherlands
20 6. Netherlands Heart Institute, Utrecht, the Netherlands
21 7. Health Data Research UK and Institute of Health Informatics, University College London,
22 London, United Kingdom
23

24 Short title: Cardiomyopathy variants in the general population

25 Corresponding author: Mimount Bourfiss; Department of Medicine, Division of Cardiology;

26 University Medical Center Utrecht; room E.02.561; Heidelberglaan 100; 3584 CX Utrecht, the

27 Netherlands. Email address: m.bourfiss-2@umcutrecht.nl Telephone number: +31-88-77570240. Fax

28 number: +31-88-7555660.

29 Journal Subjects Terms: Cardiomyopathy, Genetics, Heart failure, Hypertrophy, Magnetic Resonance

30 imaging

31 Total word count: 6993

32

33 **Abstract**

34 **Background.** Pathogenic and likely pathogenic variants associated with arrhythmogenic
35 right ventricular cardiomyopathy (ARVC), dilated cardiomyopathy (DCM) and hypertrophic
36 cardiomyopathy (HCM) are recommended to be reported as secondary findings in genome
37 sequencing studies. This provides opportunities for early diagnosis, but also fuels
38 uncertainty in variant carriers (G+), since disease penetrance is incomplete. We assessed the
39 prevalence and disease expression of G+ in the general population.

40 **Methods.** We identified pathogenic and likely pathogenic variants associated with ARVC,
41 DCM and/or HCM in 200,643 UK Biobank individuals, who underwent whole exome
42 sequencing. We calculated the prevalence of G+ and analysed the frequency of
43 cardiomyopathy/heart failure diagnosis. In undiagnosed individuals, we analysed early signs
44 of disease expression using available electrocardiography and cardiac magnetic resonance
45 imaging data.

46 **Results.** We found a prevalence of 1:578, 1:251 and 1:149 for pathogenic and likely
47 pathogenic variants associated with ARVC, DCM and HCM respectively. Compared to
48 controls, cardiovascular mortality was higher in DCM G+ (OR 1.67 [95% CI 1.04;2.59],
49 $p=0.030$), but similar in ARVC and HCM G+ ($p\geq 0.100$). Cardiomyopathy or heart failure
50 diagnosis were more frequent in DCM G+ (OR 3.66 [95% CI 2.24;5.81], $p=4.9\times 10^{-7}$) and HCM
51 G+ (OR 3.03 [95% CI 1.98;4.56], $p=5.8\times 10^{-7}$), but comparable in ARVC G+ ($p=0.172$). In
52 contrast, ARVC G+ had more ventricular arrhythmias ($p=3.3\times 10^{-4}$). In undiagnosed
53 individuals, left ventricular ejection fraction was reduced in DCM G+ ($p=0.009$).

54 **Conclusions.** In the general population, pathogenic and likely pathogenic variants associated
55 with ARVC, DCM or HCM are not uncommon. Although G+ have increased mortality and
56 morbidity, disease penetrance in these carriers from the general population remains low

57 (1.2-3.1%). Follow-up decisions in case of incidental findings should not be based solely on a
58 variant, but on multiple factors, including family history and disease expression.

59

60 **Key Words:** whole exome sequencing, genetics, arrhythmogenic right ventricular
61 cardiomyopathy, dilated cardiomyopathy, hypertrophic cardiomyopathy

62

63 **Non-standard Abbreviations and Acronyms**

64 **ACMG** – American College of Medical Genetics and Genomics

65 **ACTC1** – Actin Alpha Cardiac Muscle 1

66 **ACTN1** – Alpha-actinin 2

67 **AHA** – American Heart Association

68 **ARVC** – Arrhythmogenic right ventricular cardiomyopathy

69 **BAG3** – BAG Cochaperone 3

70 **BMI** – Body mass index

71 **CI** – Confidence interval

72 **ClinGen** – Clinical Genome Resource

73 **CMR** – Cardiac magnetic resonance imaging

74 **CSRP3** – Cysteine And Glycine Rich Protein 3

75 **DCM** – Dilated cardiomyopathy

76 **DES** – Desmin

77 **DSC2** – Desmocollin 2

78 **DSG2** – Desmoglein 2

79 **DSP** – Desmoplakin

80 **ECG** – Electrocardiography

- 81 **EDM** – End diastolic mass
- 82 **EDV** – End diastolic volume
- 83 **EF** – Ejection fraction
- 84 **ESV** – End systolic volume
- 85 **FLNC** – Filamin-C
- 86 **G+** – Genotype positive (variant carriers)
- 87 **G-** – Genotype negative (controls)
- 88 **i** – Indexed, corrected for body surface area
- 89 **JPH2** – Junctophilin 2
- 90 **JUP** – Junction Plakoglobin
- 91 **HCM** – Hypertrophic cardiomyopathy
- 92 **LMNA** – Lamin A/C
- 93 **LoF** – Loss of function
- 94 **LV** – Left ventricular
- 95 **MAF** – Minor allele frequency
- 96 **MET** – Metabolic equivalent of task minutes
- 97 **MVR** – Mass to EDV ratio
- 98 **MYBPC3** – Myosin Binding Protein C3
- 99 **MYH7** – Myosin Heavy Chain 7
- 100 **MYL2** – Myosin Light Chain 2
- 101 **MYL3** – Myosin Light Chain 3
- 102 **NEXN** – Nexilin F-Actin Binding Protein
- 103 **NGS** – Next generation sequencing
- 104 **OR** – Odds ratio

- 105 **P+** – Phenotype positive
- 106 **P-** – Phenotype negative
- 107 **PKP2** – Plakophilin 2
- 108 **PLN** – Phospholamban
- 109 **RBM20** – RNA Binding Motif Protein 20
- 110 **RV** – Right ventricular
- 111 **SCN5A** – Sodium Voltage-Gated Channel Alpha Subunit 5
- 112 **SD** – Standard deviation
- 113 **SV** – Stroke volume
- 114 **TFC** – Task force criteria
- 115 **TMEM43** – Transmembrane Protein 43
- 116 **TNNC1** – Troponin C1, Slow Skeletal And Cardiac Type
- 117 **TNNI3** – Troponin I3, Cardiac Type
- 118 **TNNT2** – Troponin T2, Cardiac Type
- 119 **TPM1** – Tropomyosin 1
- 120 **TTN** – Titin
- 121 **UKB** – UK Biobank
- 122 **VKGL** – Vereniging Klinische Genetische Laboratoriumdiagnostiek
- 123 **VCL** – Vinculin
- 124 **WES** – Whole exome sequencing
- 125
- 126
- 127
- 128

129 **Background**

130 The major inherited cardiomyopathies arrhythmogenic right ventricular cardiomyopathy
131 (ARVC), dilated cardiomyopathy (DCM) and hypertrophic cardiomyopathy (HCM) are
132 characterized by ventricular dysfunction and ventricular arrhythmias that can lead to
133 progressive heart failure and sudden cardiac death¹. ARVC is mainly caused by pathogenic
134 variants in desmosomal genes, whereas DCM and HCM are mainly caused by sarcomeric
135 gene variants². These cardiomyopathies are typically inherited in an autosomal dominant
136 fashion with incomplete penetrance and variable expressivity. As such, phenotypic
137 expression may vary greatly, even among family members or individuals carrying the same
138 pathogenic variant.

139 With the implementation of next-generation sequencing (NGS), genetic testing has
140 become an important part of routine clinical care in the diagnosis of inherited
141 cardiomyopathies³. Technical advances and commercial availability of NGS, have led to
142 more affordable and accessible genetic testing. The American College of Medical Genetics
143 and Genomics (ACMG) has developed recommendations for the reporting of incidental or
144 secondary findings unrelated to the test indication⁴. In this framework, variants in genes
145 associated with ARVC, DCM and HCM are recommended to be reported as secondary
146 findings from clinical exome and other genome sequencing tests⁴. Although this offers the
147 potential to prevent morbidity and mortality of heart failure and sudden cardiac death by
148 early treatment, it also fuels uncertainty in carriers of likely pathogenic or pathogenic
149 variants (G+) and their family members, since factors that influence disease penetrance in
150 the general population are largely unknown. More knowledge about disease penetrance of
151 these variants in an unselected population cohort is needed to determine screening
152 protocols in these individuals.

153 In this study, we aimed to assess the prevalence of pathogenic and likely pathogenic
154 variants in the general population using a set of recently curated genes for ARVC⁵, DCM⁶ and
155 HCM⁷ in two (inter)national databases⁸ (see **Figure 1** and **Table I**). In order to do so, we
156 leveraged data from the UK Biobank (UKB), a population-based cohort with whole exome
157 sequencing (WES) data available of 200,643 individuals⁹. Furthermore, we looked into the
158 UKB-reported phenotypical characteristics of these G+ and assessed the occurrence of early
159 signs of disease in undiagnosed G+ using available electrocardiography and cardiac magnetic
160 resonance (CMR) imaging data.

161

162 **Methods**

163 Ethics approval for the UKB study was obtained from the North West Centre for Research
164 Ethics Committee (11/NW/0382) and all participants provided informed consent. All data
165 and materials have been made publicly available on Github and can be accessed at
166 <https://github.com/CirculatoryHealth/Inherited-cardiomyopathies>. Full methods are
167 available in the **Supplemental Material**. Disease definitions are given in **Table II**.

168

169 **Results**

170 *Population characteristics*

171 We identified 2,207/200,643 unique G+ individuals from a total of 2,493 included individuals
172 (89%) (see **Figure 2**) classified as 1) ARVC G+ (n=347, 54% female, median age of 57 [50-64]
173 years); 2) DCM G+ (n=800, 56% female, median age of 58 [51-64] years) and; 3) HCM G+
174 (n=1,346, 54% female, median age of 56 [49-63] years). The matched control group
175 consisted of 9,972 individuals (55% females, median age of 57 [49-63] years). **Table 1** and
176 **Table III** show the baseline characteristics of the included individuals. The majority of G+

177 were of white ethnicity (ARVC 90%, DCM 96%, HCM 75%), followed by Asian (ARVC 3%,
178 DCM 1%, HCM 19%) and black ethnicity (ARVC 2%, DCM 2% and HCM 2%). This is
179 comparable to what is observed in the UKB, where the majority is of white ethnicity (94%),
180 followed by Asian (2%) and black ethnicity (2%).

181

182 *Genotypic characteristics of pathogenic and likely pathogenic variant carriers*

183 Prevalence of pathogenic and likely pathogenic variants in the general population

184 We found a prevalence of 1 ARVC G+ in 578 people in the general population (1:578 [1:521;
185 1:644]) and identified 75 variants out of the 593 (13%) pathogenic and likely pathogenic
186 variants described in ClinVar and VKGL: 13 missense and 62 loss of function (LoF) (**Table IV**).
187 As shown in **Figure 3**, most ARVC G+ harbored a pathogenic variant in *PKP2* (n=185, 53%),
188 followed by *DSP* (n=49, 14%), *DSC2* (n=42, 12%), *DSG2* (n=31, 9%), *JUP* (n=24, 7%), *DES*
189 (n=15, 4%), and *PLN* (n=1, 0.3%).

190 We found a prevalence of 1:251 [1:234; 1:269] DCM G+ and identified 216 out of
191 3,460 (6%) pathogenic and likely pathogenic variants described in ClinVar and VKGL: 80
192 missense and 136 LoF (**Table IV**). Variants in *TTN* (n=272, 34%) and *MYH7* (n=158, 20%)
193 were most prevalent among DCM G+, followed by *SCN5A* (n=59, 7%), *FLNC* (n=56, 7%), *DSP*
194 (n=49, 6%), *DES* (n=49, 6%), *LMNA* (n=42, 5%), *TNNI3* (n=35, 4%) and *TNNT2* (n=32, 4%).
195 Eight more genes with a frequency of less than 3% were identified: *BAG3*, *PLN*, *TNNC1*,
196 *ACTN2*, *RBM20*, *NEXN*, *TPM1*, and *ACTC1* (**Figure 3**).

197 We found a prevalence of 1:149 [1:141; 1:157] HCM G+ and identified 131 out of
198 1,512 (9%) pathogenic and likely pathogenic variants from ClinVar and VKGL: 98 missense
199 and 23 LoF (**Table IV**). Most individuals carried a pathogenic and likely pathogenic variant in
200 *MYBPC3* (n=723, 54%), followed by *TNNT2* (n=274, 20%), *MYH7* (n=232, 17%), and *TNNI3*

201 (n=50, 4%). A frequency of less than 3% was found in *CSRP3*, *MYL2*, *TNNC1*, *JPH2*, *TPM1*,
202 *ACTC1*, and *MYL3* (**Figure 3**). Interestingly, a variant in *TNNT2* (c.862C>T p.Arg288Cys)
203 affected 242 individuals (18%). Four of these carriers were diagnosed with heart failure, of
204 whom one also with HCM. All four heart failure patients also suffered from chronic ischemic
205 heart disease. Furthermore, a variant in *MYBPC3* (c.3628-41_3628-17del) was mainly seen in
206 individuals with an Asian ethnicity (n=237, 18% of the total HCM G+). Four of these
207 individuals were diagnosed with heart failure, of whom two also had coronary artery disease
208 and one was diagnosed with DCM, however none were diagnosed with HCM. When
209 excluding these two variants, we found a HCM G+ prevalence of 1:250 [1:234; 1:269].
210 *MYBPC3* remained the most prevalent gene (52%), whereas the *TNNT2* frequency decreased
211 to 4%.

212 The prevalence of G+ per gene for all cardiomyopathies is depicted in **Table V**.

213

214 Overlapping variants and individuals

215 Some pathogenic and likely pathogenic variants were identified in multiple cardiomyopathy
216 subtypes. First, 26 variants were described in both ARVC and DCM, affecting 53 individuals.
217 Most of these variants (n=20/26 variants, 77%) were found in *DSP* (n=37 individuals, 70%),
218 of whom one individual (3%) had heart failure and one (3%) was diagnosed with a
219 cardiomyopathy. Five variants out of 26 (19%) were found in *DES* (n=15 individuals, 28%) of
220 whom two individuals (13%) had heart failure, and one was diagnosed with both DCM and
221 HCM. One variant out of 26 (4%) was found in *PLN* (c.26G>A; p.Arg9His, NM_002667.5) in
222 one individual (2%) who was not diagnosed with a cardiomyopathy or heart failure.

223 Second, 52 variants were described in DCM and HCM, affecting 232 individuals. Most
224 of these variants (n=33/52 variants, 63%) were found in *MYH7* (n=158 individuals, 68%),

225 followed by 10 variants (19%) in *TNNT2* (n=29 individuals, 13%), 6 variants (12%) in *TNNI3*
226 (n=35 individuals, 15%), and 1 (2%) variant in *TNNC1*, *ACTC1* as well as *TPM1*. In this group
227 of 232 individuals, 9 (4%) individuals had a cardiomyopathy or heart failure diagnosis, of
228 whom 5 were diagnosed with HCM and none with DCM.

229 Furthermore, three individuals carried two pathogenic variants. Individual 1 was a
230 65-year old male, carrying variants in *MYBPC3* (c.3628-41_3628-17del, NM_000256.3) and
231 *TNNT2* (c.460C>T; p.Arg154Trp, NM_001276345.2) and was diagnosed with heart failure,
232 with underlying chronic ischemic heart disease. Individual 2 was a 65-year old male, carrying
233 variants in *FLNC* (c.7450G>A; p.Gly2484Ser, NM_001458.5) and *PKP2* (c.1867G>T;
234 p.Glu623Ter, NM_001005242.3) and was therefore included in both the ARVC as well as the
235 DCM G+ group. Individual 3 was a 64-year old male, carrying variants in *MYBPC3*
236 (c.1504C>T; p.Arg502Trp, NM_000256.3) and *MYH7* (c.5655G>A; p.Ala1885=
237 NM_000257.4). Individuals 2 and 3 were not diagnosed with a cardiomyopathy or heart
238 failure and none had CMR data available.

239 **Table VI** shows the prevalence of the cardiomyopathy variants, with and without the
240 inclusion of overlapping variants.

241

242 *Phenotypic characteristics of pathogenic and likely pathogenic variant carriers*

243 Cardiovascular risk factors

244 Hypertension, BMI, and level of activity in metabolic equivalent of task minutes (MET) per
245 week were comparable between G- and G+ for all cardiomyopathies (p≥0.055; **Table 1** and
246 **Table VII**). Diabetes was more prevalent in G+ HCM (9.2% (G-) vs 11.4% (G+), p=0.008),
247 while smoking was more prevalent in DCM G+ (41.4% vs 46.4%, p=0.007) (**Table VIII**).

248

249 Cardiovascular disease

250 As seen in **Figure 4** and **Table VIII**, compared to G-, cardiomyopathy/heart failure without
251 previous ischemic heart disease (P+, phenotype positive) was more often diagnosed in DCM
252 G+ (OR 3.66 [95% CI 2.24;5.81], $p=4.9\times 10^{-7}$) and HCM G+ (OR 3.03 [95% CI 1.98;4.56],
253 $p=5.8\times 10^{-7}$). Among DCM G+, 25 individuals (3.1%, genes: 8 *MYH7*, 8 *TTN*, 2 *BAG3*, 2 *DSP*, 2
254 *FLNC*, 1 *DES*, 1 *SCN5A* and 1 *TNNT2*) were P+. Within HCM G+, 32 individuals (4.0%, genes:
255 20 *MYBPC3*, 10 *MYH7*, 1 *TNNI3* and 1 *TNNT2*) were P+. For ARVC G+, 4 individuals (1.2%,
256 genes: 2 *DSP*, 1 *DES* and 1 *PKP2*) were P+, being comparable to G- controls (87 subjects,
257 0.8%).

258 Ventricular arrhythmias occurred more often in G+ compared to G-, reaching
259 statistical significance for ARVC (OR 6.20 [95% CI 2.30;14.38], $p=3.3\times 10^{-4}$) and DCM (OR 4.97
260 [95% CI 2.39;9.75], $p=1.9\times 10^{-5}$). Atrial arrhythmias were more prevalent among DCM G+ (OR
261 2.27 [95% CI 1.52;3.31], $p=8.2\times 10^{-5}$). Finally, all-cause mortality (OR 1.39 [95% CI 1.02;1.85],
262 $p=0.032$) and death due to a cardiovascular cause were more prevalent in DCM G+ (OR 1.67
263 [95% CI 1.04;2.59], $p=0.030$) but did not reach statistical significance for ARVC G+ and HCM
264 G+ ($p\geq 0.100$). **Figure I** depicts the overlap in cardiomyopathy, heart failure, ventricular
265 arrhythmia and ischemic heart disease diagnosis. **Figure II** depicts the forest plots when
266 excluding the more prevalent *TNNT2* and *MYBPC3* variants in HCM G+ individuals and **Figure**
267 **III** shows the association between different outcomes stratified by genes for each
268 cardiomyopathy.

269

270 *Deep phenotyping of undiagnosed pathogenic variant carriers*

271 Next, we set out to study early signs of disease in G+ without a cardiomyopathy/heart
272 failure diagnosis (P-) using registered ICD-10 codes, self-reported cardiac symptoms, and
273 abnormal ECG and CMR values.

274

275 Diagnosis and symptoms

276 Ventricular arrhythmias were more prevalent in ARVC G+P- (OR 5.85 [95% CI 1.98;14.40],
277 $p=0.001$) and DCM G+P- (OR 3.43 [95% CI 1.35;7.68], $p=0.005$) but not in HCM G+P- (OR 1.01
278 [95% CI 0.26;2.86], $p=1.000$) compared to G-P- controls. Also, atrial arrhythmias (OR 2.12
279 [95% CI 1.36;3.19], $p=7.9 \times 10^{-4}$) were more frequent in DCM G+P- compared to G-P- controls.
280 Finally, angina pectoris occurred more often in HCM G+P- (OR 1.38 [95% CI 1.01;1.85],
281 $p=0.038$), but not in ARVC G+P- and DCM G+P- ($p \geq 0.117$; **Table VII**).

282

283 Electrocardiography

284 In total, 231 out of 2,207 G+P- and 1,058 out of 9,885 G-P- had various ECG variables
285 available. None of these ECG variables differed significantly between all undiagnosed G+ and
286 control individuals (**Table VIII**).

287

288 Cardiac magnetic resonance imaging

289 CMR data was available in 225 G+P- of the 2,207 unique G+P- individuals: $n=33$ for ARVC
290 G+P-, $n=87$ for DCM G+P- and $n=130$ for HCM G+P-) and 986 of the 9,885 G-P- controls. As
291 shown in **Table IX**, G+P- and G-P- individuals with CMR data available were mostly
292 comparable. Only smoking was more prevalent among DCM G+P- compared to G-P- controls
293 (OR 1.59 [95%CI 1.00; 2.53], $p=0.041$). Outliers were observed in G-P- controls: 4 individuals
294 with a median age of 64 [60-67] years had a left ventricular ejection fraction (LVEF) below

295 40% and 3 of them were diagnosed with hypertension and 2 with acute myocardial
296 infarction in the past. In addition, 2 individuals with an age of 42 and 52 years old had an
297 RVEF below 40%. They did not suffer from hypertension and did not have any cardiac
298 diagnosis.

299 As shown in **Figure 5** and **Table VIII**, all RV ($p \geq 0.546$) and LV ($p \geq 0.052$) functional and
300 structural parameters in ARVC G+P- were comparable to G-P- controls. Three ARVC G+P-
301 individuals had an RVEDV corrected for body surface area (RVEDVi) between 100-110 mL/m²
302 for males or 90-100 mL/m² for females, meeting the minor CMR task force criteria (TFC) if
303 wall motion abnormalities were present, and two ARVC G+P- individuals had an RVEDVi
304 larger than 110 mL/m² for males or 100 mL/m² for females, meeting the major CMR TFC¹⁰.
305 In addition, ARVC G+P- had reduced inferior and posterolateral wall thickness compared to
306 controls ($p \leq 0.035$).

307 Overall, DCM G+P- and G-P- controls had comparable RV functional and structural
308 measures ($p \geq 0.048$). However, DCM G+P- had lower LVEF (57.3% [52.6, 62.8] vs. 59.5%
309 [55.3, 63.5] vs, $p=0.009$) and less negative LV peak longitudinal strain (-22.3% [-24.6, -19.86]
310 vs. -23.3% [-26.0, -21.4], $p=0.009$). Although LVEDVi was not significantly increased in DCM
311 G+P-, the LVEDV/RVEDV ratio (0.9 [0.9, 1.0] vs 1.0 [0.9, 1.1], $p=8.2 \times 10^{-4}$) and LVESVi (30.0
312 ml/m² [25.1, 35.7] vs 31.7 ml/m² [26.2, 39.8], $p=0.032$) were increased. Six individuals had
313 an LVEF below 45%, but none of the individuals met the Henry criteria for DCM (LVEF below
314 45% and LVEDVi two times the normal SD)¹¹.

315 For HCM G+P-, most RV and LV functional and dimension parameters were
316 comparable to G-P- controls ($p \geq 0.051$). Only RVEF was higher than controls (58.4% [54.2,
317 62.7] vs 59.6% [54.8, 64.0], $p=0.025$). Importantly, wall thickness was not significantly
318 different between HCM G+P- without a cardiomyopathy/heart failure diagnosis and G-P-

319 ($p \geq 0.160$). None of the G+P- individuals met HCM criteria¹² of ≥ 15 mm wall thickness, but
320 two individuals met the criteria for limited hypertrophy (13-15mm) in the presence of a
321 positive genetic test¹².

322 **Figure IV** shows a summary of all the differences tested.

323

324 Exclusion of the more prevalent *TNNT2* and *MYBPC3* variants

325 When excluding the more prevalent *TNNT2* and *MYBPC3* variants in HCM G+P- individuals,
326 the occurrence of ventricular arrhythmias (OR 1.72 [95% CI 0.44;4.89], $p=0.306$) and atrial
327 arrhythmias (OR 1.43 [95% CI 0.84;2.32], $p=0.156$) was comparable to G-P- controls.

328 However, the maximum wall thickness (8.47mm [7.59, 9.94] vs. 8.09mm [7.24, 9.01],
329 $p=0.008$) and basal anterior wall thickness (7.93mm [6.97, 9.11] vs. 7.65mm [6.81, 8.49], p
330 $=0.029$) were significantly increased in HCM G+P- compared to controls (**Table VIII**). Two
331 individuals had a maximum wall thickness between 13-15mm.

332

333 **Discussion**

334 In this study we leveraged the largest European population database including WES and
335 phenotype data to evaluate the prevalence and penetrance of previously reported
336 pathogenic and likely pathogenic variants associated with ARVC, DCM and HCM. Our study
337 has several interesting findings. First, we found a prevalence of 1:578, 1:251, and 1:149 for
338 variants previously associated with ARVC, DCM and HCM respectively. Second, 1.2% of ARVC
339 G+, 3.1% of DCM G+ and 2.6% of HCM G+ were diagnosed with a cardiomyopathy or heart
340 failure without previous chronic ischemic heart disease. Finally, 3.2% of the undiagnosed
341 ARVC G+, 1.8% of the undiagnosed DCM G+, 0.5% of the undiagnosed HCM G+ reported

342 ventricular arrhythmias or had CMR abnormalities. These results confirm the low disease
343 penetrance in G+ in the general population.

344

345 *Prevalence of pathogenic and likely pathogenic variant carriers in the general population*

346 Since rare genetic variants are the major cause of inherited cardiomyopathies , a large
347 dataset is needed to accurately identify the population prevalence of these variants.

348 Prevalence of pathogenic variants in populations has been the focus of several previous
349 studies^{4,13-15}, however they were mostly limited by the number of included individuals. At
350 time of analysis, we had access to an unprecedented number of 200,643 individuals.

351 Previously reported prevalence of ARVC G+ in the general population ranges
352 between 1:143 to 1:1,706¹³⁻¹⁵. This variability is likely to be explained by heterogeneity in
353 study populations and definitions of variant pathogenicity. For example, many previous
354 studies did not include all eight curated genes with strong or moderate disease-gene
355 association but also marked other genes (e.g. *TGFB3*) with only limited evidence as
356 associated with ARVC^{14,15}. In addition, we included both missense and LoF variants whereas
357 prior studies only included LoF variants.

358 For DCM, little is known about the prevalence of DCM causing variants in the general
359 population. Studies focusing on truncating *TTN* variants in the general population found a
360 prevalence ranging between 1:33 and 1:526^{16,17}. This wide range can partly be explained by
361 the used definition of pathogenicity. Also, disease causing truncating *TTN* variants
362 associated with DCM are known to be highly enriched in the A band. However, recently,
363 truncating variants in the distal I-band region have also been implicated in DCM¹⁸. When
364 solely focussing on *TTN* variants, we found a prevalence of only 1:735. This differs from the
365 previous studies, probably because not all *TTN* variants are reported as pathogenic or likely

366 pathogenic in ClinVar and VKGL. Including all curated DCM-associated genes, we report a
367 prevalence of 1:251.

368 For HCM, we found a prevalence ranging between 1:250 and 1:149 individuals
369 carrying a pathogenic or likely pathogenic variant, which approaches previous estimates of
370 1:164¹⁹. In a recent study, including the UKB population, a prevalence of 1:407 was
371 reported²⁰. They included 8 sarcomere-encoding genes described to be associated with
372 HCM (*ACTC1*, *MYBPC3*, *MYH7*, *MYL2*, *MYL3*, *TNNI3*, *TNNT2* and *TPM1*) and variants that
373 were described as pathogenic or likely pathogenic in ClinVar or annotated as pathogenic or
374 likely pathogenic according to ACMG criteria and filtered variants for an allele frequency of
375 0.00004. We included additional genes (*CSRP3*, *JPH2* and *TNNC1*) and pathogenic and likely
376 pathogenic variants from the VKGL database and filtered for a minor allele frequency of
377 0.001. Especially the latter is a driving force behind the higher prevalence in this study.
378 When also using a minor allele frequency (MAF) <0.00004, the prevalence of our study
379 would be 1:475, approaching the prevalence reported by de Marvao *et al.*²⁰.

380

381 *Disease expression of pathogenic and likely pathogenic variants in the general population*

382 Most information on disease penetrance in ARVC, DCM or HCM G+ is based on observations
383 in G+ relatives of cardiomyopathy patients. Previous studies have shown that 37% of ARVC
384 G+ relatives²¹ and up to 50% of HCM G+ relatives with sarcomeric variants²² show disease
385 expression during follow-up. Our findings suggest that disease penetrance in the general
386 population is much lower. We found that 1.2% of ARVC G+, 3.1% of DCM G+ and 2.6% of
387 HCM G+ in the UKB were diagnosed with a cardiomyopathy or heart failure, in the absence
388 of chronic ischemic heart disease. Our additional analysis of ventricular function and ECG in
389 undiagnosed G+ subjects also suggests a low disease penetrance. We found significantly

390 worse LVEF and strain parameters in DCM G+P- compared to controls, however none met
391 the diagnostic Henry criteria (LVEF below 45% and LVEDVi two times the normal SD)¹¹.
392 Although CMR data was only available in a subgroup of undiagnosed G+ patients, these
393 findings make it unlikely that the low penetrance found in our study arises from missed
394 diagnoses or covert disease in the G+ cohort. Furthermore, none of the G+P- individuals met
395 HCM criteria¹² of ≥ 15 mm wall thickness, two individuals did meet the criteria for limited
396 hypertrophy (13-15mm) in the presence of a positive genetic test¹². Interestingly, maximum
397 wall thickness in de Marvao *et al.*²⁰ was higher compared to ours. Although this can partly
398 be explained by the inclusion of P+ by Marvao *et al.*, this may also be explained by
399 differences in wall thickness calculation method. While Marvao *et al.* uses the absolute
400 largest wall thickness value at a single point, we have used the AHA segment with the
401 largest wall thickness (which is an average of all the single points within one AHA segment
402 to reduce random outliers). In ARVC and DCM G+P- we found a low, but significantly higher
403 prevalence of ventricular arrhythmias compared to controls (1.7% vs. 0.3% (OR 5.85 [95% CI
404 1.98;14.40]) and 1.0% vs. 0.3% (OR 3.43 [95% CI 1.35;7.68]) respectively). In ARVC, electrical
405 abnormalities are known to precede structural abnormalities²³. Therefore, these findings
406 may suggest early disease penetrance in a small subset of undiagnosed G+ individuals. The
407 discrepancy between the high disease penetrance found in G+ family members and the low
408 penetrance in the G+ general population points towards the interaction of possible other
409 (unidentified) genetic and environmental factors leading to this variation. The median age of
410 our study population was 57 [49-63] years, however inherited cardiomyopathies are
411 generally diagnosed at a younger age. For ARVC, Groeneweg *et al.* showed, in a cohort of
412 439 index-patients, that the mean age of first disease presentation is 36 ± 14 years. Most of
413 these patients presented with symptoms (95%), of whom 11% with sudden cardiac arrest²¹.

414 Likewise in DCM, the mean age of presentation is mostly between 30-50 years²⁴. Lastly, in
415 HCM a mean age at presentation of 49±16 years was shown in a cohort of 4893 patients by
416 Lorenzini *et al.*²⁵. Interestingly, although mortality rates were low, young HCM patients
417 showed a worse prognosis compared to their healthy peers, with 80% of mortality being
418 caused by sudden cardiac death²⁵. Therefore, it should be taken into account that younger
419 patients with disease expression are likely underrepresented in our study. This is not only
420 due to higher mortality and morbidity in especially ARVC and HCM, but also because
421 individuals with a diagnosed cardiomyopathy may be less likely to participate in a large-scale
422 biobank study such as the UKB.

423 Interestingly, the South Asian *MYBPC3* and the *TNNT2* variant, showed a relatively
424 high prevalence in our cohort. In total, 19% of HCM G+ was Asian and most of these
425 individuals carried the c.3628-41_3628-17del variant in the *MYBPC3* gene. Although this
426 variant is indicated as likely pathogenic in ClinVar, a previous study suggests that this variant
427 may be reclassified as benign²⁶. In our study, none of these variant carriers were diagnosed
428 with HCM. Four were diagnosed with heart failure of whom one was diagnosed with DCM.
429 This suggests that this variant is associated with heart failure in the setting of multiple forms
430 of cardiomyopathy, and not simply HCM²⁶. Secondly, the c.862C>T p.Arg288Cys variant in
431 *TNNT2* was previously found in HCM individuals but is often observed in patients with a mild
432 phenotype or in combination with other variants. These observations suggest that this
433 variant might not be a monogenic cause of severe HCM but acts in concert with other
434 variants²⁷. Interestingly, when excluding these variants from our G+P- population, a
435 significantly higher wall thickness is measured compared to controls. These two examples
436 emphasize that when pathogenic or likely pathogenic variants are identified as a secondary

437 finding, other factors, such as the specific variant and the family history are crucial for
438 follow-up decisions.

439 We also assessed gene-specific associations with the cardiovascular outcomes. *PKP2*
440 variant carriers showed a stronger association with ventricular arrhythmias (OR 11.90 [95%
441 CI 4.38; 27.86], $p = 6.4 \times 10^{-6}$) compared to heart failure (OR 1.50 [95% CI 0.48; 3.64],
442 $p=0.395$). This is in concordance with a previous study showing sustained ventricular
443 arrhythmias to be the first clinical presentation in 61% of ARVC patients²¹. During follow-up,
444 sustained arrhythmias occurred in 72% of ARVC patients, highlighting sustained arrhythmias
445 as the most important ARVC disease manifestation. On the contrary, symptomatic heart
446 failure was seen in 13% of ARVC patients²¹. In DCM G+, ventricular arrhythmias were
447 significantly more present compared to G- controls, especially in *TTN* (OR 4.49 [95% CI 1.15;
448 12.76], $p=0.016$), *DES* (OR 12.80 [95% CI 1.45; 52.55], $p=0.013$) and *LMNA* (OR 15.04 [95% CI
449 1.69; 62.32], $p=0.009$) variant carriers. A recent meta-analysis assessing predictors for
450 sustained ventricular arrhythmias, showed *PLN* and *LMNA* to be associated with
451 arrhythmogenic outcome²⁸. Although we did not have enough power to study *PLN* G+,
452 *LMNA* G+ did show significantly more ventricular arrhythmias compared to G- controls.
453 Furthermore, *BAG3* variant carriers have been associated with significant risk of progressive
454 heart failure²⁹. In our study, *BAG3* variant carriers were significantly more often diagnosed
455 with a cardiomyopathy (OR 41.18 [95% CI 4.36; 192.17], $p=0.002$). Even though more heart
456 failure cases were seen compared to G- controls, this did not reach statistical significance
457 **(Table X)**.

458 Interestingly, self-reported health-related quality of life and psychological well-being
459 of 89 asymptomatic HCM G+ were previously evaluated in a Dutch cohort and found to be
460 at least similar to the general population, which suggests that reporting incidental findings

461 will not harm psychological well-being of G+³⁰. However, frequent cardiological examination
462 of G+ and family members turning out to be carriers after cascade screening will put a
463 burden on health care and societal costs³¹. Genetic screening and cardiological examination
464 are necessary in family members of genetic cardiomyopathy patients since disease
465 expression in family members is considerable. Disease expression in the general population
466 on the other hand is low. Therefore, in case of an incidental finding, multiple factors like
467 family history, presence of symptoms, electrical and/or structural abnormalities and gene
468 and variant type should inform follow-up decisions. Further studies on the genotype-
469 phenotype associations and disease penetrance will aid in facilitating these decisions.

470

471 *Limitations*

472 Several variants are associated with more than one cardiomyopathy. This is mainly due to
473 phenotypic heterogeneity but may also be partly explained by misdiagnosis. Information is
474 submitted to ClinVar by laboratories, not by clinicians. Phenotype description might
475 therefore be less reliable. To avoid selection bias, we included variants associated with
476 multiple cardiomyopathies in both cardiomyopathy categories, possibly leading to increased
477 prevalence estimates. Although the prevalence of cardiomyopathy variants is slightly
478 affected by including or excluding overlapping variants, this did not substantially affect the
479 results and conclusions (**Table X**). Future studies should focus on reaching consensus on
480 variant-phenotype associations for the variants described in multiple cardiomyopathies to
481 avoid variation in prevalence caused by the use of different definitions. Despite recent
482 efforts to harmonize knowledge on genes associated with inherited cardiomyopathies⁵⁻⁷,
483 and guidelines for variant classification³¹, the adjudication of the clinical significance of
484 single variants can still differ between diagnostic laboratories³¹, which has led to

485 interpretation differences and difficulties to compare results among studies using different
486 criteria. This highlights the importance of a single set of criteria to ascertain clinical
487 significance of a single variant. Furthermore, not all pathogenic or likely pathogenic variants
488 are reported in these databases, especially family-specific variants and pathogenic variants
489 in non-Caucasian populations are underreported.

490 Lastly, G+P- and G-P- individuals with CMR data available were age, sex and ethnicity
491 matched and comparable in the presence of cardiovascular risk factors and diseases.
492 Interestingly, outliers in CMR values were also present in G-P- controls, which could be
493 partly explained by the presence of past myocardial infarctions. Therefore, differences in
494 cardiac function and structure between G+P+ and G-P- could be underestimated.

495

496 **Conclusion**

497 In a cohort of 200,643 individuals with WES and phenotype data we identified a prevalence
498 of pathogenic variants associated with ARVC, DCM and HCM of 1:578, 1:251 and 1:149
499 respectively. Among the identified G+ individuals, cardiomyopathy, heart failure and
500 ventricular arrhythmias were more common compared to G-. However, overall disease
501 penetrance was low (1.2-3.1%). Therefore, in case of incidental findings, decisions on
502 application of cascade screening and frequency of cardiological examination should be
503 based on multiple factors besides variant and gene type, such as family history and disease
504 expression.

505

506 **Acknowledgements**

507 This research has been conducted using the UK Biobank Resource under Application
508 Number 24711.

509 **Funding sources**

510 The work was financially supported by the Netherlands Cardiovascular Research Initiative,
511 an initiative supported by the Dutch Heart Foundation (CardioVasculair Onderzoek
512 Nederland (CVON) projects: DOUBLE-DOSE 2020B005 (AB), PREDICT2 2018-30, eDETECT
513 2015-12 (PvT, AtR and FA) and PREDICT Young Talent Program (AtR)). In addition, this work
514 was supported by the Dutch Heart Foundation (2015T058 (AtR), 2015T041 (AB) and
515 2019T045 (MvV and JvS)). Furthermore, MB is supported by the Alexandre Suerman Stipend
516 of the UMC Utrecht (2017), AtR by the UMC Utrecht Fellowship Clinical Research Talent and
517 FA by the UCL Hospitals NIHR Biomedical Research Center.

518

519 **Disclosures**

520 None

521

522 **Supplemental Material**

523 Supplemental Methods

524 Tables I-XI

525 Figures I-IV

526 References³²⁻³⁸

527

528 **References**

- 529 1. McKenna WJ, Judge DP. Epidemiology of the inherited cardiomyopathies. *Nat Rev*
530 *Cardiol.* 2021;18:22-36.
- 531 2. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a
532 diverse genetic architecture. *Nat Rev Cardiol.* 2013;10:531-547.
- 533 3. Gersh BJ, Maron BJ, Bonow RO, Dearani JA, Fifer MA, Link MS, Naidu SS, Nishimura
534 RA, Ommen SR, Rakowski H, et al. 2011 ACCF/AHA guideline for the diagnosis and
535 treatment of hypertrophic cardiomyopathy: executive summary: a report of the
536 American College of Cardiology Foundation/American Heart Association Task Force
537 on Practice Guidelines. *Circulation.* 2011;124:2761-2796.
- 538 4. Miller DT, Lee K, Chung WK, Gordon AS, Herman GE, Klein TE, Stewart DR, Amendola
539 LM, Adelman K, Bale SJ, et al. ACMG SF v3. 0 list for reporting of secondary findings
540 in clinical exome and genome sequencing: a policy statement of the American
541 College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2021;23:1381-1390.
- 542 5. James CA, Jongbloed JD, Hershberger RE, Morales A, Judge DP, Syrris P, Pilichou K,
543 Domingo AM, Murray B, Cadrin-Tourigny J, et al. International evidence based
544 reappraisal of genes associated with arrhythmogenic right ventricular
545 cardiomyopathy using the clinical genome resource framework. *Circ Genom Precis*
546 *Med.* 2021;14:e003273.
- 547 6. Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, Celeghin R, Edwards M,
548 Fan J, Ingles J, et al. Evidence-based assessment of genes in dilated cardiomyopathy.
549 *Circ.* 2021;144:7-19.
- 550 7. Ingles J, Goldstein J, Thaxton C, Caleshu C, Corty EW, Crowley SB, Dougherty K,
551 Harrison SM, McGlaughon J, Milko LV, et al. Evaluating the clinical validity of
552 hypertrophic cardiomyopathy genes. *Circ Genom Precis Med.* 2019;12:e002460.
- 553 8. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J,
554 Hoffman D, Jang W, et al. ClinVar: improving access to variant interpretations and
555 supporting evidence. *Nucleic Acids Res.* 2018;46:D1062-D1067.
- 556 9. Szustakowski JD, Balasubramanian S, Kvikstad E, Khalid S, Bronson PG, Sasson A,
557 Wong E, Liu D, Wade Davis J, Haefliger C, et al. Advancing human genetics research
558 and drug discovery through exome sequencing of the UK Biobank. *Nat Genet.*
559 2021;53:942-948.
- 560 10. Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, Calkins H, Corrado
561 D, Cox MG, Daubert JP, et al. Diagnosis of arrhythmogenic right ventricular
562 cardiomyopathy/dysplasia: proposed modification of the task force criteria.
563 *Circulation.* 2010;121:1533-1541.
- 564 11. Henry WL, Gardin J, Ware J. Echocardiographic measurements in normal subjects
565 from infancy to old age. *Circulation.* 1980;62:1054-1061.
- 566 12. Ommen SR, Mital S, Burke MA, Day SM, Deswal A, Elliott P, Evanovich LL, Hung J,
567 Joglar JA, Kantor P, et al. 2020 AHA/ACC guideline for the diagnosis and treatment of
568 patients with hypertrophic cardiomyopathy: executive summary: a report of the
569 American College of Cardiology/American Heart Association Joint Committee on
570 Clinical Practice Guidelines. *J Am Coll Cardiol.* 2020;76:3022-3055.
- 571 13. Carruth ED, Young W, Beer D, James CA, Calkins H, Jing L, Raghunath S, Hartzel DN,
572 Leader JB, Kirchner HL, et al. Prevalence and electronic health record-based

- 573 phenotype of loss-of-function genetic variants in arrhythmogenic right ventricular
574 cardiomyopathy-associated genes. *Circ Genom Precis Med*. 2019;12:e002579.
- 575 14. Haggerty CM, James CA, Calkins H, Tichnell C, Leader JB, Hartzel DN, Nevius CD,
576 Pendergrass SA, Person TN, Schwartz M, et al. Electronic health record phenotype in
577 subjects with genetic variants associated with arrhythmogenic right ventricular
578 cardiomyopathy: a study of 30,716 subjects with exome sequencing. *Genet Med*.
579 2017;19:1245-1252.
- 580 15. Hall CL, Sutanto H, Dalageorgou C, McKenna WJ, Syrris P, Futema M. Frequency of
581 genetic variants associated with arrhythmogenic right ventricular cardiomyopathy in
582 the genome aggregation database. *Eur J Hum Genet*. 2018;26:1312-1318.
- 583 16. Akinrinade O, Koskenvuo JW, Alastalo T-P. Prevalence of titin truncating variants in
584 general population. *PLoS One*. 2015;10:e0145284.
- 585 17. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, Conner L,
586 DePalma SR, McDonough B, Sparks E, et al. Truncations of titin causing dilated
587 cardiomyopathy. *N Engl J Med*. 2012;366:619-628.
- 588 18. Schafer S, De Marvao A, Adami E, Fiedler LR, Ng B, Khin E, Rackham OJ, Van Heesch
589 S, Pua CJ, Kui M, et al. Titin-truncating variants affect heart function in disease
590 cohorts and the general population. *Nat Genet*. 2017;49:46-53.
- 591 19. Bick AG, Flannick J, Ito K, Cheng S, Vasan RS, Parfenov MG, Herman DS, DePalma SR,
592 Gupta N, Gabriel SB, et al. Burden of rare sarcomere gene variants in the
593 Framingham and Jackson Heart Study cohorts. *Am J Hum Genet*. 2012;91:513-519.
- 594 20. de Marvao A, McGurk KA, Zheng SL, Thanaj M, Bai W, Duan J, Biffi C, Mazzarotto F,
595 Statton B, Dawes TJ, et al. Phenotypic expression and outcomes in individuals with
596 rare genetic variants of hypertrophic cardiomyopathy. *J Am Coll Cardiol*.
597 2021;78:1097-1110.
- 598 21. Groeneweg JA, Bhonsale A, James CA, Te Riele AS, Dooijes D, Tichnell C, Murray B,
599 Wiesfeld AC, Sawant AC, Kassamali B, et al. Clinical presentation, long-term follow-
600 up, and outcomes of 1001 arrhythmogenic right ventricular
601 dysplasia/cardiomyopathy patients and family members. *Circ Cardiovasc Genet*.
602 2015;8:437-446.
- 603 22. van Velzen HG, Schinkel AF, Baart SJ, Oldenburg RA, Frohn-Mulder IM, van
604 Slegtenhorst MA, Michels M. Outcomes of contemporary family screening in
605 hypertrophic cardiomyopathy. *Circ Genom Precis Med*. 2018;11:e001896.
- 606 23. Gomes J, Finlay M, Ahmed AK, Ciaccio EJ, Asimaki A, Saffitz JE, Quarta G, Nobles M,
607 Syrris P, Chaubey S, et al. Electrophysiological abnormalities precede overt structural
608 changes in arrhythmogenic right ventricular cardiomyopathy due to mutations in
609 desmoplakin-A combined murine and human study. *Eur Heart J*. 2012;33:1942-1953.
- 610 24. Hershberger RE, Cowan J, Morales A, Siegfried JD. Progress with genetic
611 cardiomyopathies: screening, counseling, and testing in dilated, hypertrophic, and
612 arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circ Heart Fail*.
613 2009;2:253-261.
- 614 25. Lorenzini M, Anastasiou Z, O'Mahony C, Guttman OP, Gimeno JR, Monserrat L,
615 Anastasakis A, Rapezzi C, Biagini E, Garcia-Pavia P, et al. Mortality among referral
616 patients with hypertrophic cardiomyopathy vs the general European population.
617 *JAMA Cardiol*. 2020;5:73-80.
- 618 26. Harper AR, Bowman M, Hayesmoore JB, Sage H, Salatino S, Blair E, Campbell C,
619 Currie B, Goel A, McGuire K, et al. Reevaluation of the South Asian MYBPC3 Δ 25bp

- 620 Intronic Deletion in Hypertrophic Cardiomyopathy. *Circ Genom Precis Med*.
621 2020;13:e002783.
- 622 27. Ripoll-Vera T, Gámez JM, Govea N, Gómez Y, Núñez J, Socías L, Escandell Á, Rosell J.
623 Clinical and prognostic profiles of cardiomyopathies caused by mutations in the
624 troponin T gene. *Rev Esp Card*. 2016;69:149-158.
- 625 28. Kayvanpour E, Sedaghat-Hamedani F, Amr A, Lai A, Haas J, Holzer DB, Frese KS, Keller
626 A, Jensen K, Katus HA, et al. Genotype-phenotype associations in dilated
627 cardiomyopathy: meta-analysis on more than 8000 individuals. *Clin Res Cardiol*.
628 2017;106:127-139.
- 629 29. Domínguez F, Cuenca S, Bilińska Z, Toro R, Villard E, Barriales-Villa R, Ochoa JP,
630 Asselbergs F, Sammani A, Franaszczyk M, et al. Dilated cardiomyopathy due to BLC2-
631 associated athanogene 3 (BAG3) mutations. *J Am Coll of Cardiol*. 2018;72:2471-2481.
- 632 30. Christiaans I, Van Langen IM, Birnie E, Bonsel GJ, Wilde AA, Smets EM. Quality of life
633 and psychological distress in hypertrophic cardiomyopathy mutation carriers: a
634 cross-sectional cohort study. *Am J M Genet*. 2009;149:602-612.
- 635 31. Fontes Marx M, Ataguba JE, Vries Jd, Wonkam A. Systematic Review of the Economic
636 Evaluation of Returning Incidental Findings in Genomic Research. *Public Health*
637 *Front*. 2021;9:873.
- 638 32. Strande NT, Riggs ER, Buchanan AH, Ceyhan-Birsoy O, DiStefano M, Dwight SS,
639 Goldstein J, Ghosh R, Seifert BA, Sneddon TP, et al. Evaluating the clinical validity of
640 gene-disease associations: an evidence-based framework developed by the clinical
641 genome resource. *Am J Hum Genet*. 2017;100:895-906.
- 642 33. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon
643 E, Spector E, et al. Standards and guidelines for the interpretation of sequence
644 variants: a joint consensus recommendation of the American College of Medical
645 Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*.
646 2015;17:405-423.
- 647 34. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J.
648 Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*.
649 2011;12:745-755.
- 650 35. Petersen SE, Matthews PM, Francis JM, Robson MD, Zemrak F, Boubertakh R, Young
651 AA, Hudson S, Weale P, Garratt S, et al. UK Biobank's cardiovascular magnetic
652 resonance protocol. *J Cardiovasc Magn Reson*. 2015;18:1-7.
- 653 36. Ruijsink B, Puyol-Antón E, Oksuz I, Sinclair M, Bai W, Schnabel JA, Razavi R, King AP.
654 Fully automated, quality-controlled cardiac analysis from CMR: validation and large-
655 scale application to characterize cardiac function. *JACC Cardiovasc Imaging*.
656 2020;13:684-695.
- 657 37. Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, Pennell DJ,
658 Rumberger JA, et al. Standardized myocardial segmentation and nomenclature for
659 tomographic imaging of the heart: a statement for healthcare professionals from the
660 Cardiac Imaging Committee of the Council on Clinical Cardiology of the American
661 Heart Association. *Circulation*. 2002;105:539-542.
- 662 38. R Core Team. R: A language and environment for statistical computing: Vienna,
663 Austria: R Foundation for Statistical Computing; 2020.
- 664

665

666 **Tables**

667 **Table 1: Baseline characteristics of variant carriers and controls**

668

	ARVC G+ n=347	DCM G+ n=800	HCM G+ n=1,346	Controls G- n=9,972
<i>Demographics</i>				
Female (%)	187 (54)	450 (56)	720 (54)	5,436 (55)
Age, years	57 [50- 64]	58 [51-64]	56 [49-63]	57 [49- 63]
Ethnicity (%)				
<i>White</i>	311 (90)	760 (96)	1,001 (75)	8,288 (84)
<i>Asian</i>	10 (3)	8 (1)	251 (19)	1,076 (11)
<i>Black</i>	7 (2)	12 (2)	22 (2)	164 (2)
<i>Other</i>	17 (5)	15 (2)	55 (4)	348 (4)
<i>Cardiovascular risk factors</i>				
BMI, kg/m ²	26 [24-30]	27 [24-30]	27 [24-30]	27 [24-30]
Diabetes (%)	35 (10)	62 (8)	154 (11)*	914 (9)
Hypertension (%)	116 (33)	287 (36)	475 (35)	3,420 (34)
Hypercholesterolemia (%)	86 (25)	211 (26)	369 (27)*	2,416 (24)
Ever Smoked (%)	161 (46)	371 (46)*	543 (40)	4,132 (41)
MET minutes per week, ml/kg/min	2,001 [923-3,551]	1,695 [784-3,536]	1,762 [848-3,426]	1,773 [810-3,453]
Family heart disease (%)	179 (52)*	380 (48)	623 (46)	4,458 (45)
<i>Cardiac disease/outcomes</i>				
Cardiomyopathy (%)	3 (0.9)	22 (3)**	27 (2)**	37 (0.4)
DCM (%)	2 (0.6)	9 (1)**	1 (0.1)	14 (0.1)
HCM (%)	1 (0.3)	7 (1)**	20 (2)**	8 (0.1)
Heart failure (%)	9 (3)	36 (5)**	33 (3)	182 (2)
Ventricular arrhythmias (%)	7 (2)**	13 (2)**	8 (1)	33 (0.3)

Atrial arrhythmias (%)	7 (2)	34 (4)**	32 (2)	191 (2)
Chronic ischemic heart disease (%)	35 (10)	73 (9)	93 (7)	725 (7)
Acute myocardial infarction (%)	15 (4)	27 (3)	36 (3)	298 (3)
Cardiac arrest (%)	0 (0)	6 (1)	5 (0.4)	34 (0.3)
Cardiovascular death (%)	11 (3)	24 (3)*	18 (1)	181 (2)
All cause mortality (%)	19 (6)	56 (7)*	62 (5)	513 (5)
<i>Cardiac symptoms</i>				
Cardiac problem	3 (1)	3 (0.4)	5 (0.4)	41 (0.4)
Angina pectoris	16 (5)	30 (4)	56 (4)*	312 (3)

669

670 Number (percentages) are given or median [IQR]. * p=0.001-0.05 and ** p <0.001 difference

671 compared to the control group. A more extensive overview of the baseline characteristics

672 are given in **Table III**. Individuals with a variant in *PLN*, *DES* and *DSP* variant were included in

673 both ARVC and DCM. Individuals with a variant in *MYH7*, *ACTC*, *JPH*, *TNNC1*, *TNNI3*, *TNNT2*,

674 *TPM* were included in both DCM and HCM.

675 *Abbreviations: ARVC= arrhythmogenic right ventricular cardiomyopathy; DCM= dilated*

676 *cardiomyopathy; HCM= hypertrophic cardiomyopathy; BMI= body mass index; MET=*

677 *metabolic equivalent of task.*

678

679 **Figures**

680 **Figure 1: Included curated genes per cardiomyopathy**

681 The Venn diagram of curated genes included in this study shows the overlap in genes per
682 cardiomyopathy. Unless otherwise indicated, pathogenicity of genes are classified as
683 definitive. If a superscript S or M is given, genes are classified as having a strong or
684 moderate pathogenicity respectively. In the overlapping circles, yellow, black and red colors
685 refer to ARVC, DCM, and HCM respectively. **Table I** gives an overview of the included genes
686 and pathogenicity classification per gene and abbreviation per gene.

687 *Abbreviations: ARVC= arrhythmogenic right ventricular cardiomyopathy; DCM= dilated*
688 *cardiomyopathy; HCM= hypertrophic cardiomyopathy.*

689

690 **Figure 2: Flowchart inclusion of variants**

691 Flowchart depicting the inclusion of pathogenic and likely pathogenic variants associated
692 with arrhythmogenic cardiomyopathy, dilated cardiomyopathy and hypertrophic
693 cardiomyopathy from the ClinVar⁸ and VKGL database.

694 *Abbreviations: ARVC= arrhythmogenic right ventricular cardiomyopathy; DCM= dilated*
695 *cardiomyopathy; HCM= hypertrophic cardiomyopathy; VKGL= Vereniging Klinische*
696 *Genetische Laboratoriumdiagnostiek.*

697

698 **Figure 3: Distribution of genes per cardiomyopathy**

699 Piecharts with the distribution of curated genes for A) arrhythmogenic right ventricular
700 cardiomyopathy (ARVC); B) dilated cardiomyopathy (DCM); C) hypertrophic cardiomyopathy
701 (HCM).

702 *Abbreviations of the different genes are given in Table III*

703 *Abbreviations: G+= pathogenic variant carrier.*

704

705 **Figure 4: Forest plot cardiac outcomes stratified per inherited cardiomyopathy**

706 Odds ratios and 95% confidence interval are given for the associations between cardiac
707 outcomes and ARVC, DCM, or HCM pathogenic variant carriers.

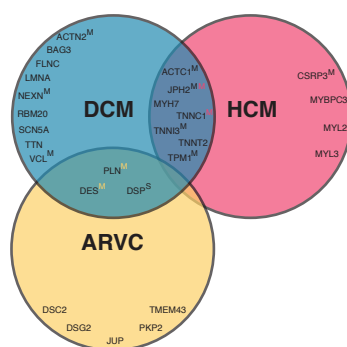
708 *Abbreviations: ARVC= arrhythmogenic right ventricular cardiomyopathy; DCM= dilated*
709 *cardiomyopathy; G+= pathogenic variant carrier; HCM= hypertrophic cardiomyopathy.*

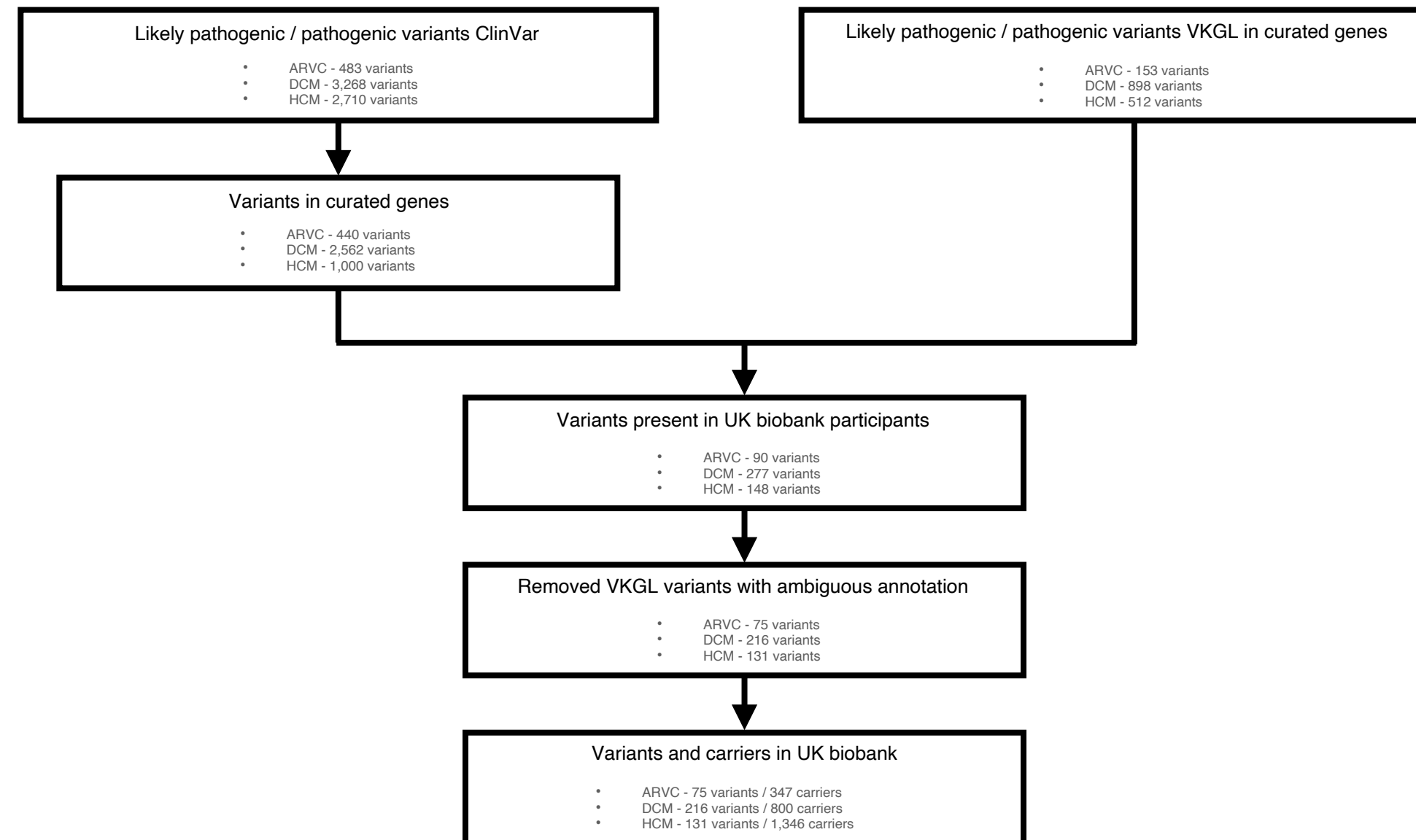
710

711 **Figure 5: CMR parameters stratified per inherited cardiomyopathy**

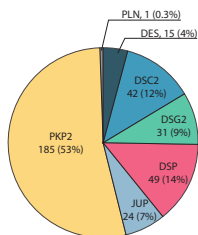
712 Boxplots of CMR parameters showing the summary statistics of CMR parameters stratified
713 by controls and individuals with a pathogenic variant associated with ARVC, DCM, or HCM.
714 Displayed summary statistics include the median, first and third quartile (lower and upper
715 box edges), and the whiskers represent values within 1.5 times the interquartile range from
716 the box edges.

717 *Abbreviations: ARVC= arrhythmogenic right ventricular cardiomyopathy; DCM= dilated*
718 *cardiomyopathy; EDVi= body surface area corrected end-diastolic volume; EF= ejection*
719 *fraction; G+= pathogenic variant carrier; HCM= hypertrophic cardiomyopathy; LV= left*
720 *ventricular; RV= right ventricular.*

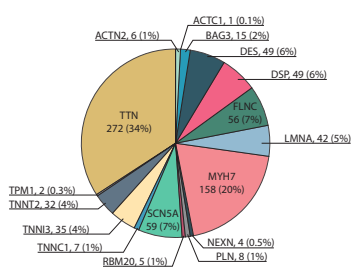




A) ARVC



B) DCM



C) HCM

